FEDERAL RESERVE BANK *of* ATLANTA **WORKING PAPER SERIES**

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

Eric Smith, Zoe Xie, and Lei Fang

Working Paper 2022-15
October 2022

Abstract: This paper investigates the US housing market from just before the Great Recession onward (2006–19) and assesses the viability of stock-flow matching in generating the observed outcomes. The paper documents that the probability that a house sells declines sharply after listing for two weeks. Moreover, the probability and associated price of a fast sale recover from the housing slump sooner, faster, and more prominently than slower sales. The simulated stock-flow matching model can mimic not only sales, prices, listings, and time-on-market but also capture the distinctions in quick and slower trades, indicating the importance of stock-flow matching for understanding housing market dynamics.

JEL classification: E30, R21, R31

Key words: housing, stock-flow matching, trading dynamics, duration dependence

https://doi.org/10.29338/wp2022-15

First draft: October 2022. The views expressed here are those of the authors and not necessarily those of the Federal Reserve Bank of Atlanta, the Federal Reserve System, the International Bank for Reconstruction and Development/World Bank, or its executive directors. Smith acknowledges the support of the Business and Local Government Data Research Centre (grant number ES/L011859/1), funded by the Economic and Social Research Council, for undertaking this work. The authors would like to gratefully acknowledge early contributions to the project from Yunjuan Liu. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Eric Smith, University of Essex, esmith@essex.ac.uk; Zoe Xie, World Bank, xiexx196@gmail.com; or Lei Fang, Federal Reserve Bank of Atlanta, leifang@gmail.com.

1. Introduction

In 2004 and 2005, house prices and sales both surged in the United States. In 2006, a sharp and dramatic downturn—the largest decline in the post-war period by some measures—triggered the onset of the Great Recession. This paper investigates house trading patterns from the slump just before the Great Recession, into the eventual recovery, and on until the start of the COVID-19 pandemic. The paper contributes to the literature by first assembling and documenting a variety of standard as well as less familiar empirical regularities. It then assesses the viability of the stock-flow matching framework in generating the volatility, co-movement, persistence, and duration dependence observed in the housing market during this time period.

As expected, sales, new listings, inventories, time-to-sale, and average price exhibit a pronounced and dramatic slump in 2006. Recovery begins in 2011 and takes off in 2012. Similarly, the probability a house sells conditioned on how long it has been for sale falls and recovers with the other market variables. Prices broken down by the time-on-the-market when the sale occurs behave in the same way.

The more novel finding is that there are clear and distinct differences between immediate trade and trade after a more prolonged spell of listing. At any given point in time during the slump and recovery cycle, the probability of a house selling within the first two weeks of being listed is markedly higher than the probability of a house selling after it has been listed between two and four weeks. Moreover, as time progresses, the probability of a quick sale and the accompanying price both recover from the slump sooner, faster, and more prominently than sales with durations longer than two weeks. Once the recovery takes hold, the hazard function thus displays substantial and growing duration dependence as the drop in the likelihood of trade after only two weeks becomes more prominent. The average price for the shortest duration is likewise out of sync with the average prices for other durations.

The challenge of accounting for the behaviour of these housing market variables in a plausible, tractable, and coherent framework is substantial. The housing market is notoriously difficult to replicate. The aim here is to consider the prospects of a specific framework—stock-flow matching—in which flows of entering traders match and exchange with the stocks 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. They check intermediaries such as real estate agencies and websites. Unlike competitive search, however, buyers and sellers in the stock-flow framework trade in precise and distinct markets, differentiated by location and other features. Although the trading platforms provide information on a wide variety of opportunities, traders look for very specific characteristics. In the housing market, multiple markets exist in a local geographic region as buyers seek a 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. If lucky, an entrant is on the short side and finds one or more options immediately available. Thus trade in this case occurs quickly. If the entrant is unlucky and is on the long side, there are no potential partners immediately at hand. In the event that no partners currently exist, 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 thus produces duration dependence in the probability of a sale.

This stock-flow microstructure of dynamic trade in differentiated goods embeds several compelling features associated with the housing market. It also provides a plausible mechanism for stochastically generating time between exchange 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 when this shock hits, more motivated sellers none the less enter faster than buyers. Inventories 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 pre-existing buyers to dramatically curtail immediate trades. When the downturn ends, buyer entry picks up but seller entry does not fully resume until inventories eventually clear. Stock and flow sellers both 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? Using a small set of parameters, simulations reveal that the model is consistent to a large extent with the familiar as well as the duration dependent housing market outcomes outlined above. It captures much of the level, variation, shape, and timing of listings, sales, inventories, time on the market and average price over the cycle. The more fundamental contribution is to establish that with only two aggregate states (slump and prosperity), the simulated stock-flow model can generate differential responses over time in short and long duration trades similar to the empirical findings in the selling probability and the prices conditioned on duration.

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 briefly outlines the stock-flow model and discusses its trading outcomes. Section 5 simulates the model and compares the simulation results with empirical findings in the data. Section 6 concludes with a discussion.

2. Related Literature

Beginning with Wheaton (1990), the application of search frictions to housing markets initially addressed three broad empirical regularities that revolved around price fluctuations.

- Housing cycles occur—there is short run positive serial correlation in prices but mean reversion in the long run (Case and Shiller 1997; Muellbauer and Murphy 1997).
- There is excess volatility in prices and quantities relative to fundamentals (Shiller 1982; Glaeser et al. 2014).
- Price and sales exhibit positive correlation, while price and seller time-on-market exhibit 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-sell, the stock of houses for sale, and, of course, house prices. Ngai and Sheedy (2020b) document that these variables are all highly volatile and highly correlated.[1] Other research documents that non-negligible residual price dispersion coexists with these regularities.[2]

[1] 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, negatively correlated with time-to-sell.

[2] See Merlo and Ortalo-Magné (2004) on UK data, Leung et al. (2006) on Hong Kong data, de Wit and van der Klaauw (2013) on Dutch data, and Drenik et al. (2019) on Spanish data, Rekkas et al. (2021) on Canadian data, and the numerous studies of hedonic price regressions.

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.[3] 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 sell respond as well to aggregate shocks and absorb some of the market-wide variation.

Novy-Marx (2009) obtains amplification and generates excess volatility through an endogenous market tightness response to aggregate demand shocks. Increased demand for houses (say from rising incomes) enables sellers to trade faster. Fewer houses are thereby available next period which 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 co-movement in prices, sales and selling probabilities across steady states. 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 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. They also allow for the turnover of existing homes. Ngai and Sheedy (2020a), Ngai and Sheedy (2020b), Moen et al. (2021), and Anenberg and Ringo (2022) likewise emphasize the importance of the moving decision - to simultaneously sell the current house and buy a new one - in propagating and amplifying external shocks. Concentrating on price volatility, Arefeva (2020) demonstrates that both the prevalent dynamic random and

[3] 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 (2020).

directed search models with bargained prices cannot explain price volatility but incorporating auctions substantially increases volatility.[4]

Given this picture of the housing search literature, the first fundamental contribution of this paper is to augment the already established empirical regularities by documenting the distinct behaviors of long-term and short-term sellers over the business cycle. The second significant contribution establishes that the stock-flow framework provides a useful model for framing and replicating this distinction 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 considered only steady states.[5] As in Head et al. (2014) and Ngai and Sheedy (2020b), the model in this paper incorporates not only endogenous entry but also aggregate booms and busts from shocks to the trading environment.

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. Labour market information regarding the number of vacant jobs and job seekers drives the movements over time in the hazard functions.[6] 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. Whereas 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) over time supports stock-flow matching.

[4] Anenberg and Bayer (2020), Caplin and Leahy (2011), and Smith (2020b) 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.

[5] See for example Wheaton (1990); Krainer (2001); Albrecht et al. (2007); Head and Lloyd-Ellis (2012); Rekkas et al. (2021).

[6] 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 (2020a) for an overview of these results.

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

This section first describes the data sources used to construct the trading patterns in the US housing market between 2006-2019. It then documents patterns in the volatility, co-movement, and persistence not only of standard 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 the studied period.

The variation, progression, and timing of the data establish the following key empirical findings:

- Aggregate data patterns of sales, new listings, inventories, days on market, and average price are consistent with a pronounced and dramatic housing slump in 2006 and a sustained recovery beginning in 2011 and taking off in 2012.
- The average duration on the market for sold houses (completed spells) is smaller than the average duration of houses unsold on the market (uncompleted spells).
- A sharp distinction exists between sales that occur quickly and those that take more than a month. In particular:
  - The probability of a house selling declines sharply after it has been listed for two weeks.
  - The probability of a quick sale and the accompanying price both recover from the slump sooner, faster, and more prominently than sales with durations longer than a month.

3.1 Data Sources

The data come from multiple complementary sources: the Case-Shiller price index, the National Association of Realtors (NAR), the on-line platform Redfin, and confidential data from CoreLogic. First, the Case-Shiller price index is a well-known repeated-sales house-price index for single family homes. Second, the NAR provides long-standing monthly time series of the total number of existing houses for sale and of the number of existing houses that sold. It does not, however, provide publicly available data on listings, days on the market, and prices. Third, the on-line 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 two weeks, and the median days-on-market for sold homes, but all data are available from only 2012 onward. Lastly, the data service CoreLogic provides detailed and comprehensive information on individual house transactions.

Because the CoreLogic data contain all variables of interest and represents the primary source of data in this paper, more detailed information is provided here. 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 onwards. 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 percent of all active listings nationwide. The analysis of CoreLogic data is restricted to 50 metropolitan areas with a sizeable population and with 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.

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 beginning of each month), days on the market (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 US metropolitan areas. The NAR and Redfin inventory measures also contain monthly sales. To be consistent with the directly observed CoreLogic measure for inventory, the end-of-month inventory measures from the NAR and Redfin are computed by subtracting houses sales from the total houses-for-sale during the month. But these end-of-month counts do not exclude houses withdrawn without a sale. In addition, although CoreLogic and Redfin provide direct listing counts, the NAR does not. Using the available data, the NAR listings are computed as total houses-for-sale in the month less the end-of-month inventories from the previous month.[7]

[7] The reported NAR and Redfin "inventory" measures count all houses for sale in a month. Subtracting sales from the total for sale count in the month leaves the inventory of unsold houses at the end of the month. Abstracting from possible withdrawals of for-sale houses, new listings in the subsequent month are then calculated as the total for sale in the subsequent month less inventory at the end of the previous month. In Figure 1, the imputed NAR listings track Redfin actual listings sufficiently well to discern patterns over the cycle with some confidence before 2012.

Redfin's reported monthly total house-for-sale count during the month less sales plus next month's new listings generates a measure for "inventories" in the subsequent period. Comparing this calculation with the reported figures generates a measure for seller withdrawals. This figure - 6.6% of houses for sale in a month - is non-trivial and helps explain the high count in the

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 estimator for the probability that a house on the market sells conditional on the DoM d is simply the proportion of the successful sales among qualifying houses for sale. In other words, it is the ratio at date t of sold houses with duration d over the sum of sold and unsold houses on the market with the same duration d.

$$K_t^d = \frac{\text{Houses with } d \text{ sold in time period } t}{\text{Houses on the market with } d \text{ in time period } t} \tag{1}$$

In the data set, dates t are months and durations d are 15 day intervals of DoM. The estimator is computed for each duration interval up to 120 days, in each sample month, for each metropolitan area. These estimates are then aggregated over cities to the national level. To control for differences in housing and market characteristics, a set of logit regressions yield a corresponding set of conditional Kaplan-Meier statistics to study patterns in the housing market. Appendix A provides the details for the construction, and shows that there is little substantive difference between the conditional and the raw Kaplan-Meier statistics (without controls).

Hedonic Sale Prices. In addition to the Case-Shiller price index, hedonic sale prices—prices with controls for house characteristics—are constructed and examined in each city by duration or time to agreed sale. Specifically, in each city, the hedonic price for house i with characteristics $X_{i,t}$ at date t is given by:

$$\ln(P_{it}) = \Omega X_{it} + \sum_d \theta_t^d D_t^d + \zeta_z + \gamma_m + \epsilon_{it}. \tag{2}$$

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 1-15 days to 106-120 days in 15 day intervals. The estimates for the duration category dummies generate prices broken down by duration conditional on average local housing characteristics. Additionally, month-of-year (γ_m) and zip code (ζ_z) fixed effects are included.

simulations. Although not shown here, this figure decreases over time as the market recovery gains pace from 2012 onward.

