International Economic Review

WILEY

ORIGINAL ARTICLE OPEN ACCESS

The Short and the Long of It: Stock-Flow Matching in the US Housing Market

Eric Smith[1] | Zoe Xie[2] | Lei Fang[3]

[1]University of Essex, Colchester, Essex, UK | [2]World Bank, Washington, District of Columbia, USA | [3]Federal Reserve Bank of Atlanta, Atlanta, Georgia, USA

Correspondence: Eric Smith (esmith@essex.ac.uk)

Received: 23 May 2023 | **Revised:** 25 July 2025 | **Accepted:** 2 March 2026

Keywords: housing cycle | stock-flow matching | trading dynamics | duration dependence

ABSTRACT

From 2006 until 2020, the probability of selling a house in the U.S. declined sharply after listing for 2 weeks. Moreover, sales within the first 2 weeks of listing ("quick sales") and sales happening afterward ("slow sales") behaved differently over the housing cycle. The probability and associated price of a quick sale recovered from the slump sooner, faster, and more prominently than a slow sale. This paper demonstrates that a calibrated stock-flow matching model not only generates quantitatively consistent sales, prices, listings, and time on the market but also captures distinctions between fast and slow sales over the housing cycle.

JEL Classification: E30, R21, R31

1 | Introduction

The most important asset for many U.S. households is their primary residence. Because of the large transaction values and market frictions involved, the speed and the price of a house sale have significant consequences for a household's economic well-being. Recessions compound these consequences as the speed with which a property is transferred from one homeowner to another slows down. To shed light on housing market dynamics, this paper investigates the trading patterns from just before the Great Recession until the start of the COVID-19 pandemic. It develops a stock-flow matching model to generate the housing market trading patterns during this period and to focus on duration-dependent selling probabilities and prices.

It is well established that in 2004 and 2005, house prices and sales both surged in the United States. A sharp and dramatic downturn followed—the largest decline in the postwar period by some measures—which triggered the onset of the Great Recession. Sales, new listings, and average price slumped, whereas inventories and time to sale exhibited substantial increases. Recovery began in 2011 and took off in 2012.

The more novel and intriguing findings from this time period involve the duration dependence of house sales. This paper documents clear and important differences between immediate trade and trade after a more prolonged spell of listing. Specifically, at any given point in time during the slump and recovery periods, the probability of a house selling within the first 2 weeks of being listed ("quick sales") is higher than the probability after it has been listed longer.[1] Sale prices exhibit similar duration dependence.[2]

Over time, the duration-specific selling probabilities and sale prices fall and recover with other housing market variables. More importantly, the probability and sale price both recover from the slump sooner, faster, and more prominently for a quick sale than for sales with longer durations. As the housing market recovers, the gap in the selling probability between a quick sale and sales with longer durations widens. The gap in sale prices evolves similarly. Taken together, the "speed of sale" is not constant—it varies with how long the house has been on the market. Moreover, the speeds by duration vary differentially with market conditions. These findings have important implications for households deciding when and how much to list and sell a house and for understanding the evolution of trading dynamics.

The challenge of accounting for the behavior of these housing market variables in a plausible, tractable, and coherent framework is substantial—the housing market is notoriously difficult to model. This paper develops a stock-flow matching model with endogenous seller entry and two aggregate states, prosperity, and slump, to generate housing market dynamics.

In the model, flows of entering traders match and exchange (using first price auctions) with the stock of previously unsuccessful traders on the other side of the market. As in competitive or directed search models, stock-flow matching (see Taylor 1995; Coles and Smith 1998; Coles and Muthoo 1998; Lagos 2000) assumes that buyers and sellers do not search randomly. Instead, market participants have a good idea about where to look for suitable partners. Unlike competitive search, however, buyers and sellers in the stock-flow framework trade in precise and distinct markets with small numbers of traders on each side and without impediments to trade. Traders look for a very specific set of characteristics so that multiple small markets exist in a local geographic region as buyers seek a particular combination of rooms, acreage, schools, amenities, and so on.

As buyers and sellers in the stock-flow model randomly come and go in each particular market, their numbers fluctuate so that traders can be on either the long or the short side of their precise market. An entrant is on the short side when she finds one or more potential partners immediately available. Without impediments to trade, transactions in this case occur quickly. An entrant is on the long side when there are no potential partners immediately at hand. In that case, the entrant becomes a part of the stock of traders and must wait to match with the flow of potential partners entering the market. The distinction between trades on the long and on the short side in stock-flow matching thus produces duration dependence in the probability of a sale, a key feature of the data.

This stock-flow microstructure of dynamic trade in differentiated goods embeds several compelling features associated with the housing market. It also provides a mechanism to stochastically generate time to sale as the long side of the market waits for the arrival of traders on the short side. To generate a cyclical slump and recovery, this paper incorporates a sudden reduction in the entry rate of buyers. Although some sellers avoid entering the market after this shock hits, more motivated sellers nonetheless enter faster than buyers. Inventories thus accumulate as listings exceed buyer entry. As sellers compete harder for scarce buyers, the likelihood of trade for any seller drops along with the price, but especially so for sellers who just entered. The drop is larger for sellers who just enter, as the inflow of new houses quickly mops up preexisting buyers to dramatically curtail immediate trades. When the downturn ends, buyer entry picks up but seller entry does not fully resume until inventory eventually clears. As a result, sellers on both the short and on the long side of the market experience faster trades and higher prices, but not in tandem as they face different trading conditions.

How well does the stock-flow model with aggregate shocks perform? The model is calibrated to the observed housing market statistics in the United States from 2006 to 2019. The calibrated model is consistent with standard housing market variables such as sales, listings, inventories, inventory-to-sales ratios, and days on the market (DoM). It also captures well the duration dependence in the selling probabilities and in sale prices. The simulated model generates similar level, variation, shape, and timing for these variables over the sample period. Consequently, the model results for these variables are highly correlated with the corresponding data series—all the correlations are above 0.76. The more fundamental contribution is that with only two aggregate states, the calibrated stock-flow model generates differential responses to aggregate market conditions in the short- and long-side selling probabilities (and in the short- and long-side prices), as observed in the data. Section 6 demonstrates that this time-varying duration dependence, embedded in the stock-flow model, is crucial in generating the observed time-series variations in inventories. This section also establishes that the model is consistent with the cross-city differences over the sample period.

Compared with the stock-flow model, random and directed search models struggle to qualitatively account for the duration dependence of selling probabilities and of sale prices. For example, random search implies that the steady-state time on the market of sold houses is the same as the time on the market for unsold houses, yet these two measures differ substantially in the data. Sellers in directed search models can lower their asking price to sell more quickly, making prices negatively correlated (not positively as observed) with the probabilities of a sale. In contrast, the core mechanics at the heart of stock-flow matching inherently build in duration dependence. Moreover, the paper demonstrates that the stock-flow matching framework can generate time-varying duration dependence in a tractable way.

The rest of the paper is organized as follows. The next section reviews the literature and provides background for the contributions of this paper. Section 3 describes the data and documents the empirical findings. Section 4 outlines the stock-flow model and discusses its trading outcomes. Section 5 calibrates the model and compares the model results with empirical findings in the data. Section 6 evaluates the quantitative implications of duration dependence, assesses variation across cities, and considers the role of public policy. Section 7 concludes.

2 | Related Literature and Contributions

Beginning with Wheaton (1990), the application of search frictions to housing markets initially addressed three broad empirical regularities that revolved around price fluctuations. First, housing cycles occur: there is a short-run positive serial correlation in prices but mean reversion in the long run (Case and Shiller 1997; Muellbauer and Murphy 1997). Second, there is excess volatility in prices and quantities relative to fundamentals (Shiller 1982; Glaeser et al. 2014). Third, price and sales exhibit a positive correlation, while price and time on the market exhibit a negative correlation (Stein 1995; Krainer 2001; Glaeser et al. 2014).

Recent research documents a wider set of stylized facts about the joint cyclical properties of key housing market variables, namely, sales, new listings, time to sale, stock of houses for sale, and, of course, house prices. Ngai and Sheedy (2024) document that these variables are all highly volatile and highly correlated.[3] Other research documents that nonnegligible residual price dispersion coexists with these regularities.[4]

The initial three and the more recent set of empirical regularities are broadly, but not completely, consistent with search models of housing markets. Han and Strange (2015) observe that while the basic pairwise random matching model can tie booms and busts in the housing market to macroeconomic volatility and thereby generate some of these stylized facts in response to external shocks, it struggles to fully explain both persistence and excess price volatility, even when the model allows for a variety of amplification mechanisms.[5] For example, Krainer (2001) specifies aggregate demand shocks and demonstrates a positive correlation in prices, sales, and speed of sale. House price volatility in Krainer (2001), however, is lower than aggregate volatility—changes in time to sale respond as well to aggregate shocks and absorb some of the market-wide variations.

Novy-Marx (2009) obtains amplification and generates excess volatility through an endogenous response of market tightness to aggregate demand shocks. Increased demand for houses (say from rising incomes) enables sellers to trade faster. Fewer houses are thereby available next period that then leads to a further price rise and amplification of the shock. A limited supply-side response generates a market tightness feedback mechanism that amplifies demand shocks while maintaining comovement in prices, sales, and selling probabilities across steady states. Ngai and Sheedy (2020), Ngai and Sheedy (2024), Moen et al. (2021), and Anenberg and Ringo (2022) emphasize the importance of the moving decision—to simultaneously sell the current house and buy a new one—in propagating and amplifying external shocks.[6] Concentrating on price volatility, Arefeva (2025) demonstrates that both the prevalent dynamic random and directed search models with bargained prices cannot explain price volatility but incorporating auctions substantially increases volatility. The stock-flow model proposed in this paper generates similar price movements as in the data but the volatility is smaller. As discussed in Subsection 4.3, lower volatility occurs because prices in search models divide a match-specific surplus but are largely silent on aggregate factors that affect the buyers' and sellers' payoffs to ownership in the same way.

Díaz and Jerez (2013) add aggregate supply shocks to the Novy–Marx insight and quantify the feedback mechanisms. They find that amplification and propagation are more pronounced in a competitive directed search environment than under random search. Head et al. (2014) likewise calibrate a directed search model, allowing for the turnover of existing homes, to explore the dynamics of house prices, sales, construction, and the entry of buyers in response to city-specific income shocks. Their model, which incorporates Wheaton (1990)'s insight of the joint buyer–seller problem, quantitatively accounts for a large share of house price variation driven by income shocks and approximately half the serial correlation in house price growth. Head et al. (2014) focus on the role of transitional dynamics of prices and construction of new homes in response to shocks.

Hedlund (2016a), Hedlund (2016b), and Garriga and Hedlund (2020) link the macrohousing literature on credit constraints and financial heterogeneity[7] with directed search to assess cyclical housing dynamics and housing stabilization policies. In these models, sellers lower their asking price to attract buyers as the seller's financial position deteriorates. Trading probabilities therefore adjust over time. The quantitative assessment using quarterly data is consistent with the findings of Merlo and Ortalo-Magné (2004) that "listing price reductions are fairly infrequent and when they occur are typically large." But their quarterly analysis does not speak to the high-frequency movements of the selling probabilities and prices emphasized here.

Given this picture of the housing search literature, the first fundamental contribution of this paper is to document the duration dependence of sale probability and of sale price, and how the duration dependence changes over the housing cycle. The second significant contribution establishes that the stock-flow framework can generate not only the variability of the standard housing market variables (sales, listings, inventories, and DoM) but also the duration dependence over the bust and boom between 2006 and 2019. A number of search and matching studies of the housing markets treat the aggregate housing stock as fixed, and/or consider only steady states.[8] Smith (2020a) demonstrates that the stock-flow matching model with endogenous seller entry can generate hot and cold spells in house sales but limits its assessment to a stationary environment without an aggregate housing cycle. The model in this paper incorporates not only endogenous entry but also aggregate shocks to the trading environment.[9]

The data sources used here, unfortunately, do not contain buyer information that has helped assess the performance of stock-flow matching in other contexts. For instance, in labor market studies, the validation of the model has primarily come from gauging the trading patterns—the quantities—on both sides of the market. Labor market information regarding the number of vacant jobs and job seekers drives the movements over time in the hazard functions.[10] Without information on house buyers, this sort of evaluation of hazard functions is not available here. On the other hand, new comparisons are available. Observed sale prices based on high-quality data cover new territory in the quantitative evaluation of stock-flow matching. While wages are not reliably available for high-frequency data flows, observed sale prices are available in the housing data and provide an alternative perspective on assessing the merits of the stock-flow approach. The evidence for prices (broken down by duration) supports stock-flow matching.

3 | Trading Patterns in the U.S. Housing Market

This section documents patterns in the volatility and comovement not only for standard housing market measures—prices, sales, listings, inventories, and time to sale—but also for outcomes broken down by duration—trading probabilities and prices by time on the market—over 2006–2019. The variation, progression, and timing of the data establish three key empirical findings. First, aggregate data patterns of sales, new listings, inventories, DoM, and average price are consistent with a pronounced and dramatic housing slump starting from 2006 and a sustained recovery beginning in 2011 and taking off in 2012. Second, the average duration on the market for sold houses (completed spells) is shorter than the average duration for unsold houses (uncompleted spells) at any given point in time. Third, a sharp distinction exists between sales that occur quickly and those that take longer. In particular, throughout the sample period, the probability of a house selling declines sharply after it has been listed for 2 weeks. Moreover, the probability and price of a quick sale both recover from the slump sooner, faster, and more prominently than sales with durations longer than a month.

3.1 | Data Sources

Confidential records from the data service CoreLogic provide detailed and comprehensive information on individual house transactions that generate the variables of interest and represent the primary source of data in this paper. CoreLogic gathers from regional realtor boards property-level information on listing and sale dates, prices, location, and housing characteristics of properties for sale or for rent across much of the United States from 2006 onward. CoreLogic standardizes these variables to improve consistency across realtor boards. The coverage of these records is rather limited for the initial years. Since then the coverage has expanded and as of 2014, the data capture around 56% of all active listings nationwide. The analysis of CoreLogic data is restricted to 50 metropolitan areas that have a sizeable population and widespread coverage of available properties. The average size of the housing stock for sale in each city over the sample period is used as the weight for the aggregation across cities. Appendix A provides more detailed data descriptions.

Given the limited coverage of CoreLogic early on, data from multiple complementary sources are also presented, including the Case–Shiller price index, the National Association of Realtors (NAR), and the online platform Redfin. The Case–Shiller price index is a well-known repeated-sales house price index for single-family homes. The NAR provides a long-standing monthly time series of the total number of existing houses for sale and the number of existing houses that sold nationwide.[11] It does not, however, provide publicly available data on listings, DoM, and prices. Lastly, the online platform Redfin provides a broader set of aggregate information on the housing market: monthly new listings counts, inventories of houses for sale, house sale counts, the proportion of houses that sell within 2 weeks, and the median DoM for sold homes, but all data are available from only 2012 onward.

3.2 | Variables of Interest

Sales, Listings, Inventories, and DoM. The CoreLogic data generate a monthly panel of listings (flows in), sales (flows out), inventories (unsold houses at the end of each month), DoM of sold houses (average duration of completed spells), and DoM of unsold houses (average duration of uncompleted spells) for single-family homes and townhouses in the most active U.S. metropolitan areas.

Kaplan–Meier Statistics. Kaplan–Meier statistics distinguish house sales by duration on the market or time to sale. In a given period t, the Kaplan–Meier statistic K_t^d measures the probability that a house on the market sells conditional on the DoM d and is therefore the proportion of the successful sales among the qualified houses for sale. In other words, it is the ratio of sold houses at date t with duration d over the sum of sold and unsold houses on the market with the same duration d:[12]

$$K_t^d = \frac{\text{Houses with duration } d\text{, sold in time period } t}{\text{Houses with duration } d\text{, sold and unsold in time period } t}.$$

In the empirical analysis, dates t are months, and durations d are 15-day intervals of DoM. The Kaplan–Meier statistic is computed for each duration interval of up to 120 days, in each sample month, for each metropolitan area. These estimates are then weighted using the housing stock for sale in each city to obtain the averages.

Hedonic Sale Prices. Hedonic sale prices—prices with controls for house characteristics—are constructed and examined in each city by duration or time to sale. Specifically, in each city, the log hedonic price for house i with characteristics X_{it} at date t is given by

$$\ln(P_{it}) = \Omega X_{it} + \sum_d \theta_t^d D_t^d + \zeta_z + \gamma_m + \epsilon_{it}. \tag{1}$$

X_{it} is a set of controls for housing characteristics, including the age of the house (and squared), numbers of bedrooms, baths and fireplaces, living area, and indicators for whether the property has a large lot, has a pool, or is in distress (short-sale, REO, or foreclosure). D_t^d represents the duration category dummy at date t for duration d, where d again goes from 0–15 days to 106–120 days in 15-day intervals. The estimated coefficients for the duration category dummies generate prices broken down by duration conditional on average local housing characteristics. Additionally, the estimation also controls for month-of-year (γ_m) and zip code (ζ_z) fixed effects.

3.3 | Empirical Findings

Comparing Overlapping Sources. The CoreLogic data overlap with the alternative data sources in some areas thereby enabling comparisons of the relative strengths and limitations of the CoreLogic information. Table 1 reports the minimum, maximum, percentage change between minimum and maximum, contemporaneous correlation with price, and standard deviation for the observed housing series. Whenever possible, statistics from both the CoreLogic data and an alternative source are included. To remove contemporaneous seasonal movements, the series in this table are seasonally adjusted using a plus and minus 6-month moving average.[13] Figures 1 and 2 supplement the information in Table 1 by plotting the time series of these variables. Note that the CoreLogic data reported here are weighted averages of the 50 metropolitan areas and hence correspond to a representative metropolitan area.

The figures and table collectively demonstrate that sales, new listings, inventory, DoM, and average price from various sources (where they overlap) are largely consistent with each other in their evolution over time, persistence,[14] and the correlations with prices. One discrepancy across data sources is worth noting. The limited coverage of the CoreLogic data in the initial years yields lower housing counts. As a result, data patterns before 2009

TABLE 1 | Descriptive statistics

Variable	Source	Min	Max	% change max-min	Corr with price	Std Dev (log)
Sales	CoreLogic	773	1277	49.2	0.202	0.142
	NAR	334	513	42.2	0.661	0.109
	CoreLogic (2009-)	912	1277	33.4	0.921	0.094
	NAR (2009-)	338	470	32.6	0.787	0.102
Listings	CoreLogic	772	1260	48.0	0.235	0.134
	NAR	285	536	61.1	0.666	0.150
	CoreLogic (2009-)	938	1260	29.4	0.952	0.093
	NAR (2009-)	285	465	47.9	0.771	0.142
Inventory	CoreLogic	1972	2911	38.5	−0.676	0.115
	NAR	1628	3736	78.6	−0.031	0.269
	CoreLogic (2009-)	1972	2911	38.5	−0.901	0.117
	NAR (2009-)	1628	3359	69.4	−0.665	0.220
Inventory/Sales	CoreLogic	1.46	3.27	76.7	−0.405	0.280
	NAR	3.71	10.83	97.8	−0.225	0.352
	CoreLogic (2009-)	1.46	3.23	75.7	−0.915	0.248
	NAR (2009-)	3.71	10.18	93.1	−0.670	0.320
DoM of Sales	CoreLogic	52.2	101.0	63.7	−0.541	0.242
	Redfin (2012-)	39.8	84.8	72.1	−0.962	0.199
DoM of Inventory	CoreLogic	78.8	133.8	51.7	−0.810	0.171
0–15 day Kaplan–Meier	CoreLogic	0.147	0.380	88.7	0.326	0.309
	Redfin (2012-)	0.305	0.398	26.7	0.890	0.087
31–45 day Kaplan–Meier	CoreLogic	0.138	0.250	57.8	0.407	0.201
Average Price	CoreLogic ($1000)	141	280	66.2	—	0.175
	Case-Shiller (Real Index)	111	169	41.1	—	0.106

Note: (1) All data are for January 2007–December 2019 unless otherwise noted. The full sample CoreLogic data are over 2006–2019 for sales and new listings and over 2007–2019 for the other variables. Redfin data are from 2012 onward and exhibit less variation than over the full cycle. NAR data (min and max) are reported in thousands. (2) All data variables are in monthly frequency. The Case–Shiller index is seasonally adjusted at source (FRED). The CoreLogic hedonic price regression includes month-of-year dummies. Other monthly variables are seasonally adjusted by taking a ±6-month moving average. (3) Inventory is the value at the end of each month. The NAR sales are for existing home sales. The NAR listings are computed from total houses for sale at the start of the month less houses sold in the month and less unsold houses at the end of the month. (4) Average prices from CoreLogic are hedonic regression dollar values. The Case–Shiller price index is the real national index, constructed using smoothed and seasonally adjusted nominal index and CPI data from FRED, indexed at January 2005 = 100. (5) CoreLogic prices are conditional statistics derived from hedonic regressions. Correlation with price uses the conditional price series for variables from CoreLogic and the Case–Shiller real index series for variables from non-CoreLogic sources. (6) CoreLogic Kaplan–Meier are unconditional statistics. The conditional Kaplan–Merier statistics are very similar to unconditional ones. Figure A.4 and Table A.1 compare the two. (7) % change max-min uses the average of max and min as the denominator. (8) The coefficient of variation has similar values as the standard deviations of logged variables.

diverge between the CoreLogic and the NAR. As seen in Panel (A) of Figure 1, CoreLogic sales and listings gradually increase from 2007 until 2009, whereas the NAR counts fall consistently from early 2006 until 2009. After 2009, although sales and listings are still lower in the CoreLogic, they exhibit comparable month-over-month changes to the NAR data as shown in Figure 1. Table 1 therefore includes sales and listings statistics from the NAR and CoreLogic for both the full sample period and a subsample from 2009 onward.[15] The table shows that for the post-2009, subsample statistics for the sales and listings are comparable between the CoreLogic and the NAR, confirming patterns in Panels (A)–(C) in Figure 1.

