The Ins and Outs of Selling Houses*

L. Rachel Ngai†
London School of Economics

Kevin D. Sheedy
London School of Economics

First draft: 14th December 2012
This version: 4th February 2015

Abstract

The number of houses for sale is as volatile as sales volume and much more volatile than house prices, yet it has received relatively little attention. What drives volatility in the number of houses for sale? Is it due to changes in the difficulty of selling houses or changes in the incentive to put houses up for sale? This paper presents evidence that both inflows and outflows are important using a variance decomposition. It then uses a search-and-matching model with both the decision of when to agree a sale (outflows) and the decision of when to put a house up for sale (inflows) to understand the behaviour of sales, listings, and prices in the housing market. Quantitatively, the model does a much better job of matching relative volatility and correlations between housing-market variables than those that abstract from the inflow decision.

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

*We thank Morris Davis, Martin Gervais, Lu Han, Allen Head, and Chris Pissarides 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 and the 2013 Spring Housing-Urban-Labor-Macro conference for their comments.

†Email: l.ngai@lse.ac.uk

1 Introduction

There is a growing literature that has adopted the idea of modelling the housing market as subject to search frictions similar to those typically assumed in the labour market. This is natural because there are many features shared by the housing and labour markets. To name but a few of these, there is substantial heterogeneity (in houses and in preferences for houses, analogous to the heterogeneity in jobs and workers in the labour market), and there is the coexistence of buyers and sellers who invest significant amounts of time and resources before a successful transaction can occur.

The first contribution of this paper is to document a set of variables that gives a better understanding of the housing market, very much in line with the approach that has been taken in the labour literature. Using the price index from the Federal Housing Finance Agency (FHFA) and sales and inventory data for existing homes from the National Association of Realtors (NAR), this paper constructs time series for prices, sales, new listings, the number of houses for sale, and the average time taken to sell for the period 1991–2012. Consistent with existing findings, there is a positive correlation between prices and sales and a negative correlation between sales and time-to-sell. These two correlations have received much attention and models have been developed to understand them. They have led research on housing to highlight the importance of search frictions and to focus on the decision of when potential buyers and sellers should stop searching (that is, when sales occur).

By taking a more complete look at the housing-market variables, this paper uncovers two overlooked facts. First, despite being a stock variable, the number of house for sale is as volatile as sales and much more volatile than prices, yet the literature has focused mainly on sales and prices. To put it in the language of labour-market variables, much attention has been paid to the behaviour of the job-finding rate (sales) and too little to the unemployment rate (number of house for sale). Second, the change in the number of houses for sale is significantly more volatile than outflows as represented by sales, which indicates there is substantial variation in inflows to the housing market. It is natural, therefore, to document more precisely the contributions of the inflow and outflow rates to the change in the stock of houses for sale. The underlying economic issue is to understand what drives fluctuations in the number of houses for sale. Is it due to changes in the difficulty of selling houses, or changes in the number of houses that are put up for sale?

An inflows and outflows decomposition of the change in the numbers of houses for sale provides an answer to that question. The method follows recent labour-market research in reassessing the 'ins and outs' of unemployment initiated by Shimer (2012) (see also Petrongolo and Pissarides, 2008, Fujita and Ramey, 2009). The decomposition presented here shows that volatility in inflow rates is as important as volatility in outflow rates in accounting for volatility in the number of houses for sale: about 50% of the volatility in houses for sale is attributed to inflows during the period 1991–2012.

The 'ins and outs' decomposition demonstrates that the decision of when to put a house up for sale is as important, if not more so, than the decision of when to stop searching for a house to buy. This feature of the data raises a number of questions, such as how to develop a model of the housing market that can be used to understand changes in inflows, unlike majority of existing models that

have treated the inflow rate as an exogenous variable. There is then the question of whether the endogeneity of inflows significantly alters the responses of other housing-market variables to shocks when compared to an analysis where the inflow rate is assumed to be exogenous.

In addressing these questions, the second contribution of this paper is to apply the endogenous moving model of Ngai and Sheedy (2014) to study the housing market dynamics. In that model, both inflows and outflows are endogenous. It is shown that a calibrated version of the model is consistent with the finding that inflows into the housing market have a volatility comparable to that of outflows from the market. Moreover, as the model nests the special cases where the inflow rate is an exogenous constant, and where the inflow rate is subject to exogenous time variation, its empirical performance can be compared against the assumption of exogenous inflows. 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 some 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.

The analysis thus points towards the importance of modelling inflows into the housing market as an endogenous decision variable. The paper shows furthermore that endogenizing the moving decision can markedly change the dynamics of the housing market (beyond simply the inflow rate) in response to shocks. Compared to an otherwise identical model with an exogenous inflow rate, the model can predict 'over-shooting', where variables adjust by more in the short run than their ultimate long-run equilibrium, and even the presence of oscillations in some cases. These findings indicate that modelling inflows as exogenous is far from an innocuous simplifying assumption.