3.3 Empirical Findings

Comparing Overlapping Sources. The CoreLogic data overlap with the alternative data sources in some areas thereby enabling comparisons for the relative strengths and limitations of the CoreLogic data. Table 1 reports the minimum, maximum, percentage change between minimum and maximum, autocorrelation, and contemporaneous correlation with price for some observed aggregate housing series. Whenever possible, variables from both the CoreLogic data and an alternative source are included.[8] To remove contemporaneous seasonal movements, the series in this table are seasonally adjusted using a plus and minus six month moving average.[9] Figures 1-5 supplement the information in Table 1 by plotting the time series of the underlying data series.

The figures and table collectively demonstrate that sales, new listings, inventory, days on market, and prices from various sources (where they overlap) are largely consistent with each other in their evolution over time, the autocorrelations, and the correlations with prices. One discrepancy across data sources is worth noting. The limited coverage of CoreLogic data in the initial years distorts its housing counts. As a result, data patterns before 2009 diverge between CoreLogic and the NAR data. As seen in Figure 1, CoreLogic sales and listings gradually increase from 2007 until 2009 whereas the NAR counts fall consistently from early 2006 until 2009. The moving averages in the bottom panel of Figure 1 make this observation particularly transparent. After 2009, although sales and listings are still lower in 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 sub-sample from 2009 onward.[10] The table shows that for the post-2009 subsample statistics for the sales and listings are comparable between CoreLogic and the NAR, confirming patterns in Figure 1. Because of the discrepancy in the pre-2009 period, the sales and listings from the NAR are preferred over the CoreLogic measures when compared with the model in Section 5.

Aggregate patterns. Table 1 reveals the variability, co-movement, and persistence 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, days on the market, the Kaplan-Meier statistics (prob-

[8]The CoreLogic series are weighted by market size for the sample period.

[9]Prices coefficients are seasonally adjusted by construction and the Kaplan-Meier hazard rates conditioned on observables from CoreLogic display little seasonality.

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

Table 1: Descriptive Statistics

Variable	Source	Min	Max	% change max-min	Auto-corr (1 lag)	Corr with price
GDP (bn USD)		15,161.8	19,202.5	23.5	–	–
Sales	NAR	334,000	512,615	42.2	0.990	0.787
	CoreLogic	789.7	1329.5	50.9	0.982	-0.204
	NAR (2009-)	338,300	470,100	32.6	0.992	0.847
	CoreLogic (2009-)	976.9	1329.5	30.6	0.975	0.915
Listings	NAR	283,308	539,615	62.3	0.973	0.750
	CoreLogic	804.3	1308.1	47.7	0.993	-0.059
	NAR (2009-)	283,300	466,200	48.8	0.976	0.795
	CoreLogic (2009-)	985.2	1308.1	28.2	0.976	0.941
Inventory	NAR	12,288,500	31,943,800	88.9	0.999	-0.558
	CoreLogic	1781.4	2654.3	39.4	0.995	-0.780
Days on Market - Sales	CoreLogic	46.5	85.8	59.4	0.992	-0.504
	Redfin (2012-)	39.8	84.8	72.1	0.996	-0.944
Days on Market - Inventory	CoreLogic	66.8	110.0	48.9	0.992	-0.837
1-15 Day Kaplan-Meier	CoreLogic	0.174	0.530	101.1	0.973	0.261
	Redfin (2012-)	0.312	0.411	27.5	0.997	0.944
31-45 Day Kaplan-Meier	CoreLogic	0.151	0.395	89.3	0.957	0.368
Average Price	Case-Shiller (Real Index)	71.0	107.7	41.1	0.998	–
	CoreLogic	130,372.7	254,987.1	64.7	0.969	–

Note: 1. All data are for Jan 2007–Dec 2019 unless otherwise noted. The full sample CoreLogic data is 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.
2. All data variables are in monthly frequency, except for GDP, which is quarterly. GDP and the Case-Shiller index are seasonally adjusted at source (FRED). Other monthly variables are seasonally adjusted by taking a $\pm$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 and end of month inventory are computed from total houses for sale less houses sold in the month as noted.
4. Average prices from CoreLogic are in dollars. 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. Kaplan-Meier and average prices from CoreLogic are conditional statistics derived from regressions following procedures outlined in Appendix A. 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.

ability to sell), and prices all vary more from peak to trough than aggregate output and also exhibit strong autocorrelations. Moreover, as expected, sales, listings, and the Kaplan-Meier statistics positively co-move with prices, while inventory and days on the market negatively co-move with prices.

The companion figures reveal that sales, new listings, inventories, and days on market exhibit a pronounced and dramatic housing slump in 2006 and a sustained recovery taking off in 2012. In Figure 1, the monthly NAR sales and listings of existing houses peak in 2006 and decline dramatically in 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 the NAR sales occurs within two years but never fully recovers to the pre-slump high levels of sales. The top panel of Figure 1 further reveals that non-seasonally adjusted listings and sales closely mirror each other following a strong seasonal pattern. At a high frequency, the inflows match outflows very closely from all three data sources. The tight relationship between inflows and outflows is a signature feature of stock-flow matching.[11]

The bottom panel of Figure 1 displays heightened fluctuations (coupled with a minor rise and fall) of sales from 2009 until midway through 2010, a pattern more clearly evident in the NAR sales series. This temporary sub-cycle or wobble (from the progression within the downturn) coincides with the onset and expiration of the components of the American Recovery and Reinvestment Act of 2009 (Indiviglio 2011) aimed at stimulating the housing market. Looking ahead, similar patterns are visible in inventories, Kaplan-Meier hazards, and prices. This temporary change is small, short-lived, and left out of the modelling analysis in Section 5.

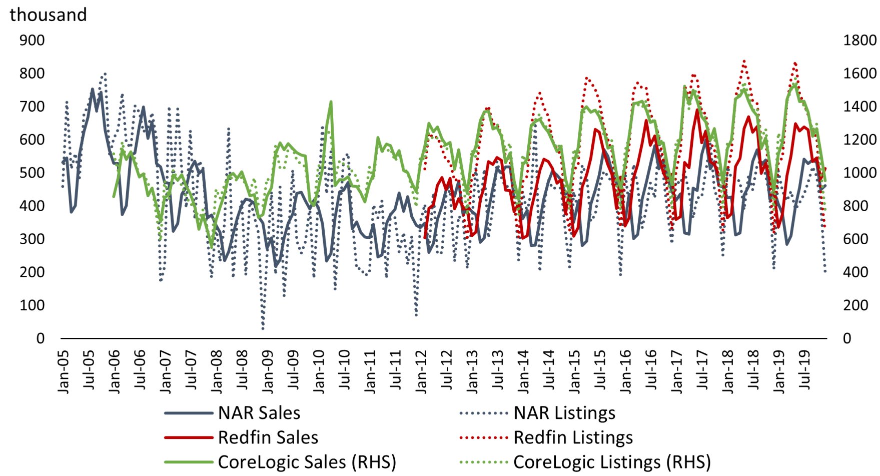
End-of-month inventories in Figure 2 accumulate (decumulate) if listings are greater (less) than sales. The NAR listings exceed sales until 2008 whereas CoreLogic listings exceed sales until mid 2011. As a result, the rise and fall of the two inventories series exhibit different timings over the cycle. 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.

Figure 3 plots days on the market (DoM) for sold and unsold houses. Redfin's seasonally adjusted series for sold houses from 2012 onward aligns closely with the CoreLogic counterpart. Given the limited availability of the Redfin series and its alignment with CoreLogic data, the CoreLogic series

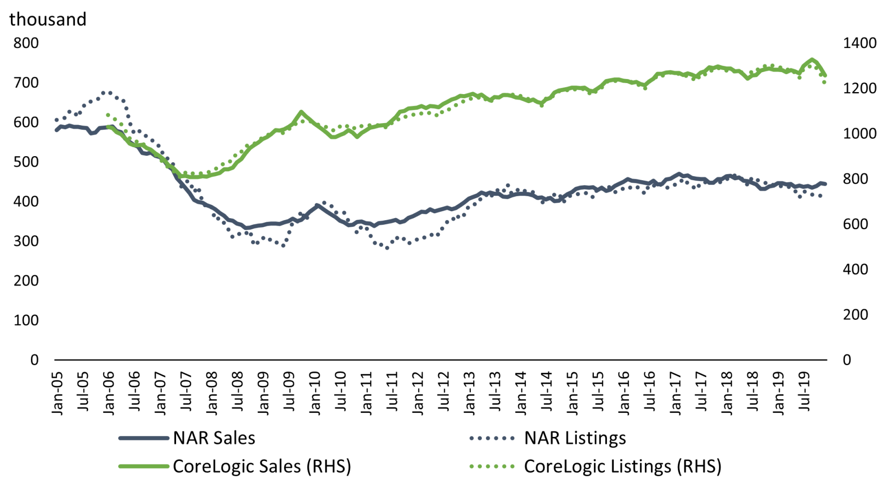
[11] Coles and Smith (1998) document a similar finding in the labor market with (lagged but less seasonal) worker flows into jobs closely linked with the new vacancies and much less so with the (smoother) stock of job vacancies.

Figure 1: Monthly sales and new listings

(A) Sales and new listings of existing homes



(B) Sales and new listings of existing homes (moving averages)



becomes focal.[12] Four observations emerge. First, 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. Second, the DoM for unsold homes steadily rises until 2012 and subsequently declines. Third, at the outset of 2007—in the midst of the slowdown—the average time on the market for sold and unsold houses nearly equal to each other, an outcome consistent with a constant (with respect to durations at that point in time) hazard function 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.

[12]The limited early coverage in CoreLogic should not dramatically affect these two series as the sample was representative.

Figure 2: Monthly inventories of existing homes

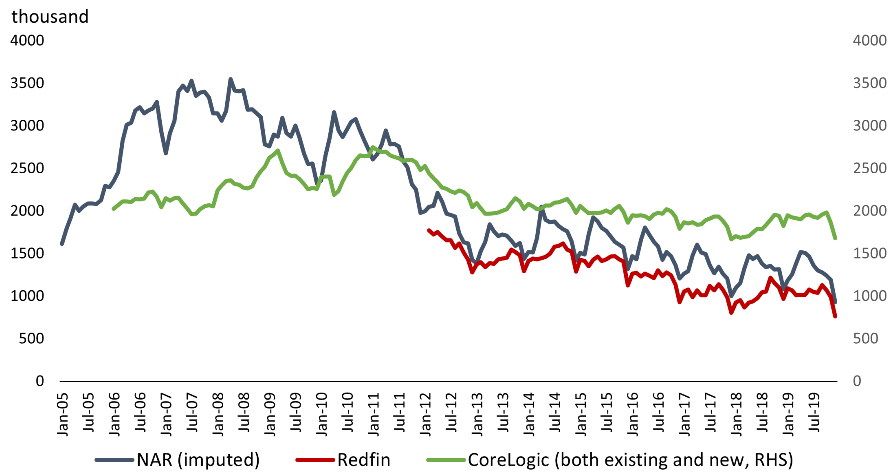
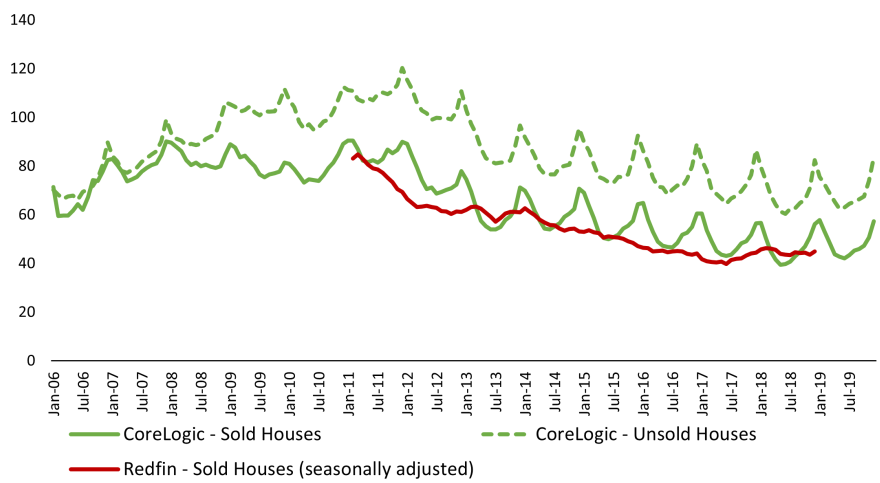