Aggregate Patterns. Table 1 and the two companion figures reveal the variability and comovement in the housing market over the sample period. As measured by the percent changes between maximum and minimum values over this period, sales, listings, inventory, DoM, the Kaplan–Meier statistics (the selling probability), and average price all vary more from peak to trough than GDP per capita, which changed by 17.6% from peak to trough during this sample period. Moreover, as expected, sales, listings, and the Kaplan–Meier statistics positively comove with prices, while inventory, the inventory-to-sales ratio, and DoM negatively comove with prices. As measured by the standard deviation of the logged series, inventory-to-sales

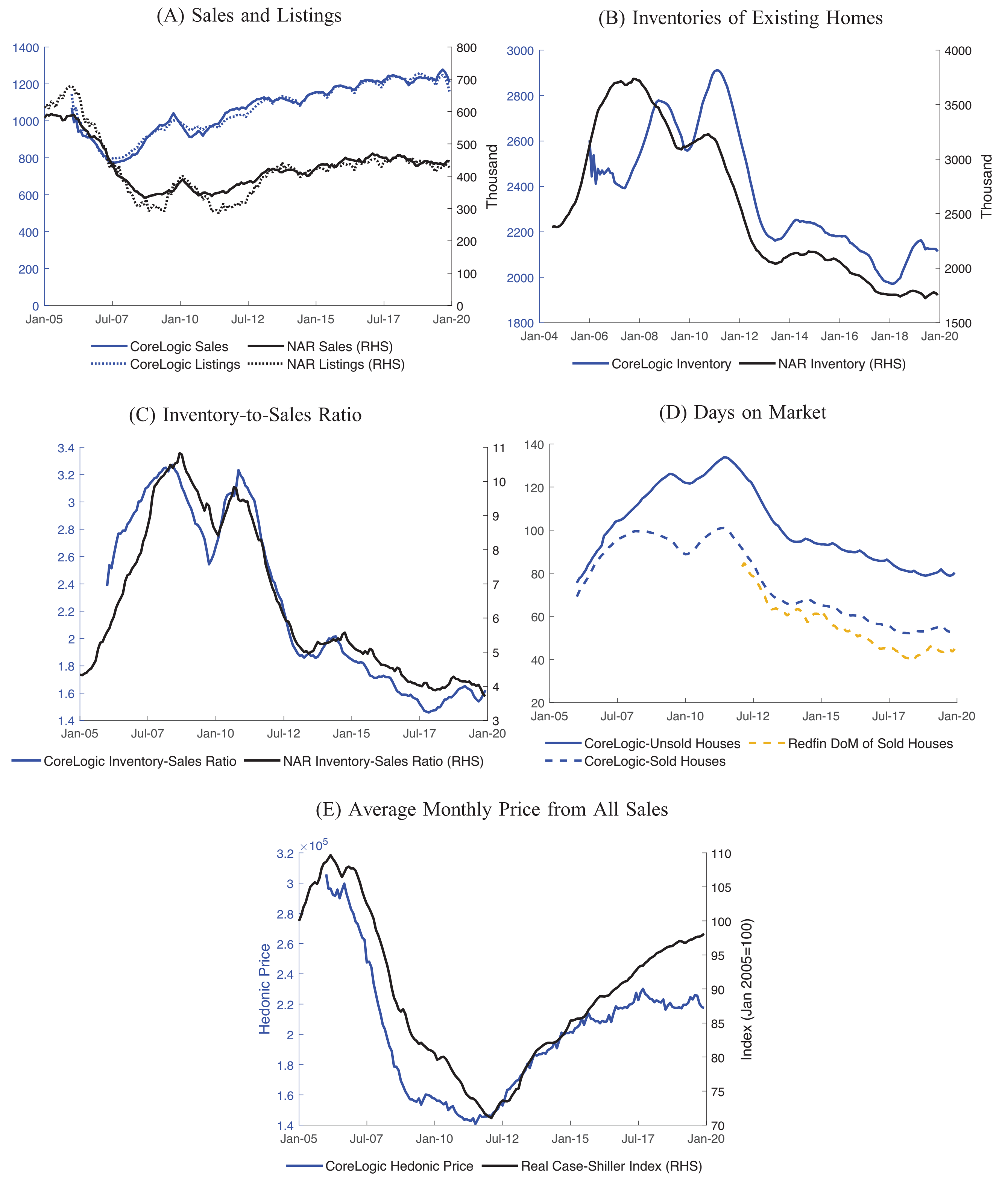

FIGURE 1 | Aggregate trading patterns. *Notes:* All series labeled "CoreLogic" are the authors' calculations based on CoreLogic data.

ratio, DoM of sales, and the CoreLogic Kaplan–Meier statistics are more volatile than the other series. Moreover, all these housing market variables are more volatile than GDP per capita, the standard deviation of which is 0.05.

Figure 1 reveals that sales, new listings, inventories, the inventory-to-sales ratio, DoM, and average price all exhibit a pronounced and dramatic housing slump starting from 2006 and a sustained recovery taking off in 2012. In the figure, the monthly sales and

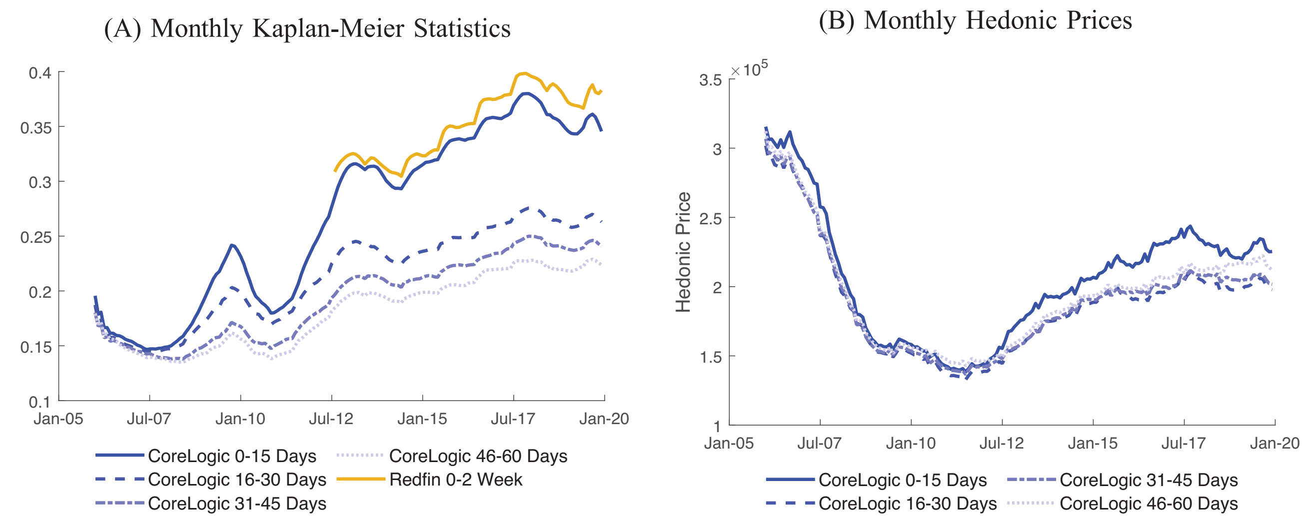

FIGURE 2 | Kaplan–Meier statistics and hedonic prices by duration. *Notes:* All series labeled "CoreLogic" are the authors' calculations based on CoreLogic data.

listings of existing houses peak in 2006 and decline dramatically from 2006 and throughout 2007. From 2008 onward, the NAR sales and listings are essentially flat (with heightened variation in 2009–2010) until a sustained recovery begins in late 2011. The subsequent recovery from 2012 forward in sales and listings occurs within 2 years, but the NAR series never fully recover to the preslump high levels of sales and listings.

Panel (A) of Figure 1 further displays a minor rise and fall of sales and listings from 2009 until midway through 2010, a pattern more pronounced in the NAR series. This temporary subcycle or wobble (from the progression within the downturn) coincides with the onset and expiration of the components of the American Recovery and Reinvestment Act (ARRA) of 2009 (Indiviglio 2011) aimed at stimulating the housing market. Similar patterns are visible in inventories, prices, and (looking ahead) the Kaplan–Meier hazards. This temporary change is small and short-lived.

End-of-month inventories plotted in Panel (B) of Figure 1 accumulate (decumulate) if listings are greater (less) than sales. The NAR listings exceed sales until 2008, whereas CoreLogic listings exceed sales until mid-2011. As a result, the rise and fall of the two inventories series exhibit different timings over the cycle. The NAR series peaks nearly two and a half years before the CoreLogic series, but its more prominent decline after 2011 coincides with the timing of the CoreLogic decline. Despite the uneven relationship in inventory levels, the CoreLogic and NAR inventory-to-sales ratios in Panel (C) mirror each other throughout the sample period.

Panel (D) in Figure 1 plots DoM for sold and unsold houses. Redfin's seasonally adjusted series for sold houses from 2012 onward aligns closely with the CoreLogic counterpart.[16] Four observations emerge. First, the DoM for unsold homes steadily rises until 2012 and subsequently declines. Second, overlooking the initial periods as inventories accumulate, DoM at the time of sale is relatively flat (declining and rising marginally with the 2009–2010 Recovery Act policies) until the sustained recovery begins in 2011. Third, at the outset of 2007—in the midst of the slowdown—the average time on the market for sold and unsold houses is nearly equal to each other, an outcome consistent with a constant (with respect to durations at that point in time) hazard function and often associated with random matching. Lastly, as time progresses from 2007, the DoM for sold houses diverges from that for unsold houses—predicting that the hazard function (at a point in time) becomes downward sloping with respect to durations as found below in the Kaplan–Meier measures.

Panel (E) in Figure 1 plots the average monthly price from all sales.[17] The average CoreLogic hedonic price is largely consistent with the shape and timing of the well-known Case–Shiller housing price index. Both series fall until late 2011 and then recover. These prices thus comove with the sales and listings in Panel (A) and also exhibit a wobble during the Recovery Act period.

Patterns by Duration. Panel (A) in Figure 2 plots the Kaplan–Meier statistics for the first four 15-day intervals along with the short Redfin series for the proportion of houses that sell within 2 weeks. The Redfin series echos the 0–15 day CoreLogic Kaplan–Meier hazard rates. Analyzing the progressions of the CoreLogic series over the sample period by duration reveals several important findings. First, for all durations, the Kaplan–Meier hazard rates fall between 2006 and 2007 and then recover. Second, in a given month, the Kaplan–Meier statistics decline by duration and the largest decline is observed soon after a house is listed on the market. As reflected in the two DoM measures for sold and unsold houses, the (within-the-month) hazard function drops distinctly after the initial interval, becoming more step-like as the 0–15 day hazard rate rises well above the others when the recovery takes hold from 2012 onward.

Third, the recoveries of the hazard rates after reaching the bottom exhibit two distinct patterns. Specifically, for durations less than 30 days, the recovery is much sooner and stronger than for longer durations. The series for the two shortest durations (0–15 days and

16–30 days) begin rising after an initial, relatively sharp decline from as early as 2007, with the most pronounced improvement appearing in the 0–15 day hazard. In contrast, the series for longer durations are flat from 2008 until roughly 2012, at which point a sustained gradual rise occurs. These patterns imply that the probabilities of a sale conditional on duration do not evolve evenly over the housing cycle. The probability of selling quickly is more volatile and responsive over the housing cycle, compared to the probability of selling after a month. Moreover, the 0–15 day duration Kaplan–Meier displays the largest apparent reaction to the Recovery Act.

Panel (B) of Figure 2 plots the evolution of the conditional prices by DoM constructed using the CoreLogic data. Closely following the evolution of average price over time, the CoreLogic hedonic price estimates by duration also fall until late 2011 and then recover while wobbling during the Recovery Act period. As such, these prices also comove with the sales and listings in Figure 1. Echoing the evolutions of the Kaplan–Meier series, prices over different durations comove over time with the price at the shortest end (0–15 days) recovering faster and stronger. For the three durations greater than 15 days, prices synchronously rise and fall over the cycle. In contrast, the 0–15 day price starts higher than these three prices but this premium falls as the market slumps. After reaching its lowest level, the 0–15 day price then rises sooner and faster, restoring the premium over longer durations as time progresses.

4 | A Stock-Flow Matching Model of the Housing Market

This section develops a stock-flow matching model to frame and assess the empirical findings in the U.S. housing market. The model builds on Smith (2020a), which develops and characterizes a stationary stock-flow housing market model with endogenous seller entry. The model here generalizes the matching framework in Smith (2020a) to an environment with both market prosperity and slump. Incorporating two aggregate states with endogenous seller entry allows the model to generate housing cycle dynamics.

4.1 | Environment

Aggregate States. At any point in time, the economy is in one of two different aggregate states—prosperity or slump—distinguished by high and low buyer entry, respectively.

Buyers and Sellers. The economy consists of many small, independent, and isolated housing markets. Buyers and sellers trade in their particular market (hereafter their submarket) in continuous time. Buyers and sellers are risk neutral and discount the future at a rate $r > 0$. A buyer derives x units of discounted total lifetime utility from home ownership. A seller incurs flow search costs $d_s\,dt$ over a period dt in the market. For simplicity, buyers do not pay flow search costs.

Buyers looking for a house enter their submarket at a state-contingent exogenous Poisson rate: let β and $\hat{\beta}$ denote buyer arrival rate during prosperous times and in a slump, respectively, where $\hat{\beta} < \beta$. Following Albrecht et al. (2007), there are two potential types of sellers: motivated and relaxed sellers each with one home for sale. In both states, motivated sellers arrive in their submarket—and have the opportunity to offer their house for sale—at Poisson arrival rate α, whereas relaxed sellers arrive at Poisson rate σ. The relaxed or motivated distinction is related to the willingness to sell. Motivated sellers are more eager to sell, which may be a result of job relocation or changes in a seller's financial situation. In contrast, relaxed sellers do not have an imminent need to sell and thus can afford to wait for the market conditions to become more favorable. For instance, households who move up the housing ladder over the life cycle are more likely to be relaxed sellers.

Upon arrival, motivated and relaxed sellers decide whether or not to sell their houses. They choose to sell if the state-contingent expected payoff exceeds their respective outside option. Relaxed sellers who choose not to sell receive a payoff normalized to zero. Motivated sellers have a lower nonparticipation payoff, $V < 0$. Sellers who decide to sell pay an up-front fixed cost $F > 0$ that is less than x.

During prosperous times, the arrival rate of motivated sellers is less than the arrival rate of buyers ($\alpha < \beta$), but the combined arrival rate of both seller types is greater than the arrival rate of buyers ($\alpha + \sigma > \beta$).[18] As potential sellers receive the opportunity to participate in the submarket more frequently than buyers, the sellers' decision to sell will balance the number of traders in the submarket over time. In a slump, the arrival rate of buyers falls below the arrival rate of motivated sellers ($\alpha > \hat{\beta}$), and so an inventory of houses for sale will accumulate even if only motivated sellers decide to sell.

If a trade takes place, the buyer and seller involved in the trade both permanently leave the submarket. An accepted bid at price P yields a payoff $x - P$ to the buyer and a payoff P to the seller. Unsatisfied traders—buyers and sellers who did not trade—remain behind to wait for the next trading opportunity. Once a seller decides to sell, she does not leave the submarket until the house is sold. In this sense, sellers cannot reoptimize.

Auctions and Bidding. A seller choosing whether or not to sell knows the outcome of the previous transaction, including the number of buyers and the date of that transaction. A potential new seller also knows whether there are any existing homes for sale, that is, if there are any unsatisfied prior sellers who entered the submarket but did not trade. This seller, however, does not know the outcome of buyer entry over the period since the last transaction occurred.

Upon entry, sellers of either type immediately hold a full-information, first-price auction for their houses regardless of the submarket conditions. The calling of an auction reveals to all agents the total number of buyers before they submit their offers. All buyers in the submarket are obliged to bid. Sellers either accept a single bid or reject them all in which case they hold another auction straightaway. If there are no buyers present to bid, the seller waits for the arrival of other traders. All sellers are committed to the submarket and unable to withdraw. As agents in the submarket become perfectly informed about existing trading conditions, there are no impediments to trade after entry.

Aggregate Shocks. To align with the observed data series documented in Section 3, the trading environment begins in a prosperous and thriving state, almost immediately suffers a slump for a period of time, and then recovers. In particular, suppose after a run of time in the prosperous state that has settled the submarkets, the economy suffers an unanticipated slump that reduces the entry of buyers. The arrival rate of house buyers falls below the arrival rate of motivated sellers but the existing stocks of buyers and sellers are unaltered. During the slump, a permanent recovery to the prosperous state occurs with Poisson arrival rate τ.

Although buyers and sellers do not anticipate the initial shock and downturn, they recognize that the slump is temporary and account for the likelihood of recovery to prosperous trading in their expected payoffs and hence trading decisions during the slump. As a result, the slump and the recovery are derived from rational, forward-looking buyers and sellers.

4.2 | Equilibrium

This subsection summarizes the trading outcomes. It first establishes that trading in the stock-flow model occurs immediately. Then, it discusses the entry decision of sellers and briefly describes the buyers' and sellers' payoff functions. The details of the derivations are included in Appendix B.

Immediate Trade. The first salient feature of the stock-flow framework is that agents realize any gains to trade in the submarket without delay. Immediate trade is a natural outcome in this environment (see also Taylor 1995; Coles and Muthoo 1998). To see it, suppose that up to some point in time, all past entrants have exhausted all possible trading opportunities so that there are either excess buyers or excess sellers but not both. Now suppose seller entry occurs when there is more than one excess buyer. Without impediments to trade, buyers on the long side of the submarket compete with each other and bid up to their reservation value at which point they are indifferent between purchasing and waiting. If a trading surplus exists, the buyers' bids exceed the seller's reservation payoff. The seller selects randomly among the set of identical, highest bids from the indifferent buyers. On the other hand, suppose that a buyer enters the submarket and there are excess sellers. This buyer is on the short side of the submarket and hence the lone buyer. As such, the buyer will bid the reservation value of one lone seller or a set of identical sellers and then select randomly among the houses for sale. In both cases, trade occurs after entry without delay.

In an immediate trade equilibrium, auctions involve either a single buyer bidding for one or more houses or a single seller offering her house to one or more buyers. Let the integer N_B and N_S be the number of buyers and sellers over the course of submarket history up to the time of the last sale. Then,

$$N = N_B - N_S$$

represents the number of known buyers remaining since the last sale if $N > 0$ and the number of sellers or inventory of houses if $N < 0$. Let D be the duration of time since the last auction. D then governs the number of potential new buyers given the arrival rates. The pair (N, D) summarizes the conditions in a submarket and therefore determines the payoffs of buyers and sellers.

Seller Entry. Potential sellers make decisions on whether to enter the submarket. Combining the two seller types with the two aggregate states of the economy gives four types of entry decisions. The entry decisions depend on the condition of the submarket (N, D). The analysis of the entry decision begins by asserting that for all (N, D) motivated sellers always enter during prosperous periods and that relaxed sellers do not enter during the slump. These assertions are subsequently checked and hold in the quantitative analysis of Section 5.

Motivated seller's entry decision during the slump depends on the inventory level. In the slump state, buyer entry is slower than the arrival of motivated sellers. Even without relaxed seller entry, motivated seller entry will over time generate an observable accumulation of inventories of houses ($N \leq 1$) and thereby lower the payoff to entry. A cutoff threshold of inventory, N, emerges. Motivated sellers enter only if the accumulated number of houses for sale in the submarket is below the cutoff level of inventory that makes the expected payoff from entry equal to their outside exogenous option V.

Relaxed seller's entry decision during prosperous periods depends on both the inventory level and duration. Over time, motivated seller entry is less than buyer entry. Thus, relaxed sellers must enter from time to time for the submarket to balance. During prosperous times, the expected payoff for relaxed sellers depends on the number of known buyers N and the potential and as yet unseen new buyers which is governed by the duration D. Relaxed seller entry occurs if this expected payoff exceeds the entry cost F.

The cutoff thresholds of inventory N and duration D for relaxed sellers during the prosperous state arise as follows. The distinction in market power between one and two buyers—that is, the difference between monopoly and Bertrand competition—determines a threshold N for relaxed seller entry during prosperous times. With just one buyer, monopolistic bidding allows the buyer to capture the trading surplus by offering the sellers' reservation sale price. If the reservation price is below the sunk fixed cost of entry, then immediately after auctions where only zero or one buyer remains, relaxed sellers will decline the opportunity to participate. In contrast, with two or more buyers engaged in Bertrand competition on the long side of their submarket, market power resides with the seller. Competitive bidding from two or more buyers can induce competing buyers to offer a profitable price and thereby induce relaxed seller entry.

The relaxed seller's entry decision in the prosperous state also depends on the duration D since the last house sale. First, if there are other houses on the submarket, that is, $N \leq -1$, relaxed sellers will not enter for any D. In this situation, when a buyer eventually arrives, that buyer will offer a monopoly price below the sunk entry cost F. Second, with two or more known bidders ($N \geq 2$) remaining after the last sale, relaxed sellers will enter for any D and receive Bertrand bids above F. Finally, if one or no buyers are known ($N = 0, 1$), then the relaxed seller's entry decision depends on D. Recall that new buyers are unobserved until there is a seller present to host an auction. Thus, for small D—before sufficient time has passed without trade—relaxed sellers expect no new buyers in the submarket, and they will not enter. For D large enough—after sufficient time has passed without trade—relaxed sellers expect new buyers have entered, and their entry resumes.

Relaxed sellers choose to enter under conditions discussed above. These conditions, based on the difference between monopoly and Bertrand bidding, are checked for the parameters used in the quantitative section and confirmed to hold in the prosperous state. The condition does not hold during the simulated slump so that relaxed sellers do not want to enter in this aggregate state.

Payoffs and Prices. The analysis characterizes state-contingent payoffs and prices in an immediate trade, monopoly-Bertrand partitioned, Markov Perfect Bayesian Equilibrium, given fully revealing auctions, transitions for N and D, and rational expectations of buyer and seller entry. The quantitative analysis verifies that the seller entry decisions are optimal and that the payoffs to all agents are consistent with the monopoly-Bertrand partition.

In a submarket with excess sellers ($N \leq -1$), sellers accept bids that make them indifferent between trading and waiting for the next auction. In this case, α alone governs the entry of sellers—relaxed sellers decide not to sell, and so, all new sellers are motivated—whereas β or $\hat{\beta}$ governs the arrival rate of buyers.