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.[1]

[1]The introduction of an endogenous moving decision is analogous to the endogenous job separation decision introduced in Mortensen and Pissarides (1994). However, the modelling of the 'separation' decision is different here. First, in Mortensen and Pissarides (1994), when a match is subject to an idiosyncratic shock, a new match quality is drawn independently of the match quality before the shock (the stochastic process is 'memory-less'). Here, idiosyncratic shocks degrade match quality, but an initially higher quality match remains of a higher quality than a

The model thus 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.

Finally, the paper allows for the possibility that the distribution of new match quality may improve when there are more houses for sale. [2] In the presence of the thick-market effect, the endogenous moving decision described above generates an additional effect that can lead to oscillations in the housing market following an aggregate shock. The reason is that the overshooting caused by endogenous moving leads homeowners with somewhat higher match quality to bring forward moving to coincide with those who remained following the lower degree of cleansing earlier on. This 'preemption effect' then leads to more cleansing of the future distribution of match quality than would otherwise occur, putting the process just described into reverse, and resulting in a bunching of activity in the housing market.[3]

The parameters of the model are calibrated to match the steady state to features of the U.S. economy. The model is then simulated following an aggregate housing demand shock that increases the utility value of owning a home. The model with the endogenous moving decision does a good job in matching the relative volatilities and correlations found in the data between sales (outflows), listings (inflows), the number of houses for sale, and the time taken to sell. The model performs best when a moderate thick-market effect is also assumed. When compared to a benchmark model with an exogenous moving decision, the endogenous moving model performs better not only in terms of the behaviour of inflows, but also in terms of the relative volatilities and correlations among the other variables. Furthermore, it is shown that a model with exogenous aggregate shocks to inflows does not perform as well in terms of these relative volatilities and correlations.

There is a strand of the literature (starting from Wheaton, 1990, and followed by many others) that studies frictions in the housing market as here with a search-and-matching model.[4] The key contribution of this paper to the literature is in studying housing-market inflows.[5] In exploring

low-quality match hit by the same shock. Second, here for simplicity, only those who receive an idiosyncratic shock make an endogenous moving decision. This allows for a tractable model of endogenous separation from a match, but one in which the distribution of the quality of existing matches plays an important role.

[2]This type of thick-market effect refers solely to an improvement in match quality, not a rise in the arrival rate of viewings. The idea that this quality effect is present in the housing market was discussed in Goodman (1993) and has been used to understand seasonal fluctuations in the housing market by Ngai and Tenreyro (2014). The modelling approach goes back to Diamond (1981) in the labour market, who assumes the distribution of mobility costs is stochastically lower when there are more firms and workers.

[3]This bunching phenomenon was also discussed in Goodman (1993).

[4]See, for example, Albrecht, Anderson, Smith and Vroman (2007), Caplin and Leahy (2011), Coles and Smith (1998), Díaz and Jerez (2012), Head, Lloyd-Ellis and Sun (2012), Krainer (2001), Ngai and Tenreyro (2014), Novy-Marx (2009), and Piazzesi and Schneider (2009).

[5]Another difference is in the modelling of the nature of the search frictions. A popular approach is to assume the search friction is the difficulty of buyers and sellers to meet each other, often modelled using a matching function

the cyclical behaviour of the housing market,[6] Díaz and Jerez (2012) 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. However, different from that paper, here the focus is on the endogenous moving decision and the search friction due to the difficulty of knowing the quality of a potential match before a viewing has occurred.[7]

The plan of the paper is as follows. Section 2 documents the data and performs the decomposition of changes in the number of houses for sale 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.[8] 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).[9] 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

As in the labour market, the mere persistence of an inventory of houses for sale and a group of potential buyers indicates the presence of search frictions in the housing market. Exploring the role

as in the canonical labour-search model (Pissarides, 2000, chapter 1), and highlighting the role of market tightness (buyer-to-seller ratio). 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.

[6] 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 (2012). For the role of home production, see recent work by Aruoba, Davis and Wright (2012).

[7] Moreover, as discussed further when the data is presented below, this paper assembles data on sales, the number of houses for sale, and time-to-sell, from the same consistent data source rather than the three different sources used by Díaz and Jerez (2012).

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

[9] Methodology and recent data are available at `http://www.realtor.org/research-and-statistics/housing-statistics`.

of search frictions involves documenting the time taken for transactions to occur (in other words, the time spent on the market, or the time-to-sell and the time-to-buy). Constructing a time series for the time-to-sell for an average house requires a time series for the number of houses available on the market. NAR provides a measure of 'months supply', 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. However, the inventories number ignores houses that come on to the market and those get sold during the month, thus it is not an exact measure of the actual number of houses on the market.

To address this issue, suppose that both sales and new listings occur uniformly during a month. This implies that a correct measure of the number of houses available for sale would be the average of the inventories at the end of the previous month and the current month. To see this formally, 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).

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. The mean is 6.4 months, compared to 6.6 for 'months supply', and the correlation coefficient is 0.99.