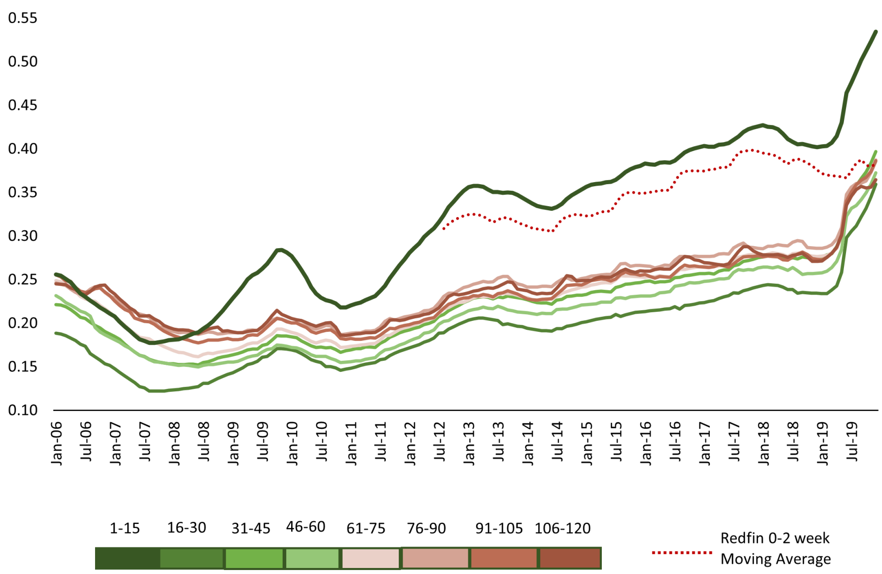
Figure 3: Days on market



Patterns by Duration. Figure 4 plots the Kaplan-Meier statistics for 15 day intervals along with the short Redfin series for the proportion of houses that sell within two weeks. The Redfin series echos the 1-15 day CoreLogic Kaplan-Meier harzard 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 exhibit a sharp decline 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 becomes clearly downward sloping after the initial interval, becoming more step-like as the 1-15 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.

Figure 4: Monthly conditional Kaplan-Meier by duration

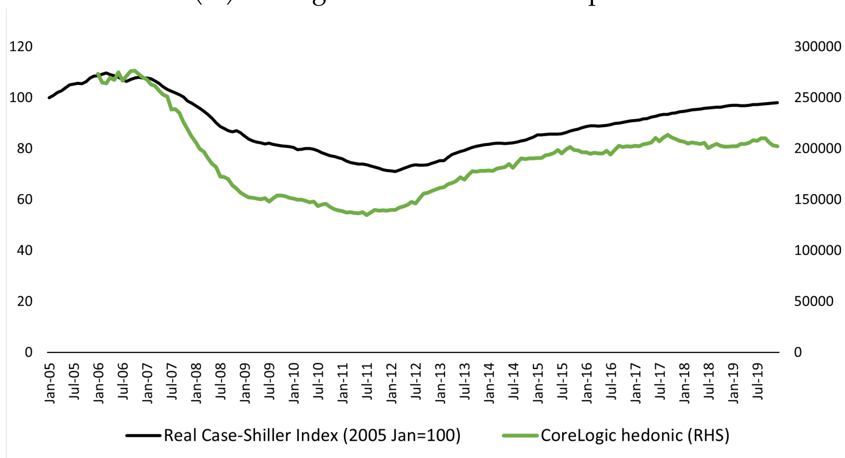


Specifically, for the durations less than 30 days, the recovery is much sooner and stronger than for longer durations. The series for the two shortest durations (1-15 days and 16-30 days) begin rising after an initial, relatively sharp decline from as early as 2007, with the most pronounced improvement appearing in the 1-15 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 business cycle. The probability of selling quickly is more volatile and responsive over the business cycle, compared to the probability of selling after a month. This difference leads to changes in the relative order of the 1-15 day hazard. Moreover, the 1-15 day duration Kaplan-Meier displays the largest apparent reaction to the Recovery Act.

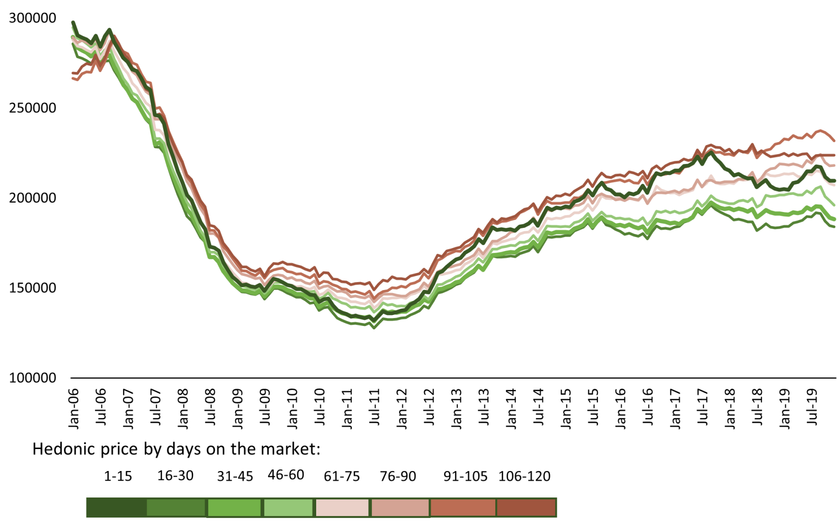
Figure 5 plots the evolution of the average conditional price and the prices by DoM constructed using the CoreLogic data. The top panel reveals that the average CoreLogic price is largely consistent with the shape and timing of the well-known Case-Shiller housing price index. The bottom panel of Figure 5 plots the CoreLogic hedonic price estimates by duration. Prices co-move with the NAR sales in Figure 1 and also exhibit a wobble during the Recovery Act period. Average price and prices for all durations fall until late 2011 and then recover. The relationship of price with hazard rates is less clear. Prices fall while the long duration Kaplan-Meier hazards are relatively flat but rise as the 1-30 day hazards begin to rise. The Kaplan-Meier hazards apparently bottom out before prices.

Figure 5: Monthly sale price statistics

(A) Average and conditional sale prices



(B) Conditional (hedonic) sale prices by duration



Echoing the evolutions of the Kaplan-Meier series, a second attribute is visible in the prices by duration.[13] Prices over different durations roughly co-move over time with the price at the shortest

[13]Both the prices by duration in the bottom panel of Figure 5 and the Kaplan-Meier statistics in Figure 4 reveal a second, intriguing relationship with time on the market. In any given month, the individual hedonic regression price coefficients and the hazard rates are U-shaped with respect to time-on-market. In every period, a prominent drop occurs after 15 days in the Kaplan-Meier hazard function. Among the higher 15+ day durations, the second interval of 16-30 days is uniformly the lowest often by a substantial margin. The 30+ days hazard rates are bunched and scrambled with longer hazard rates tending to be slightly higher.

Like the Kaplan-Meier hazards, at a given point in time the lowest prices are for the 16-30 day duration. The predicted price from an immediate sale within 1-15 days of a new listing is greater than the sale price for houses that sell within 16-30 days. The somewhat surprising feature is that this price by time-to-sale then increases with duration. After 30 days, prices gradually rise with more delayed sales, eventually overtaking the 1-15 day price between 46 and 75 days depending on the date. This observation appears somewhat at odds with the familiar observation that at a point in time prices negatively correlate with the duration on the market at the time of sale. This U-shape pattern, however, does not reflect the volume of sales. Even at the low mark in 2008, half of all homes sold within two months; by 2010 three quarters of all homes sold in the first two months; by 2018 half of all homes sold within one month which is below the figure suggest by Redfin data suggest.

end (1-15 days) recovering faster and stronger. For durations greater than 15 days, prices by duration synchronously rise and fall over the cycle, fanning out around 2010. In contrast, the 1-15 day price starts at the top of the pack but its relative position falls as all prices fall. After reaching its lowest level, the 1-15 day price then rises sooner and faster, gradually overtaking longer durations as time progresses. Moreover, in any given time period after the slump ends, the largest price change (by duration) occurs between the first and the second two-week periods of a house being on the market.

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 US housing market. The model builds on Smith (2020b), which develops and characterizes a stationary stock-flow housing market with endogenous seller entry. The model here generalizes the matching framework in Smith (2020b) to an environment with both market prosperity and slump. Incorporating two market states with endogenous seller entry provides a potential amplification mechanism of market dynamics.

4.1 Environment

Buyers and Sellers. Buyers and sellers populate a small and isolated market for a house in continuous time. Buyers and sellers are risk neutral and discount the future at rate $r > 0$. Let d_b and d_s denote the flow search costs of buyers and sellers respectively. A buyer derives x units of discounted total lifetime utility from home ownership. For sellers, the flow utility from home ownership is normalized to zero. If a trade takes place, the consummating buyer and seller both permanently leave the market. An accepted bid at price P yields a payoff $x - P$ to the buyer and transfers revenue P to the seller. Unsatisfied buyers and sellers remain behind to wait for the next trading opportunity. Idle agents waiting for a possible trade do not leave the market. They are committed to trade so there is no free disposal in this sense.

The trading environment begins in a prosperous and thriving state, almost immediately suffers a slump for a period of time, and then recovers. As time proceeds during prosperity and slump, buyers looking for a house enter the market at a state contingent constant and exogenous Poisson rate. Using "hats" to represent, when needed, the aggregate slump state, let β and $\hat{\beta}$ denote buyer arrival rates in prosperity and slump periods respectively. 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 appear (and have the opportunity to offer their house for sale) at Poisson arrival rate α whereas relaxed sellers similarly appear at Poisson rate σ.

The market entry decision of sellers, however, is partially exogenous and partially endogenous. In each state, the different sellers make different decisions on whether to take advantage of the arrival opportunity and enter the market to sell their house. In the slump, motivated sellers enter if the accumulated number of houses-for-sale in the local market is below the cutoff inventory level which makes the expected payoff from entry equal to their outside exogenous option $\hat{V}$. In the prosperous state, motivated sellers like buyers do not have a choice and automatically enter the market.[14] In contrast, relaxed sellers do not enter the market during the slump.[15] In the prosperous state, they have the option to evaluate their prospects and choose whether to enter the market and take advantage of their entry opportunity or to decline entry. Relaxed sellers who accept an opportunity to enter the market in the prosperous state pay an up-front fixed cost of entry F.

As in Smith (2020b), in a thriving or prosperous market, 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 ($\alpha + \sigma > \beta$).[16] As such, potential sellers receive the opportunity to participate in the market more frequently than buyers but the realization of seller entry is endogenous because of the existence of relaxed sellers. The relaxed seller's entry decision balances the market while the entry of motivated sellers generates (from time to time) a non-negligible stock of unsold houses in the market. In a slump, the entry of buyers falls below the entry rate of motivated sellers ($\alpha > \hat{\beta}$) thereby creating over time a growing inventory of houses for sale.

Auctions and Bidding. A seller who has the option to accept or turn down the chance to enter the market knows the outcome of the previous transaction, including the number of buyers and the date

[14] In the simulations, payoffs during prosperity are above this entry threshold.

[15] The NAR new building permit count from Figure A.4 links to some degree with σ in the simulated model. This series declines sharply when the slump hits. Permits peaked at 2263 in September 2005 when prices also peaked. By November 2005 they declined to 1535 and bottomed out at 513 in March 2009. The gradual recovery from 2012 on only surpassed 1500 in October 2019. The sharp fall and sustained low level supports to some extent the zeroing out of relaxed entry during the simulated slump

[16] 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. In this paper, sellers have a higher arrival rate than buyers but they have the option to decline the opportunity to enter the market and save the associated sunk cost of participation.

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

State Shocks. Suppose after a run of time in the prosperous state that has settled the market, the market suffers an unanticipated slump. A housing shock occurs that leaves the existing trading structure unchanged but reduces the entry of buyers. In particular, the flow in rate of house buyers falls below the flow in rate of motivated sellers but the existing stock of buyers is unaltered. When the arrival rate of buyers is temporarily below the arrival rate of motivated sellers, a permanent recovery to the previous high arrival rate of buyers occurs with Poisson arrival rate $\hat{\tau}$.

Although buyers and sellers do not anticipate the initial shock and downturn, they account for the likelihood of recovery to prosperous trading in their payoffs and hence trading decisions during the slump. Buyers and sellers recognized that the slump is temporary and rationally factor in the likelihood of a recovery in their expected payoffs. As a result, the slump and the recovery when linked with the CoreLogic data derive from rational, forward looking buyers and sellers.

