The Ins and Outs of Selling Houses: Understanding Housing-Market Volatility*

L. Rachel Ngai†
London School of Economics

Kevin D. Sheedy‡
London School of Economics

January 2020

Abstract

This paper documents the cyclical properties of housing-market variables, which are shown to be volatile, persistent, and highly correlated with each other. Is the observed volatility in these variables due to changes in the speed at which houses are sold or changes in the number of houses that are put up for sale? An inflow-outflow decomposition shows that the inflow rate accounts for almost all of the fluctuations in sales volume. The paper then shows that a search-and-matching model with endogenous moving subject to housing demand shocks performs well in explaining fluctuations in housing-market variables. A housing demand shock induces more moving and increases the supply of houses on the market, thus one housing demand shock replicates three correlated reduced-form shocks that would be needed to match the relative volatilities and correlations among key housing-market variables.

Keywords: housing market volatility; search frictions; inflow-outflow decomposition; endogenous moving.

*We thank Adam Guren, Morris Davis, Martin Gervais, Lu Han, Chris Pissarides and especially Allen Head for helpful discussions, and Thomas Doyle and Christopher Jenkins for assistance with the data. We also thank participants at the 2011 Search-and-Matching Research Group conference, the 2013 Spring Housing-Urban-Labor-Macro conference, and the 2015 Society for Economic Dynamics Annual Conferences for their comments. Rachel Ngai acknowledges support from the British Academy Mid-Career Fellowship.
†LSE, CEPR, and CfM. Email: L.Ngai@lse.ac.uk
‡LSE and CfM. Email: K.D.Sheedy@lse.ac.uk

1 Introduction

While house prices have been the focus of much research, surprisingly there has been little attempt to document systematically the joint behaviour of the key housing-market variables, namely sales, new listings, time-to-sell, and the stock of houses for sale, alongside house prices. This paper shows all these variables are highly correlated, and argues that the decision to move house explains why these variables are closely linked.

The first contribution of this paper is to assemble and document a set of stylized facts about the cyclical properties of a broad set of housing-market variables. Using sales and inventory data for existing homes from the National Association of Realtors (NAR) and the price index from the Federal Housing Finance Agency (FHFA), this paper constructs time series for sales, new listings, the number of houses for sale, the average time taken to sell, and prices for the period 1991–2012. The new findings are that listings are strongly positively correlated with sales and prices, negatively correlated with time-to-sell, and slightly more volatile than sales and significantly more volatile than prices. All the variables have high volatility, and both sales and listings have a high degree of serial correlation.

What factors can account for the volatility observed in these housing-market variables? Stock-flow accounting allows us to think of the number of houses for sale as the *stock*, the sales volume as the *outflow* and the listings as the *inflow*. Fluctuations in these variables are responses to changes in outflow and inflow rates. Is the observed volatility due to changes in the speed at which houses are sold (outflow rate), or changes in the number of houses that are put up for sale (inflow rate)? This question is answered using an inflow-outflow decomposition, a technique that has been applied in search-and-matching approaches to the labour market (Shimer, 2012, Petrongolo and Pissarides, 2008, Fujita and Ramey, 2009), with houses for sale as the equivalent of unemployment. The decomposition shows that inflow and outflow rates are of approximately equal importance in accounting for fluctuations in the stock of houses for sale.

While the stock of houses for sale is interesting because it represents the number of homeowners who anticipate a gain from selling that has yet to be realized, the duration of time they expect to remain in this state is just as important. Taken together, the measure that matters for welfare is the volume of sales. In fact, this measure is what receives most attention. Sales volume as turnover tells us how much reallocation of housing is occurring, where each reallocation results in gains from trade. Applying a similar decomposition of sales volume in terms of inflow and outflow rates shows that fluctuations are almost entirely due to changes in the inflow rate. The 'ins and outs' decomposition demonstrates that changes in the inflow rate are the dominant factor in understanding housing-market fluctuations. The inflow rate reflects individual decisions of homeowners to put their house up for sale.

The second contribution of this paper is to study housing-market dynamics using the endogenous-moving model of Ngai and Sheedy (2019). Central to the model is the idea of match quality between a house and its owner. The search friction assumed in the model is that when searching for houses, potential buyers are not sure how much they will like a house until a viewing has taken place. This is modelled as a distribution of match-specific quality, as in the matching model of Jovanovic (1979) (and adapted to the housing market by Krainer, 2001).

The buying decision is described by a cutoff rule where a sale occurs when the draw of new match quality is above a certain threshold. Match quality is persistent, so this decision process generates an endogenous distribution of existing match quality. If there were no changes in existing match quality or the outside option as determined by market conditions, homeowners would not consider moving even if they had the option to. To generate an incentive to move, the model allows for occasional idiosyncratic shocks that degrade existing match quality. If an idiosyncratic shock occurs, homeowners can decide whether the degraded match quality is still good enough for them to remain or whether they should put their current house up for sale and look for another. The cutoff rule that describes the moving decision provides a further (and more important) reason why the distribution of existing match quality is endogenous.

The parameters of the model are calibrated using the costs associated with housing transactions and the average times between inflows and outflows using U.S. data. Considering only the usual shock to housing demand (and setting its persistence to match the serial correlation of prices), the model performs remarkably well in replicating the stylized facts about relative volatility, comovement, and persistence documented earlier.

Match quality plays a crucial role in the workings of the model and its ability to explain the stylized facts with only a housing demand shock. A positive housing demand shock (raising the flow utility from all houses) raises the total surplus from a transaction and thus increases both the willingness to trade and the price. This generates a positive correlation between sales and price, which would not arise in the absence of a distribution of match quality. Due to the equilibrium distribution of match quality among existing homeowners, a persistent housing demand shock increases the incentive to invest, leading to more listings. This explains the positive correlation between listings and sales and prices, and the similar volatility of listings and sales. Furthermore, an increase in listings is essential to sustain a lasting effect on sales and thus replicate the serial correlation of sales. A rise in the sales rate alone is not sufficient because the stock of houses for sale would be quickly depleted.

The empirical success of the model is because endogenous moving means that a housing demand shock induces more moving, acting like a moving-rate shock, as well as increasing the supply of houses on the market, acting like a housing supply shock. Thus one housing demand shock replicates three correlated 'reduced-form' shocks that would be needed to match the stylized facts. Simply introducing independent shocks to moving or housing supply would not match the data because such shocks would imply a negative relationship between prices and sales.

To see this, two special cases nested in the model are considered: (i) where the inflow rate is an exogenous constant, and (ii) where the inflow rate is subject to exogenous time variation. The first of the special cases is manifestly inconsistent with the high volatility of inflows. The second case, while it can be consistent with the volatility of inflows, fails to match the relationship between the volatilities of sales, the number of houses for sale, and the time taken to sell (and several of the correlations among these variables). This is essentially because it introduces entry into the market orthogonal to variables that matter for other decisions, such as the search behaviour of buyers and sellers once they are in the market. In contrast, the response of other housing-market variables to shocks is related to changes in the inflow rate when the inflow rate is modelled as an endogenous variable.

The model features an endogenous moving rate for existing homeowners. Just as important as this is the non-random selection of those who move. This gives rise to a 'cleansing effect' on the distribution of existing match quality because those near the bottom of the distribution are more likely to move following an idiosyncratic shock. Importantly, the strength of this effect also varies endogenously over time. One consequence of this is the possibility of 'overshooting' in response to an aggregate shock. For instance, suppose an aggregate shock decreases current threshold for not moving now, which means there will be less moving, and also less 'cleansing'. Consequently, in the future, the distribution of existing match quality will be worse, making future moving more likely. These findings indicate that modelling inflows as exogenous is far from an innocuous simplifying assumption.

There is a strand of the literature (starting from Wheaton, 1990, and followed by many others as surveyed by Han and Strange, 2015) that studies frictions in the housing market as here with a search-and-matching model.[1] The key contribution of this paper to the literature is in studying housing-market inflows, with moving being exogenous in the earlier papers with the exception of Guren (2014) and Hedlund (2016), but those papers focus on price fluctuations and foreclosures.[2]

In exploring the cyclical behaviour of the housing market,[3] Díaz and Jerez (2013) is the closest to this paper in terms of its goals of examining a range of important housing-market statistics, namely, average house prices, sales, the number of houses for sale, and time-to-sell. That paper also presents a calibrated search-and-matching model to show that three correlated shocks (housing demand, housing supply, and the moving rate) are needed to account for the cyclical properties of the variables. With an endogenous moving decision, this paper shows that only one demand shock is needed. The empirical success of the model is because endogenous moving implies that a housing demand shock induces more moving, acting like a moving shock, as well as increasing the supply of houses on the market, acting like a housing supply shock. Thus one housing demand shock replicates three correlated 'reduced-form' shocks that would be needed to match the stylized facts.[4]

[1]See, for example, Albrecht, Anderson, Smith and Vroman (2007), Anenberg and Bayer (2013), Caplin and Leahy (2011), Coles and Smith (1998), Díaz and Jerez (2013), Garriga and Hedlund (2019), Head, Lloyd-Ellis and Sun (2014), Ioannides and Zabel (2017), Krainer (2001), Moen, Nenov and Sniekers (2019), Ngai and Tenreyro (2014), Novy-Marx (2009), Piazzesi and Schneider (2009), and Piazzesi, Schneider and Stroebel (2019).

[2]Another difference is in the nature of the search frictions used in the model. A popular approach is to assume that the search friction is the difficulty of buyers and sellers meeting each other, often modelled using a matching function as in the canonical labour-search model (Pissarides, 2000, chapter 1), which leads to a market tightness (the buyer-to-seller ratio) paying a key role. Instead, this paper focuses on the difficulty of knowing which houses would be a good match prior to being viewed by a buyer. This difference of approach, along with some supporting evidence, is discussed in detail when the model is presented.

[3]Davis and Heathcote (2005) is one of the first studies to look at housing and the business cycle, focusing on the role of residential investment. Another strand of the literature focuses on credit constraints. See, for example, Fisher and Gervais (2011), Iacoviello (2005), Ortalo-Magné and Rady (2005), Stein (1995), and Ungerer (2015), See Davis and Van Nieuwerburgh (2015) for a survey of housing and business cycles. For the role of home production, see recent work by Aruoba, Davis and Wright (2016).

[4]Garriga and Hedlund (2019) show that endogenous housing illiquidity can generate a positive correlation between

The plan of the paper is as follows. Section 2 presents the data and stylized facts and performs a variance decomposition of sales volume into inflow and outflow components. Section 3 presents the search-and-matching model with endogenous inflow and outflow decisions, and derives analytical results characterizing the steady state of the model. In section 4, the model is calibrated to match some key features of U.S. data. Section 5 then performs simulations of the calibrated model subject to aggregate shocks and assesses the role of the endogenous inflow decision in accounting for the joint time-series behaviour of inflows, outflows, and the number of houses for sale. Section 6 concludes.

2 Data

The Federal Housing Finance Agency (FHFA) provides a monthly repeat-sales house-price index (HPI) for single-family homes.[5] Here, the purchase-only index is used, which excludes refinancing. Data are available from January 1991. The repeat-sales index is the best available index that controls for the quality of the housing stock because it is designed to capture price changes of the same houses.

The National Association of Realtors (NAR) provides monthly estimates of sales and inventories of houses for sale at the end of each month (for existing homes including single-family homes and condominiums).[6] For consistency with the FHFA house-price index, the data for single-family homes are used, which represent about 90% of total sales of existing homes.

2.1 Summary statistics

The NAR data on house sales and inventories of unsold houses can be used to construct a measure of new listings (the number of houses put up for sale), the stock of houses for sale, and the average time taken to sell. Let N_t denote the inflow of houses that come on to the market during month t, and S_t sales (the outflow from the market) during that month. If I_t denotes the beginning-of-month t inventory (or end-of-month $t-1$) then the stock-flow accounting identity is:

$$N_t = I_{t+1} - I_t + S_t. \qquad [2.1]$$

Assuming inflows N_t and outflows S_t both occur uniformly within a month, the average number of houses U_t available for sale during month t is equal to:

$$U_t = I_t + \frac{N_t}{2} - \frac{S_t}{2} = \frac{I_t + I_{t+1}}{2}. \qquad [2.2]$$

Since the time series for inventories is quite persistent, the measure U_t of the number of houses for sale turns out to be highly correlated with inventories (the correlation coefficient is equal to 0.99).

house prices and sales.

[5]Methodology and data are available at http://www.fha.gov.

[6]Methodology and recent data are available at http://www.realtor.org/research-and-statistics/housing-statistics. The NAR data are for existing homes, so newly constructed houses are excluded.

Using the constructed series U_t for houses for sale, time-to-sell is defined as the ratio of the houses on the market to sales (U_t/S_t). This measure is highly correlated with the 'months supply' number reported by NAR, which is defined as the ratio of inventories of unsold houses at the end of the previous month divided by the number of houses sold in the current month. The mean is 6.4 months, compared to 6.6 for 'months supply', and the correlation coefficient is 0.99.

Thus, using non-seasonally adjusted monthly data on prices, sales, and inventories covering the period from January 1991 to March 2012, monthly series are constructed for new listings, the number of houses for sale, and time-to-sell. The monthly data are deseasonalized by removing the average of each month. To smooth out excess volatility due to measurement errors in the data, quarterly time series are constructed from the monthly series. Standard deviations and correlations of the quarterly series (as percentage deviations from their seasonal averages) on sales, prices, new listings, houses for sale, and time-to-sell are shown in Table 1. Standard deviations are reported relative to the volatility of sales. [7]

As is well known in the literature, there is a positive comovement between house prices and sales with a correlation of 0.72, and a negative correlation between time-to-sell and sales of -0.64, and sales volume is highly volatile compared to other macroeconomics variables such as GDP (see Díaz and Jerez (2013)). In addition to these familiar statistics, the results here show that new listings are as volatile as time-to-sell and more volatile than sales, with all three more volatile than house prices. The finding that both sales volume and new listings are highly volatile is consistent with Bachmann and Cooper (2014), who show that housing turnover is volatile using data on the flows within owner-occupied housing obtained from the Panel Study of Income Dynamics (PSID). New listings are positively correlated with sales and prices with correlation coefficients of 0.85 and 0.60 respectively, and negatively correlated with time-to-sell with a correlation of -0.62. Finally, the stock of house for sale is uncorrelated with sales volume but positively correlated with prices, which is consistent with the findings of Bachmann and Cooper (2014) and Díaz and Jerez (2013) respectively, even though the latter uses the vacancy rate from the American Housing Survey instead of inventory data from NAR. All variables are highly serially correlated as can be seen from the autocorrelation functions in Figure 1.

The observations on sales and time-to-sell illustrate the importance of search frictions for houses already on the market, where the key decision is when to leave the market (when to stop searching for buyers and sellers). This decision and the associated measure of time-to-sell have thus been the main focus of search theory applied to the housing market. However, as shown in Table 1, the observations on new listings reveal that there is substantial variation in inflows to the housing market

[7] This table is not directly comparable to findings of Díaz and Jerez (2013) because their data are detrended using the Hodrick-Prescott filter. A comparable table is provided in section A.1. Another important difference is that this paper assembles data on sales, the number of houses for sale, and time-to-sell, from the same data source rather than the three different sources used by Díaz and Jerez (2013). More specifically, in Díaz and Jerez (2013), sales data are from NAR, time-to-sell is measured only for newly built homes ('New Residential Sales' from the U.S. Census), and houses for sale are the 'vacant for sale' measure from the U.S. Census Bureau Housing Vacancy Survey. Note that the 'vacant for sale' measure includes only a small fraction of houses that are actually for sale because it excludes houses that are occupied but available for sale. According to Table 1 of NAR's methodological documentation, vacant homes are approximately 11% of all single-family homes sold.

Table 1: *Summary statistics*

	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	0.671	1.289	1.012	1.364
			Correlation coefficients		
Sales	1				
Price	0.720	1			
New listings	0.850	0.602	1		
Homes for sale	0.070	0.661	0.009	1	
Time-to-sell	−0.635	−0.004	−0.617	0.689	1

Notes: Monthly data (January 1991–March 2012), seasonally adjusted, converted to a quarterly series.

and the decision to enter the market is also very important. It is natural, therefore, to document more precisely the contributions of the inflow (listings) and outflow (sales) rates to the volatility in housing-market variables. Does this result from changes in the difficulty of selling houses, or changes in the number of houses that are put up for sale?

2.2 Decomposition of housing-market volatility

This section performs an inflow-outflow decomposition of housing-market volatility. The approach taken follows closely that of the labour-market analysis of Petrongolo and Pissarides (2008) (and see also Fujita and Ramey, 2009), with houses for sale as the equivalent of unemployment. Given the limits imposed by data availability, it is assumed that flows occur only between two states: either houses are for sale, or are occupied by their owners and not for sale. This ignores the entry of newly built houses into the market, and transitions of houses in or out of the rental market.[8]

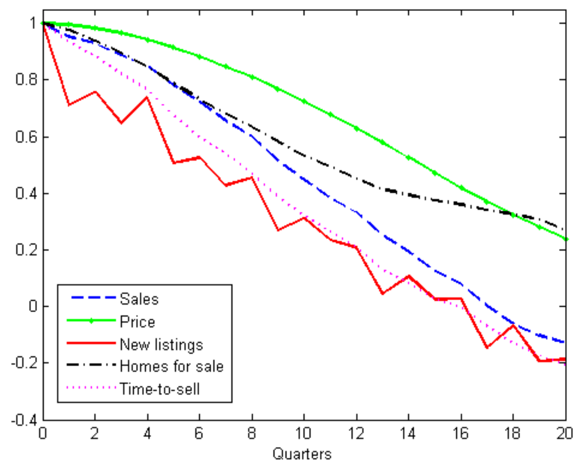
The two-state decomposition arguably still provides a good representation of overall housing market flows given that time lags in the construction of new houses are significant, and as new houses represent only a small fraction of total sales in the data. According to the 'U.S. Housing Market Conditions' report,[9] new homes sold amounted to only 6% of the number of existing homes sold in 2010 (which is consistent with the NAR sales data used in this paper), and inventories of unsold new homes were only 5% of inventories of unsold existing homes. Similar percentages are found over the time period starting in 1991 using the Census Bureau's 'New Residential Sales' data to measure the number of new homes.[10] Moreover, according to U.S. Census Bureau data,

[8]This is analogous to the two-state assumption often found in decompositions of labour-market volatility, where transitions in and out of the labour force are ignored. Houses here are the equivalent of the stock of unemployed in the labour market.

[9]U.S. Department of Housing and Urban Development, Office of Policy Development and Research, February 2011, p. 20.

[10]For the UK, the Office for National Statistics' House Price Index report (Table 28) shows that 8.7% of all mortgages were for new houses.

Figure 1: *Autocorrelation functions*



Notes: Monthly data (January 1991–March 2012), seasonally adjusted, converted to quarterly series.
Source: National Association of Realtors

the growth rate of the total housing stock was approximately the same as that of the population over the period 1991–2012 (both around 1% per year). The housing stock per person was thus approximately constant during the time period studied in this paper.

The two-states restriction also ignores transitions of houses between owner-occupation and the rental market. However, Glaeser and Gyourko (2007) argue that there is little evidence in support of significant arbitrage between the rental and owner-occupied segments of the housing market because owner-occupied homes typically have different characteristics from rental units, as is also the case for homeowners themselves in comparison to renters. More recently, Bachmann and Cooper (2014) calculate gross flows across and within the owner and renter categories using PSID data. They conclude that rental and owner-occupied markets are distinct segments owing to the dominance of moves within the same tenure category. More importantly, for the purposes of the inflows-outflows decomposition, they find that fluctuations in housing turnover are largely driven by fluctuations in the within-segment flows. Thus, omission of a rental market might not be too serious a concern for the decomposition. In any case, the available data limit the decomposition to two states.

The first step in the decomposition is to compute the quarterly inflow and outflow rates given the NAR data on sales and inventories, assuming that these rates are constant within each quarter. During time period t let the continuous-time outflow rate be s_t ('sales') and the inflow rate be n_t ('new listings'). The equations describing the total outflows S_t and inflows N_t during period t are:

$$S_t = \left(1 - e^{-s_t}\right) I_t + \int_0^1 \left(1 - e^{-s_t(1-\tau)}\right) N_{t+\tau} \mathrm{d}\tau, \qquad [2.3a]$$

and

$$N_t = \left(1 - e^{-n_t}\right)(K - I_t) + \int_0^1 \left(1 - e^{-n_t(1-\tau)}\right) S_{t+\tau} \mathrm{d}\tau, \qquad [2.3b]$$

where K is the total housing stock, assumed to be constant. These equations take into account inflows and outflows that occur within a quarter. For example, in (2.3a), the second term corresponds to flows of houses that enter the market and are sold within the quarter, analogous to the case in the labour market where a newly unemployed worker finds a job within one month. Accounting for such intra-period flows has been important in research on labour-market flows. Whether it is also important for the housing market will depend on the average flow rates, but it is important not to rule out such a possibility a priori. In any case, the second term represents houses that are available for sale, and should therefore be counted in the computation of the sales rate s_t and the resulting measure of time-to-sell.

Assuming that the inflows and outflows happen uniformly within the time interval $[t, t+\tau)$, equations (2.3a) and (2.3b) can be simplified to

$$S_t = \left(1 - e^{-s_t}\right) I_t + \left(1 - \frac{1 - e^{-s_t}}{s_t}\right) N_t, \qquad [2.4a]$$

and

$$N_t = \left(1 - e^{-n_t}\right)(K - I_t) + \left(1 - \frac{1 - e^{-n_t}}{n_t}\right) S_t. \qquad [2.4b]$$

Given data on S_t, N_t, I_t, and K, equation (2.4a) can be solved for a value of the outflow rate s_t, and equation (2.4b) for a value of the inflow rate n_t.

The data on sales S_t and inventories I_t are taken directly from NAR, and the number of houses N_t that are newly listed for sale is derived from the stock-flow accounting identity (2.1). The raw monthly data are converted to quarterly series and are seasonally adjusted. It is also necessary to have a measure of the total housing stock K. Note that as shown in equation (2.4a), the total stock K has no effect on the level of the outflow rate, and more importantly, as will be become clear later, K also has no significant effect on the decomposition of the relative importance of inflows versus outflows. The main effect of K is on the average level of the inflow rate calculated using equation (2.4b).