In the prosperous state, indifference between trading and waiting implies that the payoff to a lone seller in the submarket ($N = -1$) awaiting the arrival of buyers is given by

$$Z(-1) = \frac{1}{1 + rdt}\left[\alpha dt\, Z(-2) + \beta dt\, P(0) + (1 - \alpha dt - \beta dt)Z(-1) - d_s dt\right],$$

where $P(0)$ is the equilibrium price from an auction with one bidder for one house. With other sellers waiting for the arrival of a buyer ($N \leq -2$), the payoff is given by

$$Z(N) = \frac{1}{1 + rdt}\left[\alpha dt\, Z(N-1) + \beta dt\, Z(N+1) + (1 - \alpha dt - \beta dt)Z(N) - d_s dt\right].$$

The solutions to these difference equations are given by

$$Z(-1) = \frac{\beta P(0) - (r + \alpha(1-\lambda))d_s/r}{r + \alpha + \beta - \alpha\lambda},$$

and for $N \leq -2$

$$Z(N) = \lambda^{-N-1}[Z(-1) + d_s/r] - d_s/r,$$

where

$$\lambda = \frac{r + \alpha + \beta - [(r + \alpha + \beta)^2 - 4\alpha\beta]^{1/2}}{2\alpha}.$$

Similar logic generates the prosperous state payoff to house buyers. When relaxed sellers decide to sell, let $H(N)$ denote the prosperous state payoffs for $N \geq 1$ bidders (including the buyer) waiting for the arrival of a seller, either relaxed or motivated. After a two-bidder sale, $N = 1$ so that new relaxed sellers will decide not to sell for some period of time. When relaxed sellers decide not to sell, let $C(T_1)$ denote the expected payoff to the remaining, unsuccessful bidder. That bidder must wait either for the entry of a motivated seller or the duration $D = T_1 > 0$ before relaxed sellers decide to sell.

With only one buyer and one or more sellers in the market ($N \leq 0$), monopolistic bidding makes sellers indifferent between accepting the price and continuing the search. On the other hand, a single seller encountering more than one buyer ($N \geq 1$) will receive bids that make the buyer indifferent. In particular, two bidders will bid up to the price that makes them indifferent between buying the house and being left alone (at $N = 1$) in a submarket where relaxed sellers decide not to sell. With three or more buyers ($N \geq 2$), buyers are indifferent again between buying and waiting with one less competitor and with relaxed sellers deciding to sell.

Equilibrium price offers in the prosperous state are given by

$$P(N) = Z(N), \quad N \leq -1,$$

$$P(N) = x - C(T_1), \quad N = 1,$$

$$P(N) = x - H(N), \quad N \geq 2.$$

The equilibrium price from an auction with one bidder for one house $P(0)$ is derived in Appendix B.

During the slump without relaxed seller entry, the derivation of payoffs accounts for motivated seller entry along with lower buyer entry. Recall that motivated sellers decide not to sell once the observed inventory in their submarket reaches a cut-off threshold. Unobserved buyer entry is no longer a factor in the sellers' decision-making because of the inventory of houses for sale, which simplifies the derivation of payoffs. On the other hand, motivated seller payoffs in the slump not only account for the likelihood of returning to the prosperous state but also satisfy a value-matching condition that equates the payoffs of selling and not selling. Given the above, equilibrium prices in the slump can be derived following a similar logic as in the prosperous state.

Trading Outcome by Duration. Immediate trade from stock-flow matching generates a state-contingent, two-step, piecewise linear hazard function. As there are many independent submarkets in the economy, some newly listed houses will be on the short side with $N \geq 1$ and sell immediately at $D = 0$. Less fortunate newly listed sellers will be on the long side with $N \leq 0$ and must wait and then compete at $D > 0$ for buyers to enter the submarket. The likelihood of a sale on a given day or week is low for the long-side sellers thereby generating a two-step hazard function that declines sharply with duration or time on the submarket. Expected price similarly falls with duration.

The differences in the short- and long-side hazards and prices reflect aggregate states and different submarket-specific circumstances given by N. The short-side hazard and the associated price depend on the stock of buyers but not so on the stock of other sellers. The long-side hazard and price depend on the flow of buyers as well as the inventory of unsold homes. With lowered buyer entry during the slump, motivated sellers outnumber buyers over time. As a result, inventory tends to build up over time reenforcing the relaxed sellers' decisions not to enter the submarket. As prices are functions of the number of excess buyers or sellers, increased competition among sellers will therefore lower prices for the stock of sellers. Selling immediately on the short side becomes less likely and when it occurs, these sellers experience lower continuation payoffs. Their prices on average fall.

4.3 | Discussion

The stock-flow framework has attractive features. It provides a coherent setting where some houses are sold quickly in bidding wars and others are sold in a slower search-like fashion when no potential buyers are present and sellers wait for buyer entry. Stock-flow matching also inherently involves thin markets in which a slowly evolving history matters. Because buyers and sellers cannot rapidly change circumstances from period to period, it takes a while for the long side to evolve.[19] Gradual Poisson arrivals of buyers and sellers lead to correspondences of stock-flow models with random and directed search models. Lagos (2000) and Shimer (2007) document that aggregation over submarkets in steady states of the stock-flow models can generate the familiar random matching functions. Coles (2025) demonstrates that stock-flow matching emerges from embedding directed search with asymmetric information in a dynamic decision-making environment.

Although the parsimonious stock-flow model adopted here captures distinctive features of the trading process in the housing market, it abstracts from salient and potentially quantitatively important elements. Some omitted margins are not difficult to incorporate. For example, some heterogeneity in housing types or local amenities can be incorporated by varying parameter values across submarkets in the simulation. Other elements might not be so easy to incorporate in a stock-flow framework. For example, in the current model, the goods—that is all houses—are identical; all buyers have the same willingness to pay; there are no atypical houses or idiosyncratic tastes among buyers; the joint buyer–seller problem is absent; there are no rental markets.[20]

The comparison of prices from the model with those observed in the data is not straightforward and reflects a standard limitation of search and matching economies. In the model, the price splits the match surplus between the seller and buyer based on the current long or short position in the submarket. As this position and the match surplus evolve with changes in the aggregate state, the model price reflects the expectations of this movement. In contrast, observed prices reflect not only the way in which traders allocate the match surplus based on the gains to trade but also other factors (not captured in the model) impacting their expected payoffs from homeownership.[21] So, an aggregate factor that affects the buyers' and sellers' values in the same way would change the observed prices but not the prices in the model because it does not change the gains from trade. For example, adoption of an unanticipated property tax reduces the buyer's and the seller's values of homeownership, but it does not affect the gains from trade, because it affects the values of both parties by the same amount. As such, the property tax leaves the sales price in the model unaffected even though it alters observed prices. Likewise, as long as the buyers' and sellers' expectations for the future path of prices are consistent, changes in the expectations for resale prices would not affect the gains to trade and thus the model prices.

Perhaps more fundamentally, financing concerns and credit constraints are left out of the decision making.[22] Finance and refinancing options obviously affect trading decisions so that the number and composition of relaxed and motivated sellers could vary across the states of the economy. In addition, buyers cannot move to a different submarket in the model. The formulation of the problem where buyers can move across submarkets in a trackable way is extremely tricky. Suppose that there were several submarkets and buyers could costlessly move between these submarkets, then there would in essence be only one market. Alternatively, if there were a cost to switching submarkets, then the buyer would face a complex strategic decision, which depends strategically on the other buyers' behavior in each submarket. The curse of the dimensionality coupled with the strategic nature of a buyer's problem makes the problem difficult to solve.

5 | Quantitative Analysis

This section first calibrates the above stock-flow model to the U.S. data between 2006 and 2019. It then shows that the model predictions on the trading patterns for the housing market variables are consistent with the data for both the aggregate and duration-dependent patterns as documented in Section 3.

5.1 | Calibration

Connecting Model and Data. The model presented in Section 4 describes a small and isolated housing market. In contrast, most available housing-market statistics correspond to a larger geographic area such as a city that contains numerous small markets defined by neighborhood and housing type. As geographic-based statistics are broader than the model, the approach here is to aggregate across many model-generated submarkets and compare the aggregated statistics with the corresponding data moments. In particular, the model is simulated for many small micro- or submarkets that all share a common shock and then aggregated over these submarkets to mimic the average U.S. urban housing market.

In the simulation, the number of submarkets is set as follows. A model-generated submarket is assumed to correspond to a type of house, determined by its so-called quality class, within an elementary school district. A city with 700,000 residents, approximately the 2010 average population in the CoreLogic sample, will have 60,000 students in elementary schools, assuming equally sized cohorts of 10,000 people aged up to 70 years. The average elementary school in the United States has 479 students, which yields 125 elementary schools for a city with 700,000 residents. For comparison, Detroit and Charlotte, two cities that have roughly this number of residents, have 106 and 115 elementary schools, respectively. Focusing on the core five housing quality classes as in Smith (2020a) yields $125 * 5 = 625$ submarkets in a city with the average size of the population.[23]

Calibration Procedure. While the model is set in continuous time, the data in Section 3 are constructed in discrete time intervals. To connect the model with data, the calibration simulates the continuous time model and constructs model moments from intervals that partition continuous time into months. There are 10 parameters to calibrate: $x, r, \alpha, \sigma, \beta, \hat{\beta}, \tau, F, d_s, V$. The calibration normalizes the buyer's payoff from home ownership x to one. Mapping the model parameter rates to daily frequencies, the daily rate of time preference is set to $r = 0.05/360$, corresponding to a yearly interest rate of 5%.

The rest of the parameters are calibrated to target the housing market statistics between 2006 and 2019 in the CoreLogic data. As discussed in Section 4, the model parameters must satisfy the following restrictions:

$$0 < \alpha < \beta < \alpha + \sigma, \quad 0 < \hat{\beta} < \alpha, \quad 0 < F < x, \quad V < 0.$$

In addition, the conditions for monopoly-Bertrand entry are checked in the calibration.

The arrival rates of buyers $\hat{\beta}$ and β are calibrated to match the average daily sales in the CoreLogic data in the slump (2007) and the postslump recovery (2015–2019), respectively.[24] Average daily sales are computed as average monthly sales in the data divided by 30 days in a month and by the number of submarkets. This gives $\hat{\beta} = 0.04176$ and $\beta = 0.06453$. As a result, in the slump, a buyer in a particular submarket arrives on average once every three weeks. During recovery and prosperity, a buyer arrives on average once every 2 weeks.

The downturn in the housing market over the sample period began in early 2006 and ended in early 2011. Hence, in the simulation, the slump lasts for 60 months. The slump does not start at the same time for all cities in the data. To smooth the model implications and align with averages over cities in the data, the slump and the recovery occur in five installments and each is 2 months apart. The first one-fifth of the submarkets experience a lower entry of buyers in January 2006, the second fifth in March 2006, and the final fifth in September 2006. The recovery rolls out similarly beginning in January 2011 and ending in September 2011. Given the expected duration of the slump of 60 months, the daily recovery rate from the slump is $\tau = 1/60/30 = 0.0006$.

The arrival rate of motivated sellers affects the accumulation rate of inventories at the beginning of the slump. The daily inventory accumulation rate (due to excess motivated seller entry over buyer entry) is $\alpha - \hat{\beta} > 0$. To understand the relationship between the arrival rate of motivated sellers and the accumulation rate of inventory, it is easier to first consider the case in which the recovery rolls out at the same time for all submarkets. Assuming no submarkets are at capacity at the beginning of the slump, which is the case in the simulation, the amount of inventory accumulated over a duration of time D is given by

$$(Inventory_{t+D} - Inventory_t) = (\alpha - \hat{\beta}) \times D \quad \text{and} \quad Sales(t \text{ to } t + D) = \hat{\beta} \times D.$$

Thus, the arrival rate of motivated sellers α can be solved as[25]

$$\alpha = \hat{\beta}\left(1 + \frac{Inventory_{t+D} - Inventory_t}{Sales(t \text{ to } t + D)}\right). \tag{2}$$

Because the slump is rolled out gradually in five installments in the simulation, the calculation of the accumulated inventory and sales over the duration D, and thus α, will be similar but not exactly the same as described above. The details are included in Appendix C.2. Given the relationship between the arrival rate of motivated sellers and the accumulation rate of inventory, α can be calculated using the data for inventories and total sales for a duration D at the beginning of the slump. For robustness, two durations of D are considered: January to October 2006 and the entire year of 2006. The average of α computed from the two cases is used in the simulation.

As discussed in Section 3, the coverage of the CoreLogic data is sparse early on when the slump hits which primarily affects the inventory count during this period. For this reason, the analysis turns to the NAR inventory series when calibrating α. To account for discrepancies between the NAR and CoreLogic inventory series, the NAR series is adjusted for pending sales counted by NAR but not present in the CoreLogic measure. Appendix C.1 provides details for the adjustment procedure.

This paper models a slump as a shock to the arrival rate of buyers. The aim is to match the aggregate data while keeping the analysis transparent and reducing the calibration burden. While allowing parameters such as α to be state-contingent could better fit the data, doing so introduces challenges. In particular, α under the current calibration is determined using the changes in inventories at the start of the slump as relaxed sellers stop entering the market (Equation (2)). To calibrate a second α in the prosperous state using the change in inventories starting at the end of the slump would reflect both motivated and relaxed sellers entering the market. However, since relaxed sellers' decision to not enter is unobserved and time varying, identifying α in the prosperous state is far more complex than in the slump.

Given the parameter values for $x, r, \alpha, \beta, \hat{\beta}$, and τ, four parameters F, d_s, σ, and V remain for calibration. The two costs (F and d_s) and the daily arrival rate of relaxed sellers (σ) relate exclusively to the relaxed sellers' entry decision, whereas the last parameter V is exclusive to the slump. These two groups of parameters can therefore be calibrated sequentially. The first three parameters F, d_s, and σ are jointly calibrated to match the averages of inventory-to-sales ratio, DoM for sales,[26] and the coefficient of variation of sales between 2015 and 2019, the prosperous years during the sample period.[27]

The last unknown parameter, the external payoff V, determines the point at which motivated sellers stop entering a submarket during the slump, that is, the number of sellers in the submarket that saturates the submarket and deters further entry. As V falls, the saturation point increases and therefore the maximum inventories that can occur. Given this relationship, V can be calibrated to match the peak inventory in the NAR by iterating the model with the above parameters until the model reaches the targeted peak value. In the data, inventories peak in October 2007 and thus the 3-month average of September–November 2007 inventories are used as targets.[28]

Table 2 reports the calibrated parameter values. These values imply that relaxed sellers wait a week (7.1 days) before they are willing to enter a submarket with one known buyer and

TABLE 2 | Calibrated parameters.

Parameter meaning (daily)		Value	Targeted moments	Targets	
				Data	Model
Separately calibrated parameters					
β	Arrival rate of buyers during prosperity	0.06453	Average monthly sales post-slump	1210	1210
$\hat{\beta}$	Arrival rate of buyers during slump	0.04176	Average monthly sales in slump	783	783
τ	Expected recovery rate from slump	0.0006	Expected duration of slump	60 mos	60 mos
α	Arrival rate of motivated sellers	0.05262	Δ inventory/sales Jan–Oct 2006	0.1567	0.1567
			Δ inventory/sales Jan–Dec 2006	0.1140	0.1140
Jointly calibrated parameters					
σ	Arrival rate of relaxed sellers	0.04978	Average inventory-to-sales ratio	1.50	1.485
d_s	Seller's search cost	0.0000687	Average DoM for sales	49	44
F	Seller's fixed cost	0.01289	Coefficient of variation of sales	0.02666	0.02689
V	Outside option of motivated seller	−0.18250	Peak inventories	7492	7544

Note: The targeted moments are calculated from the CoreLogic data except for inventories that use the adjusted NAR series. See text and Appendix C.1 for more details on the calibration strategy.

over 3 weeks (23.7 days) to enter a submarket with no known buyers. The table also compares the targeted moments from the model simulation with the data, and the calibration matches the targets well.

5.2 | Model Predictions Versus Data

Using the calibrated parameters, Table 3 assesses the stock-flow model in generating the housing market trading patterns by comparing the model predictions on a set of nontargeted moments with the corresponding statistics in the data. Specifically, the moments in the comparison include the correlations between model results and data, the peak-to-trough percentage changes, the correlations of various housing market variables with prices, and the standard deviations for variables including sales, listings, inventory, inventory-to-sales ratio, DoM, Kaplan–Meier statistics, and average price. The short-side model predictions are compared with variables in the data with a duration of 15 days or less and the long side with a duration of 31–45 days.

From column (3) of Table 3, the data series and the model results are strongly correlated with all the correlations above 0.76. In particular, the correlations for inventory-to-sales ratio and the long-side Kaplan–Meier statistics are the strongest with both values around 0.9, while the slightly weaker correlations between 0.75 and 0.80 are for inventories, DoM of inventory, and prices. The model-generated correlations with prices are also quantitatively consistent with their data counterparts.[29]

From the table, the simulated peak and trough levels for the variables considered are in line with those in the data, except for average price because of the normalization of the buyers' and sellers' payoffs to owning a home in the simulation. As for the peak-to-trough percent changes, the model results are reasonably close to the data except for the average price. The last column compares the variation of the variables over the sample period, measured by the standard deviation, with the data. The model implied standard deviations are close to those values in the data. Not surprisingly, the variables with a better match to the data for peak-to-trough percent changes also match the data better in terms of the standard deviations. In addition, the model generates the variability of duration dependence: the short-side sale probability (Kaplan–Meier) is more volatile than the long side, consistent with volatility patterns in the data.

Although not reported in the table, the simulation also delivers extremely high persistence in the variables as observed in the data. The simulated autocorrelations are all above 0.95, consistent with values in the data that are all above 0.99. In addition, the model also delivers, as expected from stock-flow matching, a close correlation of 0.900 between listings and sales that aligns with the correlations of 0.954 and 0.988 for nonseasonally adjusted and seasonally adjusted series from the CoreLogic.[30]

To summarize, Table 3 demonstrates that the stock-flow matching model matches the data well along various dimensions. These results are remarkable because given the simple model structure, the simulation is able to generate these model predictions with only one shock and seven other parameters as listed in Table 2.

5.3 | Simulated Paths

Figures 3 and 4 show the simulated paths for the variables considered with the slump and recovery shock τ. The responses to the shock depict the series of events happening in the model. The arrival rates of the two types of sellers—relaxed and motivated—relative to the arrival rate of buyers govern sales as well as the levels and changes of inventories in the two aggregate states. When the slump shock hits, buyer entry immediately falls below motivated seller entry. As a result, sales and listings decline. In addition, listings exceed buyer entry and inventories

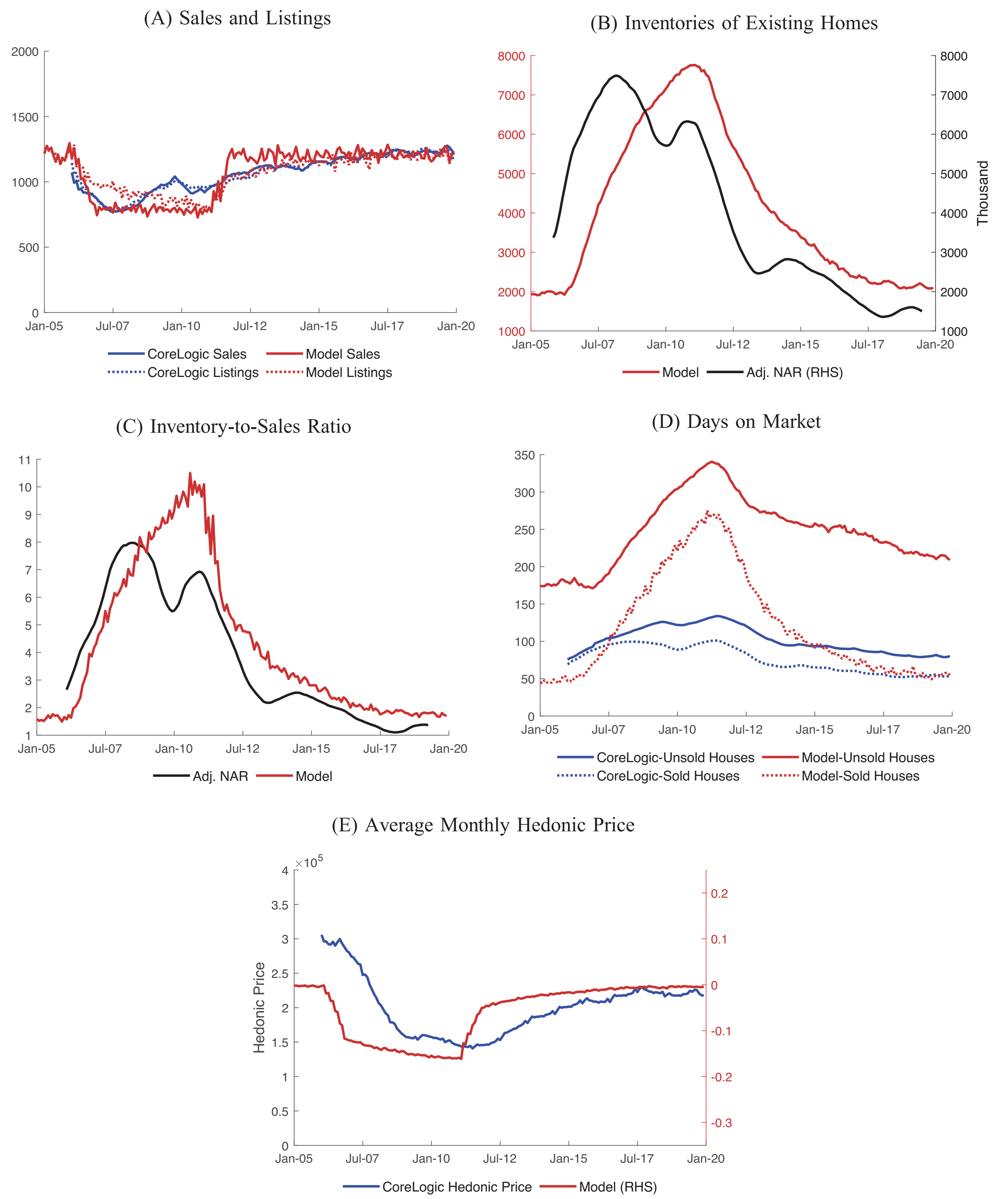

FIGURE 3 | Housing market aggregates: data versus simulated model. *Notes:* All series labeled “CoreLogic” are the authors’ calculations based on CoreLogic data.

TABLE 3 | Comparison of simulated statistics with data.