To summarize, using non-seasonally adjusted monthly data on price, 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. To compare the implied cyclical properties with those found by Díaz and Jerez (2012), the monthly data are detrended using the Hodrick-Prescott filter (with parameter 14400) and then 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 trends and 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. The statistics related to sales, prices, houses for sale, and time-to-sell are similar to those reported in Díaz and Jerez (2012). As is well known in the literature, there is a positive comovement between house prices and sales with a correlation of 0.243, and a negative correlation between time-to-sell and sales of

−0.543, and between time-to-sell and price of −0.230.[10] Perhaps more interesting is that there is also a mild positive correlation between prices and houses for sale, with the correlation coefficient equal to 0.089. Díaz and Jerez (2012) also document a similar positive correlation between price and vacancies, where vacancies are measured by 'vacant for sale' from the U.S. Census Bureau Housing Vacancy Survey.[11] Note that the measure of 'vacant for sale' includes only a small fraction of houses that are actually for sale because it excludes houses that are occupied but available for sale.[12] The positive correlation appears puzzling as it might be natural to expect a negative relationship between price and inventories (or vacancies).[13] In addition to the statistics reported in Díaz and Jerez (2012), the results here show that new listings are more volatile than sales, positivly correlated with sales and price, and negatively correlated with time-to-sell.

Table 1: *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.

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 data reveal that the decision to enter the market is also very important. Another way of seeing this is to plot time series for the level of sales and the decrease in inventories as in Figure 1. The decrease in inventories is significantly more volatile than sales (outflows from the market), which indicates there is substantial variation in inflows to the housing market according to the accounting equation [2.1].

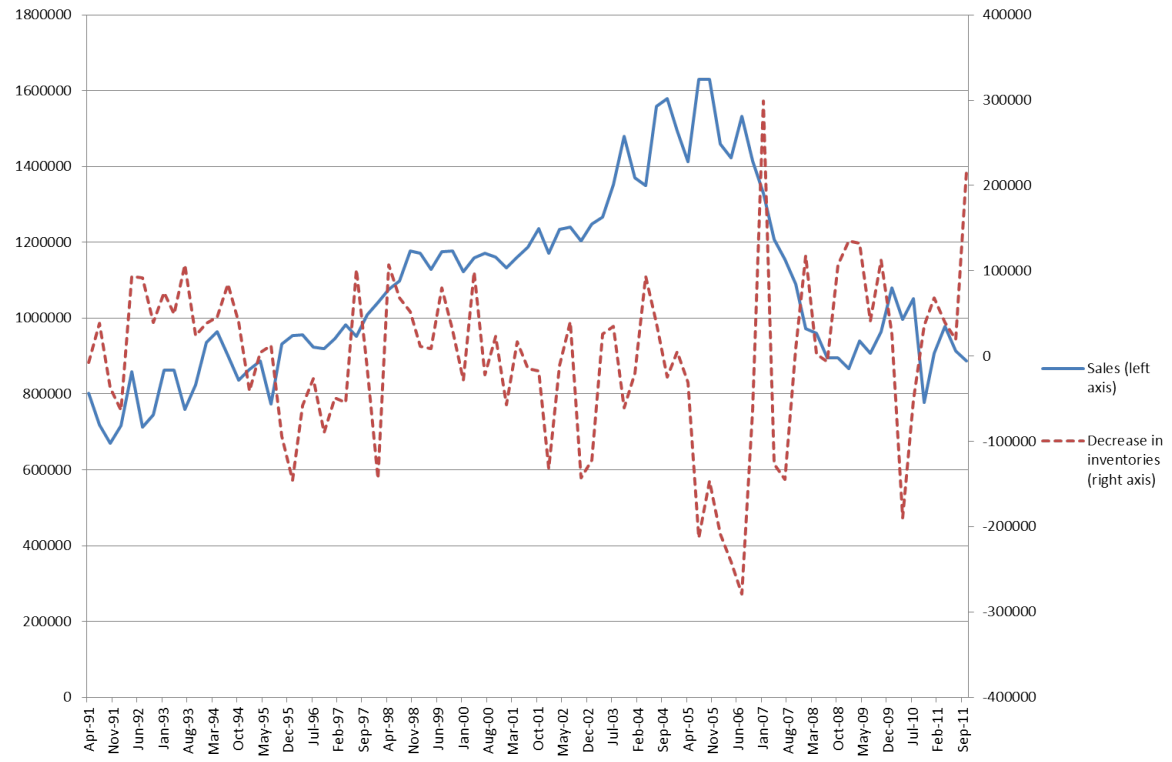
[10] Díaz and Jerez (2012) and Ungerer (2012) also find negative correlations, but using a measure of time-to-sell based only on newly built homes ('New Residential Sales' from the U.S. Census). In micro data on completed transactions, where time-to-sell can be measured directly, a positive correlation between time-to-sell and price has been found in the cross-section of sellers (Genesove and Mayer, 1997, Merlo and Ortalo-Magné, 2004, Levitt and Syverson, 2008).

[11] Using the same 'vacant for sale' index, Caplin and Leahy (2011) find that vacancies have a negative correlation with subsequent price changes.

[12] According to Table 1 of NAR's methodology documentation, vacant homes sold are approximately 11% of all single-family homes sold.