The total housing stock K should measure all houses that are either for sale or might be put up for sale, and it should be consistent with the sales and inventories data from NAR for existing single-family homes. According to Table 1.A of the U.S. Census Bureau American Housing Survey (available biannually from 1991 to 2011), single-family homes as a fraction of total housing units have been fairly stable at about 67% between 1991 to 2011, which amounts to approximately 80 million units on average. But not all of these should be counted as part of the stock K because some are houses for rent and some are vacant (for reasons other than being for sale). Taking out such houses leaves approximately 78% of single-family homes.[11] This number also includes newly built homes. Using the U.S. Census Bureau New Residential Construction data, the stock of new single-family homes for sale is about 1% of the total from 1991–2011.[12] Finally, houses owned by people older than 65 are excluded, which removes a further 19% of the housing stock.[13] The justification for this is that older people are much less likely to be active in the housing market. According to the American Housing Survey (Table 2.11), the most important reasons for moving house are changing jobs, getting married, having children, and wanting to be closer to schools. After all these adjustments are made, the resulting value of K is approximately 50 million units.

Figure 2 displays the time series for the outflow and inflow rates s_t and n_t. It is clear that both rates are highly volatile. The goal of the following decomposition is to understand their relative importance in explaining fluctuations in housing-market variables.

The law of motion for the fraction of inventories, $i_t = I_t/K$, is approximately

$$\Delta i_t \approx n_t(1 - i_t) - s_t i_t, \qquad [2.5]$$

where $n_t(1 - i_t)$ is the inflow and $s_t i_t$ is the outflow (both relative to the total stock of houses). The coefficients of variation of the inflow and outflow rates are equal to 0.27 and 0.21 respectively, suggesting that both might be relevant for understanding fluctuations in the number of houses for sale, and as a result, fluctuations in sales volume.

The method for decomposing variation in houses for sale follows the procedure used in the labour literature of approximating the unemployment rate by the steady states implied by the time-varying inflow and outflow rates. The justification is that the flow rates are sufficiently large, implying that the steady states would be reached quickly. This allows for a simple decomposition in terms of inflows and outflows by looking at the contribution of each to the change in the steady-state houses for sale (the equivalent of the unemployment rate).

While the stock of house for sale is interesting because it represents the number of homeowners who anticipate a gain from selling that has yet to be realized, the duration of time they expect to remain in this state is just as important. Taken together, the measure that matters for welfare is the volume of sales. Thus, a similar decomposition of variation in sales volume is also presented.

The inflow and outflow rates n_t and s_t imply the following steady-state fraction of houses for sale u_t^* and steady-state sales volume ($u_t = U_t/K$ has the same steady state as i_t):

$$u_t^* = \frac{n_t}{s_t + n_t}, \quad \text{and} \quad S_t^* = \frac{s_t n_t}{s_t + n_t} \qquad [2.6]$$

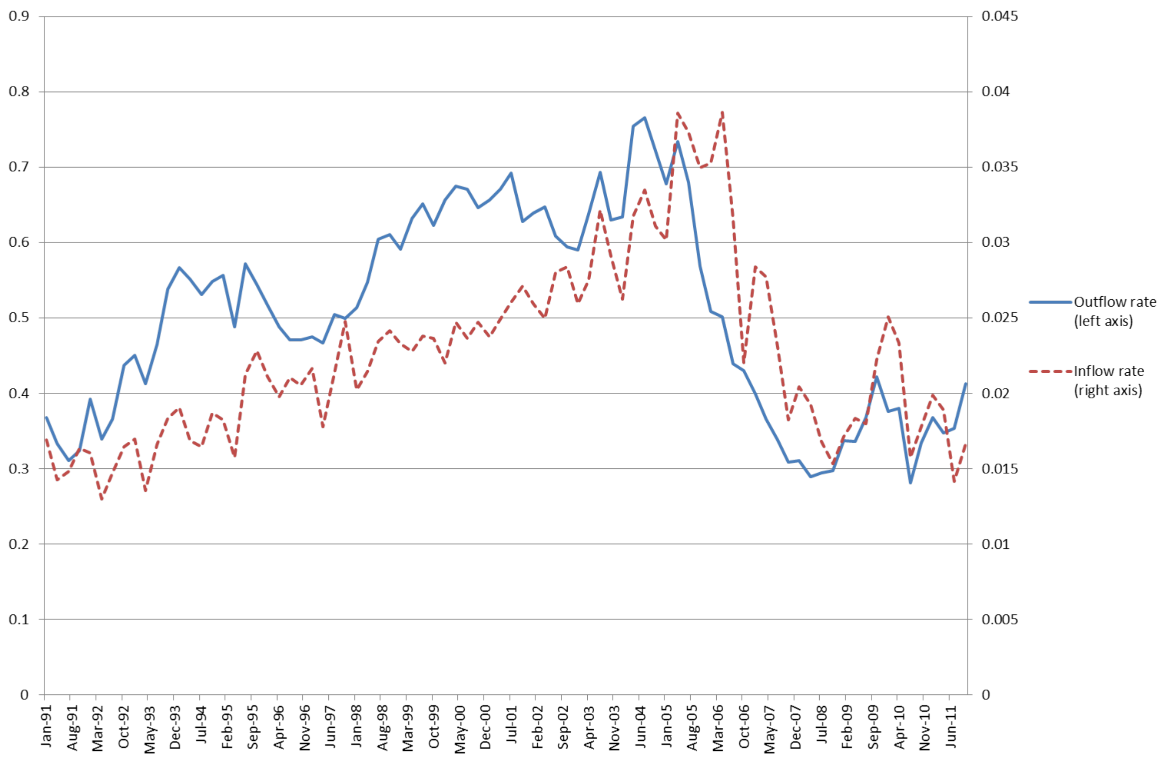
The variables u_t and u_t^* have the same mean and a correlation coefficient equal to 0.86, and the correlation coefficient between S_t and S_t^* is 0.88. Thus, the use of a decomposition based on steady states is valid.

[11] More precisely, this is the proportion of single-family homes that are either 'owner-occupied', 'vacant for sale', or 'rented-or-sold' from the eleven AHS surveys covering the period 1991–2011.

[12] Table 1.A-1 of the AHS survey also reports new construction numbers, but for a period of four years. The New Residential Construction data reports new single-family homes completed every year.

[13] This number is obtained using Table 2.9 of the AHS survey.

Figure 2: *Inflow and outflow rates*



Notes: Monthly data, seasonally adjusted, converted to quarterly series. The outflow and inflow rates are calculated using equations (2.4a) and (2.4b).
Source: National Association of Realtors

The decompositions of u_t^* and S_t^* are given by:

$$\Delta u_t^* = \Delta u_{n,t}^* + \Delta u_{s,t}^*, \quad \text{where } \Delta u_{n,t}^* \equiv (1-u_t^*)u_{t-1}^* \frac{\Delta n_t}{n_{t-1}}, \quad \text{and } \Delta u_{s,t}^* \equiv -(1-u_{t-1}^*)u_t^* \frac{\Delta s_t}{s_{t-1}};$$
$$\frac{\Delta S_t^*}{S_{t-1}^*} = \Delta\sigma_{n,t}^* + \Delta\sigma_{s,t}^*, \quad \text{where } \Delta\sigma_{n,t}^* \equiv (1-u_t^*)\frac{\Delta n_t}{n_{t-1}}, \quad \text{and } \Delta\sigma_{s,t}^* \equiv u_t^* \frac{\Delta s_t}{s_{t-1}}. \qquad [2.7]$$

The contributions of the inflow and outflow rates are assessed by calculating the regression coefficients:

$$\beta_j^u = \frac{\operatorname{Cov}\left(\Delta u_t^*, \Delta u_{j,t}^*\right)}{\operatorname{Var}\left(\Delta u_t^*\right)}, \quad j = \{n, s\} \qquad , \qquad \beta_j^S = \frac{\operatorname{Cov}\left(\frac{\Delta S_t^*}{S_{t-1}^*}, \Delta\sigma_{j,t}^*\right)}{\operatorname{Var}\left(\frac{\Delta S_t^*}{S_t^*-1}\right)}, \quad j = \{n, s\}.$$

The results are that $\beta_n^u = 0.91$ and $\beta_s^u = 0.09$ for the stock of houses for sale and $\beta_n = 0.98$ and $\beta_s = 0.02$ for sales volume. Since the decomposition depends on the actual fraction of houses for sale being close to its steady-state value, following Petrongolo and Pissarides (2008) the regression coefficients can also be calculated using only those time periods where the difference between i_t and

i_t^* as a fraction of i_t is less than 10%.[14] With this restriction, $\beta_n^u = 0.51$ and $\beta_s^u = 0.49$ for the stock of houses for sale. The restriction does not affect the results on the decomposition of sales volume.

In summary, this investigation of the data suggests that fluctuations in the number of houses that are put up for sale (inflows) are much more important than fluctuations in the number of houses sold (outflows) in understanding variation in housing-market variables. This is the first contribution of the paper. The second contribution is to show that a model with endogenous inflows as well as outflows can successfully match the cylical properties of the data presented in Table 1 with only one housing demand shock.

3 The model

This section presents a stochastic version of the endogenous moving model of Ngai and Sheedy (2019) that studies both the decisions to buy and sell houses and the decision to move house. As in the analysis of the data in section 2, the model focuses on the market for existing homes. This abstracts from new entry of homes due either to new construction or houses that were previously rented, and abstracts from the entry of first-time buyers into the market.[15]

3.1 Houses

There is an economy with a unit continuum of families and a unit continuum of houses. Each house is owned by one family (though families can in principle own multiple houses). Each house is either occupied by its owning family and yields them a stream of utility flow values, or is for sale on the market while the family searches for a buyer.[16] A family can occupy at most one house at any time. If all a family's houses are on the market for sale, the family is in the market searching for a home to buy and occupy.

Families discount the future at (continuous) rate $r > 0$. Time is indexed by t, and families make decisions at discrete time intervals τ. All units of time are measured in years throughout.

3.2 Search frictions

The housing market is subject to search frictions. First, it is time-consuming and costly for buyers and sellers to arrange viewings of houses. Let u_t denote the measure of houses available for sale and b_t the measure of buyers. Each buyer and each house can have at most one viewing in the time interval $[t, t+\tau)$.[17] For houses, this event has Poisson arrival rate $\mathcal{V}(u_t, b_t)/u_t$, where $\mathcal{V}(u, b)$ is a

[14] This subsample includes 52 out of the whole sample of 83 quarters.

[15] It is implicit in the model that families moving house might temporarily use the rental market in between selling and buying. However, there is no explicit modelling of the rental market: effectively, this is treated as a distinct segment of the housing market. As discussed in section 2.2, this view is supported by the empirical findings of Glaeser and Gyourko (2007) and Bachmann and Cooper (2014), especially when the focus is on fluctuations of housing turnover within the owner-occupied housing segment of the market.

[16] The model abstracts from the possibility that those trying to sell will withdraw from the market without completing a sale.

[17] Later, the model will be calibrated so that a discrete time period $[t, t+\tau)$ is one week ($\tau = 1/52$).

standard constant-returns meeting function (noting that not all viewings will lead to matches). For buyers, the corresponding arrival rate is $\mathcal{V}(u_t, b_t)/b_t$. During this process of search, buyers incur flow search costs τF per interval of time τ.

Given the unit measure of houses, there are $1 - u_t$ houses that are matched in the sense of being occupied by a family. As there is also a unit measure of families, there must be u_t families not matched with a house, and thus in the market to buy. This means the measures of buyers and sellers are the same ($b_t = u_t$). Given that the function $\mathcal{V}(u, b)$ features constant returns to scale, the arrival rates of viewings for buyers and sellers are then both equal to $v = \mathcal{V}(1, 1)$. This constant arrival rate summarizes all that needs to be known about the frictions in locating houses to view.

Note that there are no fluctuations in 'market tightness' (b_t/u_t), which differentiates the analysis from papers that have made market tightness central to studying housing-market fluctuations (Novy-Marx, 2009, Díaz and Jerez, 2013). Unlike the labour market, it is natural to suppose that most individuals who sell in the housing market would also want to buy, suggesting that variation in market tightness would not be the dominant factor in understanding the housing market. Furthermore, if market tightness were to change over time, with a constant-returns meeting function this would imply a negative comovement between the average time taken to buy and the average time taken to sell. There are limited data on time-to-buy, but Genesove and Han (2012) have compiled some observations using the 'Profile of Buyers and Sellers' surveys from NAR for various years from 1987–2007. Their data reveal a positive relationship between time-to-buy and time-to-sell suggesting that fluctuations in 'market tightness' alone do not provide a complete picture of the housing market.

The second aspect of the search frictions in the housing market is the heterogeneity in buyer tastes and the extent to which any given house will conform to these. The idiosyncratic utility flow value ϵ of an occupied house is match specific, that is, particular to both the house and the family occupying it.[18] When a viewing takes place, match quality ϵ is drawn from the probability distribution

$$\epsilon \sim \text{Pareto}(1; \lambda), \quad \text{where } \mathbb{P}[\epsilon \leq z] = 1 - z^{-\lambda}. \tag{3.1}$$

The Pareto distribution is chosen for analytical tractability. The minimum value of ϵ is normalized to one and the parameter λ determines the shape of the distribution. The variance of match quality is inversely related to the shape parameter λ.

The first friction can be seen as the initial step in locating houses for sale that meet a given set of objective criteria such as size, and the second friction can be seen as the time needed to view the houses to judge the match quality between the buyer and the house. The housing market features a range of intermediaries such as realtors and online advertisements that would be expected to reduce significantly the importance of the first friction by distributing information about available houses and their objective characterizes widely among potential buyers. A measure of the importance of the

[18] The model abstracts from characteristics of houses that would be equally valued by all potential buyers. This is consistent with the use of a 'repeat sales' index of house prices that controls for the characteristics of the housing stock.

second friction is the average number of viewings needed before a house can be sold (or equivalently, before a buyer can make a purchase). Ngai and Sheedy (2019) report that viewings per transaction ranges from 9 to 15 using U.S. data from Genesove and Han (2012) and UK data from the Hometrack 'National Housing Survey'.[19] The data reveal that the number of viewings per transaction is far greater than one, indicating there is substantial uncertainty about match quality prior to a viewing. Moreover, the data show that variation in time-to-sell is associated with movements in viewings-per-transaction in the same direction, and not simply due to variation in the time taken to meet buyers.

3.3 Behaviour of buyers and sellers

When a viewing occurs, ϵ is drawn and becomes common knowledge among the buyer and the seller. The value to a family of occupying a house with match quality ϵ is denoted $H_t(\epsilon)$. By purchasing and occupying this house, the buyer loses the option of continuing to search, which has present value $\beta\mathbb{E}_t B_{t+\tau}$, where B_t is the value of being a buyer at time t and $\beta = e^{-r\tau}$ is the discount factor, with $\mathbb{E}_t[\cdot]$ denoting expectations conditional on information available at time t. If the seller agrees to an offer to buy, the gain is the transaction price, and the loss is the option value of continuing to search, namely $\beta\mathbb{E}_t U_{t+\tau}$, where U_t ('unsatisfied owner') is the value of owning a house for sale. Finally, the buyer and seller face a combined transaction cost C. The total surplus $\varSigma_t(\epsilon)$ resulting from a transaction with match quality ϵ is given by

$$\varSigma_t(\epsilon) = H_t(\epsilon) - \beta\mathbb{E}_t W_{t+\tau} - C, \qquad [3.2]$$

where $W_t = B_t + U_t$ denotes the combined value of being a buyer and having a house for sale. Since $H_t(\epsilon)$ is increasing in ϵ, it is easy to see that purchases will occur if match quality ϵ is no lower than a threshold y_t, defined by $\varSigma_t(y_t) = 0$. Intuitively, given that ϵ is observable to both buyer and seller and the surplus is transferable between the two, the transactions that occur are those with positive surplus.[20] The transaction threshold y_t satisfies the following equation:

$$H_t(y_t) = \beta\mathbb{E}_t W_{t+\tau} + C. \qquad [3.3]$$

Using the distribution of ϵ in (3.1), the proportion of viewings π_t that lead to transactions is

$$\pi_t = y_t^{-\lambda}. \qquad [3.4]$$

Given the viewing rate v for both buyers and sellers, there is a probability $\nu = 1 - e^{-v\tau}$ that a buyer or a seller will make or receive a viewing in one discrete time period of length τ. Let $\varSigma_t$ denote the

[19] Hometrack data are based on a monthly survey starting in 2000. The survey is sent to estate agents and surveyors every month. It covers all postcodes of England and Wales, with a minimum of two returns per postcode. The results are aggregated over postcodes weighted by the housing stock.

[20] Some extra assumptions are implicit in this claim, namely that there is no memory of past actions, so refusing an offer yields no benefit in terms of future reputation.

average surplus from viewings that lead to sales:

$$\varSigma_t = \int_{\epsilon=y_t}^{\infty} \frac{\lambda}{y_t} \left(\frac{\epsilon}{y_t}\right)^{-(\lambda+1)} \varSigma_t(\epsilon)\mathrm{d}\epsilon. \quad [3.5]$$

The Bellman equation for the combined buyer and seller value W_t is:

$$W_t = -\tau(F + M) + \nu\pi_t\varSigma_t + \beta\mathbb{E}_t W_{t+\tau}, \quad [3.6]$$

where M is the flow cost of owning a home (incurred whether or not the owner is trying to sell). Intuitively, the first two terms capture the flow costs and benefits of being a buyer and a seller, while the final term is the continuation value.[21]

If a transaction occurs, the price $p_t(\epsilon)$ is agreed according to Nash bargaining. The surpluses of the buyer and the seller ('unsatisfied owner') are as follows, conditional on the match quality between buyer and house being ϵ:

$$\varSigma_{\mathrm{b},t}(\epsilon) = H_t(\epsilon) - \beta\mathbb{E}_t B_{t+\tau} - p_t(\epsilon) - (1-\kappa)C, \quad \text{and} \quad \varSigma_{\mathrm{u},t}(\epsilon) = p_t(\epsilon) - \beta\mathbb{E}_t U_{t+\tau} - \kappa C, \quad [3.7]$$

where κ is the fraction of the total transaction cost C borne directly by the seller. The value functions B_t of the buyer and U_t of the seller are determined by the Bellman equations:

$$B_t = -\tau F + \beta\mathbb{E}_t B_{t+\tau} + \nu \int_{\epsilon=y_t}^{\infty} \lambda\epsilon^{-(\lambda+1)}\varSigma_{\mathrm{b},t}(\epsilon)\mathrm{d}\epsilon, \quad \text{and} \quad [3.8a]$$

$$U_t = -\tau M + \beta\mathbb{E}_t U_{t+\tau} + \nu \int_{\epsilon=y_t}^{\infty} \lambda\epsilon^{-(\lambda+1)}\varSigma_{\mathrm{u},t}(\epsilon)\mathrm{d}\epsilon. \quad [3.8b]$$

The Nash bargaining solution (with bargaining power ω of the seller) implies the surplus-splitting equation $(1-\omega)\varSigma_{\mathrm{u},t}(\epsilon) = \omega\varSigma_{\mathrm{b},t}(\epsilon)$, which determines the transaction price. Hence, using the surplus equations (3.7) and the value functions (3.8a) and (3.8b), appendix A.2 derives the following formula for the transaction price for a house with match quaity ϵ to its buyer:

$$p_t(\epsilon) = \omega H_t(\epsilon) + (\kappa - \omega)C + \beta\left(\frac{\tau}{1-\beta}\right)(\omega F - (1-\omega)M). \quad [3.9]$$

The average transaction price P_t for all houses sold at time t is:

$$P_t = \omega \int_{\epsilon=y_t}^{\infty} \frac{\lambda}{y_t}\left(\frac{\epsilon}{y_t}\right)^{-(\lambda+1)} H_t(\epsilon)\mathrm{d}\epsilon + (\kappa-\omega)C + \beta\left(\frac{\tau}{1-\beta}\right)(\omega F - (1-\omega)M). \quad [3.10]$$

3.4 Behaviour of owner-occupiers

Consider a homeowner with match quality ϵ at time t. This family receives a utility flow value of $\tau\epsilon\xi_t$ during the time period $[t, t+\tau)$, where ξ_t is the exogenous economy-wide housing demand level, modelled as a change in the utility value of housing to all owner-occupiers. All homeowners

[21] The flow cost also enters the value of being a homeowner, which appears in the surplus $\varSigma_t$.

also incur a flow cost of τM during the time interval $[t, t+\tau)$ (which is also incurred by unsatisfied owners who are trying to sell). The aggregate housing demand variable ξ_t is modelled as a exogenous AR(1) process

$$\log \xi_t = (1-\rho) + \rho \log \xi_{t-\tau} + \eta_t, \quad \text{where } \ \eta_t \sim \text{i.i.d.}(0, \varsigma^2), \tag{3.11}$$

with ρ denoting serial correlation in economy-wide housing demand.

Individual match quality ϵ is a persistent variable. However, families are sometimes subject to idiosyncratic shocks that degrade match quality. These shocks can be thought of as life events that make a house less well suited to the family's current circumstances. At most one such shock occurs in the time interval $[t, t+\tau)$. The arrival of these shocks follows a Poisson process with arrival rate a. If a shock arrives, match quality ϵ is scaled down by δ ($\delta < 1$). If no shock occurs, match quality remains unchanged. Given match quality ϵ at time t, the stochastic process for match quality ϵ' at time $t+\tau$ is

$$\epsilon' = \begin{cases} \epsilon & \text{w.p. } \alpha \\ \delta\epsilon & \text{w.p. } 1-\alpha \end{cases}, \tag{3.12}$$

where $\alpha = e^{-a\tau}$ is the probability that no idiosyncratic shock is received. Following the arrival of idiosyncratic shocks, homeowners can decide whether to put their homes on the market or not. Those who do not experience a shock face a cost D if they decide to move. This cost represents the 'inertia' of families to remain in the same house.