Variable	Source	Corr with data	Min	Max	% change max-min	Corr with Price (2009 onward)	Std Dev (log)
Sales	CoreLogic	—	773	1277	49.2	0.921	0.142
	Model	0.872	727	1284	55.4	0.963	0.201
Listings	CoreLogic	—	772	1260	48.0	0.952	0.134
	Model	0.840	778	1260	47.3	0.968	0.140
Inventory	Adjusted NAR	—	1359	7492	138.6	−0.691	0.583
	Model	0.769	2081	7762	115.4	−0.886	0.456
Inventory/Sales	Adjusted NAR	—	1.10	7.97	151.4	−0.664	0.674
	Model	0.893	1.65	10.50	145.6	−0.979	0.607
DoM of Sales	CoreLogic	—	52.2	101.0	63.7	−0.966	0.242
	Model	0.810	49.0	275.4	139.6	−0.842	0.540
DoM of Inventory	CoreLogic	—	78.8	133.8	51.7	−0.981	0.171
	Model	0.774	173.1	340.6	65.2	−0.772	0.156
Short Kaplan–Meier	CoreLogic	—	0.147	0.380	88.7	0.934	0.309
	Model	0.840	0.079	0.507	146.1	0.854	0.583
Long Kaplan–Meier	CoreLogic	—	0.138	0.250	57.8	0.938	0.201
	Model	0.940	0.041	0.145	111.2	0.919	0.388
Average Price	CoreLogic ($1000)	—	141	280	66.2	—	0.175
	Model	0.794	−0.162	−0.003	193.1	—	0.062

Note: (1) All data are for January 2007–December 2019 unless otherwise noted. NAR inventory is adjusted to account for pending sales; Appendix C.1 includes the details. The adjusted NAR inventory is reported in thousands. (2) All CoreLogic variables are monthly and are seasonally adjusted by taking a ±6-month moving average. Model figures come from model simulations and are not smoothed. (3) Inventory is the 6-month moving average of the adjusted NAR value at the end of each month. The Inventory/Sales ratio uses the adjusted NAR inventories and adjusted NAR Sales series described in Appendix C.1. (4) Average CoreLogic prices are hedonic regression dollar values. (5) All correlations involving price are from 2009 onward. (6) % change max-min uses the average of max and min as the denominator. (7) CoreLogic Kaplan–Meier are unconditional statistics. The conditional Kaplan–Merier statistics are very similar to unconditional ones. Figure A.4 and Table A.1 compare the two. (8) Since the average price in the model is small and negative, log(Average Price) is approximated by the Average Price. (9) The coefficient of variation has similar values as standard deviations of logged variables.

start to accumulate in every submarket until they reach the threshold that halts motivated seller entry. As buyers become more scarce and competition among sellers becomes more fierce, the short- and long-side selling probabilities, measured by the Kaplan–Meier statistics, both decline and the DoM for sold and for unsold houses rise. These factors likewise depress prices for each duration and thus the aggregate price as well.

When the downturn ends, buyer entry picks up, inventories decline, and sales and listings increase. But relaxed seller entry in a given submarket does not resume until inventories eventually all clear in that submarket. The probability of a sale improves for both the short- and long-side Kaplan–Meier statistics but the increases in the two series are not in lock step with each other as inventories dwindle. The two DoM series fall as a result. Prices rise immediately at the time of the recovery, mirroring the rise in buyer entry. As the recovery progresses, prices, in particular, for sellers with an inventory of houses already in the market, continue to rise during the subsequent transition to the new steady state.

As shown in Figures 3 and 4, the simulated model is largely consistent with the time-series patterns in the data for the variables considered. Not surprisingly, the model cannot exactly replicate the data series. This is because other shocks that are not included in the model, such as financial shocks, could also affect the housing market. The rest of this subsection compares the simulated paths with the actual data in detail.

Panel (A) in Figure 3 plots the simulated listings and sales against the data. Targeting the average CoreLogic sales levels in the slump and the postslump recovery (with $\hat{\beta}$ and β), the magnitude of sales and listings (adjusted for market size) lines up closely with the data over the sample period. The most significant deviation occurs during the impact of the American Recovery and Reinvestment Act from May 2009 until mid-2011.[31] The associated rise and fall of the simulated inventories in Panel (B) of Figure 3 mirror the adjusted NAR inventories. When recovery occurs, inventories fall as buyer entry now exceeds motivated seller entry and relaxed sellers wait for inventories to fall. Given the close correspondence of sales and inventories with the data, it is not surprising that the inventory-to-sales ratio also fits well in Panel (C).

Panel (D) of Figure 3 plots the simulated DoM for sold and unsold houses. For houses that sell within the month and for those that remain available at the end of the month, the

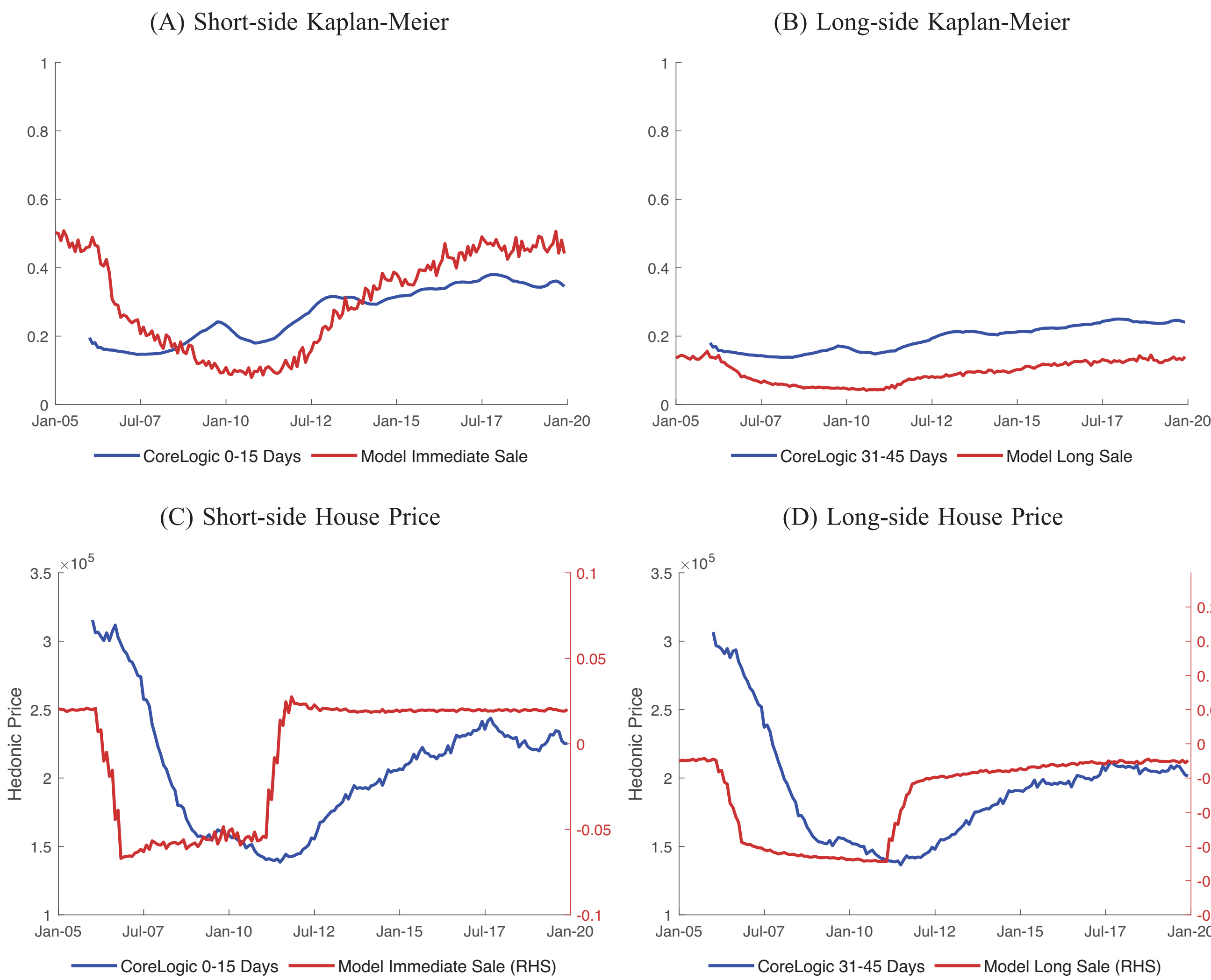

FIGURE 4 | Variables by duration: data versus simulated model. *Notes:* All series labeled "CoreLogic" are the authors' calculations based on CoreLogic data.

simulation captures the qualitative pattern found in the CoreLogic DoM data, including the distinction between the two measures, their rises in the slump and declines in the recovery, and their peaks in 2012. As noted above and documented below, stock-flow matching generates a discrete drop in the probability of selling by duration. This hazard function leads to shorter DoM for houses sold during the month (completed spells) than for the inventory of unsold houses at the end of the month (uncompleted spells) at a given point in time. The size of the drop in the selling probability by duration controls the difference between the DoM of sold and unsold houses. A higher probability of an immediate sale generates a larger difference between the two series. A random search model cannot readily generate this difference as house sales have a constant hazard function.

The simulation predicts a wider gap in DoM between the short and long durations than in the data. One contributing factor is that sellers in the simulated submarket are committed to trade and cannot withdraw. Data from Redfin suggest that approximately 6% of houses were withdrawn from the market without a sale during the recovery after 2012. The slump may well have even higher withdrawal levels. The lack of a withdrawal option raises both DoM measures in the model, especially during a slump, and generates a larger gap between different durations.

Panel (E) of Figure 3 displays the simulated average monthly house prices from all sales along with that in the CoreLogic data. Like the CoreLogic price, the model-simulated average price is U-shaped over the cycle with concurrent troughs around the end of 2011. The initial drop in prices in the model occurs, however, before the drop in the CoreLogic price, and the simulated prices rise more abruptly when the recovery begins.[32] Overall, the prices in the model are less volatile than in the data but comparisons of price levels between data and model are limited. As discussed in Section 4.3, the simulated house prices reflect the way in which agents share gains to trade, whereas the observed house prices consider more margins including future resale values. In addition, the prices from the calibrated model often become negative because the buyer's payoff from home ownership x is normalized to one and the relaxed seller's outside option is normalized to zero. Such normalization implies a negative value of the calibrated outside option V for the motivated seller, and thus, they may accept a negative price that is above their outside option. Normalizing the relaxed seller's outside option to larger values would generate nonnegative prices.

Simulated Paths by Duration. Figure 4 plots the Kaplan–Meier selling probability and prices by duration. It reveals that the progressions of the two Kaplan–Meier hazard rates in the model follow the patterns observed in the CoreLogic data. When the slump hits, both simulated hazard rates drop sharply and the decline is most pronounced in the short-side hazard rate. As a sale becomes increasingly less likely in the presence of ever-growing unsold inventories before the recovery, the hazard function continues to fall as the slump continues. More importantly, similar to the CoreLogic hazard rates, the simulated short Kaplan–Meier trading probability recovers earlier, faster, and stronger than the longer duration trading probability.

Through the sample period, the simulated long-side hazard approximately parallels the CoreLogic 31–45 Day Kaplan–Meier statistics but at a lower level. The simulated short-side Kaplan–Meier displays a similar progression as found in the data but exhibits greater variation over time and appears less parallel before the recovery. As a result, the differences in the simulated long-side and short-side model hazard rates are larger in the model than in the data during the prosperous period, especially at the beginning of the sample period.

Panel (A) of Figure 4 compares the short-side selling probability from the model with the Kaplan–Meier data for a duration of 15 days or less. This comparison offers an explanation for part of the difference between the model and data. In the model, once entered, a short-side seller finds one or more potential partners immediately available and can sell quickly. In the data, not all sales that would correspond to a short-side trade in the model occur within the first 15 days.[33] Some of the sales between 16 and 30 days could also correspond to a short-side trade. A simple adjustment, accounting for the short-side sales between 16 and 30 days, raises the maximum of the short-side Kaplan–Meier from the CoreLogic to 0.412, which compares more favorably to the simulated maximum of 0.507.[34] As the adjustment raises the minimum 0–15 day Kaplan–Meier to 0.149, the percentage change between the minimum and maximum of the short Kaplan–Meier rises to 94.2%, which is relatively close to the percentage change in the model of 111.2%. On the other hand, the higher long Kaplan–Meier in the CoreLogic may again reflect withdrawals, which improve the probability of sales among the remaining houses on the long side of the market.

Panels (C) and (D) of Figure 4 plot model-implied prices for the short and long sides along with the CoreLogic hedonic prices. In the model, prices could be negative because of the normalization of the payoffs for buyers and sellers. Although the evolution of prices is abrupt, the simulated price captures the uneven and differential response in the observed short- and long-duration prices. As in the data, the short price in the model recovers more quickly and more robustly than the longer duration price.

Taken together, the simulated results paint a supportive picture for the stock-flow matching approach with Poisson shocks for explaining the cyclical behavior of the housing market. As the correlations between the model and the data highlight, the stock-flow mechanics match up well with not only the aggregate statistics but also with prices and selling probabilities by duration. Specifically, the ability of the framework to generate distinct responses in the selling probability and prices over short and long durations not only arises naturally in this specification but also generates patterns observed in the CoreLogic data. This result is not a standard outcome of other dynamic trading formulations.

6 | Evaluation and the Role of Policy

This section considers further the quantitative contribution of the stock-flow matching model developed in Section 4. It first illustrates that time-varying duration dependence, embedded in the stock-flow model, is crucial in matching the data and then shows that the model implications are consistent with cross-sectional evidence. The section also illustrates the potential of the stock-flow model for studying the impact of policies targeting the housing market enacted during the sample period.

6.1 | Time-Varying Versus Fixed Duration Dependence

The differential responses of short- and long-side selling probabilities to market conditions are essential for understanding the evolution of inventories. To illustrate this importance, this subsection performs a counterfactual exercise that predicts an inventory series while holding the ratio of the short- and long-side Kaplan–Meier selling probabilities constant over the entire sample period. The comparison of this counterfactual inventory series with the data provides information about the importance of time-varying duration dependence in matching the data.

Let KM_t^{Short} be the Kaplan–Meier probability that a new seller finds a buyer upon entry and let KM_t^{Long} be the selling probability at which a house sells from inventories. Sales in any period t consist of the listings within the period as well as the inventories carried over from the previous period that successfully find a buyer[35]:

$$Sales_t = KM_t^{Short} \times Listings_t + KM_t^{Long} \times Inventory_{t-1}.$$

Let KM_{ss}^{Short} and KM_{ss}^{Long} be the Kaplan–Meier selling probabilities in the simulated prosperous steady state. The counterfactual experiment synchronizes the behavior of KM_t^{Short} and KM_t^{Long} over the cycle and sets the ratio of the short- and long-side KM in any period t to the ratio in the prosperous steady state:

$$KM_t^{Short}/KM_t^{Long} = KM_{ss}^{Short}/KM_{ss}^{Long}.$$

In addition, the inventory of unsold houses at end of the period to be carried over to the start of the next period consists of the inventories that did not sell this period and the new listings that did not sell:

$$\begin{aligned} Inventory_t = {} & (1 - KM_t^{Long}) \times Inventory_{t-1} \\ & + (1 - KM_t^{Short}) \times Listings_t. \end{aligned}$$

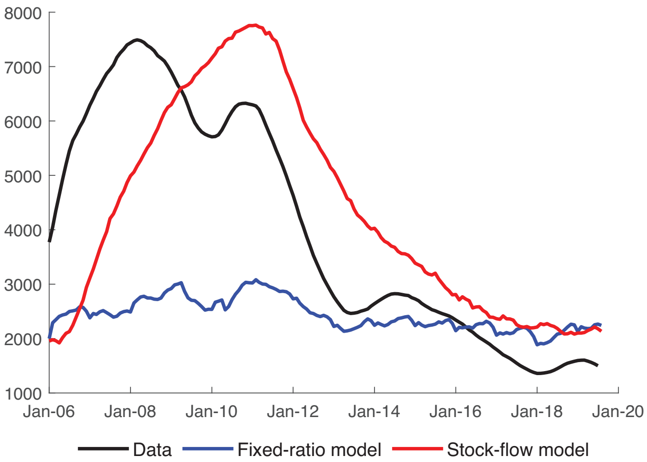

FIGURE 5 | Inventories in the fixed ratio model. *Notes:* The data for inventory are calculated from the CoreLogic data.

Using the simulated steady-state inventory as the starting value along with sales and listings data in any period t, these equations above generate counterfactual short- and long-side Kaplan–Meier probabilities and inventory for each period t.

Figure 5 shows the actual and counterfactual paths for inventory along with the path from the calibration. While the calibration gives an inventory path that varies over time as in the data, the counterfactual housing inventory does not accumulate enough during the downturn and peaks at a much lower level.[36] Thus, the large increases in inventory arise from the different cyclical behavior of selling probabilities by duration. Such time-varying duration dependence is an important feature of the stock-flow model.

6.2 | Variation Across Cities

Thus far, the analysis has focused on housing market dynamics based on statistics aggregated across U.S. cities. This subsection extends the analysis to the cross section and shows that the stock-flow model also generates housing market outcomes that are consistent with variations across cities.

More specifically, the model is simulated with city-specific β, the arrival rate of buyers during prosperity, while holding all other parameters fixed at the calibrated values. Because the difference in the arrival rate of buyers between slump ($\hat{\beta}$) and prosperity (β) drives the housing cycles in the model, comparing simulated results obtained from city-specific β's with city-level data helps assess the model's performance in generating cross-city variation over the housing cycle. In the simulation, the city-specific β is derived by setting the ratio of the arrival rates of buyers in the prosperity and slump($\beta/\hat{\beta}$) equal to the city-specific ratio of sales during the prosperity and slump periods in the data. Appendix D.1 describes the details.

Using averages over 2007–2010 for trough statistics and averages over 2015–2019 for peak statistics, Table 4 shows that the model predictions on the cross-city trough-to-peak percentage changes are positively correlated with the corresponding changes in the data for all five housing market variables. The correlation coefficients, reported in row (1), range between 0.36 and 0.61. Row (2) reports the cross-city coefficients from regressing the trough-to-peak percentage change in the model on the corresponding change in the data, with a constant. All five regression coefficients are economically sizable. They indicate that the model accounts for more than 30% of the cross-city variations in the trough-to-peak percentage changes, with the exception of the short-side Kaplan–Meier, for which the model accounts for 12%.

Rows (3) and (4) report the average trough-to-peak percentage changes across cities in the model and data, respectively. Qualitatively, the model generates the average trough-to-peak changes in the housing market. Quantitatively, the simulation generates larger trough-to-peak changes than the data for four of the five statistics. This suggests that factors other than the changes in the arrival of buyers are also relevant for generating the housing market dynamics across cities. It is nonetheless remarkable that with the change in only one parameter, β, the model can generate a significant portion of the cross-city variations in housing statistics. This consistency of the model with the cross-sectional data supports the parsimonious quantitative approach in this paper.

6.3 | First-Time Homebuyer Credit

The stock-flow matching model developed in Section 4 can be used to study the impact of policies targeting the housing market. One application is to explore the effect of the first-time homebuyer credit under the American Recovery and Reinvestment Act implemented to stimulate the housing market during the Great Recession. The policy allowed qualified buyers to claim up to $8000 toward the purchase of a home from January 2009 through April 2010 with no requirement of repayment. As the stock-flow model in this paper abstracts, among other factors, from buyers' asset holdings, the model cannot fully scrutinize the impact of the first-time homebuyer credit. However, one crucial consequence of the policy was an increase in the number of buyers in the market. As reported by the NAR, the share of first-time homebuyers among all buyers increased from 41% in 2008 to 47% and 50% in 2009 and 2010, respectively. In the context of the model, increasing the buyers' arrival rate during the slump captures this impact of the policy.

Let $\Delta\hat{\beta}$ be the increase in the buyers' arrival rate in the slump. Assuming the policy will not change the arrival rate of buyers who are not buying for their first homes, $\Delta\hat{\beta}$ can be derived as:

$$\Delta\hat{\beta} = (\hat{\beta} + \Delta\hat{\beta}) \times \text{share of first-time buyer with policy}$$
$$-\hat{\beta} \times \text{share of first-time buyer without policy.}$$

The empirical counterpart for the share of first-time buyers with policy is approximated by the average share between 2009 and 2010. There is no empirical counterpart for the share without policy. To simulate the policy effect, the average share of 2003–2007 is used to approximate the share without policy.[37] Because the slump and the recovery roll out in five installments, the policy in the experiment is also launched in five installments where $\hat{\beta}$

TABLE 4 | Impact of buyer's arrival rate: model and data comparison across cities.

	Inventory/Sales	DoM - Sales	DoM - Inventory	Short KM	Long KM
(1) Correlation	0.55	0.61	0.36	0.52	0.43
(2) Slope	0.44	0.39	0.31	0.12	0.49
	(0.19)	(0.15)	(0.23)	(0.06)	(0.29)
(3) % Change Model	−118	−90	−16	96	81
(4) % Change Data	−59	−50	−31	53	44
(3)/(4) % Explained	202	179	53	179	185

Note: The data used to derive the statistics in this table are the authors' calculations based on CoreLogic data. "Slope" is the cross-city regression coefficient of regressing the trough-to-peak percentage change in the model on the corresponding change in the data, with a constant. Standard errors from regressions are reported in parentheses. Sample comprises 14 cities that satisfy the parameter restrictions and the monopoly-Bertrand condition. The tough-to-peak percentage change uses the average of trough and peak values as the denominator.

increases to $\hat{\beta} + \Delta\hat{\beta}$ in January 2009 and returns to $\hat{\beta}$ in April 2010 for the first installment and the subsequent installment follows in every two months.

Figures D.1 and D.2 in Appendix D.2 compare the simulation results with and without the credit. Evidently, the effect of the credit is short-lived and is only relevant for the housing market during the policy period. Not surprisingly, the increase in the buyer's entry rate raises sales and listings but reduces inventories and inventories-to-sales ratio. As a result, DoM series decline for both the sold and unsold houses, and both the short- and long-side Kaplan–Meier selling probabilities and prices rise. Moreover, the responses for the short side of the market are much stronger than the long side. Although the mapping of the first-time homebuyer credit to the model is imprecise, these findings illustrate the policy's impact on the housing market through an increase in the number of buyers.