[13] This is not only a puzzle in a standard supply and demand framework. Díaz and Jerez (2012) find it remains a puzzle even in a search model with variation over time in market tightness (defined as the ratio of buyers to the number of houses for sale).

Figure 1: *Sales and the decrease in inventories*



Notes: Monthly data, seasonally adjusted, converted to quarterly series. The decrease in inventories (the change in inventories multiplied by minus one) is plotted because this would move with the sales series in the absence of fluctuations in inflows.
Source: National Association of Realtors

It is natural, therefore, to document more precisely the contributions of the inflow and outflow rates to the change in the stock of houses for sale. The underlying economic issue is to understand what drives fluctuations in the number of houses for sale. Is it due to 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

To address this question, the change in the number of houses for sale is decomposed into its inflow and outflow components. The approach taken here follows closely that of Petrongolo and Pissarides (2008) (and see also, Fujita and Ramey, 2009). 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 in or out of the rental market.[14]

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

[14]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 represent only a small fraction of total sales in the data. According to the 'U.S. Housing Market Conditions' report[15], in 2010 new homes sold amounted to only 6% of the number of existing homes sold (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 from 1991 using the Census Bureau's 'New Residential Sales' data to measure the number of new homes.[16] Moreover, according to U.S. Census Bureau data, 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). This implies that the housing stock per person was approximately constant for 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. This evidence suggests that the 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:

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

and

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

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 time interval $[t, t+\tau)$, equations

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

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

[2.3a] and [2.3b] can be simplified to

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

and

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

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 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 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.[17] 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.[18] Finally, houses owned by people older than 65 are excluded, which removes a further 19% of the housing stock.[19] 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 variation in the fraction of inventories, that is, $i_t = I_t/K$. The law of

[17] 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.

[18] 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.

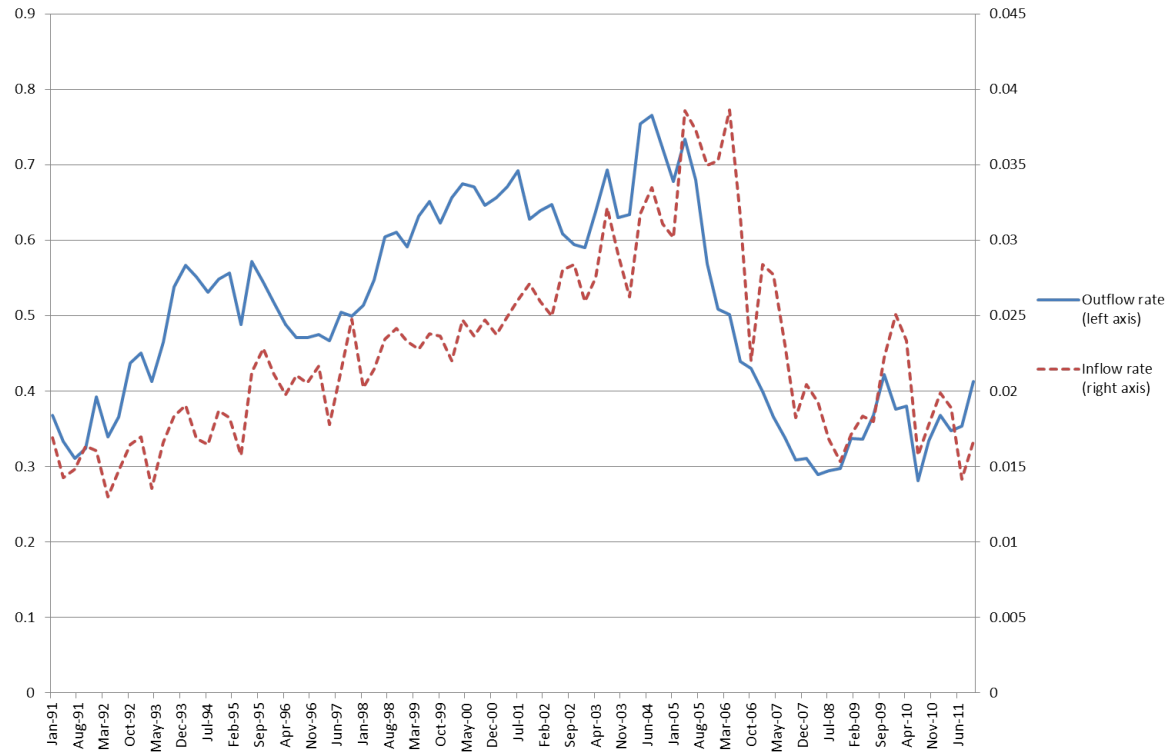
[19] This number is obtained using Table 2.9 of the AHS survey.

motion for i_t 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 inflows and outflows are equal to 0.27 and 0.21 respectively, suggesting that both are relevant for understanding fluctuations in the number of houses for sale.