The value function $H_t(\epsilon)$ for an owner-occupier is determined by the Bellman equation:

$$H_t(\epsilon) = \tau\epsilon\xi_t + \beta\alpha\mathbb{E}_t \max\{H_{t+\tau}(\epsilon) - \tau M, W_{t+\tau} - D\} + \beta(1-\alpha)\mathbb{E}_t \max\{H_{t+\tau}(\delta\epsilon) - \tau M, W_{t+\tau}\}.$$

Assuming a large moving cost in the absence of an idiosyncratic shock (the limiting case of $D \to \infty$) implies that this Bellman equation reduces to:

$$H_t(\epsilon) = \tau\epsilon\xi_t + \beta\alpha\mathbb{E}_t[H_{t+\tau}(\epsilon) - \tau M] + \beta(1-\alpha)\mathbb{E}_t \max\{H_{t+\tau}(\delta\epsilon) - \tau M, W_{t+\tau}\}. \tag{3.13}$$

When a shock to match quality is received, the owner occupier decides to move if the match quality ϵ is now below a threshold x_t defined by:

$$H_t(x_t) = W_t + \tau M. \tag{3.14}$$

If no idiosyncratic shock is received, a homeowner will always choose not to move given the simplifying assumption $D \to \infty$, so the moving decision is not fully endogenous. Only those families who receive an idiosyncratic shock will make a decision whether to move or not. However, whether those receiving idiosyncratic shocks will move depends on all relevant variables including their own idiosyncratic match quality, and current and expected future conditions in the housing market.

The model thus allows for a endogenous moving decision for those families most likely to consider moving, with a considerable gain in computational tractability by not allowing for an endogenous moving decision for those not hit by an idiosyncratic shock.

3.5 Laws of motion

In what follows, it will be assumed that the idiosyncratic shocks are sufficiently large (δ in (3.12) is low) relative to the aggregate shocks (ς in (3.11) is sufficiently small) so that the magnitude of the fluctuations in the transaction and moving thresholds y_t and x_t is small in relation to the idiosyncratic shocks. The parameters are also restricted so that an owner with match quality close to the transaction threshold y_t will always choose to move if an idiosyncratic shock is subsequently received (but not necessarily owners with higher match qualities).

The measure of properties available for sale is u_t, and this is found by determining the measure of surviving matches following the arrival of idiosyncratic shocks and after moving decisions, which is $1 - u_t$. All matches were originally drawn from the match quality distribution $\epsilon \sim \text{Pareto}(1, \lambda)$. The density function is $\lambda\epsilon^{-(\lambda+1)}$ and the probability that a draw is greater than ϵ is $\epsilon^{-\lambda}$. These matches can be characterized by a vintage $i = 1, 2, \ldots$, where i denotes the number of discrete time periods since a match formed. For vintage i matches back at time $t - \tau i$, a fraction ν of the measure of buyers met a seller and obtained an i.i.d. draw of ϵ from this distribution. This means that there were a total of $\nu u_{t-\tau i}$ of such draws.

Now consider the cohort of potential matches from draws of ϵ at time $t - \tau i$. What must be determined is how many of these matches formed at time $t - \tau i$ and survived through to time t. For these matches, let j denote the number of discrete time periods since the last idiosyncratic shock was received, where $j = 0, 1, \ldots, i - 1$. Shocks arrive with independent probability $1 - \alpha$ each time period. The distribution of the times since the last shock is geometric. The probability of each value of j is $(1 - \alpha)\alpha^j$. A further possibility is that no shocks have been received at all: this event has probability α^i.

Given that some idiosyncratic shock has been received since $t - \tau i$, it is also necessary to know the number k of other idiosyncratic shocks conditional on the most recent one arriving j periods ago, where $k = 0, 1, \ldots, i - 1 - j$. Given independent arrival times, k has a binomial distribution with parameters $i - 1 - j$ and $1 - \alpha$. The probability of each value of k is

$$\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}.$$

Let ϵ' denote the current value of match-specific quality after the sequence of idiosyncratic shocks since the first draw of match quality. Let ϵ denote the original value of match quality. If no shocks are received, $\epsilon' = \epsilon$. If at least one shock is received, $\epsilon' = \delta^{k+1}\epsilon$, where δ is the multiplicative factor applied to match quality every time an idiosyncratic shock is received.

If no shocks are received, the fraction of the original draws that form (and necessarily survive) is the probability that $\epsilon \geq y_{t-\tau i}$:

$$y_{t-\tau i}^{-\lambda}.$$

In the case that shocks have been received, a draw from the original distribution is a surviving match at time t if $\epsilon' \geq x_{t-\tau j}$, where j is the number of time periods since the last shock, and $\epsilon \geq y_{t-\tau i}$, as well as the value of match quality being always above the then prevailing destruction thresholds at any other times between $t-i$ and $t-j$ when shocks are received. Given the restriction that the idiosyncratic shocks are large relative to the aggregate shocks, the probability that all these conditions are satisfied is then simply the probability that $\epsilon' \geq x_{t-\tau j}$. Given the link between ϵ' and ϵ, this is equivalent to $\epsilon \geq \delta^{-(k+1)} x_{t-\tau j}$, where ϵ is the original draw of match quality. The Pareto distribution implies this probability is

$$(\delta^{-(k+1)} x_{t-\tau j})^{-\lambda}.$$

Putting together these observations leads to the following law of motion for the measure u_t of houses for sale:

$$u_t = 1 - \sum_{i=1}^{\infty} \Bigg\{ \alpha^i y_{t-\tau i}^{-\lambda} + \sum_{j=0}^{i-1} (1-\alpha)\alpha^j \sum_{k=0}^{i-1-j} \frac{(i-1-j)!}{k!(i-1-j-k)!} (1-\alpha)^k \alpha^{i-1-j-k} (\delta^{-(k+1)} x_{t-\tau j})^{-\lambda} \Bigg\} \nu u_{t-\tau i},$$

and using the binomial theorem, this expression simplifies to:

$$u_t = 1 - \sum_{i=1}^{\infty} \Bigg\{ \alpha^i y_{t-\tau i}^{-\lambda} + (1-\alpha)\delta^{\lambda} \sum_{j=0}^{i-1} \alpha^j (\alpha + (1-\alpha)\delta^{\lambda})^{i-1-j} x_{t-\tau j}^{-\lambda} \Bigg\} \nu u_{t-\tau i}. \qquad [3.15]$$

The sales rate s_t and sales volume S_t are:

$$s_t = \nu \pi_t, \quad \text{and} \quad S_t = s_t u_t.$$

New listings can be calculated from the stock-flow accounting identity:

$$N_t = u_t - u_{t-\tau} + S_{t-\tau}.$$

3.6 Solution in the case of no aggregate shocks

First consider the solution of the model in the case of no aggregate shocks (where $\xi_t = 1$ in (3.11) in the limiting case of $\varsigma \to 0$).[22] There is a stationary distribution of match quality ϵ over families and a steady state for the value functions, decision thresholds, and stocks and flows. As described in

[22] The model then becomes a discrete-time version of Ngai and Sheedy (2019).

detail in appendix A.3, after solving one non-linear equation for the steady-state moving threshold x, the values of all other variables in the steady state can be found analytically. The key equations needed to explain the properties of the model are given below.

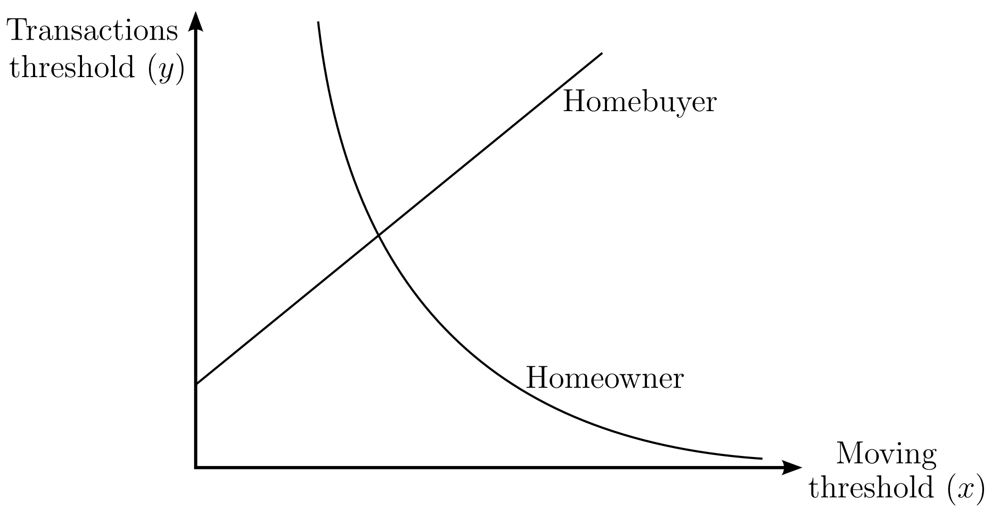
The moving and transaction thresholds x and y are determined by the pair of equations:

$$x + F = \left(\frac{\nu}{\tau}\right)\left(\frac{1}{\lambda - 1}\right)\left(\frac{\tau}{1 - \beta\alpha}\right)\left(y^{1-\lambda} + \beta\delta^{\lambda}\left(\frac{1-\alpha}{1 - \beta(\alpha + (1-\alpha)\delta^{\lambda})}\right)x^{1-\lambda}\right), \tag{3.16a}$$

$$y = \beta\alpha x + \left(\frac{1 - \beta\alpha}{\tau}\right)C. \tag{3.16b}$$

Together, equations (3.16a) and (3.16b) can be jointly solved for the thresholds x and y without reference to state variables such as the number of houses for sale or the distribution of existing match quality. Figure 3 depicts the determination of the moving and transaction thresholds as the intersection between an upward-sloping equation (3.16a) and a downward-sloping equation (3.16a). Intuitively, the upward-sloping line ties the value of a marginal homebuyer to that of a marginal homeowner together with the transaction cost (which is sunk for someone who has decided to become a buyer, but not for an existing homeowner who can choose to stay). This line is referred to as the 'homebuyer' curve. The downward-sloping curve ties the value of the marginal homeowner to the expected value of becoming a buyer. This line is referred to as the 'homeowner' curve. In (x, y) space, these two curves pin down the equilibrium values of x and y. If an equilibrium exists, it must be unique (existence can be established under weak conditions described in appendix A.3).

Figure 3: *Determination of the moving (x) and transactions (y) thresholds*



Notes: The homebuyer and homeowner curves represent equations (3.16a) and (3.16b) respectively.

Given the solution for x and y, the inflow and outflow rates and the stock of houses for sale can be determined. These flows are summarized by the following statistics that will be used to calibrate the model's parameters. Derivations of the formulas can be found in appendix A.3.

Outflows from the housing market are summarized by two statistics: viewings per sale and time to sell. Viewings per sale (denoted by V_{s}) is defined as the expected number of viewings required for a house to be sold (or equivalently in the model, the number of viewings required for a buyer to find an acceptable house):

$$V_{\mathrm{s}} = y^{\lambda}. \tag{3.17}$$

More viewings occur if the transaction threshold y is higher, or the distribution of match quality for new buyers is worse (higher λ). Time to sell (denoted by T_{s}) is defined as the expected duration (in years) for which a house will remain unsold:

$$T_{\mathrm{s}} = y^{\lambda}\left(\frac{\tau}{\nu}\right). \tag{3.18}$$

Time to sell depends on the factors that determine viewings per sale and the expected time between viewings (τ/ν).

Inflows into the housing market are summarized by two statistics: the expected duration of a new match and the average time existing homeowners have owned their homes (for reasons that will become clear, these two numbers are not the same). The expected duration of a match (denoted by T_{d}) is defined as the average time (in years) that buyers expect to remain in a newly purchased home (note that the expected duration for a particular buyer would depend on the match quality ϵ; the statistic is calculated by averaging over the distribution of match quality above the transaction threshold y):

$$T_{\mathrm{d}} = \left(\frac{\tau}{1-\alpha}\right)\left(1 + \left(\frac{\delta y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-(\alpha+(1-\alpha)\delta^{\lambda})}\right)\right). \tag{3.19}$$

This statistic depends on the probability α of not receiving an idiosyncratic shock in a given time period, the size of the shocks as controlled by δ, the distribution of match quality for new buyers as affected by λ, and the transaction and moving thresholds y and x. The first bracket in the formula is the expected time between the arrival of idiosyncratic shocks, which is the same by assumption for all homeowners. The second term reflects the proportion of homeowners who decide to remain following a shock (on average over all cohorts of homeowners). This is a more complicated formula because with an endogenous moving decision, the distribution of match quality over homeowners is correlated with the time already spent in the house. In other words, the hazard rate for moving among the population of owners is not independent of the duration of time since a house was purchased. In particular, the hazard rate will be increasing in duration because the distribution of match quality is better for new buyers than for existing owners (some of whom will have received one or more idiosyncratic shocks since first buying a house).

Given that the hazard rate is not independent of duration, the second statistic, the average time homeowners have owned their home (denoted by T_{a}) is not the same the expected duration T_{d}. The formula is:

$$T_{\text{a}} = \frac{\left(\frac{\tau}{1-\alpha}\right) + \left(\frac{\delta y}{x}\right)^{\lambda}\left(\frac{\tau}{1-(\alpha+(1-\alpha)\delta^{\lambda})}\right)\left(\frac{(2-\delta^{\lambda})(1-\alpha)}{1-(\alpha+(1-\alpha)\delta^{\lambda})}\right)}{1+\left(\frac{\delta y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-(\alpha+(1-\alpha)\delta^{\lambda})}\right)}. \qquad [3.20]$$

This formula depends on the same parameters and variables as for T_{d}, but since the hazard rate is increasing in duration, it can be shown that $T_{\text{a}} < T_{\text{d}}$.

The average transaction price for houses sold can be shown to be

$$P = \kappa C - \beta\left(\frac{\tau}{1-\beta}\right)M + \omega\left(\beta\left(\frac{\tau}{1-\beta}\right) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)\right)(x+F), \qquad [3.21]$$

which can be used to calculate the ratio of the transaction, maintenance, and search costs C, M, and F relative to the average price P, denoted by $c \equiv C/P$, $m \equiv M/P$, and $f \equiv F/P$.

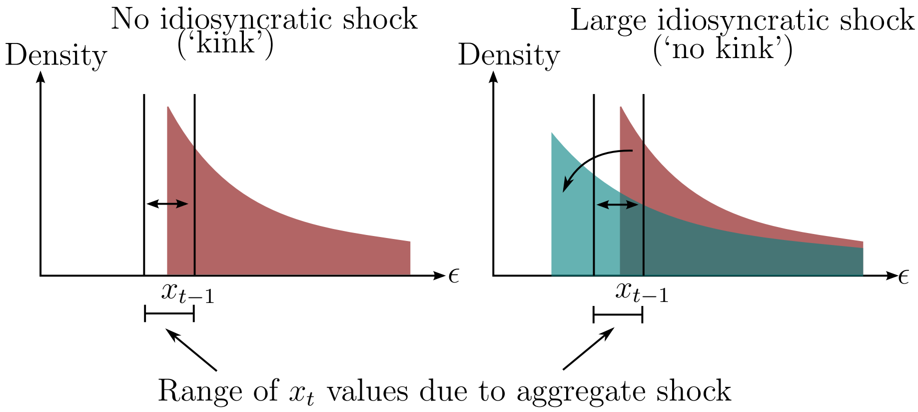
3.7 Solution with aggregate shocks

In the case of aggregate shocks, the solution of the model for aggregate variables is obtained approximately using a first-order perturbation (log linearization) around the solution with no aggregate shocks ($\varsigma = 0$) as described in section 3.6. The well-known problem of non-differentiability in models of endogenous 'lumpy' adjustments is overcome under two assumptions. First, the idiosyncratic shock is large relative to aggregate shocks, and large relative to the difference between the acceptance and moving thresholds (in steady state). Second, there is a sufficiently large moving cost faced by those who do not receive an idiosyncratic shock, so such individuals would always choose not to move.

Under these assumptions, the equations describing the equilibrium values of the aggregate variables are differentiable, and thus a perturbation method is admissible. Intuitively, this issue is illustrated in Figure 4. The left panel displays the case when there is no idiosyncratic shock. Without a large moving cost, an endogenous moving decision will imply a 'kinked' response of the overall number of homeowners who move. The idea is that if the moving threshold falls (due to aggregate shocks) then there is no change in the number of homeowners who move, unlike the case where the moving threshold rises. The right panel shows the case where the idiosyncratic shocks are large relative to changes in the moving thresholds (due to aggregate shocks). In that case there is no problem of non-differentiability. Details of the log linearizations are provided in section A.5.

In principle, solving the model requires finding the value functions for all values of match quality ϵ and the whole distribution of surviving match quality. This means the model has an infinite-dimensional state space. However, with the assumption of a Pareto distribution for new draws of ϵ, the log-linearized equations involve only the average of the value functions calculated with the steady-state distribution of match quality. Furthermore, the law of motion for the stock of houses for sale can be written in terms of a finite (and low) number of state variables. Details are provided in appendix A.5.

Figure 4: *Differentiability and idiosyncratic shocks*



3.8 The role of endogenous moving

There is special case of the parameters of the model in which the moving decision effectively becomes exogenous. If the size of the idiosyncratic shock to match quality becomes very large (the limiting case of $\delta \to 0$) then moving occurs if and only if an exogenous idiosyncratic shock is received. This provides an otherwise identical model with the exogeneity of the moving decision as the only difference.

Taking away the choice of endogenous moving implies that homeowners have the same probability of moving irrespective of existing match quality (whereas only those with relatively low match quality will move if the shock is not so large and an endogenous choice is allowed). Moreover under exogenous moving, there are both cases where homeowners that would not have moved are effectively forced to move, and homeowners that would want to move who are effectively forced to stay. These differences are crucial in understanding the role of endogenous moving in the dynamics of the model. As seen from (3.14), the decision to move equates the value of being a homeowner to the combined value of being a buyer and being a seller, where the later in turn is important to the buying decision, as shown in (3.3). Through the combined value W, the endogenous moving decision interacts with the buying decision to generate new dynamics that are absent in an exogenous moving model.

The dynamics of the model depend crucially on the distribution of match quality across all homeowners, which evolves endogenously because the threshold for moving can fluctuate. In other words, moving decisions today have consequences for future distributions of match quality. Fluctuations of the moving threshold generate a new mechanism that works through a 'cleansing' effect. More specifically, a higher current threshold for not moving implies that moving in the future becomes less likely because a greater degree of 'cleansing' has occurred. Similarly, if the current moving threshold is lower, that is, more lower-quality matches remain, there is less cleansing, so moving in the future becomes more likely. This can potentially generate overshooting.

4 Calibration

The model contains a total of 12 parameters $\{a, \delta, \lambda, v, F, M, C, \omega, \kappa, r, \tau, \rho\}$. Some parameters can be directly matched to the data, while others can be determined indirectly by choosing values that make the predictions of the model consistent with some empirical targets. U.S. data for the period 1991–2012 will be used in this calibration. Finally, for some parameters, reasonable values are directly imposed.

In the absence of aggregate shocks, the solution of the model depends only on the first 11 of the parameters, $\{a, \delta, \lambda, v, F, M, C, \omega, \kappa, r, \tau\}$. These 11 parameters will be calibrated without using any information derived from time-series variation in the data. Instead, averages in the data will be matched to the model's steady state. The final parameter is the persistent parameter for the demand shock, ρ, which is calibrated to match the persistence of houses prices. It can be shown that the model predicts that house prices have the same persistence as the housing demand shock. The first-order autocorrelation coefficient of house prices is 0.965 at a quarterly frequency. The parameter ρ is set to generate the same persistence in house prices. Results are also reported for a lower value of $\rho = 0.95$.

The length of a discrete time period τ is set to one week ($\tau = 1/52$). Given the steady state of the model is identical to Ngai and Sheedy (2019), the calibration of the remaining 10 parameters $\{a, \delta, \lambda, v, F, M, C, \omega, \kappa, r\}$ follows the same procedure as that paper, which is summarized in appendix A.4. In brief, the (annual) discount rate r is set to 7% ($r = 0.07$). Buyers and sellers are assumed to have equal bargaining power ($\omega = 0.5$).

The parameters F, M, C, and κ are calibrated to match the costs of owning a house and the costs involved in buying and selling houses, and how those costs are distributed across buyers and sellers. Let $f = F/P$, $m = M/P$, and $c = C/P$ denote the costs F, M, and C relative to the average house price P in the steady state of the model. The data provide information on costs relative to price, so calibration targets for f, m, and c are adopted that will determine F, M, and C indirectly. The determination of the price P depends in general on all the other parameters, the calibration must be done jointly with that for the remaining parameters $\{a, \delta, \lambda, v\}$. These four parameters will be calibrated using four additional empirical targets: the average time to sell a house, the average number of viewings per sale, the expected duration of a new match, and the average time home-owners have owned their homes.

The seven calibration targets used to determine the parameters $\{a, \delta, \lambda, v, C, F, M\}$ are listed in Table 2. Intuitively, the expected duration and average years since homeowners moved in provide information about the arrival rate a of idiosyncratic shocks and the size of those shocks (the parameter δ). Both a lower arrival rate and smaller idiosyncratic shocks would lead to longer expected durations of matches and a longer average time since moving.

The two parameters can be separately identified because having data on both the expected duration and the average number of years since moving provides information not only about the average hazard rate of moving, but also its dependence on duration. Furthermore, the parameters a and δ have very different effects on the hazard function. A decrease in the arrival rate a of shocks

Table 2: *Targets used to calibrate parameters*

Target description	Notation	Value
Time to sell (time to buy)	T_{s}	6.5/12
Viewings per sale (viewings per purchase)	V_{s}	10
Expected duration of ownership of a house	T_{d}	12.2
Average years since home-owners moved in	T_{a}	11
Ratio of transaction cost to average price	c	0.10
Ratio of flow search costs to average price	f	0.025
Ratio of flow maintenance costs to average price	m	0.045

Notes: The data sources for these empirical targets are discussed in section 4.

uniformly decreases the hazard rate for all durations, while a decrease in the size of the shocks also tilts the hazard function so that its slope increases. The reason is that with very large idiosyncratic shocks, the moving decision would essentially depend only on receiving one shock. With smaller idiosyncratic shocks, homeowners who start with a high match quality would require more than one shock to persuade them to move, making moving more likely for longer-duration homeowners who have had time to receive multiple shocks than for those who have moved more recently.