4.2 Trading Outcomes

Appendix B formally details the trading environment, describes the buyer's and seller's decision problems, and characterizes the trading outcomes. The main results are summarized here.

Immediate Trade. The first and critical feature of the stock-flow framework is that agents realize the gains to trade in the market 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 market 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 the highest bid, or more generally selects randomly among the set of identical, highest bids from the indifferent buyers. On the other hand, if there are excess sellers and a buyer enters the market, this buyer is on the short side of the market and hence the lone buyer. As such, the buyer will bid the reservation value of the seller or sellers who willingly accept the offer. The buyer then selects randomly among the houses for sale. In both cases trade occurs without delay after entry.

Immediate trade from stock-flow matching generates a two-step, piecewise linear hazard function. If there are many independent markets (as in the simulations below), some newly listed houses will find themselves on the short side and sell immediately. Less fortunate sellers will find themselves on the long side and must wait and then compete for buyers to enter the market. The likelihood of a sale on a given day or week is low for the long-side sellers thereby generating a two phase hazard function that declines sharply with duration or time on the market. Expected price similarly falls with duration.

The two different phases of the hazard and price functions reflect different market circumstances. The short-side phase of the hazard and associate prices depend on the stock of buyers but not so on the stock of other sellers. The long-side phase depends on the flow of buyers as well as the inventory of unsold homes. Adjustments over the business cycle in prices by short and long duration will differ accordingly. With lowered buyer entry during the slump, motivated sellers outnumber buyers in the long run. As a result, inventories will tend to build up over time re-enforcing the relaxed sellers' decisions not to enter the market. 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 during the flow phase becomes less likely and when it occurs, these sellers experience lower continuation payoffs. Their prices on average fall. This is tempered by outliers in markets that manage against the odds to build up buyers in the absence of relaxed seller entry.

Cold Spells During Prosperity. The second trading feature is that during prosperous times, some relaxed sellers decline entry and trade slows down. From time to time, cold spells occur. More specifically, relaxed sellers enter if they know or have reasonable expectation that a suitable number of bidders await in the market. In particular, sellers expect a profitable sale with two or more bidders. If one or no bidders are known or reasonably expected to be found immediately, these sellers pass on the entry option and the market slows down. Eventually sufficient time passes for relaxed sellers to expect buyer entry to have re-stocked the market and relaxed entry resumes.

The distinction in market power between one and two bidders plausibly determines the threshold state for relaxed seller entry. Monopolistic bidding allows the buyer to capture the trading surplus. If a monopolistic buyer can take advantage of the lack of competition and offers a bid (at or above the reservation sale price but) below the sunk fixed entry cost, relaxed sellers will decline the opportunity to participate immediately after auctions with zero or one bidder remaining. In contrast, with two or more Bertrand competitors on the long side of their market, market power resides with the seller. Competitive bidding from two or more buyers, creates a different scenario that can induce competing bidders to offer a profitable price. This monopoly and Bertrand distinction is plausible but obtains under explicit parameter conditions. Moreover, this condition does not hold in the simulated slump. At least three bidders would be needed. For this reason, the model and simulations assume that relaxed sellers do not enter.

5. Model Simulations

The stock-flow framework generates trading outcomes and empirical implications for comparison with the data observations described above. This section presents simulations of the above housing market model and demonstrates that the model generates trading patterns for the housing market variables over the sample period that are consistent with the data.

5.1 Primitives

The model described in section 4 represents a small and isolated housing market. In contrast, most available statistics typically 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 empirical approach here is to aggregate across many such small and isolated markets. In particular, the model is simulated for many small micro- or sub-markets that all share a common shock, and then aggregated over these submarkets to mimic a local area.

Following Smith (2020b), a submarket corresponds to a housing type in 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. Realtor data specify nine housing quality classes but not all are equally present in a school district. Focusing on the core five housing types quality classes as in Smith (2020b) yields 625 submarkets in a city with the average size of population.

The aggregated statistics for the local market are evaluated against the empirical findings in Section 3.3. This assessment considers levels, variability, correlations with prices, and autocorrelations and also emphasizes the shape and timing of events from the beginning of the slump through the recovery period before the 2020 pandemic. In continuous time, Poisson arrival rates imply that the associated waiting times between potential buyers and between potential sellers are both distributed exponentially so that more than one trade does not take place at any one instant. The organization of simulated events, however, needs to correspond to observed data from events that are recorded in discrete intervals during which more than one sale can occur. As a result, the replicated continuous time model is partitioned into and simulated for 256 intervals or discrete periods of 30 days each. The first 100 periods are dropped as the market settles around pre-crash outcomes.[17]

Summing over the month and over individual micro markets yields area-wide monthly new listings (in flows), sales (out flows), inventory for unsold homes (stocks), time on the market for sold houses (completed spells), and time on the market for the stock of inventory at the end of the month (uncompleted spells). The model also generates short and long side prices as well as short and long Kaplan-Meier trading probabilities.[18]

[17]The parameters (or some subset) could in principle be estimated using minimum distance, moment fitting methods. Such an approach poses practical difficulties. The threshold number of traders will not always satisfy monopoly-Bertrand split. At some high seller costs for F or d_s, profitable entry for sellers requires $N > 2$ buyers. Adopting an endogenous threshold for when sellers enter vastly expands the model and its computing complexity. Given this and other quantitative limitations, the approach adopted here seems adequate for demonstrating that the model can deliver outcomes broadly consistent with observed regularities.

[18]The simulation program does not follow individual houses so precise links to a wider variety of durations are unavailable. However, the simulations distinguish houses that sell quickly (very short immediate durations) from those that sell from the stock with longer stays in the market and tracks the average days on the market for these houses.

Table 2: Calibrated Parameters

Parameter	Meaning	Calibration rationale	Value
x	buyer's gain from trade	normalize	1
r	time preference rate (daily)	5% annual interest rate	0.05/360
β	arrival rate of buyers (daily)	average sales post-slump	0.071
$\hat{\beta}$	arrival rate of buyers (daily) during slump	average sales in slump	0.05
α	arrival rate of motivated sellers (daily)	see text	0.059
σ	arrival rate of relaxed sellers (daily)	see text	0.052
$\hat{\tau}$	expected market recovery rate from slump (daily)	expected duration of slump about 4.5 years	0.0006
d_b	buyer search cost	normalize	0
d_s	seller cost	ensures Assertion A.2 holds	0.000002
F	outside option for relaxed seller	see text	0.4
$\hat{V}$	outside option for motivated seller in slump	3/4 of F	0.3

Parameter Values. There are eleven parameters: $x, F, r, \alpha, \beta, \sigma, d_b, d_s, \hat{\tau}, \hat{\beta}$, and $\hat{V}$. In selecting values, the aim is not to precisely match the data. The quantitative objective is to show that the proposed stock-flow matching model, although abstracted from many complexities of the housing market, can generate patterns qualitatively similar to those found in the data as documented in Section 3.3. With this perspective in mind, normalize the buyer's gains from trade to $x = 1$ and let the daily rate of time preference be $r = 0.05/360$, corresponding to a yearly interest rate of 5%. To match the city-wide average sales in the slump and the post-slump recovery, the arrival rate of buyers is calibrated to $\hat{\beta} = 0.05$ and $\beta = 0.071$ respectively. In the slump a buyer in a particular market arrives on average once every three weeks. During recovery and prosperity a buyer arrives on average once every two weeks. In the recovery period, potential sellers arrive more than 1.5 times than the buyers: $\alpha + \sigma = 0.111$. The arrival rate for motivated sellers, $\alpha = 0.059$, is slightly more than that for the relaxed sellers, $\sigma = 0.052$, and so $\hat{\beta} < \alpha < \beta$. For every 10 buyers, this amounts to roughly 8 motivated sellers and 7 relaxed sellers.

The recovery rate from the slump $\hat{\tau} = 0.0006$ yields an expected duration of the slump of slightly more than four and a half years which is consistent with the length of the downturn in the housing market that began in 2006. The simulated slump lasts for 58 months, beginning in September 2006 and ending with a return to the prosperous state in July 2011.

The buyers' search cost is normalized to zero, $d_b = 0$. The sellers' cost is set to $d_s = 0.000002$. The outside option for a relaxed seller is set to $F = 0.4$, and the outside option for motivated sellers in the slump is set to be three quarters of F, $\hat{V} = 0.3$. Given these values, in the model simulation relaxed

Table 3: Simulated Statistics

	Source	Min	Max	% change max-min	Auto-corr (1 lag)	Corr with price
Sales	Model	873	1423	47.9	0.956	0.912
	NAR (2009-)	338,300	470,100	32.6	0.992	0.847
Listings	Model	958	1370	35.4	0.869	0.939
	NAR (2009-)	283,300	466,200	48.8	0.976	0.795
Inventory	Model	2130	8922	122.9	0.999	-0.716
	NAR	12,288,500	31,943,800	88.9	0.999	-0.558
Days on Market - Sales	Model	47.1	238.7	134.1	0.994	-0.553
	CoreLogic	46.5	85.8	59.4	0.992	-0.504
Days on Market - Inventory	Model	156.8	348.6	75.9	0.998	0.029
	CoreLogic	66.8	110.0	48.9	0.992	-0.837
Short Side Kaplan-Meier	Model	0.131	0.523	119.7	0.987	0.882
	CoreLogic - 1-15 Day	0.174	0.530	101.1	0.973	0.261
Long Side Kaplan-Meier	Model	0.042	0.145	110.3	0.990	0.879
	CoreLogic - 31-45 Day	0.151	0.395	89.3	0.957	0.368
Average Price	Model	0.303	0.384	23.6	0.994	–
	Case-Shiller (Real Index)	71.0	107.7	41.1	0.998	–

Note: 1. All data are for Jan 2007–Dec 2019 unless otherwise noted.
2. All data variables are in monthly frequency. Case-Shiller index is seasonally adjusted at source (FRED). Other monthly variables are seasonally adjusted by taking a ±6-month moving average. Model figures are not smoothed.
3. Inventory in the data is the value at the end of each month. The NAR sales are for existing home sales. The NAR listings and end of month inventory are computed from total houses for sale less houses sold in the month as noted.
4. Average prices from CoreLogic are in dollars. 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. Kaplan-Meier statistics from CoreLogic are conditional statistics derived from regressions following procedures outlined in Appendix A. 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, as in Table 1.

sellers will wait 17 days during the prosperity phase before entering a market with one known buyer and 39 days for a market with no known buyers. In the slump period, motivated buyers will not enter if inventories in the submarket exceed 26 unsold houses.

5.2 Simulation Results

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 level and changes of inventories in their two associated states. When the slump shock hits, buyer entry immediately falls below motivated seller entry. Listings exceed buyer entry. Inventories accumulate in each submarket until they reach the threshold limit that halts motivated seller entry. The short and long term hazard rates both fall as buyers become more scarce and competition among sellers becomes more fierce. These two factors likewise depress prices on aggregate and for each duration. When the downturn ends, buyer entry picks up but relaxed seller entry does not resume until inventories eventually clear in individual submarkets. Hazard rates improve but not in lock-step with each other as inventories dwindle. Prices rise immediately at the impact of the recovery, mirroring the rise in buyer entry. Prices, in particular for sellers in the inventory of houses already in the market, then continue to rise during the subsequent transition to the new steady state.

Using the set of parameter values described above to quantify these mechanics, Table 3 and Figures 6 -10 report simulated outcomes along with the preferred seasonally adjusted measure from the data. From Table 3 , the model implied peak-to-trough percentage changes are close to those in the data for the variables considered. The simulated model delivers the strong autocorrelations observed in all variables in the data. Simulated correlations with price are also consistent with their data counterparts although days on the market of unsold inventory is off target. In addition, the model further delivers, as expected from stock-flow matching, a close correlation (0.87) between listings and sales that aligns with the correlation (0.9667 in the NAR and 0.9933 in CoreLogic) in seasonally adjusted data.