7 | Conclusion

This paper analyzes the performance of the U.S. housing market from 2006 to 2019, including the housing market collapse, a prolonged slump, and the subsequent recovery and expansion right before the COVID-19 pandemic. Breaking down transactions by duration on the market, the paper establishes that recently listed sellers experience distinctly different outcomes from those sellers who have spent a month or more looking to trade. The probability of selling and the sales price in the 2 weeks immediately after listing are markedly higher. Moreover, as the market recovers from a slump, the probability of a quick sale increases sooner, faster, and more robustly than the probability of a more prolonged sale. The paper then demonstrates that a parsimonious stock-flow matching model can generate these observed outcomes.

The quantitative performance of the model is compelling. The calibrated model delivers outcomes broadly consistent with not only the observed timing and patterns of the aggregate housing market statistics but also with trading outcomes by duration—trading probabilities and prices by DoM. The calibrated model captures the distinct trading outcomes in the data of those who trade quickly and of those who take time. The ability to mimic the quantities of sales, average prices, listings, DoM, and the trading patterns by duration indicates the fundamental importance of stock-flow matching for understanding the performance of the housing market.

The quantitative analysis of this paper focuses on the aggregate U.S. housing market. Although housing markets all over the United States experienced a downturn and recovery between 2006 and 2019, the timing and severity of the downturn differed by market. As illustrated in Section 6.2, the stock-flow matching model developed in this paper can also be used to study differences in the slump and recovery by city. Further considerations of the city-specific timing and severity of the slump are left for future analysis.

Acknowledgments

Views expressed are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta, the Federal Reserve System, the IBRD/World Bank, or its executive directors. Smith acknowledges the support of the Business and Local Government Data Research Centre (grant number ES/S007156/1) funded by the Economic and Social Research Council (ESRC) for undertaking this work. We would like to gratefully acknowledge early contributions to this project from Yunjuan Liu. We would also like to thank the editors and three anonymous referees as well as numerous participants at conferences and seminars for their insightful comments.

Data Availability Statement

The data that support the findings of this study include confidential CoreLogic data. Restrictions apply to the availability of these data, which were used under license for this study. Information about the CoreLogic Data is available at: https://trestle.corelogic.com. The CoreLogic data aggregated at the national level and all other publicly available data used to construct the figures are openly available in the repository openICPSR at: https://doi.org/10.3886/E246902V1 , reference number 246902.

Endnotes

[1] Pryce and Gibb (2006) document a similar pattern for the hazard function in Scotland where sellers hold sealed bid auctions. Discounting the first 2 weeks to account for sellers specifying a closing date after they receive initial bids, the hazard peaks at 14–30 days and declines rapidly thereafter. At approximately 6 weeks, the hazard falls to one-fifth of that peak value. See also Anundsen et al. (2023) for striking evidence of the high propensity of immediate sales in Norway.

[2] An extensive literature documents a price-liquidity relationship. Cubbin (1974) and Miller (1978) are early papers revealing this link. More recently, Jiang et al. (2024) point out that "housing market liquidity appears to be a multidimensional object." They supplement the basic time on the market measure with price dispersion to capture the multidimensionality in a random matching model. In the present paper, the probabilities of a sale conditional on duration on the market capture the multiple facets that, in turn, generate dispersion in house prices.

[3] They find that listings are slightly more volatile than sales and significantly more volatile than prices. Listings are also strongly positively correlated with sales and prices, and negatively correlated with time to sale.

[4] See Merlo and Ortalo-Magné (2004) on U.K. data, Leung et al. (2006) on Hong Kong data, de Wit and van der Klaauw (2013) on Dutch data, Bierdel et al. (2023) on Spanish data, Rekkas et al. (2025) on Canadian data, and numerous studies of hedonic price regressions.

[5] Han and Strange (2015) also note that our understanding of the mechanics of housing auctions in dynamic settings with search frictions is very limited. Although bidding wars are common in practice, price determination in most matching models with frictions derives from one-to-one bargaining. The stock-flow matching approach adopted here addresses this concern. See also Arefeva (2025).

[6] Anenberg and Bayer (2020), Caplin and Leahy (2011), and Smith (2020a) all consider alternatives to the amplification of external factors. They address the same set of empirical regularities that motivate the external amplification approach but adopt a stationary environment. Anenberg and Bayer (2020) estimate a model in which the decision of homeowners to jointly sell their existing house and buy a new one creates a coordination externality that leads to an alternative explanation for endogenous booms and busts. Ngai and Tenreyro (2014) address seasonality in housing markets.

[7] Favilukis et al. (2017) and Garriga et al. (2019) attribute aggregate volatility in house prices to changes in borrowing constraints coupled with changes in capital flows, whereas Kaplan et al. (2020) cite ex-post inaccurate expectations. Although these papers generate large time-series fluctuations in house prices, they abstract from search frictions and hence cannot speak to the variations in DoM, inventory over a housing cycle, and duration dependence of sale probability and of price, which is the focus of this paper.

[8] See, for example, Wheaton (1990), Krainer (2001), Albrecht et al. (2007), Head and Lloyd-Ellis (2012), Rekkas et al. (2025), and Gabrovski and Ortego-Marti (2021).

[9] Although the economic framework incorporates novel extensions to Smith (2020a), the paper does not derive new theoretical results.

[10] Using employment spell data, Coles and Smith (1998) obtain compelling evidence in favor of stock-flow matching behavior. See also Gregg and Petrongolo (2005), Shimer (2007), Coles and Petrongolo (2008), Kuo and Smith (2009), and Andrews et al. (2013). See Smith (2020b) for an overview of these results. Gilbukh (2023) adopts the stock-flow framework to assess housing market flows.

[11] The NAR and CoreLogic data cover sales through agents. According to NAR, only 7% of sales are "For Sale By Owner" and not covered by the NAR or by CoreLogic.

[12] The conditional Kaplan–Meier statistic can be computed from a logit regression controlling for the same set of housing characteristics as in Equation (1) for the hedonic price regression. The resulting series is similar to the unconditional Kaplan–Meier statistics. See Appendix A for details.

[13] Price coefficients are seasonally adjusted by construction.

[14] Autocorrelations for the variables in Table 1 are high. For the seasonally adjusted series, the autocorrelations are all above 0.99. In the nonseasonally adjusted series, sales and listings exhibit considerable seasonality and hence the lowest autocorrelations. The nonseasonally adjusted autocorrelations for the CoreLogic sales and listings are 0.841 and 0.769, respectively. The NAR sales autocorrelation from 2007 is 0.756. The inventory-to-sales ratio is consequently the next lowest autocorrelation—CoreLogic 0.933 and NAR 0.897. The other remaining autocorrelations are all above 0.97.

[15] The NAR inventories peak in October 2009, so a post-2009 series for inventories serves little purpose.

[16] The limited early coverage in CoreLogic should not so dramatically affect the min and max of these two series as the sample was representative.

[17] Giacoletti (2021) uses CoreLogic data for California to study the importance of idiosyncratic house risk on household welfare.

[18] To maintain a well-behaved market over time, the stock-flow literature typically assumes that buyers and sellers independently enter the market one by one at the same exogenous Poisson rate. Persistent unequal entry would eventually lead toward an infinity of buyers or sellers. In this paper, sellers have a higher arrival rate than buyers but they have the option to decline the opportunity to sell and save the associated sunk cost of participation.

[19] This feature seems plausible for the housing market. Piazzesi et al. (2020) find that buyers look in narrow market segments within metropolitan areas. Han and Strange (2015) also note the relevance of thin markets in housing.

[20] Real estate agents (Hendel et al. 2009) and construction delays (Davis and Heathcote 2007; Gyourko and Saiz 2006; Head et al. 2014) are also absent. Halket and di Custoza (2015) consider the impact of the rental sector whereas Haurin (1988) and Ngai and Sheedy (2020) assess idiosyncratic and potentially privately known house characteristics. Ngai and Tenreyro (2014) include increasing returns to matching.

[21] The gains to trade x are the difference in the buyer's and the seller's expected payoffs from homeownership. The match surplus is the gains to trade less the continuation values.

[22] Stein (1995), Genesove and Mayer (2020), Ortalo-Magné and Rady (2006), Hedlund (2016a), Hedlund (2016b), and Garriga and Hedlund (2020) evaluate the role of borrowing constraints and credit frictions in the housing market.

[23] Multiple Listing Service (MLS) realtor data specify nine housing quality classes. Focusing on the five more common housing quality classes allows the calibration to capture the most representative quality class-location submarkets.

[24] The quantitative analysis focuses on sales of existing homes. The NAR provides information on building permits and sales of new houses. Incorporating this evidence would not alter the general picture. New housing counts are small relative to existing homes and broadly follow the same evolution as existing homes.

[25] At the onset of the slump, the steady-state assumption that sales equal to $\hat{\beta} \times Duration$ may undercount the actual sales. The stock of buyers declines adding to the number of sales, so this calculation is arguably a lower bound.

[26] The DoM for sales is adjusted to be consistent with inventory refinements discussed in Appendix C.

[27] The seasonally adjusted coefficient of variation of sales comes from regressing of sales (listings) (from 2013 through 2019) on month dummies and a linear trend and then dividing the root mean squared error by the mean of the dependent variable.

[28] The calibration implies that motivated sellers will not enter if inventory in the submarket exceed 18 unsold houses in the slump period.

[29] Given the measurement issues arising from the limited coverage in the early years of CoreLogic data, the comparison from 2009 onward is

preferred for all price correlations. For the same reasons, Panels (B) and (C) in Figure 3 include the NAR series.

[30] Gilbukh (2023) similarly finds a tight correspondence between listings and sales, a declining sales hazard, and a large fall in prices by duration using MLS data.

[31] During this interval, the simulations more closely follow NAR sales that stay flatter until 2011. The average across geographic areas may, of course, mask different timings across space. Individual cities may experience some abrupt changes. A number of studies emphasize the local nature of much of the time-series variation in house prices that would lead to substantial variation in performances spread around various markets. For U.S. evidence, see Abraham and Hendershott (1996), Del Negro and Otrok (2007), and Head et al. (2014). For Canada, see Allen et al. (2009). The CoreLogic data from geographic markets reveal considerable cross-sectional variation that is explored in Section 6.2. Moreover, the evolution across cities is also heterogeneous. The majority of cities follow the aggregate pattern. They exhibit a flat hazard function from 2007–2011 although precise dates, magnitudes, and details vary to some extent. After 2012, the hazard function rises for all durations, but in the first 2 weeks on the market, the rise in the hazard rate is particularly pronounced. This pattern, however, is not universal. A small group of cities does not follow this pattern and maintains a flat function hazard over the sample period.

[32] The last row of Table 3 reports that the post-2009 correlation of data with the model in prices is nearly 0.8. The correlation almost halves to 0.411 over the entire sample period from 2007. A drastic plunge in financing as mortgage rates fell from 6.46 in October 2008 to 4.79 six months later in April 2009 may have buoyed prices from falling more quickly in the data.

[33] Arefeva (2025) finds that in the Redfin data between 2009 and 2016, roughly half of all buyers in California had a competing bid. Redfin itself also reports that during March–May 2021, the proportion of houses that sell within 2 weeks rises to 0.61—above the maximum in the simulated data.

[34] Suppose that sales between 45 and 60 days result exclusively from trades with sellers on the long side so that the associated Kaplan–Meier is an accurate representation of such trades. Assuming that sales between 16 and 30 days result from a mix of trades with sellers on the long side and short side, and that the trades with sellers on the long side occur at the same rate as the 45–60-day Kaplan–Meier. The excess of the 16–30-day CoreLogic Kaplan–Meier above the 45–60-day Kaplan–Meier then provides a basic measure of the number of sellers on the short side who spill over from the 0–15 day period into the 16–30 day period. Adding this value to the 0–15 day Kaplan–Meier (multiplied by the probability of not matching during the first 15 days) yields the adjustment.

[35] Employing a similar approach, Coles et al. (2010) use the stock-flow relationship between DoM for completed and uncompleted unemployment spells to identify and estimate monthly short and long Kaplan–Meier hiring probabilities in the U.K. labor market. Their analysis addresses a number of empirical challenges from using monthly data to estimate the continuous time stock-flow model.

[36] This result is robust to alternative values of the constant ratio in short- and long-side Kaplan–Meier selling probabilities and is also robust to alternative values for the initial starting level of inventory.

[37] As reported by the NAR, the average share of first-time homebuyers is 39% between 2003 and 2007 and is 48.5% between 2009 and 2010. This gives an increase in $\hat{\beta}$ of 18%.

[38] The share of withdrawals among all listings in the sample period varies from 4% to 15%. Out of all withdrawn houses, slightly over half (56%) are ever relisted. This includes relisted immediately and also relisted after 10 years of withdrawal.

[39] The motivated seller payoffs in the prosperous state could also reach and account for this threshold. However, the flow in of buyers in the propserous state is greater than the flow in of motivated sellers, so inventories of sellers will trend away from becoming large. The decisions of relaxed buyers keep a check on an excessive number of buyers accumulating in the submarket. In the quantitative simulations, inventories do not approach this cutoff threshold, so this issue is not a substantial concern.

[40] In Norway, sellers announce a house for sale by specifying an explicit viewing period followed by an auction with a precise starting time. In Oslo, sellers typically post an advertisement on the Friday 9 days before a Sunday open house. The bidding begins Monday after the open house. Note that 45.83% of sellers agree a price within the same day with almost 27% completing within the first 2 hours after the first bid. Another 23.48% complete on the second day. See Anundsen et al. (2023) for further details.

References

Abraham, J. M., and P. H. Hendershott. 1996. “Bubbles in Metropolitan Housing Markets.” *Journal of Housing Research* 7: 191–207.

Albrecht, J., A. Anderson, E. Smith, and S. Vroman. 2007. “Opportunistic Matching in the Housing Market.” *International Economic Review* 48: 641–664.

Allen, J., D. Amano, D. P. Byrne, and A. W. Gregory. 2009. “Canadian City Housing Prices and Urban Market Segmentation.” *Canadian Journal of Economics* 42: 1132–1149.

Andrews, M. J., D. Stott, S. Bradley, and R. Upward. 2013. “Estimating the Stock-Flow Matching Model Using Micro Data.” *Journal of the European Economic Association* 11: 1153–1177.

Anenberg, E., and P. Bayer. 2020. “Endogenous Sources of Volatility in Housing Markets: The Joint Buyer-Seller Problem.” *International Economics Review* 61: 1195–1228.

Anenberg, E., and D. Ringo. 2022. “The Propagation of Demand Shocks Through Housing Markets.” *American Economic Journal: Macroeconomics* 14: 481–507.

Anundsen, A. K., A. Benedictow, E. R. Larsen, and M. Walbækken. 2023. “Bidding Behavior in Housing Auctions.” Working Paper. Available at SSRN:. https://ssrn.com/abstract=4331120.

Arefeva, A. 2025. “Housing Markets: Auctions, Granular Shocks, and Microstructure Frictions.” *Journal of Banking & Finance* 180: 107553.

Bierdel, A., A. Drenik, J. Herreño, and P. Ottonello. 2023. “Illiquid Lemon Markets and the Macroeconomy.” Working Paper No. w31711, National Bureau of Economic Research.

Caplin, A., and J. Leahy. 2011. “Trading Frictions and House Price Dynamics.” *Journal of Money, Credit and Banking* 43: 283–303.

Case, K. E., and R. J. Shiller. 1997. “The Efficiency of the Market for Single-Family Homes.” *American Economic Review* 79: 125–137.

Coles, M., P. Jones, and E. Smith. 2010. “A Picture of Stock-Flow Unemployment.” *Macroeconomic Dynamics* 14: 427–453.

Coles, M. G. 2025. “Bidding Wars in the Housing Market: When Directed Search Meets Stockflow Matching.” Working Paper. Available at SSRN:. https://ssrn.com/abstract=5531063.

Coles, M. G., and A. Muthoo. 1998. “Strategic Bargaining and Competitive Bidding in a Dynamic Market Equilibrium.” *The Review of Economic Studies* 65: 235–260.

Coles, M. G., and B. Petrongolo. 2008. “A Test Between Stock-Flow Matching and the Random Matching Function Approach.” *International Economic Review* 49: 1113–1141.

Coles, M. G., and E. Smith. 1998. “Marketplaces and Matching.” *International Economic Review* 39: 239–254.

Cubbin, J. 1974. “Price, Quality, and Selling Time in the Housing Market.” *Applied Economics* 6: 171–187.

Davis, M. A., and J. Heathcote. 2007. “The Price and Quantity of Residential Land in the United States.” *Journal of Monetary Economics* 54: 2595–2620.

Díaz, A., and B. Jerez. 2013. “House Prices, Sales, and Time on the Market: A Search-Theoretic Framework.” *International Economics Review* 54: 837–872.

de Wit, E. R., and B. van der Klaauw. 2013. “Asymmetric Information and List-Price Reductions in the Housing Market.” *Regional Science and Urban Economics* 43: 507–520.

Del, Negro, M., and C. Otrok. 2007. “99 Luftballons: Monetary Policy and the House Price Boom Across U.S. States.” *Journal of Monetary Economics* 54: 1962–1985.

Favilukis, J., S. C. Ludvigson, and S. Van Nieuwerburgh. 2017. “The Macroeconomic Effects of Housing Wealth, Housing Finance, and Limited Risk Sharing in General Equilibrium.” *Journal of Political Economy* 125: 140–223.

Gabrovski, M., and V. Ortego-Marti. 2021. “Search and Credit Frictions in the Housing Market.” *European Economic Review* 134: 103699.

Garriga, C., and A. Hedlund. 2020. “Mortgage Debt, Consumption, and Illiquid Housing Markets in the Great Recession.” *American Economic Review* 110: 1603–1634.

Garriga, C., R. Manuelli, and A. Peralta-Alva. 2019. “A Macroeconomic Model of Price Swings in the Housing Market.” *American Economic Review* 109: 2036–2072.

Genesove, D., and C. J. Mayer. 2020. “Equity and Time to Sale in the Real Estate Market.” *American Economic Review* 87: 255–269.

Giacoletti, M. 2021. “Idiosyncratic Risk in Housing Markets.” *The Review of Financial Studies* 34: 3695–3741.

Gilbukh, S. 2023. “New Listing Alert: Alternative Theory of Housing Search.” Working Paper. Available at SSRN:. https://ssrn.com/abstract=4437501.

Glaeser, E. L., J. Gyourko, E. Morales, and C. G. Nathans. 2014. “Housing Dynamics: An Urban Approach.” *Journal of Urban Economics* 81: 45–56.

Gregg, P., and B. Petrongolo. 2005. “Stock-Flow Matching and the Performance of the Labor Market.” *European Economic Review* 49: 1987–2011.

Gyourko, J., and A. Saiz. 2006. “Construction Costs and the Supply of Housing Structure.” *Journal of Monetary Economics* 46: 661–680.

Halket, J., and M. P. M. di Custoza. 2015. “Homeownership and the Scarcity of Rentals.” *Journal of Monetary Economics* 76: 107–123.

Han, L., and W. C. Strange. 2015. “Chapter 13 - The Microstructure of Housing Markets: Search, Bargaining, and Brokerage.” In *Handbook of Regional and Urban Economics*, edited by G. Duranton, J. V. Henderson, and W. C. Strange, vol. 5 of Handbook of Regional and Urban Economics, 813–886, Elsevier.

Haurin, D. 1988. “The Duration of Marketing Time of Residential Housing.” *Real Estate Economics* 16: 396–410.

Head, A., and H. Lloyd-Ellis. 2012. “Housing Liquidity, Mobility, and the Labour Market.” *Review of Economic Studies* 79: 1559–1589.

Head, A., H. Lloyd-Ellis, and H. Sun. 2014. “Search, Liquidity, and the Dynamics of House Prices and Construction.” *American Economic Review* 104: 1172–1210.

Hedlund, A. 2016a. “Illiquidity and Its Discontents: Trading Delays and Foreclosures in the Housing Market.” *Journal of Monetary Economics* 83: 1–13.

Hedlund, A. 2016b. “The Cyclical Dynamics of Illiquid Housing, Debt, and Foreclosures.” *Quantitative Economics* 7: 289–328.

Hendel, I., A. Nevo, and F. Ortalo-Magné. 2009. “The Relative Performance of Real Estate Marketing Platforms: MLS Versus FSBO Madison.com.” *American Economics Review* 99: 1878–1898.

Indiviglio, D. 2011. “After an Ugly 2010, the Housing Market Won’t Look Much Better in 2011.” The Atlantic.

Jiang, E. X., N. Kotova, and A. L. Zhang. 2024. “Liquidity in Residential Real Estate Markets.” Working Paper. Available at SSRN:. https://ssrn.com/abstract=3386353.

Kaplan, G., K. Mitman, and G. L. Violante. 2020. “The Housing Boom and Bust: Model Meets Evidence.” *Journal of Political Economy* 128: 3285–3345.

R Krainer, J. 2001. “A Theory of Liquidity in Residential Real Estate Markets.” *Journal of Urban Economics* 49: 32–53.

Kuo, M.-Y., and E. Smith. 2009. “Marketplace Matching in Britain: Evidence from Individual Unemployment Spells.” *Labour Economics* 16: 37–46.

Lagos, R. 2000. “An Alternative Approach to Search Frictions.” *Journal of Political Economy* 108: 851–873.

Leung, C. K. Y., Y. C. F. Leong, and S. K. Wong. 2006. “Housing Price Dispersion: An Empirical Investigation.” *The Journal of Real Estate Finance and Economics* 32: 357–358.

Merlo, A., and F. Ortalo-Magné. 2004. “Bargaining Over Residential Real Estate: Evidence From England.” *Journal of Urban Economics* 52: 192–216.

G Miller, N. 1978. “Time on the Market and Selling Price.” *Real Estate Economics* 6: 164–174.

Moen, E. R., P. T. Nenov, and F. Sniekers. 2021. “Buying First or Selling First in Housing Markets.” *Journal of the European Economic Association* 19: 38–81.

Muellbauer, J., and A. Murphy. 1997. “Booms and Busts in the UK Housing Market.” *Economic Journal* 107: 1701–1727.

Ngai, L. R., and K. D. Sheedy. 2020. “The Decision to Move House and Aggregate Housing-Market Dynamics.” *Journal European Economic Association* 18: 2487–2531.

Ngai, L. R., and K. D. Sheedy. 2024. “The Ins and Outs of Selling Houses: Understanding Housing-Market Volatility.” *International Economic Review* 65: 1415–1440.