There is also an intuitive connection between the parameter λ and the calibration target time-to-sell. The value of λ determines the amount of dispersion in the distribution of potential match quality, and thus the incentive to continue searching. A low value of λ indicates a high degree of dispersion, in which case families will be willing to spend longer searching for an ideal house. For the final parameter v, the average time between viewings can be found by dividing time-to-sell by viewings-per-sale, which directly provides information about the arrival rate v of viewings.

A simple method for exactly matching the seven parameters $\{a, \delta, \lambda, v, C, F, M\}$ to the seven empirical targets in Table 2 is described in appendix A.4. The parameters matching the targets and those directly calibrated are all shown in Table 3.

5 Quantitative results

This section presents the results of simulating the theoretical model described in section 3 with an endogenous moving decision when the economy is subject to aggregate shocks. Given the variation in the housing-market inflow rate documented in section 2, the aim is to study whether a model with an endogenous moving decision is able to match the size of the fluctuations in listings relative to fluctuations in other measures of market activity such as sales, prices, and the stock of houses for sale. Furthermore, does a model with an endogenous moving decision help in understanding the joint dynamics of the flows, the stock of unsold houses, and prices?

Addressing this question requires a benchmark model without an endogenous moving decision as a point of comparison. There is a special case of the parameters of the model from section 3 in which the moving decision effectively becomes exogenous. If the size of the idiosyncratic shock

Table 3: *Calibrated parameters*

Parameter description	Notation	Value
Parameters matching calibration targets		
Arrival rate of shocks	a	0.131
Size of shocks	δ	0.862
Steady-state distribution of match quality	λ	13.0
Arrival rate of viewings	v	22.8
Total transaction cost	C	0.565
Flow search costs	F	0.141
Flow maintenance costs	M	0.254
Directly chosen parameters		
Share of total transaction cost directly borne by seller	κ	1/3
Bargaining power of seller	ω	1/2
Discount rate	r	0.07
Length of discrete time period	τ	1/52
Persistent of demand shock	ρ	0.965

Notes: The parameters are chosen to match exactly the calibration targets in Table 2.

to match quality becomes very large (the limiting case of $\delta \to 0$) then moving occurs if and only if an exogenous idiosyncratic shock is received. This provides an otherwise identical model with exogeneity of the moving decision as the only difference. The exogenous moving model can be calibrated using the method described in section 4 with one modification. When the idiosyncratic shock becomes large, the expected duration of a match converges to the average number of years since moving implied by the model. Hence it is not possible to match both the expected duration (T_{d}) and the average years since moving (T_{a}) separately. The calibration target for the common value of these numbers is taken to be the average of the calibration targets of 11 and 12.2 respectively for T_{d} and T_{a} (see Table 2), that is, 11.6 years.

The aggregate shock is taken to be the housing demand shock ξ_t specified in equation (3.11) (which scales up the flow value of housing received by all homeowners), with its persistence parameter set to match the persisence of house prices. The remaining parameters of the model are chosen as described in section 4 (with the modification of this procedure for the exogenous-moving variant of the model described above). Apart from the persistence of housing demand, note that the calibration does not use any information from the time-series moments of the data documented in section 2: the only information is drawn from the average values of variables, which are matched to the model's equilibrium in the absence of aggregate shocks.

In what follows, the simulation results are obtained using a first-order perturbation of the model around its equilibrium in the absence of aggregate shocks. All aggregate variables are reported as percentage deviations from their values in the absence of aggregate shocks. The model is simulated with a discrete time period equal to one week ($\tau = 1/52$) and the results are converted to a quarterly frequency. The results are shown as impulse response functions and autocorrelation functions of sales, house prices, new listings, the number of houses for sale, and the time-to-sell following a 1% shock to housing demand. A table of relative standard deviations and correlation coefficients is also reported that can be compared to the equivalent moments in the data. To relate these results to the earlier analysis of inflow and outflow rates, note that time-to-sell is the inverse of the selling rate. The inflow rate could be similarly related to a measure of time-to-move, but since the average number of houses for sale is small relative to the total stock of houses, the inflow rate comoves almost perfectly with the level of inflows (new listings). For that reason, separate results for time-to-move are not calculated.

5.1 The benchmark case of exogenous moving decisions

The first case considered is the benchmark model with exogenous moving ($\delta = 0$). The implied relative standard deviations and correlation coefficients are displayed in Table 4 and the autocorrelation functions and impulse response functions in Figure 5 and Figure 6.

Table 4: *Exogenous moving, shock to housing demand*

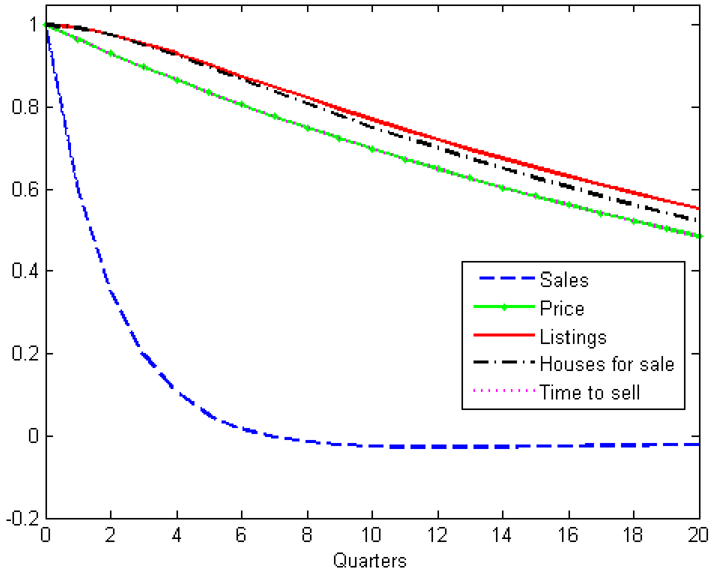
	Sales	Price	New listings	Houses for sale	Time-to-sell
			Relative standard deviations		
Sales	1	4.734	0.176	3.453	3.761
			Correlation coefficients		
Sales	1				
Price	0.429	1			
New listings	0.176	0.964	1		
Homes for sale	−0.178	−0.965	−0.999	1	
Time-to-sell	−0.429	−1.000	−0.964	0.965	1

Notes: Quarterly frequency. Parameters: $\delta = 0$.

This benchmark model faces many problems in matching the stylized facts in Table 1. Unsurprisingly, the model with exogenous moving predicts that the volatility of new listings is tiny relative to sales, while empirically, new listings is more volatile than sales. Listings also have a perfect negative correlation with houses for sale in the model, but are almost uncorrelated in the data. As a result, listings and houses for sale have same correlations (in absolute value) with other variables. The correlation with sales is also much lower than found in the data. The problem is simply that listings are proportional to the previous number of homeowners not trying to sell because a fraction of these homeowners receive an idiosyncratic shock that leads them automatically to try to sell irrespective of market conditions. Thus listings can only vary as a reflection of changes in house for sale, and by a much smaller amount.

The problems faced by the model with exogenous moving extend beyond simply the behaviour of new listings. Relative to the data, the relative volatility of sales volume is far too low and this

Figure 5: *Exogenous moving, autocorrelation functions, shock to housing demand*



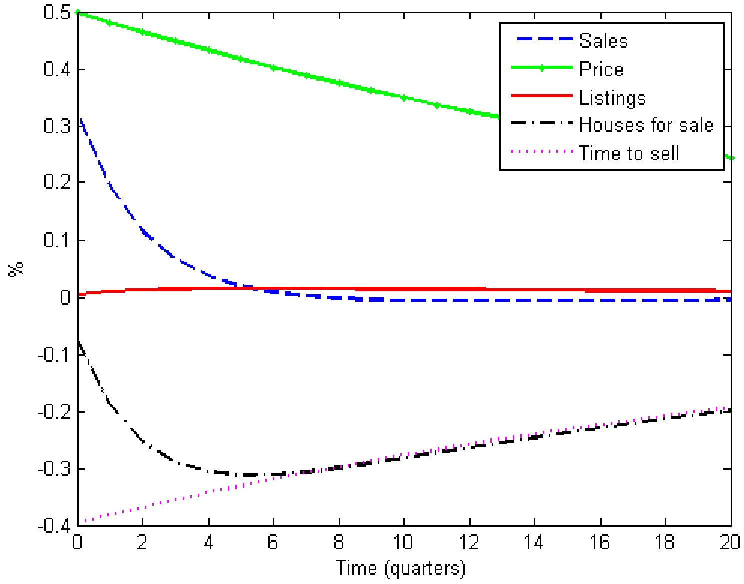
Notes: Quarterly frequency. Parameters: $\delta = 0$.

variable has very little persistence. The reason for this failing can be seen in the impulse response functions in Figure 6. While the shock to the demand for housing pushes up sales, with no possibility of significant inflows, these sales quickly deplete the stock of houses for sale, persistently reducing the stock of properties on the market. This then offsets the effect of the demand shock on sales because fewer sales take place when few properties are available, even if the selling rate remains high (and so time-to-sell remains persistently shorter). Thus the increase in sales volume is short lived, as seen from the autocorrelation function in Figure 5. Because there is no margin for more than the usual number of homeowners to enter the market as sellers, the shift in demand leads to excessive volatility in the number of houses for sale and the time taken for sales to occur.

The predicted correlation of sales and listings, and sales and prices are both too low compared to the data. It is important to note that the presence of a distribution of new match quality plays a crucial role in obtaining a positive correlation between sales and prices with only a housing demand shock. When a house is viewed by a potential buyer, new match quality is drawn from a probability distribution, and there is a transaction threshold at which the buyer is willing to trade. A positive housing demand shock (increasing the flow utility from all houses) raises total surplus from a transaction and thus increases both the willingness to trade and prices, which gives rise to the positive correlation between sales and price. This correlation would be negative in the absence of a distribution of new match quality, as in the model of Díaz and Jerez (2013).

Finally, the benchmark model also performs poorly in terms of its implications for house prices. The model-predicted house price series is too volatile (all fluctuations are due to demand shocks and the exogenous-moving model features no significant supply response).

Figure 6: *Exogenous moving, impluse respopnse function to a housing demand shock*



Notes: Quarterly frequency. Parameters: $\delta = 0$.

One natural question is whether adding an aggregate shock to the moving rate in the exogenous-moving model can resolve the problems discussed above. This modification allows the inflow rate to have exogenous time variation. This would clearly improve the model's performance in matching the volatility of listings. The question is whether this would also help in matching the joint time-series behaviour of the other variables of interest.

To address this point, consider a version of the exogenous-moving model ($\delta = 0$) where instead of shocks to housing demand ($\xi_t = 1$ replaces equation (3.11)), the exogenous aggregate shock is to the *probability* of receiving an idiosyncratic shock. The arrival rate a of these idiosyncratic shocks (which lead automatically to moving in the exogenous-moving variant of the model) is now subject to exogenous time variation. The results are reported in appendix A.6. In brief, this version of the model does indeed feature listings that have a similar volatility to sales, which is not too far from the data. However, the model struggles to match other key relative volatilities and correlations found in the data (the model predicts the wrong sign for the correlation coefficients of prices with all other variables). Intuitively, the problem stems from the fact that this version of the model introduces entry into the housing market orthogonal to the factors that matter for transactions decisions.

The conclusion here is consistent with the findings of Díaz and Jerez (2013) that in a search model with exogenous moving, three correlated shocks (housing demand, housing supply, and the moving rate) are needed to account for the observed cyclical properties of the data. The next section shows that the endogenous moving model can match the data with only one demand shock because endogenous moving means that a housing demand shock induces more moving, acting like a moving shock, as well as increasing the supply of houses on the market, acting like a housing supply shock.

Thus one housing demand shock replicates the three correlated 'reduced form' shocks that would otherwise be needed to match the stylized facts.

5.2 Endogenous moving decisions

Now consider the model with the endogenous moving decision restored. This means that δ is set to a positive number as discussed in section 4. The resulting relative volatilities and correlations are displayed in Table 5, the autocorrelation functions in Figure 7, and the impulse response functions in Figure 8.

Table 5: *Endogenous moving, shock to housing demand*

	Sales	Price	New listings	Houses for sale	Time-to-sell
			Relative standard deviations		
Sales	1	1.822	0.880	1.363	2.223
			Correlation coefficients		
Sales	1				
Price	0.919	1			
New listings	0.877	0.927	1		
Homes for sale	−0.764	−0.957	−0.868	1	
Time-to-sell	−0.919	−1.000	−0.927	0.957	1

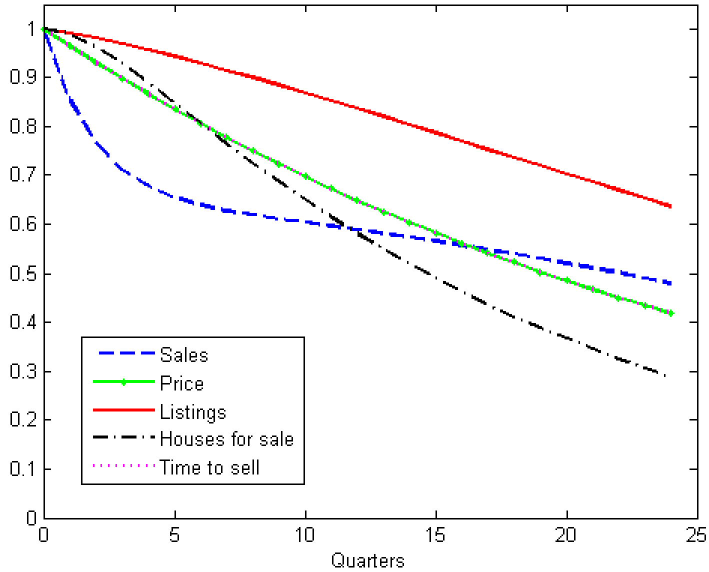
Notes: Quarterly frequency. Parameters: $\delta = 0.862$.

It can be seen that the relative volatilities predicted by the model move closer to what is found in the data: listings become substantially more volatile; houses for sale and time-to-sell are less volatile (relative to sales) than before. Prices are also now less volatile, which goes in the direction of matching the data. The model with endogenous moving also performs much better in terms of the strongly positive correlation between listings and sales and between sales and price, both now close to the data. Finally, fluctuations in sales volume are now much more persistent as in the data.

Match quality plays a crucial role in the workings of the model and its ability to explain the stylized facts with only a housing demand shock. As discussed earlier, the existence of a distribution of new match quality is crucial in obtaining a positive correlation between sales and prices with only a housing demand shock. The equilibrium distribution of match quality among existing homeowners, on the other hand, is important for the positive correlation between sales and listings. Homeowners' match quality is assumed to be a persistent variable subject to occasional idiosyncratic shocks. At any point in time, there is an endogenous distribution of match quality across existing homeowners with a moving threshold below which the owner will choose to move house, which can be seen as an investment in match quality. A persistent housing demand shock increases the incentive to invest, leading to more listings.[23] This explains the positive correlation between listings and sales and

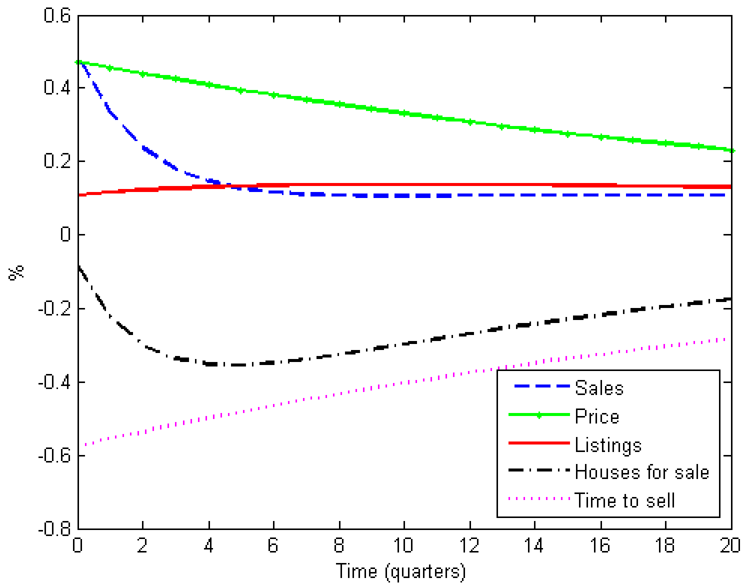
[23]The prediction of an increase in moving following an increase in housing demand is consistent with findings of

Figure 7: *Endogenous moving, autocorrelation graph, shock to housing demand*



Notes: Quarterly frequency. Parameters: $\delta = 0.862$.

Figure 8: *Endogenous moving, impluse response function to a housing demand shock*



Notes: Quarterly frequency. Parameters: $\delta = 0.862$.

prices, and how listings can have a similar volatility to sales. Less obvious is that an increase in

Bachmann and Cooper (2014) that "changing residence appears to be something that happens in times of greater economic activity".

listings is essential to sustain a lasting effect on sales and thus replicate the serial correlation of sales. A rise in the sales rate alone is not sufficient because the stock of houses for sale would be quickly depleted.

The empirical success of the model is because endogenous moving means that a housing demand shock induces more moving, acting like a moving rate shock, and increases the supply of houses on the market, acting like a housing supply shock. Thus one housing demand shock replicates three correlated reduced-form shocks that are needed to match the stylized facts. Introducing independent shocks to moving or housing supply would not match the data because such shocks would generate a negative relationship between prices and sales.

Some problems remain, though. New listings and houses for sale are still negatively correlated (albeit by less), as are house prices and the number of houses for sale. Intuitively, the model seems to suggest that there is not a strong incentive for marginal homeowners to take advantage of the option to enter the housing market. Finally, note that the overshooting discussed earlier is masked in the results because the exogenous shock is highly persistent. Results for a less persistent demand shock with ρ set to 0.95 (matching the persistence of fluctuations in real GDP) are reported in appendix A.7. In this case, there is a small degree of overshooting in the impulse response of sales. The other predictions are similar to Table 5.[24] The effects of re-calibrating the model such that sellers have all the bargaining power (ω is set to 1) are reported in appendix A.8. This case is equivalent to sellers making a take-it-or-leave-it offer. It improves the predictions of the model regarding the relative volatility of prices, listings, and time-to-sell, but worsens the predictions for the stock of houses for sale.

6 Conclusions

This paper has assembled a set of stylized facts about the cyclical properties of house prices, sales, listings, the stock of houses for sale, and the time taken to sell, and the relationships among these variables. It has documented the role of variation in both outflows (sales) and inflows (new listings) in accounting for fluctuations in the volumes of sales. Evidence has been presented to show that the inflow rate contributes around 98% of the variation in sales volume, while the outflow rate only contributes around 2%. This is important because existing search models of the housing market have emphasized only the endogeneity of the outflow decision, treating the inflows as exogenous.

This paper has presented and calibrated a search-and-matching model with endogenous moving. Simulations of the model were presented and compared to the empirical volatilities and correlations among sales, listings, the number of houses for sale, the time taken to sell, and house prices. The model performed well in matching the cyclical fluctuations in the housing market. The importance of incorporating an endogenous moving decision is not confined solely to accounting for the volatility of inflows, but also important for matching the full set of relative volatilities and correlations of

[24]It is interesting to note that in this less persistent case listings do not rise on impact. In fact, listings can fall for demand shocks with even lower persistence. The intuition is that given frictions in the housing market, shocks must have sufficient persistence for marginal homeowners to exercise their option of moving.

housing-market variables.