Aggregate Patterns. Figure 6 plots the simulated listings and sales against the NAR data. Targeting the average sales levels in the slump and the post-slump recovery (with $\hat{\beta}$ and β), the magnitude of sales and listings (adjusted for market size) line up closely with the data over the sample period. The timing of simulated changes in sales and listings, however, is overly abrupt. The simplicity of the shock specification makes the fall and eventual recovery more stark than the NAR aggregated data.[19]

The associated rise and fall of simulated inventories in Figure 7 mirrors CoreLogic inventories, timing the peak around 2012 that marks the end of the slump. When recovery occurs, inventories fall as buyer entry now exceeds motivated seller entry and relaxed sellers wait for inventories to fall. On the other hand, the peak in the simulated inventory levels overshoots the Corelogic inventory counts. At

[19] 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 which would lead to substantial variation in performances spread around various markets. For US 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. 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 two weeks on the market, the rise in the hazard rate is particularly pronounced. This pattern, however, is not universal. A group of cities do not follow this pattern and maintain a flat function hazard over the sample period. These cities include Cleveland, Corona, Detroit, Jacksonville, Miami, Minneapolis, St Louis, St Paul and Tucson.

Figure 6: Sales and listings: Data vs Model simulated

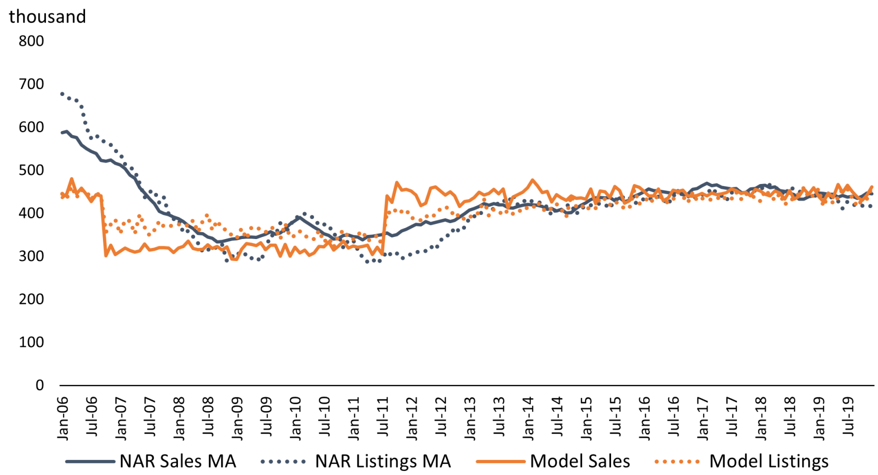
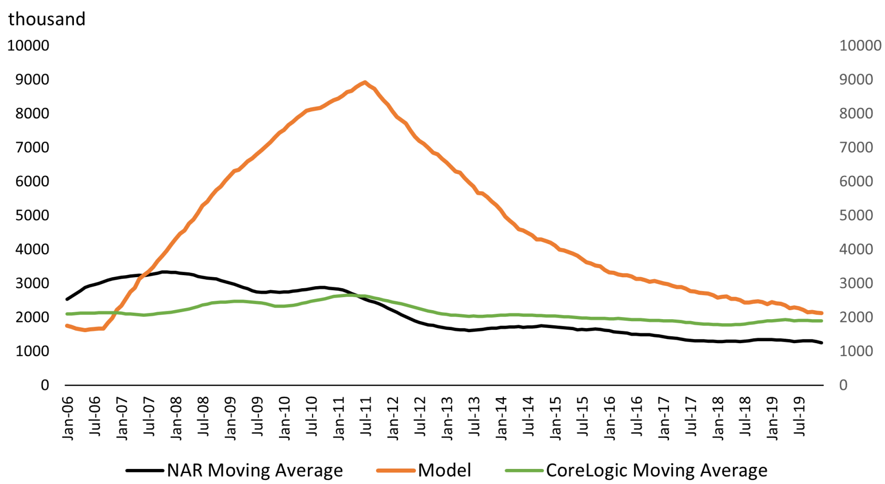


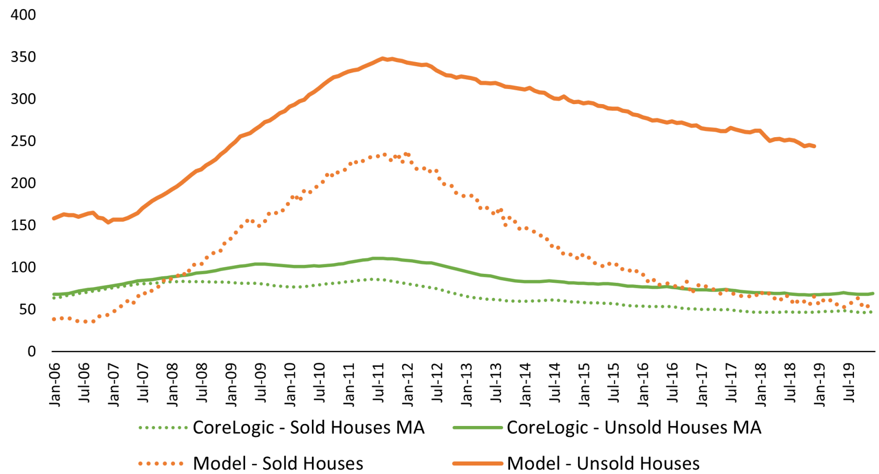
Figure 7: Inventories of existing homes: Data vs Model simulated



the peak, simulated inventories rise considerably higher - two to three times higher relative to sales - than the peak in the Corelogic data.

Several factors can be responsible for the sharp rise in inventories in the model. High inventories in the simulated model are partially due to a high inventory threshold after which motivated sellers refuse entry. In the simulations, this threshold is 26 houses which is highly saturated given the precise nature of a submarket. Inventories also overshoot due to the assumption that sellers in the market are committed to trade and cannot withdraw without a sale in order to rent or otherwise retain the house. Data from Redfin suggest that approximately 6% of houses are withdrawn from the market without a sale during the recovery after 2012. The slump may well have higher withdrawal levels. It may also be the case that real estate investment trusts (REITS) and we-buy-any-house companies purchase

Figure 8: Days on the Market: Data vs Model simulated



inventories once an individual house price becomes sufficiently low.

Figure 8 plots the simulated days on the market. For houses that sell within the month and for those that remain available at the end of the month, the simulations capture the qualitative pattern found in the CoreLogic DoM data—the distinction between the two measures, their rises in the slump and declines in the recovery, and their peaks in 2012. In addition, the simulated DoM for sold houses is close to the Corelogic DoM at the end of the recovery, at around 47 days. The peak DoM of sold homes, however, is two to three times higher than in the data. Likewise, simulated DoM for unsold houses exceeds the Corelogic data for all dates by roughly similar margins. The same factors that can lead to the overstatement of the simulated inventories can similarly drive the higher simulated DoMs during the slump and early recovery. Some of this excess may also be due to the Recovery Act inducing entry (as well as sales) from 2009 to early 2011.

As noted above and documented below, because of short sales from the inflow of new houses, stock-flow matching generates a discrete drop in the probability to sell. 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). The size of the drop in the selling probability controls the difference between the DoM of sold and unsold houses. A higher probability of an immediate sale generates a larger difference in the two series. A random search model cannot readily generate this difference as house sales have a constant hazard function.

Patterns by Duration. Table 3 establishes that the model can generate the sharp drop in the probability of selling between houses on the short and long sides. The simulated lows for the short and long Kaplan-Meier hazard rates are 0.131 and 0.042 respectively and their corresponding highs rise above 0.5 and nearly 0.15 respectively. These figures correspond well to the data. The simulated low and high for the short Kaplan-Meier hazards are close to the conditional Corelogic 1-15 day hazards while the values for the long Kaplan-Meier hazards are roughly one-third of the Corelogic 31-45 day hazards.[20]

Figure 9 not only exhibits the hazard function over time but also reveals that the progressions of the two simulated Kaplan-Meier hazard rates follow the patterns observed in the CoreLogic data. When the slump hits, both simulated hazard rates drop sharply but the decline is most pronounced in the short term hazard rate. As a sale becomes increasing 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. During the recovery phase, the model-simulated short hazard rate tracks the Corelogic 1-15 day hazard closely and the simulated long hazard appears approximately parallel to the Corelogic 31-45 day hazard but at a lower level.

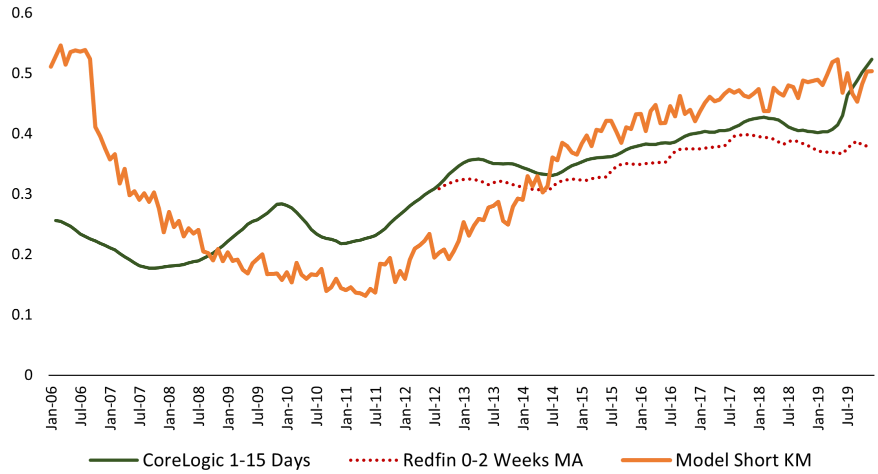
Despite the good fit of the model for the overall patterns, the timing of the recovery for the simulated hazard rates differs from the data. They decline throughout the downturn whereas the Corelogic counterparts begin to rise in 2008, fall again in 2010, and then recover in 2011. Housing market policies such as those embedded in the Recovery Act are likely responsible for exaggerating if not inducing the different timings found in the the model and in the data.

The last two rows of Table 3 report the statistics for housing prices. The percent difference in the peak and trough of the average price across all sales is 24% in the model, about half of the difference in the Case-Shiller price index. Figure 10 compares the simulated evolution of stock-flow house prices with the data. The top panel shows that similar to the Case-Shiller price index the model-simulated average price is U-shaped over the cycle with concurrent troughs around the end of 2011. The initial progression of prices in the model, an abrupt fall when the shock hits, is, however, more exaggerated

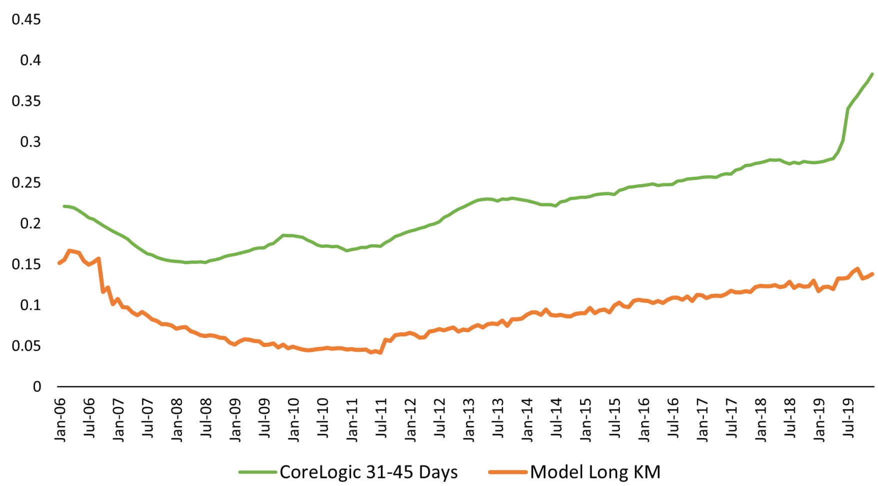
[20] One possible reason for the low levels of the long hazards might be that the model does not allow withdrawn of listings. Another mitigating factor is mis-classified short side trades. The 0-15 days may not capture all short side sales. Arefeva (2020) finds that in Redfin data between 2009 and 2016, 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 two weeks rises 0.61 - the high in the simulated data.

Figure 9: Kaplan-Meier Hazards: Data vs Model simulated

(A) Short-side Kaplan-Meier



(B) Long-side Kaplan-Meier



than in the Case-Shiller index. This exaggeration is because the on-and-off nature of the Poisson shocks generates an overly pronounced reaction.[21] This abrupt change in the simulated price is the reason for the low correlation between DoM for unsold houses and prices.

The lower panel of Figure 10 plots model-implied prices by duration along with CoreLogic hedonic prices. Although the evolution is still abrupt, the simulated price captures to some extent 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.