Figure 2: *Inflow and outflow rates*



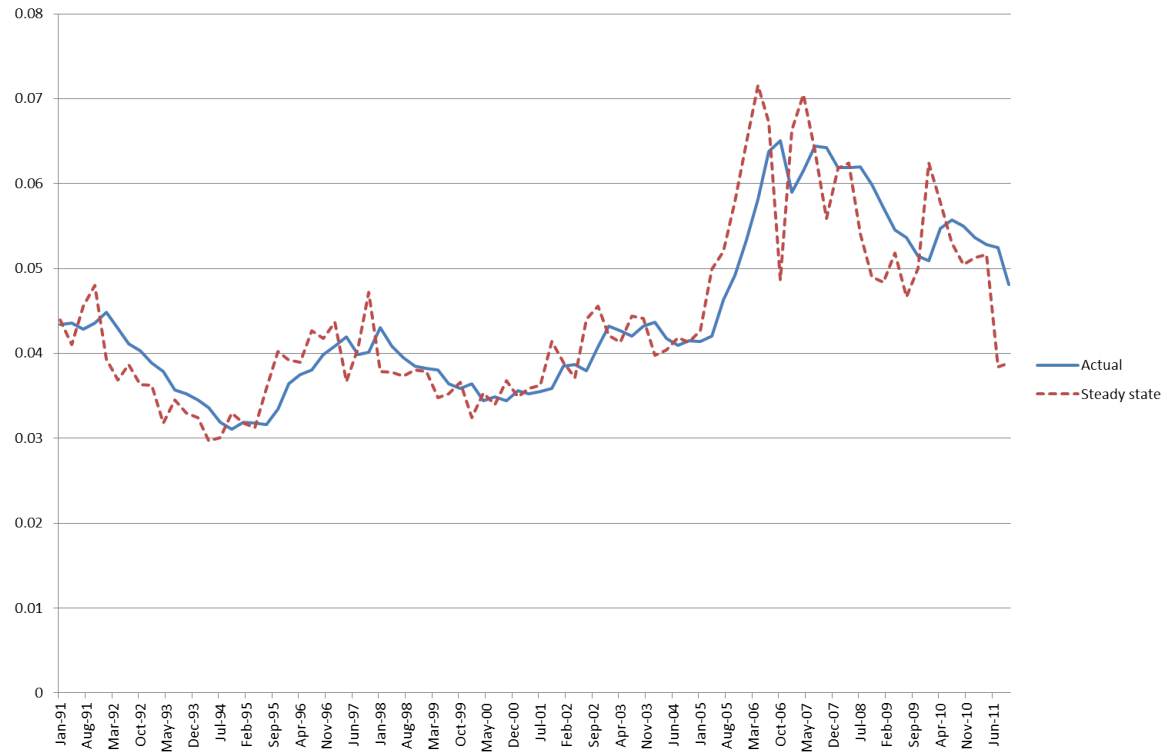
Notes: Monthly data, seasonally adjusted, converted to quarterly series. The outflow and inflow rates are calculated using equations [2.4a] and [2.4b].
Source: National Association of Realtors

The method for decomposing variation in the number of houses for sale follows the procedure used in the labour literature of approximating the unemployment rate by the steady state implied by the inflow and outflow rates. The justification is that the flow rates are sufficiently large, implying that the steady state 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 fraction of houses for sale (the equivalent of the unemployment rate). The inflow and outflow rates n_t and s_t imply the following steady-state fraction of inventories i_t^*:

$$i_t^* = \frac{n_t}{n_t + s_t}. \qquad [2.6]$$

Figure 3 plots the time series for i_t and i_t^*. These variables have the same mean and a correlation coefficient equal to 0.86. The graph indicates that i_t remains fairly close to i_t^*, so it is possible to apply the decomposition developed in the labour literature to the housing market.

Figure 3: *Homes for sale as a fraction of the total stock*



Notes: Monthly data, seasonally adjusted, converted to quarterly series. The steady state is calculated using equation [2.6].
Source: National Association of Realtors

Formally, the decomposition is given by:

$$\Delta i_t^* = \Delta i_{n,t}^* + \Delta i_{s,t}^*, \quad \text{where} \quad \Delta i_{n,t}^* \equiv \frac{\Delta n_t}{n_{t-1}}(1 - i_t^*)i_{t-1}^*, \quad \text{and} \quad \Delta i_{s,t}^* \equiv -\frac{\Delta s_t}{s_{t-1}}(1 - i_{t-1}^*)i_t^*. \quad [2.7]$$

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

$$\beta_j = \frac{\text{Cov}\left(\Delta i_t^*, \Delta i_{j,t}^*\right)}{\text{Var}\left(\Delta i_t^*\right)}, \quad j = \{n, s\}.$$

The results are that $\beta_n = 0.91$ and $\beta_s = 0.09$. Since the method 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 Δi_t and Δi_t^* as a fraction of i_t is less than 10%.[20] This leads to $\beta_n = 0.51$ and $\beta_s = 0.49$.

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

In summary, this investigation of the data suggests that fluctuations in the number of houses that are put up for sale is as important, if not more important, in understanding overall housing market fluctuations than fluctuations in the number of houses sold. This is the first contribution of the paper. The second contribution is to build a model with endogenous inflows as well as outflows, in contrast to other models in the housing literature where the inflow decision is often taken as exogenous.