References

Albrecht, J., Anderson, A., Smith, E. and Vroman, S. (2007), "Opportunistic matching in the housing market", *International Economic Review*, **48(2)**:641–664 (May). 3

Anenberg, E. and Bayer, P. (2013), "Endogenous sources of volatility in housing markets: The joint buyer-seller problem", working paper 18980, NBER. 3

Aruoba, S. B., Davis, M. A. and Wright, R. (2016), "Homework in monetary economics: Inflation, home production, and the production of homes", *Review of Economic Dynamics*, **21**:105–124 (July). 3

Bachmann, R. and Cooper, D. (2014), "The ins and arounds in the U.S. housing market", working paper 14–3, Federal Reserve Bank of Boston. 5, 7, 11, 29

Caplin, A. and Leahy, J. (2011), "Trading frictions and house price dynamics", *Journal of Money, Credit and Banking*, **43(s2)**:283–303 (October). 3

Case, K. E. and Shiller, R. J. (2003), "Is there a bubble in the housing market?", *Brookings Papers on Economic Activity*, **2003(2)**:299–342. 42

Coles, M. G. and Smith, E. (1998), "Marketplaces and matching", *International Economic Review*, **39(1)**:239–254 (February). 3

Davis, M. A. and Heathcote, J. (2005), "Housing and the business cycle", *International Economic Review*, **46(3)**:751–784 (August). 3

Davis, M. A. and Van Nieuwerburgh, S. (2015), "Housing, finance, and the macroeconomy", in G. Duranton, J. V. Henderson and W. C. Strange (eds.), *Handbook of Regional and Urban Economics*, vol. 5, North Holland, chap. 12, pp. 753–811. 3

Díaz, A. and Jerez, B. (2013), "House prices, sales, and time on the market: A search-theoretic framework", *International Economic Review*, **54(3)**:837–872 (August). 3, 5, 12, 26, 27, 33, 42

Fisher, J. D. M. and Gervais, M. (2011), "Why has home ownership fallen among the young?", *International Economic Review*, **52(3)**:883–912 (August). 3

Fujita, S. and Ramey, G. (2009), "The cyclicality of separation and job finding rates", *International Economic Review*, **50(2)**:415–430 (May). 1, 6

Garriga, C. and Hedlund, A. (2019), "Mortgage debt, consumption, and illiquid housing markets in the Great Recession", *American Economic Review*, forthcoming. 3

Genesove, D. and Han, L. (2012), "Search and matching in the housing market", *Journal of Urban Economics*, **72(1)**:31–45 (July). 12, 13, 42

Genesove, D. and Mayer, C. J. (1997), "Equity and time to sell in the real estate market", *American Economic Review*, **87(3)**:255–269 (June). 42

Ghent, A. (2012), "Infrequent housing adjustment, limited participation, and monetary policy", *Journal of Money, Credit and Banking*, **44(5)**:931–955 (August). 41

Glaeser, E. L. and Gyourko, J. (2007), "Arbitrage in housing markets", working paper 13704, NBER. 7, 11

Guren, A. M. (2014), "The causes and consequences of house price momentum", manuscript, Boston University. 3

Han, L. and Strange, W. C. (2015), "The microstructure of housing markets: Search, bargaining, and brokerage", in G. Duranton, J. V. Henderson and W. C. Strange (eds.), *Handbook of Regional And Urban Economics*, vol. 5, North Holland, chap. 13, pp. 813–886. 3

Head, A., Lloyd-Ellis, H. and Sun, H. (2014), "Search, liquidity, and the dynamics of house prices and construction", *American Economic Review*, **104(4)**:1172–1210 (April). 3

Hedlund, A. (2016), "Illiquidity and its discontents: Trading delays and foreclosures in the housing market", *Journal of Monetary Economics*, **83**:1–13 (October). 3

Hendel, I., Nevo, A. and Ortalo-Magné, F. (2009), "The relative performance of real estate marketing platforms: MLS versus FSBOMadison.com", *American Economic Review*, **99(5)**:1878–1898 (December). 42

Iacoviello, M. (2005), "House prices, borrowing constraints, and monetary policy in the business cycle", *American Economic Review*, **95(3)**:739–764 (June). 3

Ioannides, Y. M. and Zabel, J. E. (2017), "Vacancies in housing and labor markets", working paper, Tufts University. 3

Jovanovic, B. (1979), "Job matching and the theory of turnover", *Journal of Political Economy*, **87(5)**:972–990 (October). 2

Krainer, J. (2001), "A theory of liquidity in residential real estate markets", *Journal of Urban Economics*, **49(1)**:32–53 (January). 2, 3

Levitt, S. D. and Syverson, C. (2008), "Market distortions when agents are better informed: The value of information in real estate transactions", *Review of Economics and Statistics*, **90(4)**:599–611 (November). 42

Merlo, A. and Ortalo-Magné, F. (2004), "Bargaining over residential real estate: evidence from England", *Journal of Urban Economics*, **56(2)**:192–216 (September). 42

Moen, E., Nenov, P. and Sniekers, F. (2019), "Buying first or selling first? Buyer-seller decisions and housing market volatility", *Journal of the European Economic Association*, forthcoming. 3

Ngai, L. R. and Sheedy, K. D. (2019), "The decision to move house and aggregate housing-market dynamics", *Journal of the European Economic Association*, forthcoming. 1, 11, 13, 17, 22, 41

Ngai, L. R. and Tenreyro, S. (2014), "Hot and cold seasons in the housing market", *American Economic Review*, **104(12)**:3991–4026 (December). 3

Novy-Marx, R. (2009), "Hot and cold markets", *Real Estate Economics*, **37(1)**:1–22 (Spring). 3, 12

Ortalo-Magné, F. and Rady, S. (2005), "Housing market dynamics: On the contribution of income shocks and credit constraints", *Review of Economic Studies*, **73(2)**:459–485 (April). 3

Petrongolo, B. and Pissarides, C. A. (2008), "The ins and outs of European unemployment", *American Economic Review*, **98(2)**:256–262 (May). 1, 6, 10

Piazzesi, M. and Schneider, M. (2009), "Momentum traders in the housing market: Survey evidence and a search model", *American Economic Review*, **99(2)**:406–411 (May). 3

Piazzesi, M., Schneider, M. and Stroebel, J. (2019), "Segmented housing search", *American Economic Review*, forthcoming. 3

Pissarides, C. A. (2000), *Equilibrium Unemployment Theory*, MIT Press, 2nd edn. 3

Poterba, J. M. (1991), "House price dynamics: The role of tax policy and demography", *Brookings Papers on Economic Activity*, **1991(2)**:143–203. 41

Quigley, J. M. (2002), "Transaction costs and housing markets", in T. O'Sullivan and K. Gibb (eds.), *Housing Economics and Public Policy*, Blackwell, chap. 4, pp. 56–64. 41

Shimer, R. (2012), "Reassessing the ins and outs of unemployment", *Review of Economic Dynamics*, **15(2)**:127–148 (April). 1

Stein, J. C. (1995), "Prices and trading volume in the housing market: A model with down-payment effects", *Quarterly Journal of Economics*, **110(2)**:379–406 (May). 3

Ungerer, C. (2015), "Monetary policy, hot housing markets and leverage", working paper, Federal Reserve Board. 3

Wheaton, W. C. (1990), "Vacancy, search, and prices in a housing market matching model", *Journal of Political Economy*, **98(6)**:1270–1292 (December). 3

A Appendices

A.1 Stylized facts using detrended data

To compare the cyclical properties of the data with those found by Díaz and Jerez (2013), the monthly data are detrended using the Hodrick-Prescott filter (with parameter 14400) and are then deseasonalized by removing the average of each month. To smooth out excess volatility due to measurement error in the data, quarterly time series are constructed from the monthly series. The summary statistics are presented in Table 6.

Standard deviations and correlations of the quarterly series (as percentage deviations from their trends and seasonal averages) on sales, prices, new listings, houses for sale, and time-to-sell are shown in Table 6. Standard deviations are reported relative to the volatility of sales. The statistics related to sales, prices, houses for sale, and time-to-sell are similar to those reported in Díaz and Jerez (2013) and broadly consistent with those presented in Table 1. To highlight a few differences compared to Table 1, the positive correlations between house prices and sales, new listings and sales, and new listings and prices are all weaker with correlation coefficients of 0.24, 0.57, and 0.26 respectively. The negative correlations between sales and time-to-sell, and new listings and time-to-sell, are weaker with a correlation coefficients of −0.54 and −0.33 respectively.

Table 6: *Summary statistics (with detrending)*

	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	0.180	2.071	0.858	1.503
			Correlation coefficients		
Sales	1				
Price	0.243	1			
New listings	0.566	0.256	1		
Homes for sale	−0.238	0.089	−0.005	1	
Time-to-sell	−0.543	−0.230	−0.334	0.560	1

Notes: Monthly data (January 1991–March 2012), HP filtered and seasonally adjusted, converted to a quarterly series.

A.2 Transaction prices

Using equation (3.7), a buyer matched at time t with a house of match quality ϵ buys at the following price $p_t(\epsilon)$:

$$p_t(\epsilon) = \omega H_t(\epsilon) + \beta \mathbb{E}_t[(1-\omega)U_{t+\tau} - \omega B_{t+\tau}] + (\kappa - \omega)C. \qquad \text{[A.2.1]}$$

Using the Bellman equations (3.8a) and (3.8b) for B_t and U_t:

$$(1-\omega)U_t - \omega B_t = \tau(\omega F - (1-\omega)M) + \beta \mathbb{E}_t[(1-\omega)U_{t+\tau} - \omega B_{t+\tau}],$$

which can be solved to deduce that

$$(1-\omega)U_t - \omega B_t = \left(\frac{\tau}{1-\beta}\right)(\omega F - (1-\omega)M).$$

Substituting this into (A.2.1) yields the formula in (3.9) for the transaction price for a house with match quality ϵ to its buyer.

A.3 Steady state

The steady state of the model is defined by the absence of aggregate shocks ($\xi_t = 1$).

Equation (3.14) implies that $H(x) = W + \tau M$. Since $H(\delta\epsilon) < W + \tau M$, the Bellman equation (3.13) with $\epsilon = x$ implies:

$$H(x) = \tau x + \beta\alpha(H(x) - \tau M) + \beta(1-\alpha)W.$$

Substituting $H(x) = W + \tau M$ into this equation allows the value of W to be found:

$$W = \left(\frac{\tau}{1-\beta}\right)(x - M). \qquad \text{[A.3.1]}$$

Using $H(x) = W + \tau M$, an expression for $H(x)$ is also obtained:

$$H(x) = \left(\frac{\tau}{1-\beta}\right)(x - \beta M). \qquad \text{[A.3.2]}$$

Equation (3.3) for the acceptance threshold implies $H(y) = C + \beta W$. Combining this with (A.3.1) yields a formula for $H(y)$:

$$H(y) = C + \beta\left(\frac{\tau}{1-\beta}\right)(x - M). \tag{A.3.3a}$$

It is assumed that idiosyncratic shocks are sufficiently large that they will cause some homeowners to move. This requires that $\delta y < x$, since new matches form only if $\epsilon \geq y$. Since (3.14) implies $H(x) = W + \tau M$, it follows that $H(\delta y) < W + \tau M$, and hence using the Bellman equation (3.13) with $\epsilon = y$:

$$H(y) = \tau y + \beta\alpha(H(y) - \tau M) + \beta(1-\alpha)W.$$

Substituting the expression for W in (A.3.1) and simplifying:

$$H(y) = \left(\frac{\tau}{1-\beta\alpha}\right)y + \beta\left(\frac{1-\alpha}{1-\beta\alpha}\right)\left(\frac{\tau}{1-\beta}\right)x - \beta\left(\frac{\tau}{1-\beta}\right)M. \tag{A.3.3b}$$

Equating the two formulas for $H(y)$ in (A.3.3a) and (A.3.3b) leads to:

$$y = \alpha\beta x + \left(\frac{1-\beta\alpha}{\tau}\right)C. \tag{A.3.4}$$

At the steady state, the Bellman equation (3.6) becomes:

$$W = -\tau(F + M) + \beta W + \nu\pi\varSigma,$$

which can be solved for W:

$$W = -\left(\frac{\tau}{1-\beta}\right)(F + M) + \left(\frac{\nu}{1-\beta}\right)\pi\varSigma.$$

Substituting this into equation (A.3.1) yields the following expression for the average surplus $\varSigma$:

$$x + F = \left(\frac{\nu}{\tau}\right)\pi\varSigma. \tag{A.3.5}$$

To derive another equation for the surplus $\varSigma$, define the following function $\varPsi(z)$ for an arbitrary value of $z \leq y$:

$$\varPsi(z) = \int_{\epsilon=z}^{\infty}\frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)}(H(\epsilon) - H(z))\mathrm{d}\epsilon. \tag{A.3.6}$$

Since $\varSigma(y) = 0$ and (3.2) states that $\varSigma(\epsilon) = H(y) - \beta W - C$, the surplus can be written as $\varSigma(\epsilon) = H(\epsilon) - H(y)$. Comparing equations (3.5) and (A.3.6):

$$\varPsi(y) = \varSigma = \int_{\epsilon=y}^{\infty}\frac{\lambda}{y}\left(\frac{\epsilon}{y}\right)^{-(1+\lambda)}(H(\epsilon) - H(y))\mathrm{d}\epsilon. \tag{A.3.7}$$

At the steady state, the Bellman equation (3.13) for an owner-occupier is:

$$H(\epsilon) = \tau\epsilon + \beta\alpha(H(\epsilon) - \tau M) + \beta(1-\alpha)\max\{H(\delta\epsilon) - \tau M, W\}.$$

Taking any $z \geq y$, since $\delta y < x$, it must be the case that $\delta z < x$, and hence $H(\delta z) < W + \tau M$. Therefore, $H(z)$ is:

$$H(z) = \tau z + \beta\alpha(H(z) - \tau M) + \beta(1-\alpha)W.$$

Using $H(x) = W + \tau M$, the difference between the value functions $H(\epsilon)$ and $H(z)$ is:

$$H(\epsilon) - H(z) = \tau(\epsilon - z) + \beta\alpha(H(\epsilon) - H(z)) + \beta(1-\alpha)\max\{H(\delta\epsilon) - H(x), 0\}. \tag{A.3.8}$$

Substituting this expression into the definition of $\Psi(z)$ from (A.3.6) yields:

$$\Psi(z) = \int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} \tau(\epsilon - z)\mathrm{d}\epsilon + \beta\alpha\Psi(z) + \beta(1-\alpha)\int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} \max\{H(\delta\epsilon) - H(x), 0\}\mathrm{d}\epsilon. \tag{A.3.9}$$

Observe that the first integral on the right-hand side can be written as:

$$\int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} (\epsilon - z)\mathrm{d}\epsilon = \frac{1}{\lambda - 1} z. \tag{A.3.10}$$

To find the second integral on the right-hand side of (A.3.9), make the change of variable $\epsilon' = \delta\epsilon$ as follows:

$$\int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} \max\{H(\delta\epsilon) - H(x), 0\}\mathrm{d}\epsilon = \int_{\epsilon'=\delta z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{\delta z}\right)^{-(1+\lambda)} \max\{H(\epsilon') - H(x), 0\}\frac{\mathrm{d}\epsilon'}{\delta}.$$

Breaking the range of integration down into two parts, $[\delta x, x)$ and $[x, \infty)$, and noting that $H(\epsilon) < H(x)$ for all $\epsilon < x$:

$$\int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} \max\{H(\delta\epsilon) - H(x), 0\}\mathrm{d}\epsilon = \int_{\epsilon=\delta z}^{x} \frac{\lambda}{\delta z}\left(\frac{\epsilon}{\delta z}\right)^{-(1+\lambda)} \max\{H(\epsilon) - H(x), 0\}\mathrm{d}\epsilon$$
$$+ \int_{\epsilon=x}^{\infty} \frac{\lambda}{\delta z}\left(\frac{\epsilon}{\delta z}\right)^{-(1+\lambda)} \max\{H(\epsilon) - H(x), 0\}\mathrm{d}\epsilon = \left(\frac{\delta z}{x}\right)^{\lambda} \int_{\epsilon=x}^{\infty} \frac{\lambda}{x}\left(\frac{\epsilon}{x}\right)^{-(1+\lambda)} (H(\epsilon) - H(x))\mathrm{d}\epsilon.$$

Together with (A.3.10), this equation shows that (A.3.9) implies

$$\Psi(z) = \left(\frac{\tau}{1-\beta\alpha}\right)\left(\frac{z}{\lambda - 1}\right) + \beta\varDelta\left(\frac{1-\alpha}{1-\beta\alpha}\right)\left(\frac{z}{x}\right)^{\lambda}\Psi(x), \tag{A.3.11}$$

for all $z \leq y$, where $\varDelta \equiv \delta^{\lambda}$ is defined.

Evaluating (A.3.11) at $z = x$ (valid since $x < y$) yields an expression for $\Psi(x)$:

$$\Psi(x) = \left(\frac{\tau}{1-\beta\sigma_x}\right)\left(\frac{x}{\lambda - 1}\right), \tag{A.3.12}$$

where $\sigma_x \equiv \alpha + (1-\alpha)\varDelta$ is defined. Then evaluating (A.3.11) at $z = y$ and substituting the expression for $\Psi(x)$ from (A.3.12):

$$\Psi(y) = \left(\frac{y}{\lambda - 1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(1 + \beta\varDelta\left(\frac{y}{x}\right)^{\lambda-1}\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\right). \tag{A.3.13}$$

Noting that equation (3.4) implies that $\pi = y^{-\lambda}$ in the steady state, by substituting (A.3.13) into (A.3.7) and then (A.3.5):

$$x + F = \left(\frac{\nu}{\tau}\right)\left(\frac{1}{\lambda - 1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(y^{1-\lambda} + \beta\varDelta\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)x^{1-\lambda}\right). \tag{A.3.14}$$

The steady-state thresholds x and y are the solution of the simultaneous equations (A.3.4) and (A.3.14). Each equation defines a relationship between x and y. Equation (A.3.4) implies a positive relationship between x and y, while equation (A.3.14) implies a negative relationship between x and y. If a solution exists, it must then be unique. Since (A.3.4) implies x is positive when $y = 0$, and because (A.3.14) implies $y \to 0$ as $x \to \infty$, while x tends to positive number when $y \to \infty$, it follows that a unique solution $x > 0$ and $y > 0$ exists. However, the equations are only meaningful if $y > 1$ and $\delta y < x$. The solution features

$y > 1$ if and only if:

$$\left(\frac{1-\left(\frac{1-\beta\alpha}{\tau}\right)C}{\beta\alpha}\right)+F < \left(\frac{\nu}{\tau}\right)\left(\frac{1}{\lambda-1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(1+\beta\Delta\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\left(\frac{1-\left(\frac{1-\beta\alpha}{\tau}\right)C}{\beta\alpha}\right)^{1-\lambda}\right).$$

In addition, it is necessary to verify that $\delta x < y$.

Given the solutions for x and y, the value functions W, $H(x)$, and $H(y)$ can be found using (A.3.1), (A.3.2), and (A.3.3a). The average surplus Δ can be found using (A.3.5). Using $H(\epsilon) = H(y) + \Sigma(\epsilon)$, the steady-state value of H can be found using $H = H(y) + \Sigma$:

$$H = C + \beta\left(\frac{\tau}{1-\beta}\right)(x-M) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)(x+F), \qquad \text{[A.3.15]}$$

where $\pi = y^{-\lambda}$. The steady-state average price is then found using equation (3.10):

$$P = \omega H + (\kappa-\omega)C + \beta\left(\frac{\tau}{1-\beta}\right)(\omega F - (1-\omega)M),$$

and by substituting the expression for H from (A.3.15):

$$P = \kappa C - \beta\left(\frac{\tau}{1-\beta}\right)M + \omega\left(\beta\left(\frac{\tau}{1-\beta}\right) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)\right)(x+F). \qquad \text{[A.3.16]}$$

To derive the mean-min ratio $P/p(y)$ measure of price dispersion, start by noting that (3.9) implies that the minimum price $p(y)$ is

$$p(y) = \omega H(y) + (\kappa-\omega)C + \beta\left(\frac{\tau}{1-\beta}\right)(\omega F - (1-\omega)M),$$

which can be combined with the expression for $H(y)$ from (A.3.3a) to obtain:

$$p(y) = \kappa C - \beta\left(\frac{\tau}{1-\beta}\right)M + \beta\left(\frac{\tau}{1-\beta}\right)\omega(x+F).$$

Now divide both sides by the average price P:

$$\frac{p(y)}{P} = \kappa c - \beta\left(\frac{\tau}{1-\beta}\right)m + \beta\left(\frac{\tau}{1-\beta}\right)\omega\left(\frac{x}{P}+f\right). \qquad \text{[A.3.17]}$$

Dividing both sides of the equation (A.3.16) for the average price by P and rearranging it to deduce the following:

$$\omega\left(\frac{x}{P}+f\right) = \frac{1-\left(\kappa c - \beta\left(\frac{\tau}{1-\beta}\right)m\right)}{\beta\left(\frac{\tau}{1-\beta}\right)+\frac{1}{\pi}\left(\frac{\tau}{\nu}\right)}.$$

Substituting this formula into (A.3.17) and simplifying leads to:

$$\frac{p(y)}{P} = \frac{\beta\left(\frac{\tau}{1-\beta}\right)+\frac{1}{\pi}\left(\frac{\tau}{\nu}\right)\left(\kappa c - \beta\left(\frac{\tau}{1-\beta}\right)m\right)}{\beta\left(\frac{\tau}{1-\beta}\right)+\frac{1}{\pi}\left(\frac{\tau}{\nu}\right)}. \qquad \text{[A.3.18]}$$

The mean-min ratio is therefore given by:

$$\frac{P}{p(y)} = \frac{1 + \beta\pi\left(\frac{\nu}{\tau}\right)\left(\frac{\tau}{1-\beta}\right)}{\left(\kappa c - \beta\left(\frac{\tau}{1-\beta}\right)m\right) + \beta\pi\left(\frac{\nu}{\tau}\right)\left(\frac{\tau}{1-\beta}\right)}. \tag{A.3.19}$$

Turning now to the stocks and the flows, at the steady state, equation (3.15) for the unsold stock of houses becomes:

$$u = 1 - \nu u \sum_{i=1}^{\infty}\left\{y^{-\lambda}\alpha^i + x^{-\lambda}(1-\alpha)\varDelta\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j}\right\}, \tag{A.3.20}$$

noting the definitions $\varDelta = \delta^\lambda$ and $\sigma_x = \alpha + (1-\alpha)\varDelta$. Observe that

$$\begin{aligned}(1-\alpha)\varDelta\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j} &= (1-\alpha)\varDelta\sigma_x^{i-1}\sum_{j=0}^{i-1}\left(\frac{\alpha}{\sigma_x}\right)^j = (1-\alpha)\varDelta\sigma_x^{i-1}\left(\frac{1-(\alpha/\sigma_x)^i}{1-(\alpha/\sigma_x)}\right)\\ &= \frac{(1-\alpha)\varDelta\sigma_x}{\sigma_x-\alpha}\frac{\sigma_x\sigma_x^{i-1}}{\sigma_x^i}(\sigma_x^i-\alpha^i) = \sigma_x^i - \alpha^i,\end{aligned} \tag{A.3.21}$$

using $\sigma_x - \alpha = (1-\alpha)\varDelta$. This equation can then be used to simplify the terms that appear in (A.3.20):

$$\begin{aligned}&\sum_{i=1}^{\infty}\left\{y^{-\lambda}\alpha^i + x^{-\lambda}(1-\alpha)\varDelta\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j}\right\}\\ &\qquad = y^{-\lambda}\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^\lambda(\sigma_x^i-\alpha^i)\right\} = \pi\left(\frac{\alpha}{1-\alpha} + \left(\frac{y}{x}\right)^\lambda\left(\frac{\sigma_x}{1-\sigma_x} - \frac{\alpha}{1-\alpha}\right)\right),\end{aligned} \tag{A.3.22}$$

noting that $\pi = y^{-\lambda}$. Now simplify the following term using $\sigma_x - \alpha = (1-\alpha)\varDelta$:

$$\frac{\sigma_x}{1-\sigma_x} - \frac{\alpha}{1-\alpha} = \frac{\sigma_x(1-\alpha)-\alpha(1-\sigma_x)}{(1-\alpha)(1-\sigma_x)} = \frac{\sigma_x-\alpha}{(1-\alpha)(1-\sigma_x)} = \frac{(1-\alpha)\varDelta}{(1-\alpha)(1-\sigma_x)} = \frac{\varDelta}{1-\sigma_x}.$$

Substituting this into (A.3.22) and then into (A.3.20) leads to the following equation:

$$u = 1 - \nu u\left\{\pi\left(\frac{\alpha}{1-\alpha} + \frac{\varDelta}{1-\sigma_x}\left(\frac{y}{x}\right)^\lambda\right)\right\}.$$

This can be solved explicitly for the steady-state stock u of unsold houses:

$$u = \frac{1}{1 + \pi\nu\left(\frac{\alpha}{1-\alpha} + \varDelta\left(\frac{y}{x}\right)^\lambda\frac{1}{1-\sigma_x}\right)}. \tag{A.3.23}$$

The steady-state selling probability is $s = \nu\pi$, where $\pi = y^{-\lambda}$. The number of sales in a discrete time period of length τ is then $S = su$, where u is taken from (A.3.23). At the end of a time period, a total of $1 - u + S = 1 - (1-s)u$ homes are owned by families not trying to sell them. The steady-state moving probability in a discrete time period is n, and total inflows are $N = n(1-(1-s)u)$. Inflows are equal to outflows in steady state, so $S = N$, which requires $su = n(1-(1-s)u$. Solving this equation for u yields a formula in terms of the inflow and outflow probabilities:

$$u = \frac{1}{1 + s\left(\frac{1}{n}-1\right)}.$$

Since $s = \nu\pi$, comparing this equation with (A.3.23) leads to the following expression for the inflow probability:

$$n = \frac{1-\alpha}{1+\mathit{\Delta}\left(\frac{y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-\sigma_x}\right)}. \tag{A.3.24}$$

Viewings result in a successful sale with probability π, and the outcome of each viewing is independent because the draws of match-specific quality ϵ are independent. The probability that a house is sold with ℓ viewings is therefore $\pi(1-\pi)^{\ell-1}$. This implies the expected number of viewings per sale is:

$$V_{\mathrm{s}} = \sum_{\ell=1}^{\infty} \ell\pi(1-\pi)^{\ell-1} = \frac{1}{\pi}. \tag{A.3.25}$$

Since a viewing is received with probability ν, the probability of a successful sale in one discrete time period (of length τ years) is s. Thus, the probability that a house sells after ℓ time periods is $s(1-s)^{\ell-1}$. It follows that the expected selling time in years is:

$$T_{\mathrm{s}} = \tau\sum_{\ell=1}^{\infty} \ell s(1-s)^{\ell-1} = \frac{\tau}{s}.$$

Since $s = \nu\pi$, the formula for time-to-sell can be written as:

$$T_{\mathrm{s}} = \left(\frac{\tau}{\nu}\right)\frac{1}{\pi} = \left(\frac{\tau}{\nu}\right)V_{\mathrm{s}}. \tag{A.3.26}$$

Finally, the steady-state survival rate and hazard function for new matches are derived. In steady state, all cohorts of new matches have match-specific quality $\epsilon \sim \text{Pareto}(y;\lambda)$ because a $\text{Pareto}(1;\lambda)$ distribution truncated at $\epsilon = y$ remains a Pareto distribution with shape parameter λ. The group that receives an idiosyncratic shock but decides not to move has a distribution of match quality $\epsilon \sim \text{Pareto}(x;\lambda)$ because the distribution is truncated at $\epsilon = x$. After receiving further shocks, the distribution of surviving match quality remains $\epsilon \sim \text{Pareto}(x;\lambda)$.