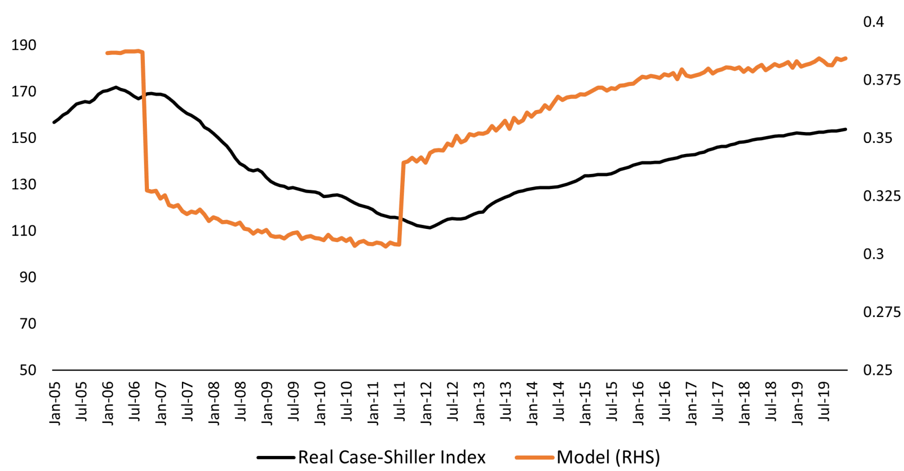
Taken together, the simulated results paint a supportive picture for the stock-flow matching ap-

[21] Again, note that average prices can mask the reality of particular local market responses in the timings if the areas are uncoordinated.

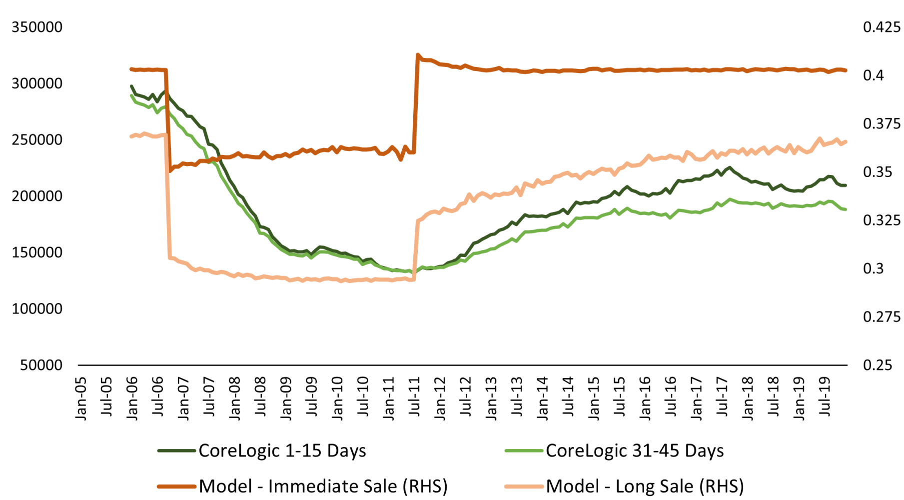
proach with Poisson shocks for explaining the cyclical behaviour of the housing market. The general mechanics match up well with not only the aggregate statistics but also with prices and sales by duration. Specifically, the ability of the framework to generate distinct responses in the selling probability and prices over short and long durations not only arises naturally in this specification but also plausibly generates patterns observed in the Corelogic data. This result is not a standard outcome of other dynamic trading formulations.

Figure 10: Prices by Duration: Data vs Model simulated

(A) Average house prices



(B) House prices by duration



6. Discussion

This paper analyses the recent performance of the aggregate US housing market. The paper details the remarkable collapse in 2006, the prolonged slump into 2011, the vigorous recovery beginning in 2012, and the continued expansion until the onset of the COVID 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. There is a markedly higher probability of selling and receiving a higher price in the two weeks immediately after listing a house for sale. Moreover, these advantages recover sooner, faster and more robustly as recovery occurs. The paper then demonstrates that a parsimonious stock-flow matching model can generate these observed outcomes.

The stock-flow framework has appealing features. For example, stock-flow matching provides a coherent setting for generating some houses being sold in bidding wars and others in a slower search-like fashion when no bidders are present and sellers wait for buyer entry. Stock-flow matching also inherently involves thin markets in which a slowly evolving history matters. Buyers and sellers cannot change circumstances rapidly period to period. It takes a while for the long side to evolve.[22] On the other hand, although the parsimonious structure highlights critical components, the particular stock-flow model adopted here abstracts from salient and quantitatively important elements of housing markets, many of which might not fit easily in with a stock-flow framework. For example, houses and willingness to pay are the same; there is no ex-ante or ex-post heterogeneity among buyers or sellers; there are no rental markets; the joint buyer-seller problem is absent.[23] Perhaps more fundamentally, there are no movements across different submarkets. In practice, although house sellers tend to be immobile and might not differ substantially in how much they like the sales price, buyers who find stiff competition for a particular type of house in a confined neighborhood might expand their horizons.

Despite these abstractions, the quantitative performance of the model is compelling. Under viable

[22] 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.

[23] Allowing for heterogeneity among buyers or sellers as found in many pairwise matching models such as Arefeva (2020) would require keeping track of the number of every type of trader, a burdensome computation. Likewise, borrowing constraints and credit frictions (Stein 1995, real estate agents (e.g. see Hendel et al. 2009), construction delays (see Davis and Heathcote 2007; Gyourko and Saiz 2006; and Head et al. 2014); Genesove and Mayer 2020; Ortalo-Magné and Rady 2006; and Hedlund 2016) do not exist. Halket and di Custoza 2015 considers the impact of the rental sector whereas Haurin 1988 and Ngai and Sheedy (2020a) assess idiosyncratic and potentially privately known house characteristics. Ngai and Tenreyro 2014) include increasing returns to matching.

parameterizations, simulations deliver outcomes broadly consistent with not only the observed timing and pattern of familiar aggregates but also with trading outcomes by duration—trading probabilities and prices by time on the market. The simulated model captures the distinct trading outcomes seen in the data of those who trade quickly and of those that take time. The ability to mimic the known quantities of sales, average prices, listings and time on the market and to correspond to the trading patterns by duration indicates the fundamental importance of stock-flow matching for understanding the performance of the housing market.

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.

Arefeva, A. (2020): "Auctions, microstructure noise, and weekly patterns in housing markets," Working Paper.

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. 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.

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.

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.

Drenik, A., H. J., and P. Ottonello (2019): “Measuring information frictions: Evidence from capital markets,” Mimeograph.

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

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

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*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Elsevier, vol. 5 of *Handbook of Regional and Urban Economics*, 813–886.

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. (2016): “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 husing market won’t look much better in 2011,” *The Atlantic*.

Krainer, J. R. (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.

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 (2020a): "The decision to move house and aggregate housing-market dynamics," Working Paper.

——— (2020b): "The Ins and Outs of Selling Houses: Understanding Housing-Market Volatility," Working Paper.

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.

Rekkas, M., R. Wright, and Y. Zhu (2021): "How Well Does Search Theory Explain Housing Prices?" Working Paper.

Shiller, R. J. (1982): "Consumption, Asset Markets, and Macroeconomic Fluctuations," NBER Working Paper 0838.

Shimer, R. (2007): "Mismatch," *American Economic Review*, 97, 1074–1101.

Smith, E. (2020a): "High and low activity spells in housing markets," *Oxford Research Encyclopedia of Economics and Finance*.

——— (2020b): "Stock-flow models of market frictions and search," *Review of Economic Dynamics*, 36, 1–28.

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

Taylor, C. R. (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.

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

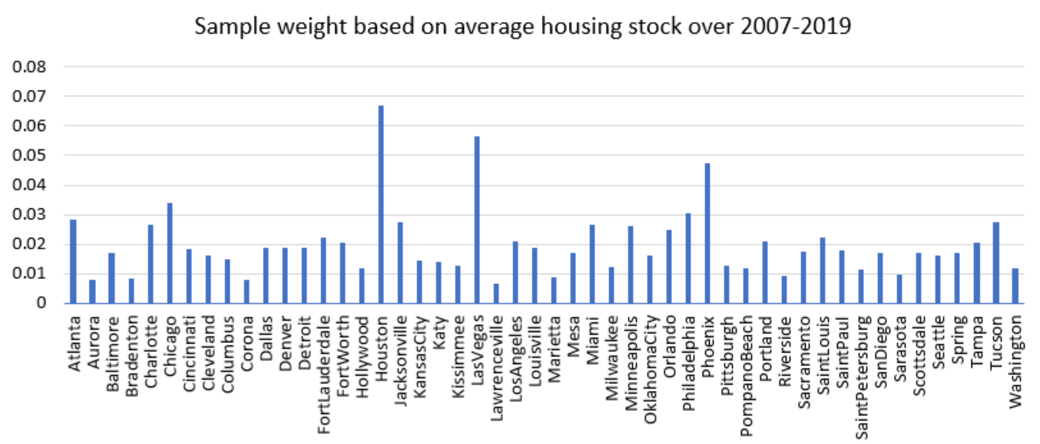
Appendix for
"The US Housing Market: A Stock-Flow Matching Perspective"

A. Data and Variable Construction

Data and sample. The empirical analysis uses the CoreLogic's Multiple Listing Service (MLS) data for the years 2006–2019 which 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 percent 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 the duplicate listings through listing dates and property house. Dropping the duplicate listings gives the new listing to sales ratio close to the ratio in the NAR data.

A sample is then selected consisting of 50 metropolitan markets with sizeable population and enough listings over the sample period. The total sample size across 50 markets and over 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 weight in the sample.

Figure A.1: Cities and sample weights



For the Kaplan-Meier statistics and price series, both the raw statistics and conditional (hedonic) series are constructed. In the paper, the conditional series are focal since they control for differences in housing and market characteristics. But the raw series do not deviate meaningfully from the conditional series. The methodology of constructing these series are explained in detail below. Table A.1 compares the raw and conditional series and

Table A.1: Comparison of conditional and unconditional statistics

Variable	Source	Min	Max	% change max-min	Auto-corr (1 lag)	Corr with price[a]
1-15 Day Kaplan-Meier	CoreLogic raw	0.162	0.386	81.7	0.990	0.269
	CoreLogic cond.	0.174	0.530	101.1	0.973	0.261
31-45 Day Kaplan-Meier	CoreLogic raw	0.158	0.271	52.7	0.991	0.409
	CoreLogic cond.	0.151	0.395	89.3	0.957	0.368
Average Price	CoreLogic raw	106,130.1	205,374.3	63.7	0.980	–
	CoreLogic cond.	130,372.7	254,987.1	64.7	0.969	–

Note: All data variables are in monthly frequency and seasonally adjusted by taking a ±6-month moving average. [a]Correlation with price uses the conditional price series for the conditional Kaplan-Meier statistics and the unconditional price series for the unconditional Kaplan-Meier.

show that they exhibit similar patterns.

Unconditional 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 1-15 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:

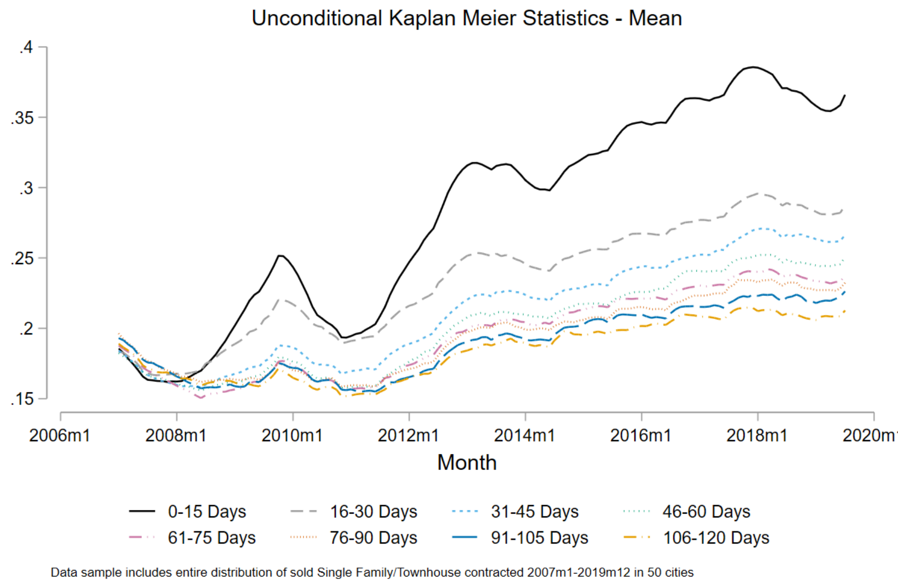
- 1-15 duration: A house is coded 0 if it is listed in the first 1-15 days of the month; and 1 if it also goes under contract in the month and 1-15 days after listing; missing otherwise.
- 16-30 duration: A house is coded 0 if it is listed 1-15 days before the month; and 1 if it also goes under contract 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 1-15 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 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 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.2 shows the time series for these unconditional statistics, averaged across cities.