Ngai, L. R., and S. Tenreyro. 2014. “Hot and Cold Seasons in the Housing Market.” *American Economic Review* 104: 3991–4026.

Novy-Marx, R. 2009. “Hot and Cold Markets.” *Real Estate Economics* 37: 1–22.

Ortalo-Magné, F., and S. Rady. 2006. “Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints.” *Review of Economic Studies* 73: 459–485.

Piazzesi, M., M. Schneider, and J. Stroebel. 2020. “Segmented Housing Search.” *American Economic Review* 110: 720–759.

Pryce, G., and K. Gibb. 2006. “Submarket Dynamics of Time to Sale.” *Real Estate Economics* 34: 377–415.

Rekkas, M., L. Jiang, R. Wright, and Y. Zhu. 2025. “Price Dispersion and Price Stickiness in a Competitive Search Model of Housing Markets.”*Economic Theory* 1–36. https://doi.org/10.1007/s00199-025-01659-z.

J Shiller, R. 1982. “Consumption, Asset Markets, and Macroeconomic Fluctuations.” *Carnegie-Rochester Conference Series on Public Policy* 17: 203–238.

Shimer, R. 2007. “Mismatch.” *American Economic Review* 97: 1074–1101.

Smith, E. 2020a. “High and Low Activity Spells in Housing Markets.” *Review of Economic Dynamics* 36: 1–28.

Smith, E. 2020b. “Stock-Flow Models of Market Frictions and Search.” In *Oxford Research Encyclopedia of Economics and Finance*, edited by A. Banerjee and J. H. Hamilton. Oxford University Press.

C Stein, J. 1995. “Prices and Trading Volume in the Housing Market: A Model With Down-Payment Effects.” *Quarterly Journal of Economics* 110: 379–406.

R Taylor, C. 1995. "The Long Side of the Market and the Short End of the Stick: Bargaining Power and Price Formation in Buyers', Sellers' and Balanced Markets." *Quarterly Journal of Economics* 110: 837–855.

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

Appendix A: Data and Variable Construction

Data and Sample. The CoreLogic's MLS data for the years 2006–2019 contain property-level listings that are on the market for sale or for rent. The sample consists of single-family houses and townhouses listed for sale. The data set includes a large amount of information on each listing, but most importantly, it includes information on the geographic location, listing and closing dates, prices, and detailed characteristics of the listing. CoreLogic obtains the data directly from regional boards of realtors and provides standardized fields for some key variables to improve consistency across realtor boards. As of 2014, the data cover around 56% of all active listings nationwide. Occasionally, the same house listing appears multiple times in the data set, possibly due to listing agents entering duplicate listings. We identify duplicate listings through listing dates and property addresses. Dropping the duplicate listings gives the new listing-to-sales ratio close to the ratio in the NAR data. A small percentage of houses are relisted soon after being withdrawn. When a house is withdrawn and quickly relisted within 15 days, the two observations are concatenated using the first listing date. If a relisting occurs between 16 and 180 days, the relisting is dropped.[38] Relisting after 180 days begins a new observation. This adjustment has little effect other than to raise DoM marginally.

A sample is then selected consisting of 50 metropolitan markets with a sizeable population and enough listings over the sample period. The total sample size across 50 markets from 2006–2019 is over 11 million transactions. The average size of housing stock in each city is computed over the sample period and used as the weight when aggregating across cities. Figure A.1 shows the cities and relative weights in the sample. Figure A.2 displays the series in Figure 1 (excluding hedonic prices) without seasonal adjustment, that is, without taking the 12-month moving average. Figure A.3 extends Figure 2 to display the Kaplan–Meier statistics and prices by 15-day durations out to 120 days.

Kaplan–Meier Statistics. The Kaplan–Meier statistics are defined over 15-day intervals for houses that *go under contract* after being on the market for 0–15 days, 16–30 days, …, 106–120 days. Crucially, the statistics are constructed based on the time that a house goes under contract, which avoids the lengthy period between contract and closing. For each month, define:

- 0–15 duration: A house is coded 0 if it is listed in the first 0–15 days of the month; and 1 if it also goes under contract in the month and 0–15 days after listing; missing otherwise.
- 16–30 duration: A house is coded 0 if it is listed 0–15 days before the month; and 1 if it also goes under contract in the month and 16–30 days after listing; missing otherwise.
- 31–45 duration: A house is coded 0 if it is listed 16–30 days before the month; and 1 if it also goes under contract in the month and 31–45 days after listing; missing otherwise.…
- 106–120 duration: A house is coded 0 if it is listed 91–105 days before the month; and 1 if it also goes under contract in the month and 106–120 days after listing; missing otherwise.

Houses are counted for the relevant Kaplan–Meier only if they are not potentially right or left-censored. For example, for the 0–15 day Kaplan–Meier, houses that enter in the second half of the month with less than 15 days in the month remaining are coded as missing. Houses that entered in the previous month with 20 days remaining and were not sold in that month are also coded as missing for subsequent month's 16–30 day hazard as they only qualify 10 out of a possible 15 days for a 16–30 day hazard. The Kaplan–Meier statistic for each duration in each month is then computed as the average across all houses coded 0 or 1 for the month and duration. Figure A.3 shows the time series for the full set of these unconditional statistics, averaged across cities.

Conditional Kaplan–Meier statistics. The conditional Kaplan–Meier statistics is constructed using coefficients from the following logit regression, separately for each city:

$$K_{it} = \Omega_j X_{it} + \sum_d \theta^d_{jt} D^d_{it} + \sum_d \gamma^d_{jm} D^d_{it} + \zeta_z + \epsilon_{it}, \tag{A.1}$$

where X_{it} is a vector of housing characteristics for house i at the time (month-year) t and includes the age of the house (and squared), numbers of bedrooms, baths and fireplaces (all winsorized), living area, indicators for whether the property has a large lot, has a pool, is in distress (short-sale, REO, or foreclosure); ζ_z is fixed effect for zip code to control for within-city geographical variations (e.g., schools, access to transportation). The dependent variable K_{it} is an indicator for whether

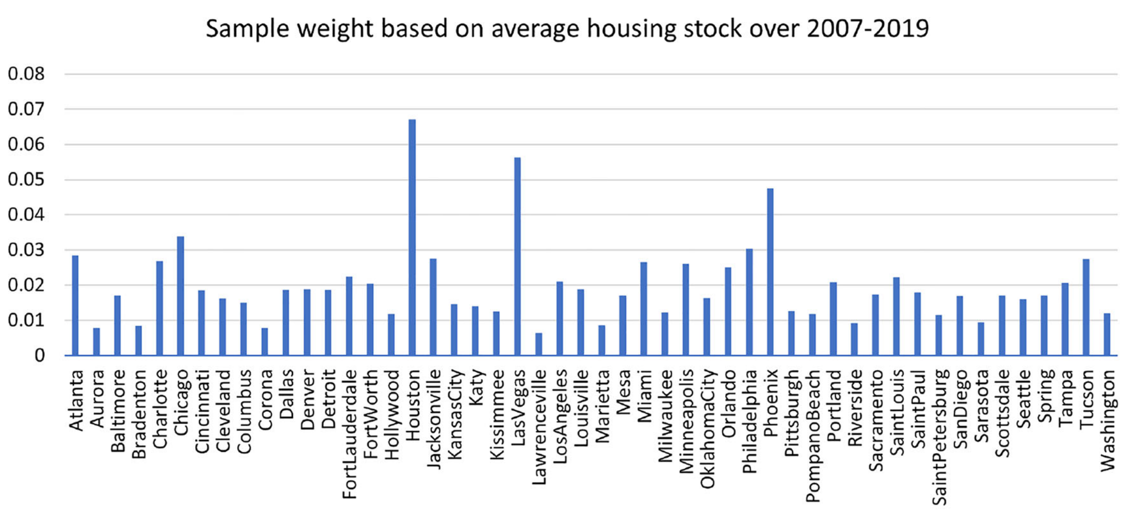

FIGURE A.1 | Cities and sample weights. *Notes:* Sample weights are calculated based on average housing stock over 2007–2019 from the CoreLogic data.

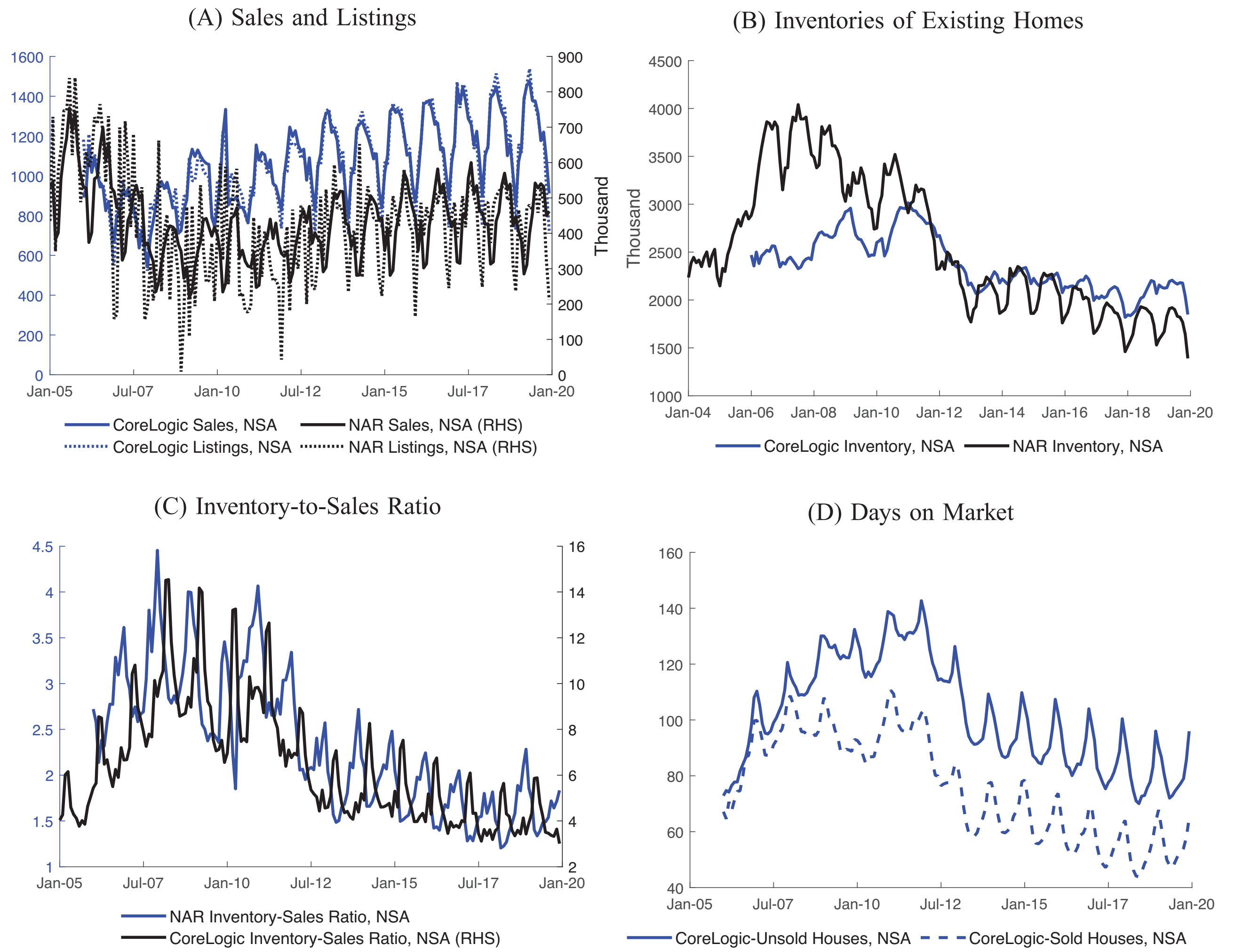

FIGURE A.2 | Nonseasonally adjusted (NSA) aggregate trading patterns. *Notes:* All series labeled "CoreLogic" are the authors' calculations based on CoreLogic data.

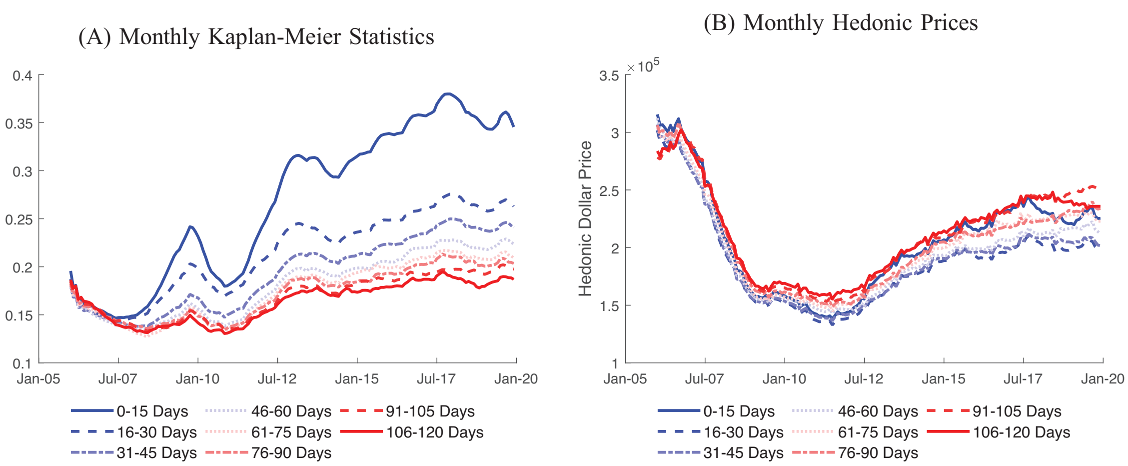

FIGURE A.3 | Complete set of Kaplan–Meier statistics and hedonic prices by duration. *Notes:* The plotted series are the authors' calculations based on CoreLogic data.

TABLE A.1 | Comparison of conditional and unconditional Kaplan–Meier.

Variable	Source	Corr with model	Min	Max	% change max-min	Corr with price (2009 onward)	Std Dev (log)
Short Kaplan–Meier	Unconditional	0.840	0.147	0.380	88.7	0.934	0.309
	Conditional	0.834	0.148	0.443	99.8	0.928	0.317
Long Kaplan–Meier	Unconditional	0.940	0.138	0.250	57.8	0.938	0.201
	Conditional	0.935	0.127	0.316	85.6	0.917	0.226

Note: The reported statistics are the authors' calculations based on CoreLogic data. (1) Data for unconditional Kaplan–Meier is January 2007–December 2019, and for conditional Kaplan–Meier is January 2007–June 2019. (2) The unconditional Kaplan–Meier is seasonally adjusted by taking a ±6-month moving average. (3) % change max-min uses the average of max and min as a denominator.

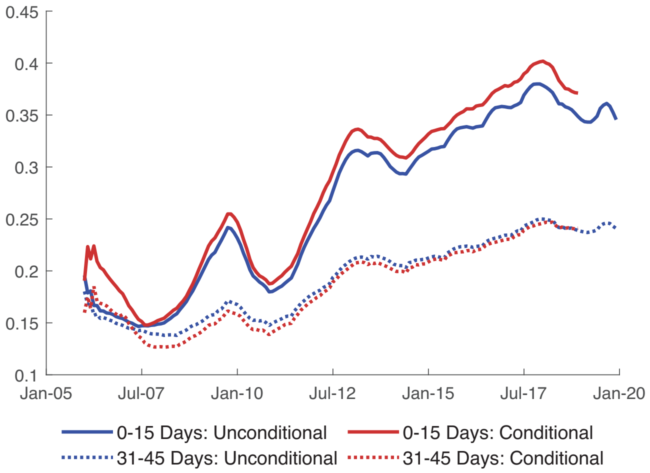

FIGURE A.4 | Unconditional versus conditional Kaplan–Meier. *Notes:* The plotted series are the authors' calculations based on CoreLogic data.

a house is sold or not at time t. The set of main independent variables D^d_{it} represents the duration category dummy in month t for duration d defined over 15-day intervals. The set of estimated parameters θ^d_{jt} is allowed to vary by month t to capture the variation in the effect of duration over time and is estimated separately for each city j. The set of parameters γ^d_{jm} varies by month of the year m and city j and captures any seasonality effect in the variation of D^d_{jt}. The conditional statistic for each city j in month t is then computed as

$$\hat{K}_{jt} = \left[1 + e^{-(\theta^d_{jt} + \Omega_j \bar{X}_j + \bar{\zeta}_z)}\right]^{-1},$$

where $\bar{X}_j$ is the median value of the housing characteristic variable for city j; $\bar{\zeta}_j$ is the median zip code fixed effects among houses in the city; θ^m_{jt} and Ω_j are parameters estimated from (A.1). Essentially, the conditional Kaplan–Meier is the city-level average probability that a house on the market for a certain duration sells in a month, controlling for housing characteristics and zip code. Houses that were listed in late 2019 and went under contract in 2020 were likely exposed to the shock of the pandemic, which is outside the consideration of the analysis here. Including those houses in the estimation sample would affect the estimated coefficients on the duration category dummy and thus apply the pandemic effect to all the conditional estimates during 2006–19. For this reason, the estimation of the conditional statistics excludes houses that went under contract in 2020, and the conditional Kaplan–Meier is estimated up to June 2019.

Figure A.4 and Table A.1 compare the unconditional and conditional Kaplan–Meier statistics for the short (0–15 day) and long (31–45 day) durations. The conditional statistics take into account variations in housing characteristics (captured by X_{it} in Equation (A.1)) and house locations (captured by zipcode). Nonetheless, Table A.1 shows that the conditional statistics are very close to the unconditional ones both in the scale of variation, variability, and correlation with prices. Figure 3 shows that on both the short and long sides, the conditional and unconditional Kaplan–Meier statistics comove.

Appendix B: The Stock-Flow Model with Endogenous Entry and Aggregate Shocks

This appendix derives equilibrium outcomes in one of the many independent submarkets of the stock-flow model described in Section 4. For added generality, the analysis here allows for positive buyer flow search costs $d_b > 0$. These costs are set equal to zero in the text and in the quantitative analysis.

As discussed in Section 4 and derived formally in Coles and Muthoo (1998), buyer and seller decisions involve immediate trade when feasible. This result leads to following:

Assertion A.1. *Immediate trade occurs so that the submarket never simultaneously has waiting buyers and waiting sellers.*

To establish the transition process for (N, D), define seller entry as the arrival of a seller who decides to sell. Given buyer entry $i \in \mathbb{N}$ since the previous auction and a vanishing interval length $dt \to 0$, transitions in the submarket for $N \geq 0$ follow

$$\begin{aligned}\Gamma(N,D) &= (N, D+dt) \mid \textit{No Seller Entry}\\ &= (N', 0) \quad \mid \textit{Seller Entry}\\ &= (N', dt) \quad \mid \textit{Auction with 1 Seller},\end{aligned}$$

where $N' = N + i - 1$. If seller entry does not take place, D evolves incrementally with time, whereas N remains unaltered as buyer entry i remains obscured. If seller entry occurs, the duration is reset to $D = 0$ and N is updated before the bidding occurs to include not only revealed buyer entry i but also the seller's own entry. Seller entry triggers an auction knowing the updated N'. The seller immediately receives all potential bids, if any. If there are buyers present, the seller either accepts one or rejects all bids. Time then proceeds incrementally.

For $N \leq -1$, entry of either a buyer or seller also resets $D = 0$. Seller entry prompts an auction albeit without a buyer present in the submarket in which case inventories increase by one house and time proceeds incrementally. Buyer entry leads to a single bidder auction for the inventory of $-N$ houses. Transitions in the submarket for $N \leq -1$ follow

$$\begin{aligned}\Gamma(N,D) &= (N, D+dt) && \mid \textit{No Buyer or Seller Entry}\\ &= (N-1, dt) && \mid \textit{Seller Entry}\\ &= (N+1, 0) && \mid \textit{Buyer Entry}\\ &= (N+1, dt) && \mid \textit{Auction with } 1 \textit{ buyer}.\end{aligned}$$

B.1 | Prosperity

The analysis of equilibrium outcomes now turns to the prosperous state and closely follows Smith (2020a). Without trading delays, time periods between observed sales result from waiting times between seller entry (their arrival and deciding to sell) when there are excess buyers, or between buyer entry if there are excess sellers. When relaxed sellers turn down selling opportunities in the prosperous state, these waiting times become prolonged. Let the decision of relaxed sellers to sell therefore define hot and cold phases of submarket activity in the prosperous state.

Definition. During prosperity trading, a submarket is hot when relaxed sellers accept the option to sell, and is cold when relaxed sellers decide not to sell.

Suppose that the transaction price falls as the number of bidders declines. Since the submarket becomes less profitable for sellers with fewer bidders, relaxed potential sellers will decide not to sell (given a sufficiently high F) for some period of time after an auction that leaves the number of remaining buyers below a zero expected profit threshold that reflects the competitiveness of bidding. The proposed outcome is that relaxed sellers decide to sell with two or more known available bidders but relaxed sellers decide not to sell for some period of time following an auction with one remaining buyer who bids monopolistically not competitively.

Assertion A.2 (Monopoly–Bertrand partition). *The submarket is hot for all $N \geq 2$. For $N = 1$ and $D = dt$, the submarket is cold.*

This assertion implies that the submarket becomes cold immediately following an auction with two or fewer bidders. The simulations check that this assertion holds in the quantitative assessment of the model.

To revitalize expected profit and relaxed seller entry so that cold submarkets eventually become hot, buyer entry must replenish the pool of bidders. Immediately after an auction, the number of potential bidders is known with certainty. As time proceeds, random buyer entry occurs. If no houses are available for sale, no auctions occur and potential sellers do not observe buyer entry. Since buyers enter at Poisson rate β, buyers and sellers (who have yet to decide to sell) have common beliefs that the probability of i buyer entrants over a duration D since the last auction is

$$\pi_i(D) = \frac{e^{-\beta D}(\beta D)^i}{i!}$$

As time proceeds without trade, the number of potential buyers grows stochastically. Expected prices rise. For submarkets with zero or one known buyer, the payoff to deciding to sell eventually becomes profitable. Relaxed seller entry resumes and the submarket becomes hot. The selling process coupled with Assertion A.2 implies that cold submarkets with $N \in \{0,1\}$ transition to a hot submarket after endogenous durations $D = T_0$ and $D = T_1$, respectively.