Starting from the distribution $\epsilon \sim \text{Pareto}(x;\lambda)$, the probability that $\epsilon' = \delta\epsilon$ remains above x is $\mathit{\Delta} = \delta^{\lambda}$. If no idiosyncratic shock is received, no moves will occur. Hence for a cohort that has already received one shock and survived, the unconditional survival probability in one discrete time period is $\sigma_x = \alpha + (1-\alpha)\mathit{\Delta}$. For the cohort starting with distribution $\epsilon \sim \text{Pareto}(y;\lambda)$ and that has previously received no idiosyncratic shocks, the survival probability following the first shock is $(\delta y/x)^{\lambda}$. The unconditional survival probability is therefore $\sigma_y = \alpha + (1-\alpha)\Delta(y/x)^{\lambda}$ for this group until the first shock is received.

Let ψ_ℓ denote the proportion of the new cohort that still survive after ℓ discrete time periods ($\psi_0 = 1$), namely, the survival function. In each time period, the probability of no idiosyncratic shock is α. After ℓ periods, the proportion of matches that have received no idiosyncratic shock at any point is α^{ℓ}. Given that σ_x is the survival probability for those who have previously received a shock while σ_y is the survival probability for those who have received no prior shock, the iteration to calculate the survival function is $\psi_\ell = \sigma_y\alpha^{\ell-1} + \sigma_x(\psi_{\ell-1} - \alpha^{\ell-1})$, which implies:

$$\psi_\ell - \alpha^{\ell} = (\sigma_y - \alpha)\alpha^{\ell-1} + \sigma_x(\psi_{\ell-1} - \alpha^{\ell-1}).$$

Iterating this sequence leads to $\psi_0 - \alpha^0 = 0$, $\psi_1 - \alpha^1 = \sigma_y - \alpha$, $\psi_2 - \alpha^2 = (\sigma_y - \alpha)(\alpha + \sigma_x)$, and the general formula:

$$\psi_\ell - \alpha^{\ell} = (\sigma_y - \alpha)\left(\alpha^{\ell-1} + \alpha^{\ell-2}\sigma_x + \cdots + \alpha\sigma_x^{\ell-2} + \sigma_x^{\ell-1}\right). \tag{A.3.27}$$

Summing the geometric series implies:

$$\psi_\ell - \alpha^\ell = (\sigma_y - \alpha)\sigma_x^{\ell-1}\left(\frac{1-\left(\frac{\alpha}{\sigma_x}\right)^\ell}{1-\frac{\alpha}{\sigma_x}}\right) = \left(\frac{\sigma_y - \alpha}{\sigma_x - \alpha}\right)\left(\sigma_x^\ell - \alpha^\ell\right) = \left(\frac{y}{x}\right)^\lambda \left(\sigma_x^\ell - \alpha^\ell\right),$$

and substituting this into (A.3.27) yields the survival function:

$$\psi_\ell = \alpha^\ell + \left(\frac{y}{x}\right)^\lambda (\sigma_x^\ell - \alpha^\ell). \tag{A.3.28}$$

The survival function sequence ψ_ℓ defines the hazard function ζ_ℓ through the equation $\psi_\ell = (1-\zeta_\ell)\psi_{\ell-1}$. The probability that a new match will survive for exactly ℓ discrete time periods of length τ is $\zeta_\ell\psi_{\ell-1}$. Using $\zeta_\ell\psi_{\ell-1} = \psi_{\ell-1} - \psi_\ell$, the expected duration T_d in years of a new match is:

$$T_\text{d} = \tau\sum_{\ell=1}^\infty \ell\zeta_\ell\psi_{\ell-1} = \tau\sum_{\ell=1}^\infty \ell(\psi_{\ell-1} - \psi_\ell) = \tau\left(\sum_{\ell=0}^\infty (\ell+1)\psi_\ell - \sum_{\ell=0}^\infty \ell\psi_\ell\right) = \tau\sum_{\ell=0}^\infty \psi_\ell. \tag{A.3.29}$$

The formula in (A.3.28) can be used to sum the survival function sequence:

$$\sum_{\ell=0}^\infty \psi_\ell = \frac{1}{1-\alpha} + \left(\frac{y}{x}\right)^\lambda\left(\frac{1}{1-\sigma_x} - \frac{1}{1-\alpha}\right) = \left(\frac{1}{1-\alpha}\right)\left(1 + \mathit{\Delta}\left(\frac{y}{x}\right)^\lambda\left(\frac{1-\alpha}{1-\sigma_x}\right)\right), \tag{A.3.30}$$

noting that $\sigma_x - \alpha = (1-\alpha)\mathit{\Delta}$. This is substituted into (A.3.29) to obtain the expected duration of a match:

$$T_\text{d} = \left(\frac{\tau}{1-\alpha}\right)\left(1 + \mathit{\Delta}\left(\frac{y}{x}\right)^\lambda\left(\frac{1-\alpha}{1-\sigma_x}\right)\right). \tag{A.3.31}$$

Comparing this to equation (A.3.24), the expected duration can also be written in terms of the unconditional inflow probability n:

$$T_\text{d} = \frac{\tau}{n}. \tag{A.3.32}$$

The stationary distribution of homeowner tenure (how long since they purchased their houses) is denoted by ϕ_ℓ. The distribution of tenure can be derived from the survival function ψ_ℓ by noting that $\phi_\ell = (1-\zeta_\ell)\phi_{\ell-1}$. Since $\psi_\ell = (1-\zeta_\ell)\psi_{\ell-1}$ and $\psi_0 = 1$, this implies that $\phi_\ell = \psi_\ell\phi_0$. Given that $\sum_{\ell=0}^\infty \phi_\ell = 1$, it follows that:

$$\phi_\ell = \frac{\psi_\ell}{\sum_{j=0}^\infty \psi_j}. \tag{A.3.33}$$

The average probability of moving ζ is given by:

$$\zeta = \sum_{\ell=1}^\infty \zeta_\ell\phi_{\ell-1} = \sum_{\ell=1}^\infty (\phi_{\ell-1} - \phi_\ell) = \phi_0 = \frac{1}{\sum_{\ell=0}^\infty \psi_\ell},$$

noting that $\zeta_\ell\phi_{\ell-1} = \phi_{\ell-1} - \phi_\ell$ and the expression for ϕ_0 from (A.3.33). From (A.3.30) it can be seen that $\zeta = n$, where n is the unconditional inflow probability. The average tenure T_a is defined as follows:

$$T_\text{a} = \tau\sum_{\ell=1}^\infty \ell\phi_{\ell-1}.$$

To calculate T_{a}, the expression for ψ_ℓ in (A.3.28) can be used to obtain:

$$\sum_{\ell=1}^{\infty} \ell\psi_{\ell-1} = \frac{1}{(1-\alpha)^2} + \left(\frac{y}{x}\right)^{\lambda}\left(\frac{1}{(1-\sigma_x)^2} - \frac{1}{(1-\alpha)^2}\right)$$

Noting that $\phi_\ell = \psi_\ell \phi_0$ and using the expression for ϕ_0 from (A.3.33), the equation above implies:

$$T_{\mathrm{a}} = \frac{\tau\sum_{\ell=1}^{\infty}\ell\psi_{\ell-1}}{\sum_{\ell=0}^{\infty}\psi_\ell} = \frac{\frac{\tau}{1-\alpha} + \left(\frac{y}{x}\right)^{\lambda}\left(\frac{\tau}{1-\sigma_x}\right)\left(\frac{(1-\alpha)^2-(1-\sigma_x)^2}{(1-\alpha)(1-\sigma_x)}\right)}{1+\varDelta\left(\frac{y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-\sigma_x}\right)}.$$

By simplifying the expression, the following formula for T_{a} is obtained:

$$T_{\mathrm{a}} = \frac{\left(\frac{\tau}{1-\alpha}\right) + \varDelta\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\tau}{1-\sigma_x}\right)\left(\frac{2-\alpha-\sigma_x}{1-\sigma_x}\right)}{1+\varDelta\left(\frac{y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-\sigma_x}\right)}. \qquad \text{[A.3.34]}$$

This completes the derivation of the steady state.

A.4 Calibration method

First, some of the parameters are set directly. These are the length τ of a time period, the discount rate r, the bargaining power ω of the seller, the share κ of the transaction costs borne by the seller, and the persistence parameter ρ of the demand shock. This leaves seven parameters to be determined indirectly. These are the arrival rate a of idiosyncratic shocks, the size δ of those shocks, the arrival rate v of viewings, the transaction cost C, the flow search costs F, the flow maintenance cost M, and the parameter λ describing the steady-state distribution of match quality.

Identifying these seven parameters requires seven empirical targets. These are the average time-to-sell T_{s}, the average number of viewings V_{s} per sale, the expected duration T_{d} of owning a newly purchased house, the average number of years T_{a} since all owners bought their houses, the ratio c of transaction cost to price, the ratio f of flow search costs to price, and the ratio m of flow maintenance costs to price. Empirical evidence on these targets is presented in Ngai and Sheedy (2019), which is summarized below.

Following Poterba (1991), the flow cost M of owning a house is set so that in equilibrium it is 4.5% of the average house price ($m = 0.045$). This cost is made up of a 2.5% maintenance cost and a 2% property tax. The maintenance cost is interpreted as the cost required perpetually to maintain a house in the same physical condition as when it was first purchased. The value of 2.5% from Poterba (1991) is used a benchmark for this maintenance cost.

The costs incurred in buying and selling houses comprise the one-off transactions cost C and the flow costs of search F. For the transaction costs, Quigley (2002) estimates the total costs as being in the range 6–12% in the U.S., with about 3–6% being the realtor's fee paid by the seller. Ghent (2012) summarizes recent research and uses a total transaction cost of 13.1%, where 5.1% is the realtor's fee borne by the seller. In light of these findings, the total transaction cost C is set so that it is 10% of the price ($c = 0.1$), and the share κ of these costs borne by the seller is set to be 1/3.

For the flow cost parameter F, unfortunately there are almost no estimates of the flow costs of searching. The approach taken here is to base an estimate of F on the opportunity cost of the time spent searching. Assuming one house viewing entails the loss of a day's income, the value of F can be calibrated by adding up the cost of making the expected number of viewings. In the model, time-to-buy is equal to time-to-sell, so buyers will incur search costs $T_{\mathrm{s}}F$ per housing transaction on average, where T_{s} denotes time-to-sell. With viewings-per-sale equal to the average number of viewings made by a buyer, the total search cost should be equated to $V_{\mathrm{s}}I/365$, where V_{s} denotes average viewings-per-sale and I denotes average annual income. Thus, the calibration assumes $T_{\mathrm{s}}F = V_{\mathrm{s}}I/365$, and by dividing both sides by PT_{s}, this implies an equation for $f \equiv F/P$:

$$f = \frac{1}{365} \frac{I}{P} \frac{V_\mathrm{s}}{T_\mathrm{s}}.$$

Using a house price to income ratio of 2 as a reasonable average value (Case and Shiller, 2003) together with the values of $T_\mathrm{s} = 6.5/12$ and $V_\mathrm{s} = 10$ discussed below, the ratio of the flow cost of search to the average price is calibrated to be 2.5% ($f = 0.025$). Note that 2.5% should be interpreted as the hypothetical cost of spending a whole year searching.

The construction of the measure of time-to-sell using data from the National Association of Realtors (NAR) on sales and inventories (for existing single-family homes) was described in section 2. The average of this variable over the period 1991-2012 is 6.4 months, while the average of the 'months supply' produced by NAR is 6.6 months (the difference between these numbers is discussed in section 2).

Previous research on housing markets has used a variety of sources for data on time-to-sell, and there is a considerable dispersion in these estimates. The average value of time-to-sell is crucial in quantifying the importance of search frictions in the housing market. Using the 'Profile of Buyers and Sellers' survey collected by NAR, Genesove and Han (2012) report that for the time period 1987–2008, the average time-to-sell is 7.6 weeks, the average time-to-buy is 8.1 weeks, and the average number of homes visited by buyers is 9.9. They also discuss other surveys that have reported similar findings.

However, the estimates of time-to-sell and time-to-buy derived from survey data are likely to be an underestimate of the actual time a new buyer or seller would expect to spend in the housing market. The reason is that the survey data include only those buyers and sellers who have successfully completed a house purchase or sale, while the proportion of buyers or sellers who withdraw from the market (at least for some time) without a completed transaction is substantial. Genesove and Mayer (1997) estimate the fraction of withdrawals at 50%, and Levitt and Syverson (2008) report a withdrawal rate of 22%. In comparing the efficiency of different platforms for selling properties, Hendel, Nevo and Ortalo-Magné (2009) explicitly control for withdrawals and report a time-to-sell of 15 weeks (using the Multiple Listing Service for the city of Madison).[25]

An alternative approach to estimating time-to-sell that does not face the problem of withdrawals is to look at the average duration of the time for which a home is vacant using data from the American Housing Survey. In the years 2001–2005, the mean duration for a vacancy was 7–8 months. However, that number is likely to be an overestimate of the expected time-to-sell because it is based on houses that are 'vacant for sale'. Houses that are for sale but currently occupied would not be counted in this calculation of average duration. A further approach that avoids the problems of withdrawals is to look at the average time taken to sell newly built houses. Díaz and Jerez (2013) use the Census Bureau 'New Residential Sales' report to find that the median number of months taken to sell a newly built house is 5.9 (for the period 1991–2012). This is only slightly shorter than the average of the time-to-sell number constructed using NAR data on existing single-family homes, but there is reason to believe that newly built homes should sell faster than existing homes owing to greater advertising expenditure and differences in the target groups of buyers.

In summary, most studies find that average time-to-sell is less than three months in cases where there is a potential withdrawal bias that is not controlled for. Most studies that are not subject to this bias, or attempt to control for it, find times-to-sell of more than four months. Since the dynamic predictions of the model will be compared to variables constructed from the NAR sales and inventories data, a measure of time-to-sell consistent with this data will be used. The calibration target is a time-to-sell of 6.5 months (the average of the NAR 'months supply' number and the time-to-sell number derived from the NAR data), hence $T_\mathrm{s} = 6.5/12$. The calibration target for viewings per sale is set to 10 ($V_\mathrm{s} = 10$) on the basis of the studies discussed above.

The remaining calibration targets are for the number of years a buyer expects to remain in the same

[25]For the U.K., Merlo and Ortalo-Magné (2004) obtain data from four real estate agencies that contain 780 completed transaction histories between 1995–1998 for Greater London and for South Yorkshire. They report an average time-to-sell of 11 weeks, but this number does not control for withdrawals, which they find occur at a rate of 25% in their data. They also report an average of 9.5 viewings per transaction for a sub-sample of 199 properties in their data.

house (expected duration), and the average number of years existing home-owners have lived in their current houses (average years since moving). Note that these two numbers are not necessarily the same when the hazard rate of moving is not independent of the time already spent in a house. An estimate of both expected duration and average years since moving can be derived from the data in Table 2.9 (Year Householder Moved into Unit) of the American Housing Survey, which gives a frequency distribution for the time since owners moved into their homes. The data are supplied in 5-year bins for durations of less than 40 years, and in 10-year bins for longer durations.[26] In calculating the expected duration and the average number of years since moving, the frequency in each bin is assumed to be equally distributed within the bin. As in the calculation of the total housing stock in section 2, elderly owners (over 65 years) are removed from the data because such individuals are less likely to consider moving. Using the 2005 survey, the average years since moving is found to be 11 years ($T_{\mathrm{a}} = 11$).

The expected duration is found from the same data by calculating the hazard function for moving house consistent with the frequency distribution of the years since the home-owner moved to their current house (this assumes that the empirical distribution is the stationary distribution implied by the hazard function). The method leads to an estimate of expected duration of 12.2 years ($T_{\mathrm{d}} = 12.2$). Note that this is very close to the expected duration of 11.9 years consistent with the average inflow rate found in section 2 (and which was derived from an independent data source). That the expected duration is longer than the average number of years since moving is consistent with the model's prediction of a hazard rate for moving that is increasing in time spent in a house.

Finally, given the seven empirical targets, $\{T_{\mathrm{s}}, V_{\mathrm{s}}, T_{\mathrm{d}}, T_{\mathrm{a}}, c, f, m\}$, the numerical approach to finding the parameters that match the targets is to search over values of one parameter, the arrival rate a of idiosyncratic shocks. Given this parameter, all other parameters can be obtained from analytical expressions. Finally, a criterion that the correct value of a must satisfy is derived. The method is therefore to search over values of a until the one satisfying the criterion is found.

The method starts from known values of τ, r, ω, and κ. There is a guess for the value of a. The coefficients β and α are calculated using $\beta = e^{-r\tau}$ and $\alpha = e^{-a\tau}$. Now define a variable variable $\wp$ as follows:

$$\wp \equiv \mathit{\Delta} \left(\frac{y}{x}\right)^{\lambda} \left(\frac{1-\alpha}{1-\sigma_x}\right). \tag{A.4.1}$$

Given the calibration target T_{d} and equation (A.3.31), $\wp$ can be found using:

$$\wp = \left(\frac{1-\alpha}{\tau} T_{\mathrm{d}}\right) - 1. \tag{A.4.2}$$

Now observe using equation (A.4.1) that the expression for T_{a} in (A.3.34) can be written in terms of $\wp$ as:

$$T_{\mathrm{a}} = \frac{(1+\wp)\left(\frac{\tau}{1-\alpha}\right) + \wp\left(\frac{\tau}{1-\sigma_x}\right)}{1+\wp},$$

which can be rearranged as follows:

$$\frac{\tau}{1-\sigma_x} = \left(\frac{1+\wp}{\wp}\right)\left(T_{\mathrm{a}} - \left(\frac{\tau}{1-\alpha}\right)\right).$$

[26] The first bin requires special treatment because it covers a five-year interval that does not generally coincide with the survey year, and because the survey itself is conducted in the middle of the year (between mid-April and mid-September during a survey year). For example, in 2005, the first bin starts in the survey year so this bin effectively covers only one tenth of the time spanned by the other bins. The frequency in the first bin is scaled up accordingly.