Conditional Kaplan-Meier statistics. The conditional Kaplan-Meier statistics are constructed using coefficients

Figure A.2: CoreLogic: Monthly unconditional Kaplan-Meier statistics by duration



from the following logit regression, separately for each city:

$$I_{it} = \Omega X_{it} + \sum_d \theta_t^d D_{it}^d + \sum_d \gamma_m^d D_{it}^d + \zeta_z + \epsilon_{it} \tag{3}$$

where X_{it} is a vector of housing characteristics for house i at 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 I_{it} is an indicator for whether a house is coded 1 (or 0) in the definition of the unconditional Kaplan-Meier statistics – this makes sure that the same sample of houses are used for construction of the unconditional and conditional statistics. The set of main independent variables D_{it}^d represent the duration category dummy in month t for duration d defined over 15 day intervals. The set of estimated parameters θ_t^d is allowed to vary by month to capture the variation in the effect of duration over time, while γ_m^d captures any seasonality effect in the variation of D_{it}^d. The conditional statistic for each city in month t is then computed as:

$$\hat{K}_t = \left[1 + e^{-(\theta_t^d + \Omega \bar{X} + \bar{\zeta}_z)}\right]^{-1} \tag{4}$$

where $\bar{X}$ is the median value of the housing characteristic variable for each city; $\bar{\zeta}_z$ is the median zip code fixed among houses in the city; θ_t^d and Ω are parameters estimated from (3). 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.

Conditional price statistics. The conditional price statistics are estimated based on a sample of sales prices, after trimming the extreme (top and bottom 5%) values. The price coefficient is estimated from:

$$\ln(P_{it}) = \Omega X_{it} + \sum_d \theta_t^d D_t^d + \zeta_z + \gamma_m + \epsilon_{it} \tag{5}$$

where γ_m is month fixed effects to control for seasonality in prices, in addition to the housing characteristics and zip code fixed effects. The conditional hedonic price estimates are then:

$$\hat{P}_t = e^{\theta_t^d + \Omega \bar{X} + \bar{\zeta}_z} \tag{6}$$

Additional time series. 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. As seen in Figure A.4 these counts are small relative to exiting homes and broadly follow the same evolution as existing homes.

Another familiar housing series is the houses-for-sale to houses-sold ratio. HUD provides this series and its evolution over time corresponds to the ratios from CoreLogic and Redfin. The Corelogic inventory-sales ratio looks plausible relative to the other two sources and roughly fits within the HUD ratio of total houses for-sale to total houses sold in Figure A.3. The monthly HUD House for Sale to Houses Sales ratio in Figure A.3 peaks at 12.2 in January 2009 before declining to 4.0 in January 2013 where after it has risen slowly. The NAR reported ratio of "Inventory-to-Sales" at least from 2005 until 2015 closely tracks in levels and over time the HUD measure highly suggesting that inventories from the NAR and Redfin measure all houses for sale in a month, not the number observed for sale at the end of the month as calculated using the Corelogic sample. This distinction partially accounts for the lower inventories to sales ratios from Corelogic. Moreover, from June 2007 until the end of the sample period the trend line for the NAR imputed listings is relatively flat given the observed volatility which aligns with the Corelogic observation.

The model generates price dispersion that varies over the bust and boom of the market. Comparing these with the distribution residual prices at different phases of the market - before the crash, the bottom of the downturn, the middle of the recovery, it is tempting to link the the Corelogic dispersion to the dispersion in the model. Given the challenges of accurately specifying and measuring price dispersion, we prefer to be cautious. There are many ways to measure dispersion. Moreover the observed dispersion in cities is far from uniform across cities. More careful analysis of the data is required. Likewise, as the within city pattern of dispersion is very mixed across cities a more thorough evaluation of the Corelogic data is needed for comparison. Nonetheless, model dispersion exhibits potential for looking at dispersion over the cycle, an assessment not found in the literature.

Figure A.3: Inventory-to-Sales Ratio

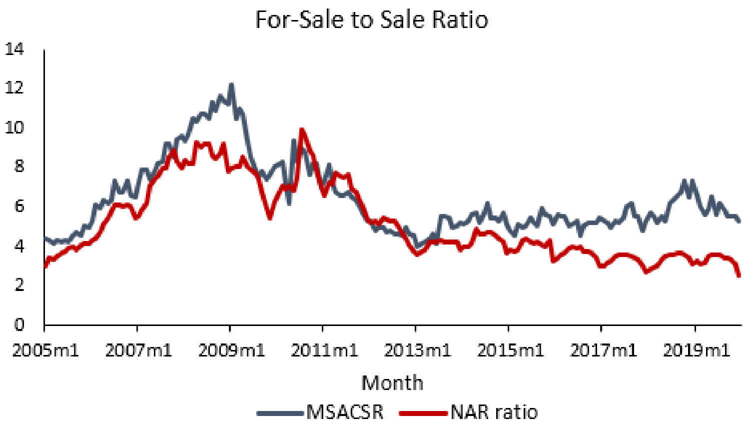
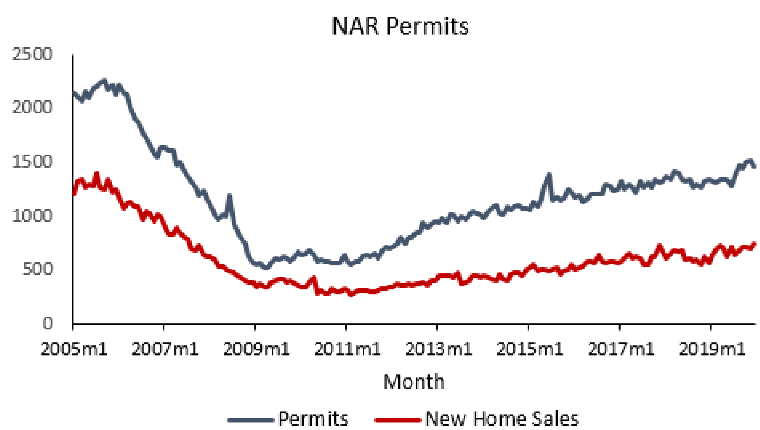


Figure A.4: The NAR Housing Permits



B. The Stock Flow Model with Endogenous Entry and Aggregate Shocks

Let $r > 0$ be the discount rate. Let d_b and d_s denote buyer and seller search costs respectively. A buyer derives $x = 1$ units of discounted total lifetime utility from home ownership. For sellers, the flow utility from home ownership is normalized to zero. If a trade takes place, the consummating buyer and seller both permanently leave the market. An accepted bid at price P yields a payoff $x - P$ to the buyer and transfers revenue P to the seller. Unsatisfied buyers and sellers remain behind to wait for the next trading opportunity. Idle agents waiting for a possible trade do not leave the market. They are committed to trade so there is no free disposal in this sense.

Prosperity

The analysis of the prosperous state closely follows Smith (2020b). The parameter β represents the inflow rate of buyers in this state. The parameters α and σ denote the inflow rates of motivated and relaxed sellers respectively. To keep the market well behaved during prosperous trade, these rates satisfy $\alpha < \beta$ and $\alpha + \sigma > \beta$.[24] Relaxed sellers do not have the option to wait or return at some other time. As a result, the payoff to a seller who decides to turn down the trading opportunity is zero. The fixed cost of entry is $F > 0$ hence the potential gains from trade between a potential seller and an existing buyer are $x - F > 0$.

Since buyers enter at Poisson rate β, buyers and sellers have common beliefs that the probability of i entrants over a duration D since the last auction is

$$\pi_i(D) = \frac{e^{-\beta D}(\beta D)^i}{i!} \tag{7}$$

As discussed in Coles and Muthoo (1998), buyers and seller decisions also involve immediate trade:

Assertion A.1. Immediate trade occurs so that the market never simultaneously has unsatisfied buyers and unsatisfied sellers.

In an immediate trade equilibrium with fully revealing auctions, the integer, N, denoting the known participants remaining from the last sale and a duration $D \geq 0$ denoting the time since the last seller entry and auction summarize the (sub)state of the submarket within the aggregate state of prosperity. In particular, if over the course of market history up to the time of the last sale, a total of N_B buyers and N_S sellers had entered the submarket, then

$$N = N_B - N_S$$

represents the number of known buyers (sellers) remaining after a sale for $N > 0$ ($N < 0$).

Given buyer entry $i \in \mathbb{N}$ since the previous auction and a vanishing interval length $dt \to 0$, transitions in the

[24] The standard specification in the stock-flow literature is that buyer and seller entry rates are equal and exogenous so that market balances over time. Persistent unequal entry would eventually lead towards an infinity of buyers or sellers.

submarket follow

$$\begin{aligned}\Gamma(N,D) &= (N, D+dt) \mid No\, Seller\, Entry \\ &= (N^{'}, 0) \qquad \mid Seller\, Entry \\ &= (N^{'}, dt) \qquad \mid Seller\, Acceptance\, or\, Rejection\end{aligned} \tag{8}$$

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 phase to include not only revealed buyer entry i but also the seller's own entry.

Without trading delays, periods between observed sales result from waiting times between seller entry when there are excess buyers, or between buyer arrivals if there are excess sellers. When relaxed sellers turn down entry opportunities, these waiting times become prolonged. Let the entry decision therefore define hot and cold states.

Definition. A state for a submarket is hot when relaxed sellers accept the option to enter, and is cold when potential sellers decide not to participate.

Suppose the transaction price falls as the number of bidders declines. Since the submarket becomes less profitable for sellers with fewer bidders, potential relaxed sellers will decline entry (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 which reflects the competitiveness of bidding. The proposed outcome is that relaxed sellers enter in any state with two or more known available bidders but entry ceases for some period of time following an auction with one remaining buyer who bids monopolistically not competitively.

Assertion A.2. (Monopoly - Bertrand Partition) The market is hot for all states where $N \geq 2$. For $N = 1$ and $D = dt$ the market is cold.

This assertion implies that the market becomes cold immediately following an auction with two bidders. The simulations check that this threshold is valid.

To revitalize expected profit and relaxed seller entry so that cold states 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 or become available, no auctions occur and potential sellers do not observe buyer entry. As time proceeds without trade, the number of potential buyers grows stochastically. Expected prices rise. For markets with zero or one known buyer, trade in the market eventually becomes profitable, relaxed seller entry resumes and the market becomes hot. The selling process coupled with Assertion A2 implies that cold markets transition to a hot market from states with $N \in \{0, 1\}$ after endogenous durations $D = T_1$ and $D = T_0$ respectively.

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 market inactivity without any (motivated) seller entry. After some length of time, seller expectations of (unobserved) buyer entry eventually improve enough to induce entry of relaxed sellers. 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 the market and hence resets the entry decision of potential 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. In this case, the market remains cold - no relaxed seller entry - until the stock of available homes is sold and then followed by an appropriate cold period of duration T_0 to replenish buyers.

Entry of motivated sellers occurs during both active and inactive markets. Relaxed sellers will not enter markets with excess supply until all of the previous sellers who entered consummate trades. Motivated sellers, however, may enter to cause excess supply. Really cold markets, those with excess sellers, remain cold until balance is restored. 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 markets will experience having excess sellers. A non-negligible inventory of unsold homes can build up.

Consider the following sequence of events. After a sufficient duration from the time of an auction with only one bidder, suppose no buyers have entered and relaxed seller entry gets switched on. If a seller then enters before a buyer, there are excess sellers ($N = -1$). The unsuccessful seller is the only agent currently in the market so no sale takes place. As potential relaxed entrants observe an existing seller already in the market and the absence of any sale with excess sellers, they will turn down entry until the existing house is sold. Trade eventually takes place when a buyer arrives in which case the market transits to $N = 0$ and $D = dt$. After such a sale, the market is empty with no houses for sale and no bidders wanting to buy. Thus, immediately after that trade, relaxed entrants continue to decline trading opportunities until after a further duration T_0. Hence, A2 implies the set of hot states 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 states are cold.