For $N \in \{0,1\}$, the transition from a cold phase to a hot phase comes about in one of two ways. The cold phase may conclude after a sufficiently long period of submarket inactivity without any (motivated) seller entry. After some length of time, either $D = T_0$ or $D = T_1$, seller expectations of (unobserved) buyer entry eventually improve enough to induce relaxed sellers to try to sell. In addition, during the cold phase, a motivated seller might enter and trigger an auction. The outcome of this auction reveals the number of buyers who have entered during the cold phase of this submarket and hence resets the decision to sell of relaxed sellers. If the auction reveals a sufficient number of remaining bidders ($N \geq 2$), entry of relaxed sellers becomes re-activated. If not, the waiting decision resets itself to the beginning of the cold phase conditional on the number of bidders. If seller entry occurred but no sale followed, the inventory of available homes builds up.

Relaxed sellers will not enter submarkets with excess supply until all of the previous sellers who entered consummate trades. Entry of motivated sellers, however, occurs during both hot and cold phases. Even though motivated sellers enter at a slower rate than buyers, from time to time the realization of the entry processes will be such that more motivated sellers than buyers enter and cold submarkets will experience having excess sellers. Really cold submarkets, those with excess sellers $N \leq -1$, remain cold until all inventories are sold and enough time has elapsed to revitalize seller entry such that $N = 0$ and $D > T_0$. Hence, A.2 implies the set of hot submarkets is given by

$$\Omega = \{(N \geq 2, D \geq 0) \cup (N = 1, D \geq T_1) \cup (N = 0, D \geq T_0)\}.$$

All other (N, D) are cold.

From the transition function Γ, $D = 0$ implies that an auction takes place. If $N \geq 0$, there is one house for sale and $N + 1$ buyers bidding. If buyer entry occurred and $N < 0$, then there is one bidder and $-N + 1$ houses for sale. Let $P(N)$ denote the buyers' symmetric price bids in these auctions. Let $H(N)$ represent the payoff to a house buyer from being in a hot submarket with $N \geq 1$ bidders (including the buyer) waiting for the arrival of a seller. Let $C(T; N \geq 1, D)$ represent the expected payoff to a buyer in a cold submarket who must wait the remaining duration $T > 0$ before the submarket becomes hot again and the entry of relaxed sellers resumes. For $N \leq -1$, let $Z(N)$ denote the sellers' payoff of waiting in a cold submarket for the next buyer entry and subsequent auction.

After a seller accepts a bid in an auction with two buyers, one buyer remains in the submarket. The submarket becomes cold. The buyer who does not purchase the house receives expected payoff $C(T_1; N = 1, D = 0)$ where again T_1 denotes the duration that this seller must wait after the sale before seller entry resumes. From Ω, the submarket is hot for $N \geq 2$. Waiting buyers are absent for $N \leq 0$ making $C(T_1; N = 1, D = 0)$ the only cold phase payoff relevant to the buyer.

To solve the auction stages, note that the seller's acceptance payoff is strictly increasing in price. Hence, the seller's best response strategy in an auction with one bidder has the reservation property. In auctions where $N \leq -1$, the lone bidder optimally offers the two or more sellers the payoff to waiting $P(N) = Z(N)$. With two or more bidders ($N \geq 1$) price offers are derived from the hot and cold submarket payoffs $H(N)$ and $C(T_1)$ where to ease notation, the understood $N = 1, D = 0$ in $C(T_1)$ is now dropped. Following Taylor (1995) for $N \geq 1$, the buyers are now the ones who become indifferent between paying $P(N)$ and waiting for the next seller to enter, whether in a cold submarket for $P(1)$ or in hot submarket for $P(N)$ with $N \geq 2$. These monopolistic and competitive bidding scenarios yield the following characterization of prices offers:

$$\begin{aligned}P(N) &= Z(N), \quad N \leq -1\\ P(N) &= x - C(T_1), \quad N = 1\\ P(N) &= x - H(N), \quad N \geq 2.\end{aligned}$$

For $N = 0$, a lone buyer will optimally offer the lone seller the reservation price $P(0)$ that equates the seller's payoff of accepting this offer with the seller's payoff to waiting the next dt to repeat holding an auction. Note that sellers are willing to accept these $P(N)$ bids.

For $N = 1$ and $N = 2$, a buyer's expected payoffs to waiting in a hot submarket can be written as

$$H(1) = \frac{1}{1+rdt}\left[(\alpha+\sigma)dt(x - P(0)) + \beta dt\, H(2) + (1-(\alpha+\sigma+\beta)dt)H(1) - d_b dt\right],$$

$$H(2) = \frac{1}{1+rdt}\left[(\alpha+\sigma)dt\, C(T_1) + \beta dt\, H(3) + (1-(\alpha+\sigma+\beta)dt)H(2) - d_b dt\right].$$

For $N \geq 3$,

$$H(N) = \frac{1}{1+rdt}[(\alpha+\sigma)dt\, H(N-1) + \beta dt\, H(N+1) + (1-(\alpha+\sigma+\beta)dt)H(N) - d_b dt].$$

The solution to these difference equations is given by

$$H(N) = \eta^{N-2}[H(2) + d_b/r] - d_b/r,$$

$$H(2) = \frac{(\alpha+\sigma)C(T_1) - (r+\beta(1-\eta))d_b/r}{r+\alpha+\sigma+\beta(1-\eta)},$$

$$H(1) = \frac{\beta H(2) + (\alpha+\sigma)(x-P(0)) - d_b}{r+\alpha+\sigma+\beta},$$

where

$$\eta = \frac{r+\alpha+\sigma+\beta - [(r+\alpha+\sigma+\beta)^2 - 4(\alpha+\sigma)\beta]^{1/2}}{2\beta} < 1.$$

The payoff to waiting in a cold phase is more involved. With probability $\alpha e^{-\alpha t}dt$, a motivated seller enters the submarket during the cold phase after a duration t and triggers an auction with the existing buyers and any other buyers who might have entered during the cold phase up to time t. In this environment, the buyer's expected payoff in a cold submarket can be written as

$$C(T_1) = \int_0^{T_1} e^{-(r+\alpha)t}\left[-d_b + \alpha\sum_{i=0}^{\infty}\pi_i(t)[x - P(i+1)]\right]dt + e^{-(r+\alpha)T_1}\sum_{i=0}^{\infty}\pi_i(T_1)H(i+1).$$

Solving gives

$$C(T_1) = \frac{\Psi_1(x-P(0)) + \Psi_3 d_b/r}{1-\Psi_2},$$

where

$$\Psi_1 = A_1 + A_2,$$
$$\Psi_2 = (\eta - D_4)A_3 + D_4A_4 - D_4A_5 + A_6,$$
$$\Psi_3 = D_5A_3 - D_5A_4 + D_5A_5 - A_7 - A_8,$$

$$A_1 = \frac{\alpha(1-e^{-(r+\alpha+\beta)T_1})}{r+\alpha+\beta},$$

$$A_2 = e^{-(r+\alpha+\beta)T_1}D_1,$$

$$A_3 = \frac{\alpha\beta[1-e^{-(r+\alpha+\beta)T_1}(1+(r+\alpha+\beta)T_1)]}{\eta(r+\alpha+\beta)^2},$$

$$A_4 = \frac{\alpha(1-e^{-(r+\alpha+\beta(1-\eta))T_1})}{\eta^2(r+\alpha+\beta(1-\eta))},$$

$$A_5 = \frac{\alpha(1-e^{-(r+\alpha+\beta)T_1})}{\eta^2(r+\alpha+\beta)},$$

$$A_6 = e^{-(r+\alpha+\beta)T_1}\left(D_2 + \frac{e^{\beta\eta T_1}-1}{\eta}\right)D_4,$$

$$A_7 = e^{-(r+\alpha+\beta)T_1}\left[rD_2D_5 + rD_3 + \frac{e^{\beta\eta T_1}-1}{\eta}(rD_5 - 1) + e^{\beta T_1} - 1\right],$$

$$A_8 = -(1-e^{-(r+\alpha)T_1}) + \frac{\alpha(1-e^{-(r+\alpha+\beta)T_1})}{r+\alpha+\beta},$$

$$D_1 = \frac{\alpha+\sigma}{r+\alpha+\sigma+\beta},$$

$$D_2 = \frac{\beta}{r+\alpha+\sigma+\beta},$$

$$D_3 = \frac{1}{r+\alpha+\sigma+\beta},$$

$$D_4 = \frac{\alpha+\sigma}{r+\alpha+\sigma+\beta(1-\eta)},$$

$$D_5 = \frac{r+\beta(1-\eta)}{r+\alpha+\sigma+\beta(1-\eta)}.$$

Like buyers in submarkets with excess buyers, sellers in submarkets with other houses accept bids that make them indifferent between trading and waiting for the next auction, $P(N) = Z(N)$ for all $N \leq -1$. Since α alone governs the arrival rate of sellers in this situation, all of them motivated when there are excess sellers, the payoff to a lone seller in the submarket ($N = -1$) awaiting for the arrival of buyer is given by

$$Z(-1) = \frac{1}{1+rdt}[\alpha dt\, Z(-2) + \beta dt\, P(0) + (1-\alpha dt - \beta dt)Z(-1) - d_s dt].$$

With other sellers waiting the arrival of a buyer ($N \leq -2$), the payoff $Z(N)$ is given by

$$Z(N) = \frac{1}{1+rdt}\left[\alpha dt\, Z(N-1) + \beta dt\, Z(N+1) + (1-\alpha dt - \beta dt)Z(N) - d_s dt\right].$$

The solution to these difference equations is given by

$$Z(-1) = \frac{\beta P(0) - (r+\alpha(1-\lambda))d_s/r}{r+\alpha+\beta-\alpha\lambda}$$

and for $N \leq -2$

$$Z(N) = \lambda^{-N-1}[Z(-1) + d_s/r] - d_s/r,$$

where

$$\lambda = \frac{r+\alpha+\beta - [(r+\alpha+\beta)^2 - 4\alpha\beta]^{1/2}.}{2\alpha}$$

If the auction has one bidder for one house, the buyer's offer $P(0)$ makes the seller again indifferent between waiting and accepting. As buyer entry in the next instant will leave one of the two buyers in a cold submarket whereas seller entry in the next instant will leave one house left in inventory, the equilibrium bid with one buyer and one seller satisfies

$$P(0) = \frac{1}{1+rdt}\left[\alpha dt\, Z(-1) + \beta dt\,(x - C(T_1)) + (1-\alpha dt - \beta dt)P(0) - d_s dt\right].$$

Substitution for $Z(-1)$ gives

$$P(0) = \frac{\beta(r+\alpha+\beta-\alpha\lambda)[x - C(T_1)] - [(r+\alpha)(r+\alpha(1-\lambda)) + r\beta]d_s/r}{(r+\alpha+\beta-\alpha\lambda)(r+\alpha+\beta) - \alpha\beta}.$$

Plugging in for $C(T_1)$ yields

$$P(0) = \frac{B_1(1-\Psi_1-\Psi_2)x - B_1\Psi_3 d_b - (1-\Psi_2)B_2 d_s/r}{1-\Psi_2-B_1\Psi_1},$$

where

$$B_1 = \frac{\beta(r+\alpha+\beta-\alpha\lambda)}{(r+\alpha+\beta)(r+\alpha+\beta-\alpha\lambda)-\alpha\beta},$$
$$B_2 = \frac{(r+\alpha)(r+\alpha(1-\lambda))+r\beta}{[(r+\alpha+\beta)(r+\alpha+\beta-\alpha\lambda)-\alpha\beta]}.$$

Now consider the relaxed seller's decision to enter the submarket. For $N \geq 0$, the seller's revenue before deciding whether or not to sell is uncertain since unobserved buyer entry can occur between trades. If $N \geq 1$, a potential house seller knows a sale will occur immediately—there is at least one buyer in the submarket—but not the transaction price. Following a cold phase with one buyer of duration T_1 without any seller entry, expected revenue less the sunk fee F for a seller contemplating whether to sell in the submarket is given by

$$R(N=1,T_1) = \sum_{i=0}^{\infty} \pi_i(T_1)P(i) - F.$$

If $i \geq 0$ buyers arrived during D to replenish the submarket, the anticipated price $P(i)$ accounts for both i revealed buyers and the seller's own entry.

Similarly, following a cold phase of duration T_0 with no known buyer in the submarket and without any seller entry, expected revenue less the fee F for a seller deciding whether to sell in the submarket is given by

$$R(N=0,T_0) = \pi_0(T_0)Z(-1) + \sum_{i=1}^{\infty} \pi_i(T_0)P(i-1) - F.$$

If no buyer arrived during the cold phase, the seller must wait with associated payoff $Z(-1)$. If one buyer arrived since the last auction, seller entry balances the submarket at $N = 0$. The lone bidder offers $P(0)$ and the transaction takes place immediately. Likewise for two or more buyers with the number of bidders all coming from new buyers. By Assertion A.2, revenue is positive and relaxed seller entry occurs for $N \geq 2$. For $N \leq -1$, expected revenue $R(-N, D)$ to seller entry is negative.

These two expected net revenues determine the corresponding durations T_1 and T_0. Suppose first that there is one remaining buyer from the last auction, $N = 1$. If no new buyers have entered since the previous auction ($i = 0$), seller entry decreases N and the lone buyer bids the price $P(0)$. With two buyers (one old and one new), the seller receives $P(1) = x - C(T_1)$ where the unsuccessful buyer expects sellers to delay entry for the period of duration T_1. With three or more buyers ($i \geq 2$), the price offered and paid is $P(i) = x - H(i)$. Plugging in these outcomes along with the Poisson probabilities $\pi_i(t)$, expected profit becomes

$$R(1,D) = e^{-\beta D}P(0) + \beta D e^{-\beta D}[x - C(T_1)] + \sum_{i=2}^{\infty} \frac{(\beta D)^i e^{-\beta D}}{i!}[x - H(i)] - F.$$

Entry occurs if and only if $R(1,D) \geq 0$.

When there are no buyers remaining from the last auction ($N = 0$), entry occurs if and only if the corresponding cold phase duration T_0 is sufficiently long so that $R(0,D) \geq 0$. Plugging in again for prices $P(N)$ as well as for $Z(-1)$ delivers the following result.

Lemma 1. *For sellers aware of only one or zero known buyers, the critical cold phase duration of delayed entry T_1 and T_0, respectively, solves*

$$R(1,T_1) = e^{-\beta T_1}P(0) + \beta T_1 e^{-\beta T_1}[x - C(T_1)] + \sum_{i=2}^{\infty} \frac{(\beta T_1)^i e^{-\beta T_1}}{i!}[x - H(i)] - F = 0$$

and

$$R(0,T_0) = e^{-\beta T_0}\frac{\beta P(0)}{r+\beta} + \beta T_0 e^{-\beta T_0}P(0) + \frac{\beta T_0^2}{2}e^{-\beta T_0}(x - C(T_1)) + \sum_{i=3}^{\infty} \frac{(\beta T_0)^i e^{-\beta T_0}}{i!}[x - H(i)] - F = 0.$$

The buyer and seller payoff functions determine price offers. Price offers determine revenue. Revenue for $N \in \{0, 1\}$ determines entry. The last step is to verify that the monopoly-Bertrand partition in A.2 is valid. Given prices, payoffs, and cold phases, sellers would enter monopolistic ($N \leq 1$) submarkets and the submarket would remain hot if discounted expected net revenue is positive immediately following an auction that left one buyer remaining in the submarket:

$$R(1; dt) = P(0) - F > 0.$$

Conversely, submarkets with two known remaining buyers would become cold and entry would not occur if the expected price did not cover fixed costs:

$$P(1) = x - C(T_1) \leq F.$$

Smith (2020a) derives explicit restrictions on parameters for zero search costs. Here, it suffices to check the simulated values.

B.2 | Slump

The only change in the economic environment during the slump is that the arrival rate of buyers falls to $\hat{\beta} < \alpha$. Agents expect this aggregate slump state to end with Poisson arrival rate τ in which case the arrival rate of buyers returns indefinitely to the prosperity level $\beta > \alpha$.

The proposed equilibrium outcome is that relaxed sellers decide not to sell during the slump. When the slump hits, some submarkets will have existing known buyers that potentially make relaxed entry profitable. To confirm this proposed outcome, the quantitative analysis simply checks that at the start of the slump, the new equilibrium price given the number of existing buyers is less than F.

The determination of the remaining buyer offers and of waiting buyer and waiting seller payoffs in the slump follow the same logic as in the prosperous state. Use "hats" to represent, when needed, the aggregate slump state. Without relaxed seller entry during the slump, the waiting buyer payoffs for $N \geq 2$ become

$$\hat{H}(N) = \frac{1-\tau dt}{1+r dt}\left[\alpha dt\, \hat{H}(N-1) + \beta dt\, \hat{H}(N+1) + (1 - \alpha dt - \beta dt)\hat{H}(N)\right] + \tau dt\, H(N) - d_b dt,$$

which can be rewritten as

$$\beta\hat{H}(N+2) - (r+\tau+\alpha+\beta)\hat{H}(N+1) + \alpha\,\hat{H}(N-1) = -\tau H(N+1) + d_b = -\tau\eta^{N-1}[H(2) + d_b/r] + (r+\tau)d_b/r.$$

The solution to this difference equation is given by

$$\hat{H}(N) = \hat{\eta}^{N-2}C_b + A_b\eta^{N-2}[H(2) + d_b/r] - d_b/r,$$

where

$$\hat{\eta} = \frac{r+\tau+\alpha+\beta-[(r+\tau+\alpha+\beta)^2-4\alpha\beta]^{1/2}}{2\beta} < 1,$$

$$A_b = -\tau\eta/[\beta\eta^2-(r+\tau+\alpha+\beta)\eta+\alpha],$$

$$C_b = \hat{H}(2) - A_b(H(2)+d_b/r) - d_b/r.$$

The solution specifies $\hat{H}(2)$ as opposed to $\hat{H}(1)$ to avoid confusion with the role of $C(T_1)$.

For $N = 2$, a buyer's expected payoff

$$\hat{H}(2) = \frac{1-\tau dt}{1+rdt}\left[\alpha dt\,\hat{H}(1)+\beta dt\,\hat{H}(3) + (1-\alpha dt-\beta dt)\hat{H}(2)\right] + \tau dt\,H(2) - d_b dt$$

can be rewritten as

$$\beta\hat{H}(3) - (r+\tau+\alpha+\beta)\hat{H}(2) + \alpha\,\hat{H}(1) = -\tau H(2) + d_b = -\tau[H(2)+d_b/r] + (r+\tau)d_b.$$

Plugging in for $\hat{H}(3)$ gives

$$\hat{H}(2) = \frac{\alpha\hat{H}(1)}{r+\tau+\alpha+\beta(1-\hat{\eta})} + G_1,$$

where

$$G_1 = \frac{(\tau-\beta(\hat{\eta}-\eta)A_b)(H(2)+d_b/r) - (r+\tau+\beta(1-\hat{\eta}))d_b/r}{r+\tau+\alpha+\beta(1-\hat{\eta})}.$$

To keep the analysis tractable assume that on the rare occasion when the recovery occurs with one known buyer in the submarket, buyers anticipate a cold phase for the duration T_1 and payoff $C(T_1)$. In the quantitative section, T_1 is only 7.1 days. The seller's outside option does not directly impact the derivation of buyer payoffs but the payoffs indirectly adjust with the determination of the balanced trade price $\hat{P}(0)$. For $N = 1$, a buyer's expected payoffs in a hot submarket

$$\hat{H}(1) = \frac{1-\tau dt}{1+rdt}\left[\alpha dt\,(x-\hat{P}(0)) + \beta dt\,\hat{H}(2) + (1-\alpha dt-\beta dt)\hat{H}(1)\right] + \tau dt\,C(T_1) - d_b dt$$

can be rewritten as

$$\beta\hat{H}(2) - (r+\tau+\alpha+\beta)\hat{H}(1) = -\alpha(x-\hat{P}(0)) - \tau C(T_1) + d_b.$$

Plugging in for $\hat{H}(2)$ gives

$$\frac{\alpha\beta\hat{H}(1)}{r+\tau+\alpha+\beta(1-\hat{\eta})} - (r+\tau+\alpha+\beta)\hat{H}(1) = -\alpha(x-\hat{P}(0)) - \tau C(T_1) - \beta G_1 + d_b,$$

where

$$G_1 = \frac{(\tau-\beta(\hat{\eta}-\eta)A_b)(H(2)+d_b/r) - (r+\tau+\beta(1-\hat{\eta}))d_b/r}{r+\tau+\alpha+\beta(1-\hat{\eta})}.$$

Rearranging and solving gives

$$\hat{H}(1) = G_2\left\{\alpha(x-\hat{P}(0)) + \tau C(T_1) + \beta G_1 - d_b\right\},$$

where

$$G_2 = \frac{r+\tau+\alpha+\beta(1-\hat{\eta}))}{(r+\tau+\alpha+\beta(1-\hat{\eta}))(r+\tau+\alpha+\beta)-\alpha\beta}.$$

Note that

$$\hat{P}(1) = x - \hat{H}(1) = x - \alpha G_2\hat{P}(0) - G_2\{\alpha x + \tau C(T_1) + \beta G_1 - d_b\}.$$

Payoffs for motivated sellers during the slump account for their outside option. A critical threshold value of inventory of unsold houses in the submarket triggers nonentry by the incoming motivated sellers. To find this critical value, consider the solution method to the second order difference equations governing seller payoffs. In general, two roots and two "integration" constants solve these equations. With unbounded inventory, the solution involves setting the unknown constant associated with the explosive root to zero to avoid unbounded payoffs. With the addition of an outside option capping entry, the constant associated with the explosive root becomes a nonzero object so that the solution joins up with the outside option V minus the fixed cost F. In particular, recalling that the inventory for sale corresponds to negative values of N, an upper inventory bound requires $N \geq \bar{N}$. The two root solution not only satisfies the initial condition (determined by balanced trade), but also the condition that at $\bar{N} \leq 0$, the payoff equals $V - F$.[39]

The seller payoffs in the slump for $N \leq -2$

$$\hat{Z}(N) = \frac{1-\tau dt}{1+rdt}\left[\alpha dt\,\hat{Z}(N-1) + \beta dt\,\hat{Z}(N+1) + (1-\alpha dt-\beta dt)\hat{Z}(N)\right] + \tau dt\,Z(N) - d_s dt$$

can be rewritten as

$$\alpha\hat{Z}(N-2) - (r+\tau+\alpha+\beta)\hat{Z}(N-1) + \beta\,\hat{Z}(N) = -\tau Z(N-1) + d_s = -\tau\lambda^{-N}[Z(-1)+d_s/r] + (r+\tau)d_s.$$