By substituting the expression for $\wp$ from (A.4.2) into the above:

$$\frac{\tau}{1-\sigma_x} = \left(\frac{T_{\mathrm{a}} - \left(\frac{\tau}{1-\alpha}\right)}{T_{\mathrm{d}} - \left(\frac{\tau}{1-\alpha}\right)} \right) T_{\mathrm{d}},$$

which can be rearranged to yield an equation for σ_x:

$$\sigma_x = 1 - \left(\frac{\tau}{T_{\mathrm{d}}}\right) \left(\frac{T_{\mathrm{d}} - \left(\frac{\tau}{1-\alpha}\right)}{T_{\mathrm{a}} - \left(\frac{\tau}{1-\alpha}\right)} \right). \tag{A.4.3}$$

Since $\sigma_x = \alpha + (1-\alpha)\mathit{\Delta}$, it follows that $\mathit{\Delta}$ can be obtained from knowledge of α and σ_x using:

$$\mathit{\Delta} = \frac{\sigma_x - \alpha}{1-\alpha}. \tag{A.4.4}$$

With values of α, σ_x, $\mathit{\Delta}$, and $\wp$, equation (A.4.1) implies an expression for $(y/x)^\lambda$:

$$\left(\frac{y}{x}\right)^\lambda = \frac{\wp}{\mathit{\Delta}} \left(\frac{1-\sigma_x}{1-\alpha}\right). \tag{A.4.5}$$

The next step is to divide both sides of equation (A.3.16) by the average price P:

$$1 = \kappa c - \beta\left(\frac{\tau}{1-\beta}\right) m + \omega\left(\beta\left(\frac{\tau}{1-\beta}\right) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)\right)\left(\frac{x}{P} + f\right),$$

using the definitions $c \equiv C/P$, $m \equiv M/P$, and $f \equiv F/P$. By substituting the expression in equation (A.3.26):

$$1 = \kappa c - \beta\left(\frac{\tau}{1-\beta}\right) m + \omega\left(\beta\left(\frac{\tau}{1-\beta}\right) + T_{\mathrm{s}}\right)\left(\frac{x}{P} + f\right),$$

which can be solved for x/P as follows:

$$\frac{x}{P} = \frac{1 - \kappa c + \beta\left(\frac{\tau}{1-\beta}\right) m}{\omega\left(\beta\left(\frac{\tau}{1-\beta}\right) + T_{\mathrm{s}}\right)} - f. \tag{A.4.6}$$

Making use of the definition of c, divide both sides of equation (A.3.4) by P to obtain:

$$\frac{y}{P} = \alpha\beta\frac{x}{P} + \left(\frac{1-\beta\alpha}{\tau}\right) c, \tag{A.4.7}$$

and this yields the value of y/P given the value of x/P found from (A.4.6). Since $y/x = (y/P)/(x/P)$, the ratio y/x is now known. Together with the value of $(y/x)^\lambda$ obtained from (A.4.5), λ is found using the identity:

$$\lambda = \frac{\log\left(\frac{y}{x}\right)^\lambda}{\log\frac{y}{x}}. \tag{A.4.8}$$

There is one further equation that must hold. Using $\pi = y^{-\lambda}$, equation (A.3.14) can be written as follows:

$$x + F = \pi\left(\frac{\nu}{\tau}\right)\left(\frac{1}{\lambda-1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(y + \beta\mathit{\Delta}\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\left(\frac{y}{x}\right)^\lambda x\right),$$

and dividing both sides by P and using the definition of f and the expression for T_{s} from (A.3.26):

$$\frac{x}{P}+f=\frac{1}{T_{\mathrm{s}}}\left(\frac{1}{\lambda-1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(\frac{y}{P}+\beta\mathit{\Delta}\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\left(\frac{y}{x}\right)^{\lambda}\frac{x}{P}\right). \quad \text{[A.4.9]}$$

The guess for a is correct if and only if this equation holds given the values of the variables appearing in the equation conditional on the guess for a. The numerical component of the calibration method is therefore to search over values of a until one is found where (A.4.9) is satisfied.

Once the value of a (and hence $\alpha = e^{-a\tau}$) is known, the other parameters can be obtained as follows. The arrival rate of viewings can be found using (A.3.26) and the definition $\nu = 1 - e^{-v\tau}$:

$$v=-\frac{1}{\tau}\log\left(1-\tau\frac{V_{\mathrm{s}}}{T_{\mathrm{s}}}\right). \quad \text{[A.4.10]}$$

The value of λ is found using equation (A.4.8) via the steps described above. The value of δ can be obtained using the definition $\mathit{\Delta} = \delta^{\lambda}$ and the expression for $\mathit{\Delta}$ in (A.4.4):

$$\delta=\mathit{\Delta}^{\frac{1}{\lambda}}. \quad \text{[A.4.11]}$$

To find the values of parameters C, F, and M, it is necessary to obtain the value of the average price P, which requires knowledge of x and y. Using equation (A.3.25) and $\pi = y^{-\lambda}$ it follows that:

$$y=V_{\mathrm{s}}^{\frac{1}{\lambda}}.$$

Using the value of the ratio y/x derived earlier, the value of x can be obtained from the identity $x = y/(y/x)$. It is then possible to obtain the parameter F by rearranging equation (A.3.14) and using $\pi = y^{-\lambda}$ and (A.3.26), and the values of σ_x and $\mathit{\Delta}$ found in (A.4.3) and (A.4.4):

$$F=\frac{1}{T_{\mathrm{s}}}\left(\frac{1}{\lambda-1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(y+\beta\mathit{\Delta}\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\left(\frac{y}{x}\right)^{\lambda}x\right)-x. \quad \text{[A.4.12]}$$

With F, the price P can be obtained from $P = F/f$ by rearranging the definition of f. Now that P is known, the parameters C and M are given by $C = cP$ and $M = mP$. This completes the calibration routine.

A.5 Log linearizations

The Bellman equation (3.6) for the combined buyer-seller value W_t has the following log-linear approximate form:

$$\mathsf{W}_t=\beta\mathbb{E}_t\mathsf{W}_{t+\tau}+(1-\beta)\left(\frac{x+F}{x-M}\right)(\pi_t+\Sigma_t), \quad \text{[A.5.1]}$$

where the steady-state values in (A.3.1) and (A.3.5) have been used to derive the coefficients. There exists an $\eta > 0$ such that for any $\epsilon < y + \eta$, the value of ϵ' after a shock is less than x_t for a given bound on fluctuations in x_t. For these ϵ values, the Bellman equation (3.13) reduces to:

$$H_t(\epsilon)=\tau\epsilon\xi_t+\beta\alpha\mathbb{E}_t[H_{t+\tau}(\epsilon)-\tau M]+\beta(1-\alpha)\mathbb{E}_tW_{t+\tau},$$

which can be iterated forwards to deduce:

$$H_t(\epsilon)=\sum_{\ell=0}^{\infty}(\beta\alpha)^{\ell}\mathbb{E}_t[\tau\epsilon\xi_{t+\tau\ell}-(\beta\alpha)\tau M+\beta(1-\alpha)W_{t+\tau(\ell+1)}]. \quad \text{[A.5.2]}$$

Since $\epsilon = x_t$ satisfies the restriction required to use (A.5.2), equations (3.14) and (A.5.2) imply:

$$W_t + \tau M = \sum_{\ell=0}^{\infty} (\beta\alpha)^{\ell} \mathbb{E}_t[\tau x_t \xi_{t+\tau\ell} - (\beta\alpha)\tau M + \beta(1-\alpha) W_{t+\tau(\ell+1)}],$$

which can be log linearized as follows (using the formula for the steady-state value of W in (A.3.1)):

$$\left(1 - \frac{M}{x}\right)\left(\frac{\tau}{1-\beta}\right) \mathsf{W}_t = \tau \sum_{\ell=0}^{\infty} (\beta\alpha)^{\ell} \mathbb{E}_t \left[\mathsf{x}_t + \xi_{t+\tau\ell} + \beta(1-\alpha)\left(1 - \frac{M}{x}\right)\left(\frac{\tau}{1-\beta}\right) \mathsf{W}_{t+\tau(\ell+1)}\right].$$

This implies the following recursive equation:

$$\left(1 - \frac{M}{x}\right)\left(\frac{1}{1-\beta}\right)(\mathsf{W}_t - \beta\alpha \mathbb{E}_t \mathsf{W}_{t+\tau}) = \left(\frac{1}{1-\beta\alpha}\right)(\mathsf{x}_t - \beta\alpha \mathbb{E}_t \mathsf{x}_{t+\tau}) + \xi_t$$
$$+ \beta(1-\alpha)\left(1 - \frac{M}{x}\right)\left(\frac{1}{1-\beta}\right)\mathbb{E}_t \mathsf{W}_{t+\tau},$$

which simplifies to:

$$\left(1 - \frac{M}{x}\right)\left(\frac{1-\beta\alpha}{1-\beta}\right)(\mathsf{W}_t - \beta \mathbb{E}_t \mathsf{W}_{t+\tau}) = \mathsf{x}_t - \beta\alpha \mathbb{E}_t \mathsf{x}_{t+\tau} + (1-\beta\alpha)\xi_t.$$

Substituting the approximated Bellman equation (A.5.1) for W_t into the above leads to:

$$\mathsf{x}_t = \beta\alpha \mathbb{E}_t \mathsf{x}_{t+\tau} - (1-\beta\alpha)\xi_t + (1-\beta\alpha)\left(1 + \frac{F}{x}\right)(\pi_t + \Sigma_t), \qquad \text{[A.5.3]}$$

which can be written as:

$$\mathbb{E}_t\left[\left(\frac{\mathbb{I} - \beta\alpha\mathbb{F}}{1-\beta\alpha}\right)\mathsf{x}_t\right] = \left(1 + \frac{F}{x}\right)(\pi_t + \Sigma_t) - \xi_t. \qquad \text{[A.5.4]}$$

It is also the case that $\epsilon = y_t$ meets the requirement ($\epsilon < y + \eta$) needed to apply (A.5.2). Together with (3.3) this implies:

$$\beta \mathbb{E}_t W_{t+\tau} + C = \sum_{\ell=0}^{\infty} (\beta\alpha)^{\ell} \mathbb{E}_t[\tau y_t \xi_{t+\tau\ell} - (\beta\alpha)\tau M + \beta(1-\alpha) W_{t+\tau(\ell+1)}],$$

which has the following log-linear form:

$$\left(1 - \frac{M}{x}\right)\left(\frac{\tau}{1-\beta}\right)\beta \mathbb{E}_t \mathsf{W}_{t+\tau} = \tau \sum_{\ell=0}^{\infty} (\beta\alpha)^{\ell} \mathbb{E}_t \left[\frac{y}{x}\mathsf{y}_t + \frac{y}{x}\xi_{t+\tau\ell} + \beta(1-\alpha)\left(1 - \frac{M}{x}\right)\left(\frac{\tau}{1-\beta}\right) \mathsf{W}_{t+\tau(\ell+1)}\right].$$

Putting the equation into recursive form:

$$\left(1 - \frac{M}{x}\right)\left(\frac{1}{1-\beta}\right)\beta \mathbb{E}_t[\mathsf{W}_{t+\tau} - \beta\alpha \mathsf{W}_{t+2\tau}] = \left(\frac{1}{1-\beta\alpha}\right)\left(\frac{y}{x}\right)(\mathsf{y}_t - \beta\alpha \mathbb{E}_t \mathsf{y}_{t+\tau}) + \left(\frac{y}{x}\right)\xi_t$$
$$+ \beta(1-\alpha)\left(1 - \frac{M}{x}\right)\left(\frac{1}{1-\beta}\right)\mathbb{E}_t \mathsf{W}_{t+\tau},$$

which simplifies to:

$$\left(\frac{x}{y}\right)\left(1 - \frac{M}{x}\right)\left(\frac{1-\beta\alpha}{1-\beta}\right)\beta\alpha \mathbb{E}_t[\mathsf{W}_{t+\tau} - \beta \mathsf{W}_{t+2\tau}] = \mathsf{y}_t - \beta\alpha \mathbb{E}_t \mathsf{y}_{t+\tau} + (1-\beta\alpha)\xi_t.$$

Substituting the Bellman equation (A.5.1) leads to the following equation for y_t:

$$\mathsf{y}_t = \beta\alpha\mathbb{E}_t\mathsf{y}_{t+\tau} - (1-\beta\alpha)\xi_t + \beta\alpha(1-\beta\alpha)\left(1+\frac{F}{x}\right)\left(\frac{x}{y}\right)\mathbb{E}_t[\pi_{t+\tau} + \Sigma_{t+\tau}], \qquad [\text{A.5.5}]$$

which can be written as:

$$\mathbb{E}_t\left[\left(\frac{\mathbb{I}-\beta\alpha\mathbb{F}}{1-\beta\alpha}\right)\mathsf{y}_t\right] = \beta\alpha\left(1+\frac{F}{x}\right)\left(\frac{x}{y}\right)\mathbb{E}_t[\mathbb{F}(\pi_t + \Sigma_t)] - \xi_t. \qquad [\text{A.5.6}]$$

In what follows, define the function $\Psi_t(z)$ as follows:

$$\Psi_t(z) \equiv \int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)}(H_t(\epsilon) - H_t(z))\mathrm{d}\epsilon. \qquad [\text{A.5.7}]$$

Analogous to the definition of the expected surplus in (3.5), define the expected surplus $\bar{\Sigma}_t$ using the steady-state distribution of ϵ:

$$\bar{\Sigma}_t \equiv \int_{\epsilon=y}^{\infty} \frac{\lambda}{y}\left(\frac{\epsilon}{y}\right)^{-(1+\lambda)} \Sigma_t(\epsilon)\mathrm{d}\epsilon. \qquad [\text{A.5.8}]$$

Equations (3.2) and (3.3) imply that $\Sigma(\epsilon) = H_t(\epsilon) - H_t(y_t)$, and (A.5.2) implies:

$$H_t(y_t) - H_t(y) = \tau(y_t - y)\sum_{\ell=0}^{\infty}(\beta\alpha)^{\ell}\mathbb{E}_t\xi_{t+\tau\ell}.$$

Thus, it follows from (A.5.8) that

$$\bar{\Sigma}_t = \Psi_t(y) - \tau(y_t - y)\sum_{\ell=0}^{\infty}(\beta\alpha)^{\ell}\mathbb{E}_t\xi_{t+\tau\ell},$$

which can be log linearized as follows (noting that $\Psi(y) = \bar{\Sigma} = \Sigma$, and the expression for Σ in (A.3.5)):

$$\bar{\mathsf{\Sigma}}_t = \Psi_t(y) - \nu\pi\left(\frac{y}{x+F}\right)\left(\frac{1}{1-\beta\alpha}\right)\mathsf{y}_t. \qquad [\text{A.5.9}]$$

By substituting the Bellman equation (3.13) into (A.5.7) for a value of $z < y + \eta$ (which implies $H_{t+\tau}(\delta z) - \tau M < W_{t+\tau}$):

$$\Psi_t(z) = \int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)}\tau(\epsilon - z)\xi_t\mathrm{d}\epsilon + \beta\alpha\mathbb{E}_t\left[\int_{\epsilon=z}^{\infty}\frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)}(H_{t+\tau}(\epsilon) - H_{t+\tau}(z))\mathrm{d}\epsilon\right]$$
$$+ \beta(1-\alpha)\mathbb{E}_t\left[\int_{\epsilon=z}^{\infty}\frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)}\left(\max\left\{H_{t+\tau}(\delta\epsilon) - \tau M, W_{t+\tau}\right\} - W_{t+\tau}\right)\mathrm{d}\epsilon\right].$$

For the first term, equation (A.3.10) is used; the second term can be written in terms of $\Psi_{t+\tau}(z)$ using the definition (A.5.7); and the integral in the third term is analysed by making the change of variable $\epsilon' = \delta\epsilon$:

$$\Psi_t(z) = \frac{\tau z\xi_t}{\lambda - 1} + \beta\alpha\mathbb{E}_t\Psi_{t+\tau}(z)$$
$$+ \beta(1-\alpha)\mathbb{E}_t\left[\int_{\epsilon'=\delta z}^{\infty}\frac{\lambda}{\delta z}\left(\frac{\epsilon}{\delta z}\right)^{-(1+\lambda)}\left(\max\left\{H_{t+\tau}(\epsilon') - \tau M, W_{t+\tau}\right\} - W_{t+\tau}\right)\mathrm{d}\epsilon'\right].$$

Simplifying the final term using the definition of $\varDelta \equiv \delta^{\lambda}$:

$$\varPsi_t(z) = \frac{\tau z \xi_t}{\lambda - 1} + \beta\alpha \mathbb{E}_t \varPsi_{t+\tau}(z) + \beta(1-\alpha)\varDelta\left(\frac{z}{x}\right)^{\lambda} \mathbb{E}_t\left[\int_{\epsilon=\delta z}^{\infty} \frac{\lambda}{x}\left(\frac{\epsilon}{x}\right)^{-(1+\lambda)} \max\left\{(H_{t+\tau}(\epsilon) - \tau M - W_{t+\tau}, 0\right\}) \mathrm{d}\epsilon\right],$$

and using the fact that the value function $H_{t+\tau}(\epsilon)$ is increasing in ϵ and the definition of x_t in (3.14), which implies $\tau M + W_{t+\tau} = H_{t+\tau}(x_{t+\tau})$:

$$\varPsi_t(z) = \frac{\tau z \xi_t}{\lambda - 1} + \beta\alpha \mathbb{E}_t \varPsi_{t+\tau}(z) + \beta(1-\alpha)\varDelta\left(\frac{z}{x}\right)^{\lambda} \mathbb{E}_t\left[\int_{\epsilon=x_{t+\tau}}^{\infty} \frac{\lambda}{x}\left(\frac{\epsilon}{x}\right)^{-(1+\lambda)} (H_{t+\tau}(\epsilon) - H_{t+\tau}(x_{t+\tau})) \mathrm{d}\epsilon\right]. \quad \text{[A.5.10]}$$

Note that since $x_{t+\tau} < y + \eta$, equation (A.5.2) can be used to deduce that:

$$H_{t+\tau}(x_{t+\tau}) - H_{t+\tau}(x) = \tau(x_{t+\tau} - x)\sum_{\ell=0}^{\infty}(\beta\alpha)^{\ell}\mathbb{E}_{t+\tau}\xi_{t+\tau(\ell+1)},$$

and together with the definition in (A.5.7), equation (A.5.10) can be written as follows:

$$\begin{aligned}\varPsi_t(z) = {} & \frac{\tau z \xi_t}{\lambda - 1} + \beta\alpha \mathbb{E}_t \varPsi_{t+\tau}(z) + \beta(1-\alpha)\varDelta\left(\frac{z}{x}\right)^{\lambda}\mathbb{E}_t\varPsi_{t+\tau}(x) \\ & - \beta(1-\alpha)\varDelta\left(\frac{z}{x}\right)^{\lambda}\tau\mathbb{E}_t\left[(x_{t+\tau} - x)\sum_{\ell=0}^{\infty}(\beta\alpha)^{\ell}\xi_{t+\tau(\ell+1)}\right] \\ & - \beta(1-\alpha)\varDelta\left(\frac{z}{x}\right)^{\lambda}\mathbb{E}_t\left[\int_{\epsilon=x}^{x_{t+\tau}} \frac{\lambda}{x}\left(\frac{\epsilon}{x}\right)^{-(1+\lambda)}(H_{t+\tau}(\epsilon) - H_t(x_{t+\tau}))\mathrm{d}\epsilon\right]. \quad \text{[A.5.11]}\end{aligned}$$

Log linearizing this equation at $z = y$ (noting that $\varSigma = \varPsi(y)$, and using the expression for $\varSigma$ in (A.3.5) and that for $\varPsi(x)$ in (A.3.12)):

$$\begin{aligned}\Psi_t(y) = {} & \left(\frac{\nu\pi}{\lambda-1}\right)\left(\frac{y}{x+F}\right)\xi_t + \beta\alpha\mathbb{E}_t\Psi_{t+\tau}(y) + \left(\frac{\nu\pi}{\lambda-1}\right)\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\beta(1-\alpha)\varDelta}{1-\beta\sigma_x}\right)\mathbb{E}_t\Psi_{t+\tau}(x) \\ & - \nu\pi\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\beta(1-\alpha)\varDelta}{1-\beta\alpha}\right)\mathbb{E}_t \mathrm{x}_{t+\tau}. \quad \text{[A.5.12]}\end{aligned}$$

Log linearizing equation (A.5.11) now at $z = x$:

$$\Psi_t(x) = (1-\beta\sigma_x)\xi_t + \beta\sigma_x\mathbb{E}_t\Psi_{t+\tau}(x) - (\lambda-1)(1-\beta\sigma_x)\left(\frac{\beta(1-\alpha)\varDelta}{1-\beta\alpha}\right)\mathbb{E}_t \mathrm{x}_{t+\tau}, \quad \text{[A.5.13]}$$

which can be written as:

$$\mathbb{E}_t\left[\left(\frac{\mathbb{I} - \beta\sigma_x\mathbb{F}}{1-\beta\sigma_x}\right)\Psi_t(x)\right] = \xi_t - (\lambda-1)\left(\frac{\beta(1-\alpha)\varDelta}{1-\beta\alpha}\right)\mathbb{E}_t[\mathbb{F}\mathrm{x}_t]. \quad \text{[A.5.14]}$$

By substituting equation (A.5.6) into (A.5.9):

$$\begin{aligned}\mathbb{E}_t\left[(\mathbb{I} - \beta\alpha\mathbb{F})(\mathbb{I} - \beta\sigma_x\mathbb{F})\bar{\Sigma}_t\right] = {} & \mathbb{E}_t\left[(\mathbb{I} - \beta\alpha\mathbb{F})(\mathbb{I} - \beta\sigma_x\mathbb{F})\Psi_t(y)\right] \\ & - \nu\pi\left(\frac{y}{x+F}\right)\mathbb{E}_t\left[\left(1+\frac{F}{x}\right)\left(\frac{x}{y}\right)\beta\alpha\mathbb{F}(\mathbb{I} - \beta\sigma_x\mathbb{F})(\pi_t + \Sigma_t) - (\mathbb{I} - \beta\sigma_x\mathbb{F})\xi_t\right],\end{aligned}$$

and then substituting from (A.5.12):

$$\mathbb{E}_t\left[(\mathbb{I}-\beta\alpha\mathbb{F})(\mathbb{I}-\beta\sigma_x\mathbb{F})\bar{\Sigma}_t\right] = \left(\frac{\nu\pi}{\lambda-1}\right)\left(\frac{y}{x+F}\right)\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\xi_t]$$
$$+\left(\frac{\nu\pi}{\lambda-1}\right)\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t\left[\left(\frac{\mathbb{I}-\beta\sigma_x\mathbb{F}}{1-\beta\sigma_x}\right)\mathbb{F}\Psi_t(x)\right] - \nu\pi\beta\alpha\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\mathbb{F}(\pi_t+\Sigma_t)]$$
$$+\nu\pi\left(\frac{y}{x+F}\right)\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\xi_t] - \nu\pi\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\beta(1-\alpha)\mathit{\Delta}}{1-\beta\alpha}\right)\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\mathbb{F}\mathsf{x}_t].$$