In state $(N, D = 0)$, an auction takes place. If $N \geq 0$ there is one house for sale and $N + 1$ buyers bidding. If $N < 0$ there is one bidder and $-N + 1$ houses for sale. Let $P(N)$ denote the buyers' symmetric price bids in this auction. The seller's payoff from acceptance is simply the price offered. If the buyer purchases a competitor's house, the unsuccessful seller or sellers receive $Z(N)$, their expected payoff of waiting in the market for the next auction.

Let $H(N)$ represent the payoff to a house buyer from being in a hot market 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 market who must wait the remaining duration $T > 0$ before the market becomes hot again and the entry of relaxed sellers resumes. By Assertion A2 a market with $N > 1$ is active whereas buyers are absent for $N \leq 0$. The only relevant cold market payoff for characterizing trade therefore occurs at $N = 1$.

After a seller accepts a bid in an auction with two buyers, one buyer remains in the market. The market becomes cold. The buyer who does not purchase the house receives expected payoff $C(T_1; N = 1, D = 0)$ where T_1 denotes the duration that this seller must wait after the sale before seller entry resumes. For all $D \geq T_1$, the state $(N = 1, D)$ is hot. The duration of a cold spell depends on the number N of excess buyers remaining from the last trade or seller entrance (if no trade occurred) hence the subscript notation with $N = 1$. From the buyers perspective, the market is hot if $N \geq 2$ and non-existent for $N \leq 0$ making $C(T_1)$ payoff relevant to the buyer.

Before entry occurs, for $N \geq 0$ the seller's revenue 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 market - but not the transaction price. Following a cold phase with one buyer of duration T_1 without any seller entry, expected revenue less the entry fee for a seller contemplating entry into the market 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 market, 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 market and without any seller entry, expected revenue less the entry fee for a seller contemplating entry into the market 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 payoff $Z(-1)$. If one buyer arrived since the last auction, the market is balanced at $N = 0$ with seller entry. The lone bidder offers $P(0)$ and the transaction takes place immediately. Likewise for two or more bidders with the number of bidders all coming from new buyers.

By Assertion A2, revenue is positive and relaxed seller entry occurs for $N \geq 2$. For $N \leq -1$, relaxed sellers do not enter until the existing house sells - expected revenue $R(-1, D)$ to seller entry is negative.

To solve the auction stages recursively for $N \geq 0$, note that the seller's acceptance payoff is strictly increasing in price. Hence, the seller's best response strategy in an auction has the reservation property. In state $(N, 0)$, the seller accepts $P(N)$ if and only if $P(N) \geq Z(N)$.

Now consider the bidding phase and suppose first that $N = 1$ and $D \geq T_1$, so that the market has become hot after a cold spell. If a seller enters and no buyer entry had occurred during the cold spell, N shifts and the ensuing auction proceeds with $N = 0$. With one bidder making an offer for the one house for sale, the buyer optimally bids the seller's reservation price, $P(0) = Z(0)$.

With two or more bidders ($N \geq 1$) price offers are derived from the hot and cold market payoffs $H(N)$ and $C(T_1)$ where to ease notation the understood state $N = 1, D = 0$ in $C(T_1)$ is now dropped. Following Taylor (1995) for $N \geq 1$, the buyers become indifferent between paying $P(N)$ and waiting for the next seller to enter, whether in a cold market for $P(1)$ or in hot market for $P(N)$ with $N \geq 2$.

A similar logic (but without cold spells) applies for $N \leq 0$ where a single bidders makes offers to one or more sellers. In this auction, the buyer's offer makes the seller indifferent between waiting and accepting,

$$P(N) = Z(N)$$

These monopolistic and competitive bidding scenarios yield the following characterization of prices offers.

Lemma 1. Price offers are given by :

$$P(N) = Z(N) \qquad N \leq 0$$

$$P(1) = x - C(T_1)$$

$$P(N) = x - H(N) \qquad N \geq 2$$

For $N = 1$ and $N = 2$, a bidder's expected payoffs in a hot market 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 $N \geq 3$

$$\begin{aligned} H(N) \quad &= \quad \frac{1}{1 + rdt}\left[(\alpha + \sigma)dt\, H(N - 1)\right. \\ &\qquad \left. + \beta dt\, H(N + 1) + (1 - (\alpha + \sigma + \beta)dt)H(N) - d_b dt\right] \end{aligned}$$

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 market during the cold period after a duration t and triggers an auction with the existing bidders and any other buyers who might have entered during the cold period up to time t. In this environment, the buyer's expected payoff in a cold market can be written as

$$\begin{aligned} C(T_1) &= \int_0^{T_1} e^{-(r+\alpha)t}[-d_b + \alpha \sum_{i=0}^{\infty} \pi_i(t)[x - P(i+1)]]dt \\ &\quad + e^{-(r+\alpha)T_1} \sum_{i=0}^{\infty} \pi_i(T_1) H(i+1) \end{aligned}$$

Solving gives

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

where

$$\begin{aligned} \Psi_1 &= A_1 + A_2 \\ \Psi_2 &= (\eta - D_4)A_3 + D_4 A_4 - D_4 A_5 + A_6 \\ \Psi_3 &= D_5 A_3 - D_5 A_4 + D_5 A_5 - A_7 - A_8 \end{aligned}$$

$$\begin{aligned} A_1 &= \frac{\alpha(1 - e^{-(r+\alpha+\beta)T_1})}{r + \alpha + \beta} \\ A_2 &= e^{-(r+\alpha+\beta)T_1} D_1 \\ A_3 &= \frac{\alpha\beta[1 - e^{-(r+\alpha+\beta)T_1}(1 + (r+\alpha+\beta)T_1)]}{\eta(r+\alpha+\beta)^2} \\ A_4 &= \frac{\alpha(1 - e^{-(r+\alpha+\beta(1-\eta))T_1})}{\eta^2(r+\alpha+\beta(1-\eta))} \\ A_5 &= \frac{\alpha(1 - e^{-(r+\alpha+\beta)T_1})}{\eta^2(r+\alpha+\beta)} \\ A_6 &= e^{-(r+\alpha+\beta)T_1}(D_2 + \frac{e^{\beta\eta T_1} - 1}{\eta})D_4 \\ A_7 &= e^{-(r+\alpha+\beta)T_1}[rD_2D_5 + rD_3 + \frac{e^{\beta\eta T_1} - 1}{\eta}(rD_5 - 1) + e^{\beta T_1} - 1] \\ A_8 &= -(1 - e^{-(r+\alpha)T_1}) + \frac{\alpha(1 - e^{-(r+\alpha+\beta)T_1})}{r+\alpha+\beta} \end{aligned}$$

$$\begin{aligned} D_1 &= \frac{\alpha + \sigma}{r + \alpha + \sigma + \beta} \\ D_2 &= \frac{\beta}{r + \alpha + \sigma + \beta} \\ D_3 &= \frac{1}{r + \alpha + \sigma + \beta} \\ D_4 &= \frac{\alpha + \sigma}{r + \alpha + \sigma + \beta(1-\eta)} \\ D_5 &= \frac{r + \beta(1-\eta)}{r + \alpha + \sigma + \beta(1-\eta)} \end{aligned}$$

It can be established that sellers accept the buyers' bids in Lemma 1, i.e. $P(N) \geq Z(N)$ for all $N \geq 0$. These prices coupled with the cold state payoff $C(T_1)$ in turn determine bidder payoffs $H(N)$ in a hot market. For $N = 1$ and $D > 0$, the buyer is alone in the market waiting for the arrival of a seller. Motivated sellers will enter whereas relaxed seller entry occurs after the transition from a cold period without buyer entry. Two such transitions occur. One transition occurs after a duration $D > T_1$ following a two bidder auction during which no buyer entry occurred. The second occurs after duration $D > T_0$ that follows an auction with one bidder during which no buyer entry occurred.

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

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

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

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

From time to time, entry from one side or the other of the market will occur such that the auction has one bidder and one seller. In this auction, the buyer's offer again makes the seller indifferent between waiting and accepting. Given that buyer entry or seller entry will shift the market, 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

$$
\begin{aligned}
B_1 &= \frac{\beta(r+\alpha+\beta-\alpha\lambda)}{(r+\alpha+\beta)(r+\alpha+\beta-\alpha\lambda)-\alpha\beta} \\
B_2 &= \frac{(r+\alpha)(r+\alpha(1-\lambda))+r\beta}{[(r+\alpha+\beta)(r+\alpha+\beta-\alpha\lambda)-\alpha\beta]}
\end{aligned}
$$

Now consider a potential seller contemplating the market in the entry phase. This relaxed seller's decision depends on whether the expected revenue outweighs the cost of entry. If the expected revenue following seller entry is less than the up-front fee F, relaxed sellers decline entry. They accept otherwise.

From Assertion A2, relaxed sellers accept entry opportunities and markets are hot for all states in which $N \geq 2$. Moreover, these sellers decline entry and markets are cold for $N = -1$. When $N \in \{0, 1\}$, relaxed sellers turn down entry over some period of time after a sale. For $N = 1$ entry resumes after a period T_1. Relaxed sellers turn down entry in states $(N = 1, D < T_1)$ but the market becomes hot at $(N = 1, D = T_1)$ and remains hot for all $D \geq T_1$. For $N = 0$ and $D = dt \to 0$, T_0 denotes the duration of the subsequent cold spell.

Expected net revenues or profit 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 bidders (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 bidders $(i \geq 2)$, 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 spell 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 2. For sellers aware of only one or zero known bidders, the critical cold spell duration of delayed entry T_1 and T_0 respectively solve

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

and

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

The last step is to verify that the monopoly-Bertrand partition in A2 is valid. Given prices, payoffs and cold spells, sellers would enter monopolistic ($N \leq 1$) markets and the market would remain hot if discounted expected net revenue is positive immediately following an auction that left one bidder remaining in the market:

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

Conversely, markets with two known remaining bidders 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 (2020b) derives explicit restrictions on parameters for zero search costs. Here, it suffices to check the simulated values.

Slump

The determination of payoffs and the solution approach for outcomes in the slump follow the same logic as in the prosperous state with two exceptions. First, relaxed seller entry shuts down thereby ruling out the cold state in a submarket. There is no hot/cold state distinction in a submarket for seller entry during the slump so revenue is omitted. Second, inventories of unsold houses become bounded. In the prosperous state, the inventory of unsold homes is potentially unbounded but this possibility is not a substantial concern. The flow in of buyers 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 market. During a slump, however, the flow in of motivated sellers exceeds the low flow in of buyers. Although the end of the slump reverses this outcome, it is reasonable to limit the number of houses for sale in a given submarket during a downturn by capping the entry of (motivated) sellers at some level. As motivated buyers will not enter a submarket with sufficiently many unsold homes, suppose that during the slump period, motivated sellers obtain an outside payoff $\hat{V}$ if they do not enter. (This option is not available in the boom. Inventories rarely if ever reach the boundary associated with the slump. As sellers will be more willing to enter a crowded market in a boom, this omission appears unimportant and nonbinding.)

Agents expect the aggregate slump state to end with Poisson arrival rate $\hat{\tau}$. During the slump, the parameters $\hat{\beta} < \alpha$ and $\sigma = 0$ represent the inflow rate of buyers and relaxed sellers. The inflow of motivated sellers α is unchanged. On the rare occasion when the recovery occurs with one known buyer in the (sub)market, buyers anticipate a cold market for the duration T_1 and payoff $C(T_1)$ as specified below. 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 $P(0)$.

For $N \geq 2$, the buyer payoffs become

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

which can be re-written as

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

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

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

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

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

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 bidder's expected payoffs

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

can be re-written as

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

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

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

where

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

For $N=1$, a bidder's expected payoffs in a hot market

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

can be re-written as

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

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

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

where

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

Rearranging and solving gives

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

where

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

Note that

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

Seller payoffs during the slump must account for their outside option. In this case, a critical threshold value of inventories of unsold houses in the submarket exists and triggers non-entry 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 inventories as in the prosperous period, 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 non-zero object so that the solution joins up with the outside option $\hat{V}$. In particular, recalling that inventories for sale are negative values of the state 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 $\hat{V}$.

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

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

can be re-written as

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

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

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

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

and from the particular solution

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

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

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

and

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

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

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

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

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

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

can be re-written as

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

Plugging in

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

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 $P(0)$.

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

where

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

Balanced trade implies

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

can be re-written as

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

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

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

or

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

where Ψ_{10}, Ψ_{11}, Ψ_{12}, Ψ_{13} are as above and

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

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

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

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

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