3 Model

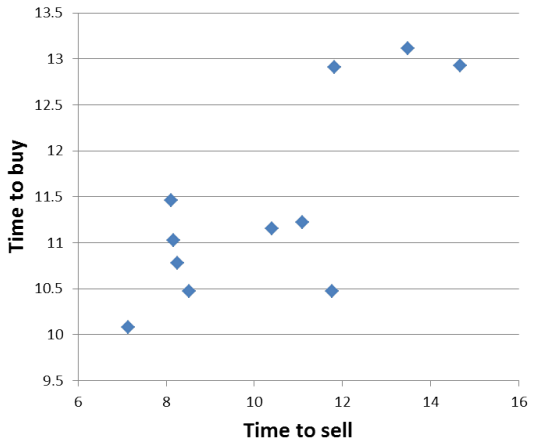
This section presents a discrete time version of the endogenous moving model of Ngai and Sheedy (2014) that studies both the decisions to buy and sell and the decision to move house. As in the analysis of the data in section 2, the model focuses on the market for existing homes. This abstracts from new 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.

There are broadly two kinds of search frictions in the housing market: the difficulty of buyers and sellers meeting each other, and the difficulty for buyers of knowing which properties would be a good match prior to viewing them. The first friction might be modelled using a meeting function. With a stable constant-returns meeting function, this friction would not be time varying unless there were changes in the ratio of buyers to sellers in the market. However, 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 this ratio would not be the dominant factor in understanding the housing market. Furthermore, if the ratio of buyers and sellers 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 is limited data on time-to-buy, but Genesove and Han (2012) have compiled some using the 'Profile of Buyers and Sellers' surveys from NAR for various years from 1987–2007. Figure 4 shows a scatterplot of time-to-buy and time-to-sell for these years. The positive relationship in the figure suggests that fluctuations in 'market tightness' alone do not provide a complete picture of the housing market.

The second friction relates to the number of properties that buyers would need to view before an ideal property is found. The ideal property for a family is not easily determined simply by knowing objective features such as the number of bedrooms. What is meant by ideal is a good match between the idiosyncratic preferences of the buyer and the idiosyncratic characteristics of the house for sale. This type of search friction can be modelled as a stochastic match-specific quality that only becomes known to a buyer when a house is actually viewed.

A measure of the importance of the second friction is the average number of viewings needed before a house can be sold (or equivalently, before a buyer can make a purchase), referred to here as viewings-per-transction. Genesove and Han (2012) also report data on the number of homes visited, which is plotted against the time-to-sell data in the left panel of Figure 5. In the UK, monthly data on time-to-sell and viewings-per-sale are available from the Hometrack 'National Housing

Figure 4: *Time-to-buy and time-to-sell*



Notes: Annual data (years in sample: 1987, 1989, 1991, 1993, 1995, 2001, and 2003–2007). Time on the market is measured in weeks.
Source: Genesove and Han (2012)

Survey'.[21] A scatterplot of this data is shown in the right panel of Figure 5. Both scatterplots illustrate 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.[22] Viewings-per-transaction can fluctuate either because buyers' threshold for an acceptable house changes (in terms of its match-specific quality) or the range of available houses changes (the distribution of match-specific quality available). The model explores the implications of both of these. To allow for the second possibility, a 'thick-market effect' is introduced where the distribution of match-specific quality is assumed to be stochastically better when there are more houses available for sale.

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.[23] 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.

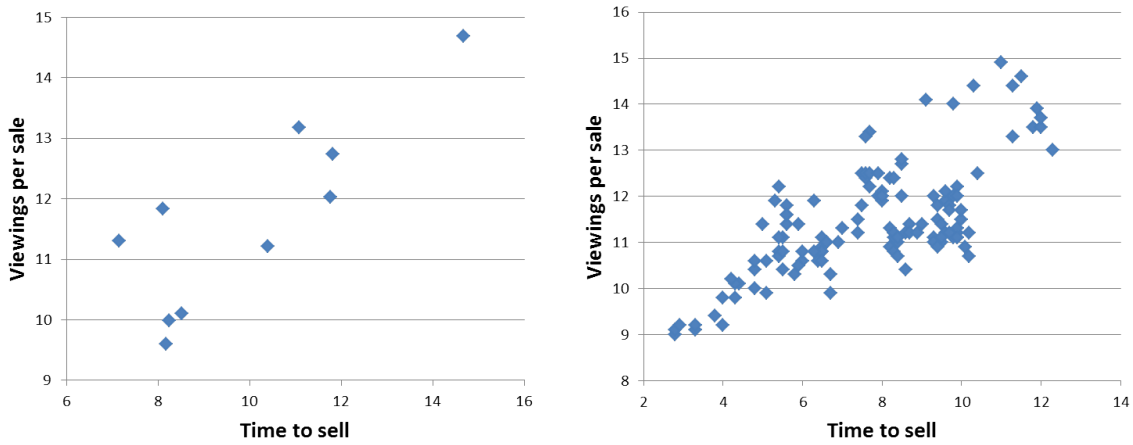
Families discount the future at (continuous) rate $r > 0$. Time is indexed by t, and families make

[21] 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.