For the next step, substitute from (A.5.14)

$$\mathbb{E}_t\left[(\mathbb{I}-\beta\alpha\mathbb{F})(\mathbb{I}-\beta\sigma_x\mathbb{F})\bar{\Sigma}_t\right] = \nu\pi\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{y}{x+F}\right)\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\xi_t]$$
$$+\left(\frac{\nu\pi}{\lambda-1}\right)\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t[\mathbb{F}\xi_t] - \nu\pi\beta\alpha\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\mathbb{F}(\pi_t+\Sigma_t)]$$
$$-\nu\pi\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\beta(1-\alpha)\mathit{\Delta}}{1-\beta\alpha}\right)\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t[\mathbb{F}^2\mathsf{x}_t]$$
$$-\nu\pi\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\beta(1-\alpha)\mathit{\Delta}}{1-\beta\alpha}\right)\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\mathbb{F}\mathsf{x}_t],$$

and by collecting terms and noting that $\sigma_x-(1-\alpha)\mathit{\Delta}=\alpha$:

$$\mathbb{E}_t\left[(\mathbb{I}-\beta\alpha\mathbb{F})(\mathbb{I}-\beta\sigma_x\mathbb{F})\bar{\Sigma}_t\right] = \nu\pi\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{y}{x+F}\right)\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\xi_t]$$
$$+\left(\frac{\nu\pi}{\lambda-1}\right)\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t[\mathbb{F}\xi_t] - \nu\pi\beta\alpha\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\mathbb{F}(\nu_t+\pi_t+\Sigma_t)]$$
$$-\nu\pi\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t\left[\left(\frac{\mathbb{I}-\beta\alpha\mathbb{F}}{1-\beta\alpha}\right)\mathbb{F}\mathsf{x}_t\right].$$

Finally, substitute equation (A.5.4) into the final term to obtain:

$$\mathbb{E}_t\left[(\mathbb{I}-\beta\alpha\mathbb{F})(\mathbb{I}-\beta\sigma_x\mathbb{F})\bar{\Sigma}_t\right] = \nu\pi\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{y}{x+F}\right)\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\xi_t]$$
$$+\left(\frac{\nu\pi}{\lambda-1}\right)\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t[\mathbb{F}\xi_t] - \nu\pi\beta\alpha\mathbb{E}_t[(\mathbb{I}-\beta\sigma_x\mathbb{F})\mathbb{F}(\pi_t+\Sigma_t)]$$
$$-\nu\pi\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\left(1+\frac{F}{x}\right)\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t[\mathbb{F}(\pi_t+\Sigma_t)] + \nu\pi\left(\frac{x}{x+F}\right)\left(\frac{y}{x}\right)^{\lambda}\beta(1-\alpha)\mathit{\Delta}\mathbb{E}_t[\mathbb{F}\xi_t],$$

and by simplifying this expression:

$$\mathbb{E}_t\left[(\mathbb{I}-\beta\alpha\mathbb{F})(\mathbb{I}-\beta\sigma_x\mathbb{F})\bar{\Sigma}_t\right] + \nu\pi\mathbb{E}_t\left[\left\{\alpha(\mathbb{I}-\beta\sigma_x\mathbb{F}) + (1-\alpha)\mathit{\Delta}\left(\frac{y}{x}\right)^{\lambda}\right\}\beta\mathbb{F}(\pi_t+\Sigma_t)\right] =$$
$$\nu\pi\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{x}{x+F}\right)\mathbb{E}_t\left[\left\{\left(\frac{y}{x}\right)(\mathbb{I}-\beta\sigma_x\mathbb{F}) + \left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\mathit{\Delta}\beta\mathbb{F}\right\}\xi_t\right]. \quad \text{[A.5.15]}$$

Equation (3.4) for the acceptance probability can be log linearized as follows:

$$\pi_t = -\lambda\mathsf{y}_t. \quad \text{[A.5.16]}$$

The equation for the expected surplus implies the following log-linear equation:

$$\bar{\Sigma}_t = \pi_t + \Sigma_t. \quad \text{[A.5.17]}$$

The log linearization of the law of motion (3.15) for houses for sale is:

$$\mathsf{u}_t = \lambda\nu y^{-\lambda}\sum_{i=1}^{\infty}\alpha^i \mathsf{y}_{t-\tau i} + \lambda\nu x^{-\lambda}\left\{(1-\alpha)\Delta\sum_{i=1}^{\infty}\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j}\mathsf{x}_{t-\tau j}\right\} - \nu\sum_{i=1}^{\infty}\left\{\alpha^i y^{-\lambda} + (1-\alpha)\Delta x^{-\lambda}\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j}\right\}\mathsf{u}_{t-\tau i}, \quad \text{[A.5.18]}$$

noting the definitions $\Delta = \delta^\lambda$ and $\sigma_x = \alpha + (1-\alpha)\Delta$. The terms in x_t in this expression can be simplified by a change in the order of summation:

$$\lambda\nu x^{-\lambda}\left\{(1-\alpha)\Delta\sum_{i=1}^{\infty}\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j}\mathsf{x}_{t-\tau j}\right\} = \lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta\left\{\sum_{j=1}^{\infty}\sum_{i=0}^{j-1}\alpha^i\sigma_x^{j-(i+1)}\mathsf{x}_{t-\tau i}\right\}$$
$$= \lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta\sum_{i=0}^{\infty}\alpha^i\left(\sum_{j=i+1}^{\infty}\sigma_x^{j-(i+1)}\right)\mathsf{x}_{t-\tau i} = \lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\Delta\left(\frac{1-\alpha}{1-\sigma_x}\right)\sum_{i=0}^{\infty}\alpha^i\mathsf{x}_{t-\tau i}, \quad \text{[A.5.19]}$$

which uses $\pi = y^{-\lambda}$. The terms in u_t from (A.5.18) can be written in the following simpler form

$$\nu\sum_{i=1}^{\infty}\left\{\alpha^i y^{-\lambda} + (1-\alpha)\Delta x^{-\lambda}\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j}\right\}\mathsf{u}_{t-\tau i}$$
$$= \nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta\sigma_x^{i-1}\sum_{j=0}^{i-1}\left(\frac{\alpha}{\sigma_x}\right)^j\right\}\mathsf{u}_{t-\tau i}$$
$$= \nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta\sigma_x^{i-1}\frac{\sigma_x}{(\alpha+(1-\alpha)\Delta)-\alpha}\frac{(\sigma_x^i-\alpha^i}{\sigma_x^i}\right\}\mathsf{u}_{t-\tau i}$$
$$= \nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i-\alpha^i)\right\}\mathsf{u}_{t-i}. \quad \text{[A.5.20]}$$

Substituting (A.5.19) and (A.5.20) into (A.5.18) implies:

$$\mathsf{u}_t = \lambda\nu\pi\sum_{i=1}^{\infty}\alpha^i\mathsf{y}_{t-\tau i} + \lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\Delta\left(\frac{1-\alpha}{1-\sigma_x}\right)\sum_{i=0}^{\infty}\alpha^i\mathsf{x}_{t-\tau i} - \nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i-\alpha^i)\right\}\mathsf{u}_{t-\tau i}. \quad \text{[A.5.21]}$$

Note the following standard results in terms of the lag operator $\mathbb{L}$:

$$(\mathbb{I}-\alpha\mathbb{L})\sum_{i=1}^{\infty}\alpha^i\mathsf{y}_{t-i} = \alpha\mathbb{L}\mathsf{y}_t;$$
$$(\mathbb{I}-\alpha\mathbb{L})\sum_{i=0}^{\infty}\alpha^i\mathsf{x}_{t-i} = \mathsf{x}_t;$$
$$(\mathbb{I}-\sigma_x\mathbb{L})\sum_{i=1}^{\infty}\sigma_x^i\mathsf{u}_{t-i} = \sigma_x\mathbb{L}\mathsf{u}_t.$$

Multiplying both sides of (A.5.21) by $(\mathbb{I} - \alpha\mathbb{L})(\mathbb{I} - \sigma_x\mathbb{L})$:

$$(\mathbb{I} - \alpha\mathbb{L})(\mathbb{I} - \sigma_x\mathbb{L})\mathsf{u}_t = \lambda\nu\pi(\mathbb{I} - \sigma_x\mathbb{L})\alpha\mathbb{L}\mathsf{y}_t + \lambda\nu\pi\left(\frac{y}{x}\right)^\lambda \varDelta\left(\frac{1-\alpha}{1-\sigma_x}\right)(\mathbb{I} - \sigma_x\mathbb{L})\mathsf{x}_t$$
$$- \nu\pi\left((\mathbb{I} - \sigma_x\mathbb{L})\alpha\mathbb{L} + \left(\frac{y}{x}\right)^\lambda((\mathbb{I} - \alpha\mathbb{L})\sigma_x\mathbb{L} - (\mathbb{I} - \sigma_x\mathbb{L})\alpha\mathbb{L})\right)\mathsf{u}_t.$$

Making some simplifications to this expression yields the following recursive equation for the law of motion:

$$\left\{(\mathbb{I} - \alpha\mathbb{L})(\mathbb{I} - \sigma_x\mathbb{L}) + \nu\pi\left(\alpha(\mathbb{I} - \sigma_x\mathbb{L}) + \left(\frac{y}{x}\right)^\lambda(1-\alpha)\varDelta\mathbb{I}\right)\right\}\mathsf{u}_t$$
$$= \lambda\nu\pi\alpha(\mathbb{I} - \sigma_x\mathbb{L})\mathbb{L}\mathsf{y}_t + \lambda\nu\pi\left(\frac{y}{x}\right)^\lambda \varDelta\left(\frac{1-\alpha}{1-\sigma_x}\right)(\mathbb{I} - \sigma_x\mathbb{L})\mathsf{x}_t.$$

The selling rate is $s_t = \nu\pi_t$, which has the following log-linear form:

$$\mathsf{s}_t = \pi_t. \tag{A.5.22}$$

Total sales are $S_t = s_t u_t$, or in log-linear terms:

$$\mathsf{S}_t = \mathsf{s}_t + \mathsf{u}_t. \tag{A.5.23}$$

The log linearization of the stock-flow accounting identity implies that new listings are given by:

$$\mathsf{N}_t = \frac{1}{s}(\mathsf{u}_t - \mathsf{u}_{t-\tau}) + \mathsf{S}_{t-\tau}. \tag{A.5.24}$$

A.6 Aggregate moving rate is exogenous stochastic process

Consider a version of the exogenous-moving model ($\delta = 0$) where instead of shocks to housing demand ($\xi_t = 1$ replaces equation (3.11)), the exogenous aggregate shock is to the *probability* of receiving an idiosyncratic shock. The arrival rate a of these idiosyncratic shocks (which lead automatically to moving in the exogenous-moving variant of the model) is now time varying according to:

$$\log a_t = (1-\rho)\log a + \rho\log a_{t-1} + \eta_t, \quad \text{where } \eta_t \sim \text{i.i.d.}(0, \varsigma^2), \tag{A.6.1}$$

and where a is the average arrival rate over time. The persistence of this shock (controlled by ρ) is set to 0.965 as in the benchmark model (which matches the persistence of prices).

Comparing the empirical evidence in Table 1 to the results of the model in Table 7, it can be seen that this version of the model does indeed feature listings that have a similar volatility to sales, which is not too far from the data (where listings are approximately 30% more volatile than sales). However, this version of the model struggles to match other key features of the data. The number of houses for sale is now too stable relative to sales (about one third as volatile as sales, compared to approximately equal volatilities in the data). Similarly, time-to-sell is much less volatile according to the model than it is in the data, and the fluctuations in prices are tiny in comparison to those in the data. Other problems include time-to-sell being negatively correlated with homes for sale (contrary to the data) and prices being negatively correlated with sales, listings, and the number of houses for sale (all contrary to the data).

Intuitively, the problem with this version of the model stems from the fact that it introduces entry into the housing market orthogonal to the factors that matter for transactions decisions. Thus, there is no incentive for buyer behaviour to change significantly, so the selling rate and hence time-to-sell do not adjust significantly. Since the selling rate changes by little and because the average time-to-sell is fairly short in relation to the duration of the aggregate shock, sales follow a very similar path to listings, so the stock of houses for sale changes by little.

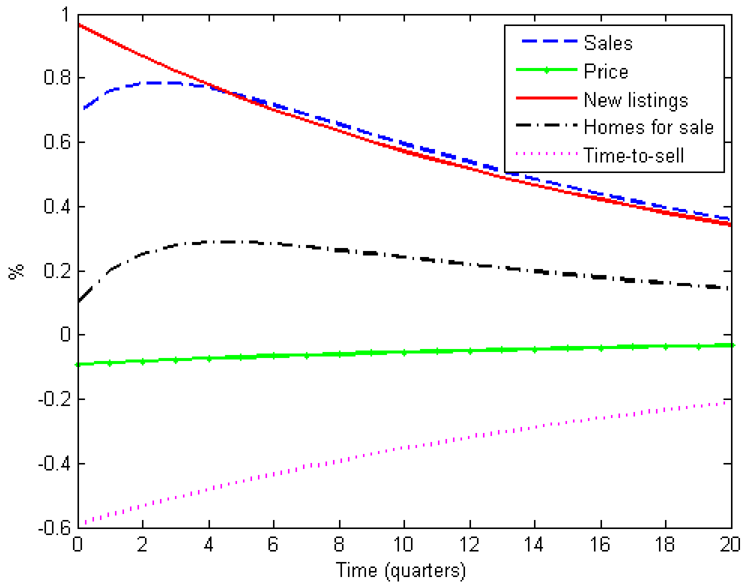
This analysis shows that while a model with exogenous shocks to moving probabilities can generate fluctuations in listings approximately equal in magnitude to fluctuations in sales, such a model does not easily

Table 7: *Exogenous moving, shock to moving probability*

	Sales	Price	New listings	Houses for sale	Time-to-sell
			Relative standard deviations		
Sales	1	0.138	1.003	0.080	0.920
			Correlation coefficients		
Sales	1				
Price	−1.000	1			
New listings	1.000	−1.000	1		
Homes for sale	0.991	−0.989	0.989	1	
Time-to-sell	−1.000	1.000	−1.000	−0.989	1

Notes: Quarterly frequency. Parameters: $\delta = 0$.

Figure 9: *Exogenous moving, impluse response function to a moving rate shock*



Notes: Quarterly frequency. Parameters: $\delta = 0$.

match both the relative volatilities of the number of houses for sale and time-to-sell, and the correlations among these variables.

A.7 Endogenous moving with a less persistent housing demand shock

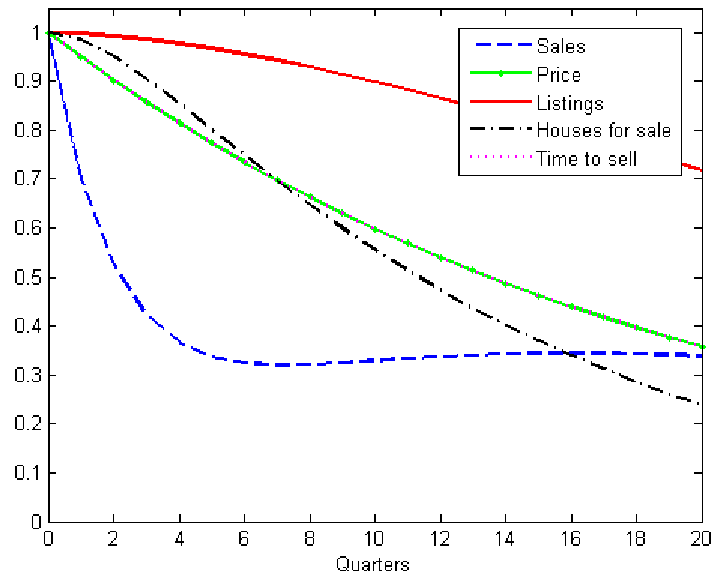
This section reports the results for a less persistent housing demand shock with $\rho = 0.95$. The relative volatilities and correlations are given in Table 8, the autocorrelation functions in Figure 10, and the impulse response functions in Figure 11.

Table 8: *Endogenous moving, less persistent shock to housing demand*

	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	1.655	0.724	2.099	2.721
			Correlation coefficients		
Sales	1				
Price	0.734	1			
New listings	0.557	0.707	1		
Homes for sale	−0.475	−0.946	−0.651	1	
Time-to-sell	−0.734	−1.000	−0.707	0.946	1

Notes: Quarterly frequency. Parameters: $\delta = 0.862$, $\rho = 0.95$.

Figure 10: *Endogenous moving, autocorrelation functions with less persistent shock to housing demand*



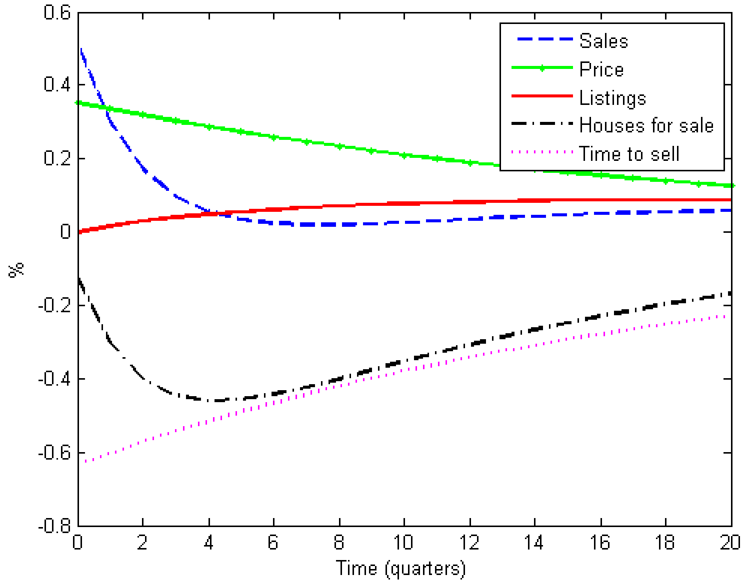
Notes: Quarterly frequency. Parameters: $\delta = 0.862$, $\rho = 0.95$.

A.8 Sellers make take-it-or-leave-it offers

This section reports the results for the case where sellers make take-it-or-leave-it offers, namely $\omega = 1$. The main effects of setting $\omega = 1$ are changes in the cost parameters. A higher bargaining power of sellers implies higher steady-state hosue prices, thus higher costs C, F, and M are needed to match the same cost targets as percentage of prices.

The resulting relative volatilities and correlations are displayed in Table 9, the autocorrelation functions in Figure 12, and the impulse response functions in Figure 13. This variant of the model improves on the predictions for the relative volatilities of prices, listings, and time-to-sell, but does worse for the stock of house for sale. All variables become highly persistent and highly correlated with one another (in some

Figure 11: *Endogenous moving, impluse response functions with a less persistent housing demand shock*



Notes: Quarterly frequency. Parameters: $\delta = 0.862$, $\rho = 0.95$.

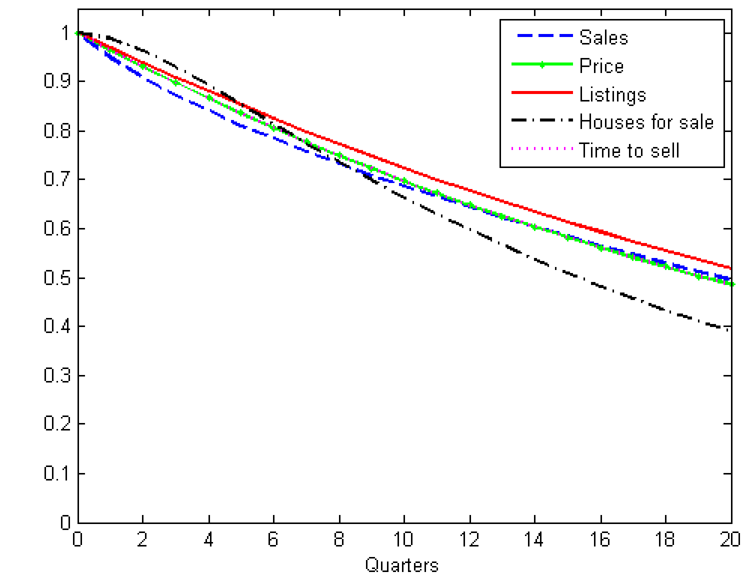
cases, much more than found in the data).

Table 9: *Endogenous moving, shock to housing demand with take-it-or-leave-it offers*

	Sales	Price	Listings	Houses for sale	Time to sell
			Relative standard deviations		
Sales	1	0.929	0.979	0.228	1.215
			Correlation coefficients		
Sales	1				
Price	0.998	1			
Listings	0.996	0.999	1		
Houses for sale	−0.929	−0.953	−0.952	1	
Time to sell	−0.998	−1.000	−0.999	0.953	1

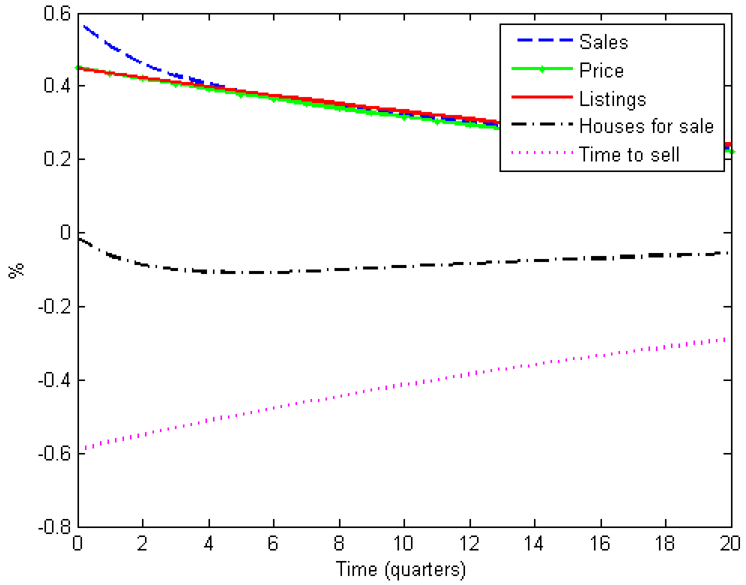
Notes: Quarterly frequency. Parameters: $\omega = 1$.

Figure 12: *Endogenous moving, autocorrelation functions, shock to housing demand with take-it-or-leave-it offers*



Notes: Quarterly frequency. Parameters: $\omega = 1$.

Figure 13: *Endogenous moving, impluse response functions to a housing demand shock with take-it-or-leave-it offers*



Notes: Quarterly frequency. Parameters: $\omega = 1$.