For unknown constants, C_{s1} and C_{s2}, the two root solution to this difference equation is given by

$$\hat{Z}(N) = \hat{\lambda}_1^{-N-1}C_{s1} + \hat{\lambda}_2^{-N-1}C_{s2} + A_s\lambda^{-N-1}[Z(-1)+d_s/r] - d_s/r,$$

where the two roots are

$$\hat{\lambda}_1 = \frac{r+\tau+\alpha+\beta-[(r+\tau+\alpha+\beta)^2-4\alpha\beta]^{1/2}}{2\alpha} < 1,$$

$$\hat{\lambda}_2 = \frac{r+\tau+\alpha+\beta+[(r+\tau+\alpha+\beta)^2-4\alpha\beta]^{1/2}}{2\alpha} > 1,$$

and from the particular solution

$$A_s = -\tau\lambda/[\alpha\lambda^2 - (r+\tau+\alpha+\beta)\lambda + \beta].$$

Note that slightly differently from buyers who have to account for a transition to prosperity with a cold phase

$$\hat{Z}(-1) = C_{s1} + C_{s2} + \Psi_{10} - d_s/r$$

and

$$\hat{Z}(\bar{N}) = \hat{\lambda}_1^{-\bar{N}-1}C_{s1} + \hat{\lambda}_2^{-\bar{N}-1}C_{s2} + \lambda^{-\bar{N}-1}\Psi_{10} - d_s/r = V - F,$$

where $\Psi_{10} = A_s[Z(1) - d_s/r]$. These two equations demonstrate that C_{s2} is a linear function of C_{s1} that, in turn, is linear in $Z(-1)$ yielding

$$C_{s1} = \frac{\hat{Z}(-1)}{1-\Psi_{11}} - \frac{\Psi_{12}+\Psi_{10}-d_s/r}{1-\Psi_{11}},$$

$$C_{s2} = \frac{-\Psi_{11}\hat{Z}(-1)}{1-\Psi_{11}} - \frac{\Psi_{11}(\Psi_{12}+\Psi_{10}-d_s/r)}{1-\Psi_{11}} + \Psi_{12},$$

where

$$\Psi_{11} = \left(\frac{\hat{\lambda}_1}{\hat{\lambda}_2}\right)^{\bar{N}-1},$$

$$\Psi_{12} = -\left(\frac{\hat{\lambda}_1}{\hat{\lambda}_2}\right)^{\bar{N}-1}\Psi_{10} + \frac{V - F + d_s/r}{\hat{\lambda}_2^{\bar{N}-1}}.$$

For $N = -1$, a seller's expected payoff

$$\hat{Z}(-1) = \frac{1-\tau dt}{1+rdt}\left[\alpha dt\,\hat{Z}(-2) + \beta dt\,\hat{P}(0) + (1-\alpha dt - \beta dt)\hat{Z}(-1)\right] + \tau dt\, Z(-1) - d_s dt$$

can be rewritten as

$$\alpha\hat{Z}(-2) - (r+\tau+\alpha+\beta)\hat{Z}(-1) + \beta\,\hat{P}(0) = -\tau Z(-1) + d_s.$$

Plugging in

$$\hat{Z}(-2) = \hat{\lambda}_1 C_{s1} + \hat{\lambda}_2 C_{s2} + \lambda\Psi_{10} - d_s/r = \Psi_{13}\hat{Z}(1) + \Psi_{14},$$

where

$$\Psi_{13} = \frac{(\hat{\lambda}_1 - \Psi_{11}\hat{\lambda}_2)}{1-\Psi_{11}},$$

$$\Psi_{14} = -(\hat{\lambda}_1 - \Psi_{11}\hat{\lambda}_2)\frac{\Psi_{12}+\Psi_{10}-d_s/r}{1-\Psi_{11}} + \hat{\lambda}_2\Psi_{12} + \lambda\Psi_{10} - d_s/r$$

reveals that $\hat{Z}(-1)$ is a linear function in not only C_{s1} and C_{s2} but also in $P(0)$.

$$\hat{Z}(-1) = \frac{\beta\hat{P}(0)}{r+\tau+\alpha+\beta-\alpha\Psi_{13}} + \Psi_{15},$$

where

$$\Psi_{15} = \frac{\tau Z(-1) - d_s + \alpha\Psi_{14}}{r+\tau+\alpha+\beta-\alpha\Psi_{13}}.$$

Balanced trade implies

$$\hat{P}(0) = \frac{1-\tau dt}{1+rdt}\left[\alpha dt\,\hat{Z}(-1) + \beta dt\,(x - \hat{H}(1)) + (1-\alpha dt-\beta dt)\hat{P}(0)\right] + \tau dt\, P(0) - d_s dt$$

can be rewritten as

$$\alpha\,\hat{Z}(-1) - (r+\tau+\alpha+\beta)\hat{P}(0) + \beta(x - \hat{H}(1)) = -\tau P(0) - d_s.$$

Plugging in for the linear functions of $\hat{Z}(-1)$ and $x - \hat{H}(1)$ gives

$$\left\{\frac{\alpha\beta}{r+\tau+\alpha+\beta-\alpha\Psi_{13}} - r+\tau+\alpha+\beta+\alpha\beta G_2\right\}\hat{P}(0) = -\alpha\Psi_{15} - \beta x + \beta G_2[\alpha x + \tau C(T_1) + \beta G_1 - d_b] - \tau P(0) + d_s$$

or

$$\hat{P}(0) = (-\alpha\Psi_{15} + \Psi_7)/\Psi_{16},$$

where $\Psi_{10}, \Psi_{11}, \Psi_{12}, \Psi_{13}$ are as above and

$$\Psi_7 = -\beta x + \beta G_2[\alpha x + \tau C(T_1) + \beta G_1 - d_b] - \tau P(0) + d_s,$$

$$\Psi_{16} = \frac{\alpha\beta}{r+\tau+\alpha+\beta-\alpha\Psi_{13}} - r+\tau+\alpha+\beta+\alpha\beta G_2.$$

The last step determines $\bar{N}$ from V. The above second-order difference equation for the sellers satisfies the initial and terminal conditions given an $\bar{N}$. The solution imposes the decision not to sell at this level of unsold houses. A seller who decides to enter at $\bar{N}$ when others are not receives

$$\tilde{Z}(\bar{N}) = \frac{\beta\hat{Z}(\bar{N}-1) + \tau Z(\bar{N}) - d_s}{r+\beta}.$$

The optimal $\bar{N}$ follows from iterating for increasing $\bar{N}$ until $\hat{Z}(\bar{N}) > \tilde{Z}(\bar{N})$

Appendix C: Quantitative Analysis

C.1 | Refining NAR and CoreLogic Inventories

The quantitative evaluation of the model primarily uses the time series derived from CoreLogic. The glaring limitation is that CoreLogic has sparse coverage early on when the slump hits. For this reason, the quantitative evaluation turns to the NAR inventories series during the slump.

To account for discrepancies between NAR (in thousands) and CoreLogic inventory series, the first adjustment to the NAR inventories is to marry up the levels of the two series. Note, in particular, that the level of CoreLogic Sales from post 2015 (when coverage is no longer sparse) is roughly 2.8 times larger than the level of NAR Sales of Existing Homes (in thousands) during the same period. Looking at inventories, however, it then initially seems peculiar that NAR Inventories of Existing Homes nearly equal CoreLogic Inventories. Indeed, the levels and pattern of the two series track each other closely from 2011 onward.

Pending sales, which are included in the NAR inventory series but not in the CoreLogic series, reconcile the inequality-in-sales levels with equality-in-inventories discrepancy. NAR sales consistently peak in August, whereas CoreLogic sales consistently peak in May. Based on this observation, it appears that it takes around 3 months to complete a transaction after agreeing to exchange. Subtracting 2.5 * mean(NAR Sales post 2015) from mean (NAR Inventory post 2015) gives a ratio of 2.5. (A factor of around 2.7 gives equal ratios. A simple Google search states that pending house sales take 30–60 days so the factor used is rounded down slightly.) To adjust for pending sales, forward sales from end-of-month inventories up to 75 days within the next 3 months are subtracted from the NAR inventory series.

A second adjustment designed to conceptually align both CoreLogic and NAR inventory series with the model outputs accounts for very recent listings. In particular, both observed inventories include houses very recently listed and hence quite likely still on the short side that inflates the target for the model. It takes time from listing to show a house, make and revise bids, and ultimately accept an offer. Suppose that it takes 7.5 days or half of the 15 day period interval used for the Kaplan–Meier measures in the paper.[40] Subtracting a quarter of the month's listings from the end of month inventory lowers the CoreLogic Inventory-Sales ratio from 1.74 to 1.50.

NAR unfortunately does not provide a listings series. To impute new listings, the adjustment takes the difference in (pending sales adjusted) inventories less imputed sales—sales from 75 to 105 days ahead of the current month. Correcting as done with the CoreLogic data for the initial period of no sales by subtracting 7.5 days worth of imputed new listings gives a new NAR series for inventory. The NAR adjusted series has a lower level, so the last step scales up Inventories to match CoreLogic levels from 2015 onward.

Validation: To check these adjustments, note that the adjusted NAR post 2015 Inv-Sales ratio equals 1.49. The CoreLogic post-2015 Inv-Sales ratio is 1.50. Moreover, the pending-sales-adjusted NAR series (without the 7.5 day correction) also matches the CoreLogic ratio before adjusting for the recent listings within 7.5 days. The NAR and CoreLogic inventory series also become more highly correlated. The adjusted NAR inventories after 2011 correlation with CoreLogic inventories post 2011 equals 0.935, whereas the unadjusted NAR correlation with CoreLogic is 0.908.

C.2 | Calibration of α

Because the recovery rolls out gradually in five installments and each is 2 months apart, α can be derived similarly as in Equation (6). In the first case, when calibrating to the rise in inventories between January to October 2006, α can be derived as

$$\text{Period 1:} \quad (Inv_{t+D} - Inv_t) = 0.2(\alpha - \hat{\beta}) * D$$
$$Sales(\text{t to t+D}) = (0.2\hat{\beta} + 0.8\beta)D,$$
$$\text{Period 2:} \quad (Inv_{t+2D} - Inv_{t+D}) = 0.4(\alpha - \hat{\beta}) * D$$
$$Sales(\text{t+D to t+2D}) = (0.4\hat{\beta} + 0.6\beta)D,$$
$$\text{Period 3:} \quad (Inv_{t+3D} - Inv_{t+2D}) = 0.6(\alpha - \hat{\beta}) * D$$
$$Sales(\text{t+2D to t+3D}) = (0.6\hat{\beta} + 0.4\beta)D,$$
$$\text{Period 4:} \quad (Inv_{t+4D} - Inv_{t+3D}) = 0.8(\alpha - \hat{\beta}) * D$$
$$Sales(\text{t+3D to t+4D}) = (0.8\hat{\beta} + 0.2\beta)D,$$
$$\text{Period 5:} \quad (Inv_{t+5D} - Inv_{t+4D}) = (\alpha - \hat{\beta}) * D$$
$$Sales(\text{t+4D to t+5D}) = \hat{\beta}D.$$

Summing up and solving out for D yields

$$\frac{Inv_{t+5D} - Inv_t}{Sales(\text{t to t+5D})} = \frac{3(\alpha - \hat{\beta})}{3\hat{\beta} + 2\beta},$$
$$\alpha = \hat{\beta} + \left(\hat{\beta} + \frac{2}{3}\beta\right)\frac{Inv_{t+5D} - Inv_t}{Sales(\text{t to t+5D})}.$$

In the second case, when calibrating to the rise in inventories between January and December 2006, an extra period of the same length D is needed:

$$\text{Period 6:} \quad (Inv_{t+6D} - Inv_{t+5D}) = (\alpha - \hat{\beta}) * D$$
$$Sales(\text{t+5D to t+6D}) = \hat{\beta}D.$$

Summing up and solving out for D yields

$$\frac{Inv_{t+5D} - Inv_t}{Sales(\text{t to t+5D})} = \frac{4(\alpha - \hat{\beta})}{4\hat{\beta} + 2\beta},$$
$$\alpha = \hat{\beta} + \left(\hat{\beta} + \frac{1}{2}\beta\right)\frac{Inv_{t+6D} - Inv_t}{Sales(\text{t to t+6D})}.$$

C.3 | Targeting Days on the Market

Accounting for the delay in time to sale after listing affects DoM in a straightforward way. Subtracting 7.5 days from the observed CoreLogic DoM for sales gives 49.5 days.

Validation: To evaluate the DoM average, consider two potential methods for imputing DoM. In the CoreLogic data, the Kaplan–Meier by duration is relatively flat after 30 days, so the mean long-side Kaplan–Meier (KM^{Long}) times the (adjusted) inventory number equals the number of sales from the "old" stock, that is houses on the market for more than 15 days. The flat Kaplan–Meier by duration implies that these sales are random draws from the stock and hence have a randomly drawn DoM for unsold house in the month. Short-side sales as determined by the short side Kaplan–Meier (KM^{Short}) have a DoM equal to zero. Accounting for two 15-day intervals in the month gives

$$First\ Imputed\ DoM\ of\ sold\ houses =$$
$$KM^{Long} * Inv * DoMUnsold/Sales = 47.2.$$

A second approach is to subtract the sales from new listings each day from the average daily sales to give daily sales from the stock. To compute daily measures, first compute the instantaneous hazard p from the 0–15 Kaplan–Meier statistic

$$1 - \exp(-p(T - 7.5)) = KM^{Short} => p = ln(1 - KM^{Short})/(T - 7.5),$$

where $T = 15$. Since there are now only 7.5 days for listings to match, the number of sales in a day from the stock is

$$Sales/30 - 7.5 * [1 - exp(-p)] * Listings/30.$$

Multiplying this figure by the DoM of unsold houses and dividing by average daily sales gives the second imputed time on market for sold houses:

$$Second\ Imputed\ DoM\ of\ sold\ houses = 50.1$$

The target of 49 days used in the calibration is close to these measures.

Note that adjusting for no trading in the initial 7.5 days raises the probability of matching in the next 7.5. If the time in the flow changes from 7.5 days (the first period plus half the second time period) to 1.5*7.5 = 11.25, the average adjusted KM^{Short} from 2015 onward becomes 0.477 that is very close to the latest steady state $KM^{Short} = 0.4709$ figure from the calibrated model. The minimum rises to 0.2115 and the maximum becomes 0.5117.

Appendix D: Variation across Cities and the Role of Policy

D.1 | Construction of City-Specific β

Let β_i be the arrival rate of buyers during prosperity in city i. β_i is derived by setting the ratio of the arrival rates of buyers in the prosperity and slump in the model ($\beta_i/\hat{\beta}$) equal to the ratio of sales during the prosperity and slump periods for the same city in the data, where $\hat{\beta}$ is the calibrated arrival rate of buyers during slump for the average city (Table 2):

$$\beta_i = \frac{\text{average sales during prosperity period}_i}{\text{average sales during slump period}_i} \times \hat{\beta}.$$

To compute β_i for each city, let prosperity be the time period 2015–2019. Let slump be the 6 months prior to the month with minimum housing sales during 2007–2019 for each city. Two types of cities are dropped when constructing β_i. First, the minimum-sales month in five cities occurs at the beginning of the city sample in January 2007. Since the CoreLogic data are sparse at the beginning of the sample, the minimum-sales month may be misidentified for these cities. Second, some cities had "double dips" with the minimum-sales month after 2010 and an earlier and smaller dip before 2010. This phenomenon is attributed to the American Recovery and Reinvestment Act, which ended in April 2010 (see Section 6.3) and had different impacts on different cities. Without the housing programs, there would only be one dip. Eliminating these two types of cities from the 50 cities in the sample gives 30 cities.

A feasible β_i has to satisfy the parameter restriction $\alpha < \beta_i < \alpha + \sigma$, and 23 of the 30 cities satisfy this condition. In addition, the simulated outcomes must satisfy the monopoly-Bertrand condition discussed in Section 4, which depends on all city parameters (except V) in the model. Of the 23 cities, 14 cities pass this restriction. The statistics reported in Table 4 are computed using these 14 cities. Note that using the 6 months prior to the month with minimum housing sales (rather than the 2007–2010 period) to construct trough statistics from the CoreLogic data gives similar results as shown in Table 4.

D.2 | Policy Experiment

Figures D.1 and D.2 plot the simulation results with and without the first-time homebuyer credit.

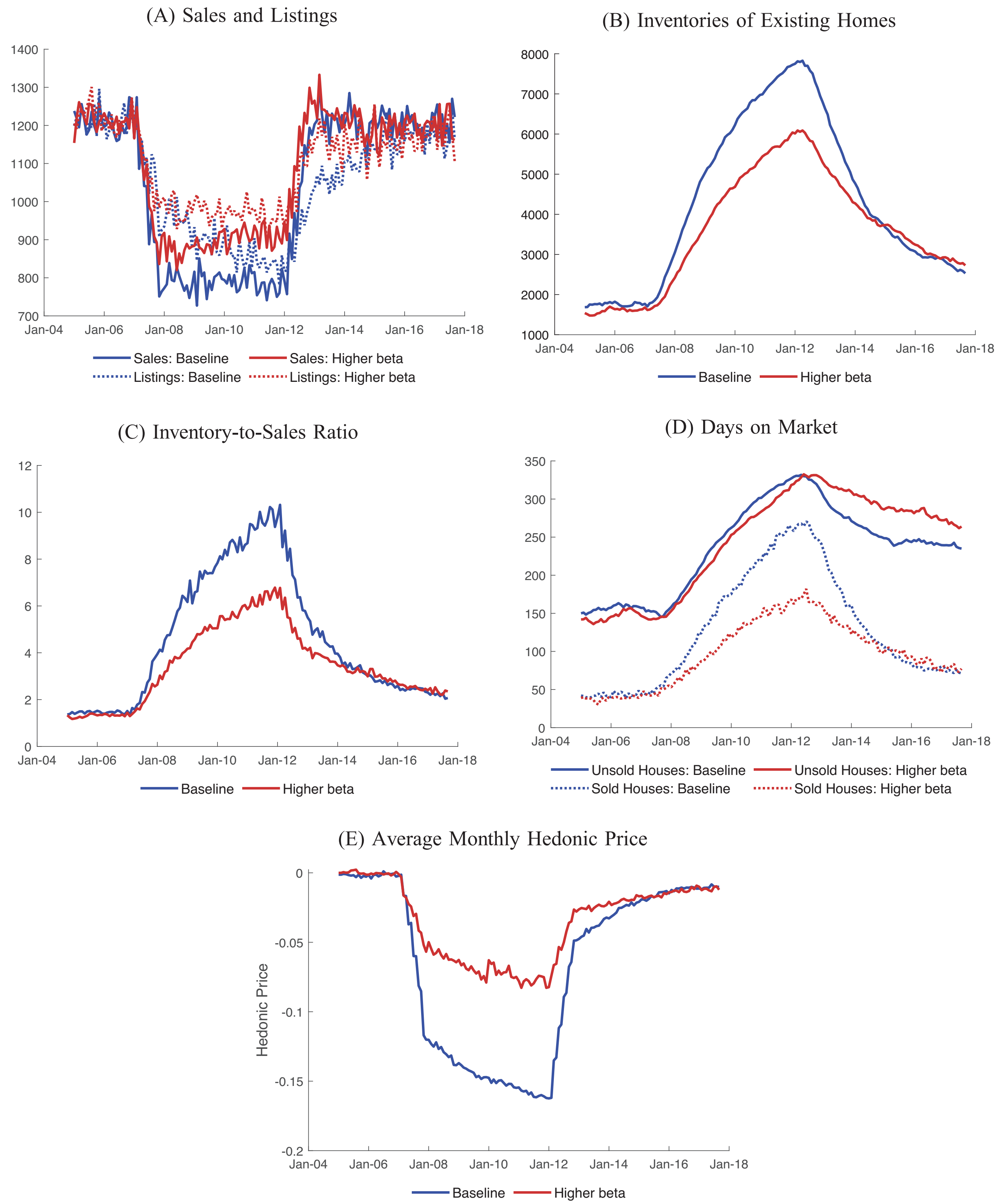

FIGURE D.1 | Housing market aggregates: impact of the first-time homebuyer credit.

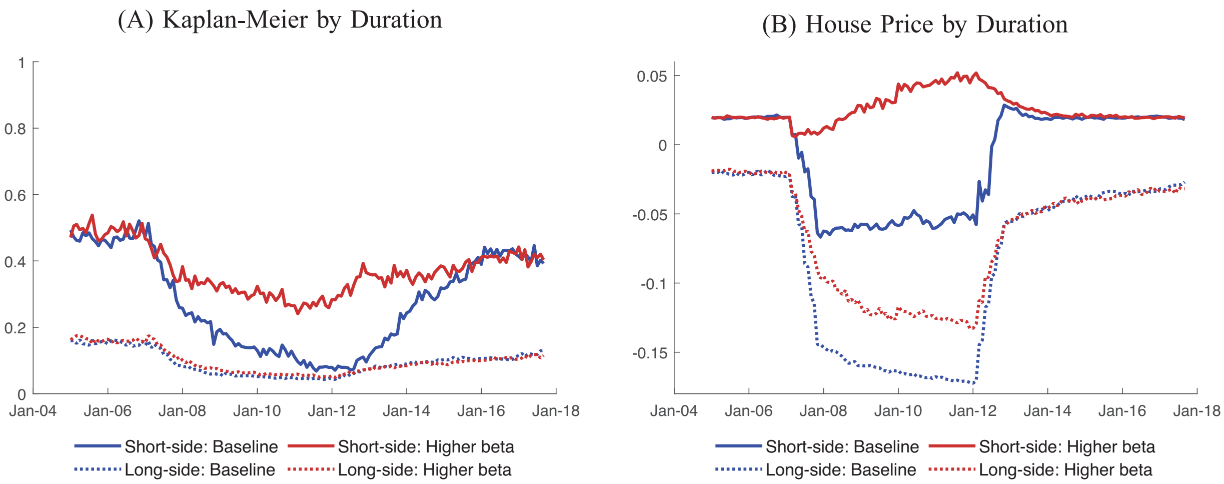

FIGURE D.2 | Variables by duration: impact of the first-time homebuyer credit.