[22] It is important to note that viewings-per-transaction are far greater than one, indicating that there is substantial uncertainty about match quality prior to a viewing.

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

Figure 5: *Viewings-per-transaction and time-to-sell*



Notes: Left panel, U.S. data, annual frequency (years in sample: 1989, 1991, 1993, 1995, 2001, and 2003–2007); Right panel, U.K. data, monthly frequency (March 2002–October 2012). Time-to-sell is measured in weeks.
Source: U.S. data, Genesove and Han (2012); U.K. data, Hometrack (`www.hometrack.co.uk`)

decisions at discrete time intervals τ. All units of time are measured in years throughout.

3.2 Search behaviour

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 $[t, t+\tau)$.[24] For houses, this event has Poisson arrival rate $\mathcal{V}(u_t, b_t)/u_t$, where $\mathcal{V}(u, b)$ is a 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. There are no fluctuations in 'market tightness' (b_t/u_t). This differentiates the analysis from papers that have made market tightness central to studying housing-market fluctuations (Novy-Marx, 2009, Díaz and Jerez, 2012).

The second aspect of the search frictions in the housing market is the heterogeneity in buyer

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

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.[25] When a viewing takes place, match quality ϵ is drawn from the probability distribution

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

The Pareto distribution is chosen for analytical tractability. The distribution may change over time in line with market conditions, as explained below. Such changes are captured by variation in the shape parameter λ_t (which satisfies $\lambda_t > 1$). The minimum value of ϵ is normalized to one.

The distribution of new draws of match quality is assumed to be affected by the measure u_t of properties on the market. The assumption is that more properties on the market (and more buyers) make high-quality matches more likely. In particular, higher u_t implies the distribution of ϵ first-order stochastically dominates that for a lower u_t. For the Pareto distribution, first-order stochastic dominance is equivalent to a lower value of the parameter λ_t. Formally, λ_t is determined by

$$\frac{\lambda_t}{\lambda_t - 1} = \Lambda u_t^{\theta}, \tag{3.2}$$

for parameters Λ (average match quality independent of the number of houses for sale) and θ (controlling the sensitivity of match quality to the number of houses for sale), where $\lambda_t/(\lambda_t - 1)$ is the expected value of ϵ at time t. The strength of the thick-market effect is increasing in θ, where the value of this parameter can be interpreted as the elasticity of expected match quality with respect to the stock of houses for sale. It is not essential to the model that θ is positive. When θ is zero, there is no thick-market effect and λ_t is a constant.

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 resulting from a transaction is given by

$$\Sigma_t(\epsilon) = H_t(\epsilon) - \beta \mathbb{E}_t W_{t+\tau} - C, \tag{3.3}$$

where $W_t = B_t + U_t$ denotes the combined value of being a buyer and having a house for sale. Since

[25] 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.

$H_t(\epsilon)$ 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 $\Sigma_t(y_t) = 0$. 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.[26] The transaction threshold y_t satisfies the following equation:

$$H_t(y_t) = \beta \mathbb{E}_t W_{t+\tau} + C. \tag{3.4}$$

Using the distribution of ϵ in [3.1], the proportion of viewings that lead to transactions is π_t:

$$\pi_t = y_t^{-\lambda_t}. \tag{3.5}$$

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 Σ_t denote the average surplus from viewings that lead to sales:

$$\Sigma_t = \int_{\epsilon=y_t}^{\infty} \frac{\lambda_t}{y_t} \left(\frac{\epsilon}{y_t}\right)^{-(\lambda_t+1)} \Sigma_t(\epsilon) \mathrm{d}\epsilon. \tag{3.6}$$

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

$$W_t = -\tau(F + M) + \nu \pi_t \Sigma_t + \beta \mathbb{E}_t W_{t+\tau}, \tag{3.7}$$

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.[27]

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:

$$\Sigma_{\mathrm{b},t}(\epsilon) = H_t(\epsilon) - \beta \mathbb{E}_t B_{t+\tau} - p_t(\epsilon) - (1-\kappa)C, \quad \text{and} \quad \Sigma_{\mathrm{u},t}(\epsilon) = p_t(\epsilon) - \beta \mathbb{E}_t U_{t+\tau} - \kappa C, \tag{3.8}$$

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

$$B_t = -\tau F + \beta \mathbb{E}_t B_{t+\tau} + \nu \int_{\epsilon=y_t}^{\infty} \lambda_t \epsilon^{-(\lambda_t+1)} \Sigma_{\mathrm{b},t}(\epsilon) \mathrm{d}\epsilon, \quad \text{and} \tag{3.9a}$$

$$U_t = -\tau M + \beta \mathbb{E}_t U_{t+\tau} + \nu \int_{\epsilon=y_t}^{\infty} \lambda_t \epsilon^{-(\lambda_t+1)} \Sigma_{\mathrm{u},t}(\epsilon) \mathrm{d}\epsilon. \tag{3.9b}$$

The Nash bargaining solution (with bargaining power ω of the seller) implies the surplus-splitting equation $(1-\omega)\Sigma_{\mathrm{u},t}(\epsilon) = \omega \Sigma_{\mathrm{b},t}(\epsilon)$, which determines the transaction price. Hence, using the surplus

[26] 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.

[27] The flow cost also enters the value of being a homeowner, which appears in the surplus Σ_t.

equations [3.8] and the value functions [3.9a] and [3.9b], the Appendix derives the following formula for the transaction price for a house with match quaity ϵ to its buyer:

$$p_t(\epsilon) = \omega H_t(\epsilon) + (\kappa - \omega)C + \beta \left(\frac{\tau}{1-\beta} \right) (\omega F - (1-\omega)M). \qquad [3.10]$$

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

$$P_t = \omega \int_{\epsilon=y_t}^{\infty} \frac{\lambda_t}{y_t} \left(\frac{\epsilon}{y_t} \right)^{-(\lambda_t+1)} H_t(\epsilon) d\epsilon + (\kappa - \omega)C + \beta \left(\frac{\tau}{1-\beta} \right) (\omega F - (1-\omega)M). \qquad [3.11]$$

3.3 Behaviour of owner-occupier

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 $\{\xi_t\}$ is the exogenous economy-wide housing demand level, modelled as a change in the utility value of housing to all owner-occupiers. All homeowners 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 stochastic process:

$$\log \xi_t = \rho \log \xi_{t-\tau} + \eta_t, \quad \text{where } \eta_t \sim \text{i.i.d.}(0, \varsigma^2). \qquad [3.12]$$

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

$$\epsilon' = \begin{cases} \epsilon & \text{w.p. } \alpha \\ \delta\epsilon & \text{w.p. } 1-\alpha \end{cases}, \qquad [3.13]$$

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:

$$H_t(\epsilon) = \tau\epsilon\xi_t + \beta\alpha\mathbb{E}_t[H_{t+\tau}(\epsilon) - \tau M] + \beta(1-\alpha)\mathbb{E}_t \max\{H_{t+\tau}(\delta\epsilon) - \tau M, W_{t+\tau}\}. \quad [3.14]$$

Thus 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:

$$H_t(x_t) = W_t + \tau M. \quad [3.15]$$

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.

3.4 Laws of motion

In what follows, it will be assumed that the idiosyncratic shocks are sufficiently large (δ in [3.13] is low) relative to the aggregate shocks (ς in [3.12] 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 $1 - u_t$. These matches can be characterized by a vintage $i = 1, 2, \ldots$. Vintage i matches were all originally drawn from the match quality distribution $\epsilon \sim \text{Pareto}(1, \lambda_{t-i})$. The density function is $\lambda_{t-i}\epsilon^{-(\lambda_{t-i}+1)}$ and the probability that a draw is greater than ϵ is $\epsilon^{-\lambda_{t-i}}$. Back at time $t-i$, 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 νu_{t-i} of such draws.

Now consider the cohort of potential matches from draws of ϵ at time $t - i$. What must be determined is how many of these matches formed at time $t - i$ and survived through to time t. For these matches, let j denote the time 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 $t - i$, 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 $\epsilon \geq y_{t-i}$:

$$y_{t-i}^{-\lambda_{t-i}}.$$

In the case that shocks have been received, a draw from the original distribution is a surviving match at time t if $\epsilon' \geq x_{t-j}$, where j is the time since the last shock, and $\epsilon \geq y_{t-i}$, 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 $\epsilon' \geq x_{t-j}$. Given the link between ϵ' and ϵ, this is equivalent to $\epsilon \geq \delta^{-(k+1)} x_{t-j}$, where ϵ is the original draw of match quality. The Pareto distribution at time $t-i$ implies this probability is

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

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

$$u_t = 1 - \sum_{i=1}^{\infty} \left\{ \alpha^i y_{t-i}^{-\lambda_{t-i}} + \sum_{j=0}^{i-1} (1-\alpha)\alpha^j \sum_{k=0}^{i-1-j} \frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k \alpha^{i-1-j-k} (\delta^{-(k+1)} x_{t-j})^{-\lambda_{t-i}} \right\} \nu u_{t-i}. \quad [3.16]$$

3.5 Solution: No aggregate shocks

First consider the solution of the model in the case of no aggregate shocks (where $\xi_t = 1$ in the limiting case of $\varsigma \to 0$ in [3.12]). 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 detail in appendix A.1, 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.

Conditional on the value of λ describing the distribution of match quality for new buyers, the moving threshold x and the transaction threshold y are determined by the pair of equations:

$$x + F = \left(\frac{\nu}{\tau}\right)\left(\frac{1}{\lambda - 1}\right)\left(\frac{\tau}{1 - \beta\alpha}\right)\left(y^{1-\lambda} + \beta\delta^{\lambda}\left(\frac{1-\alpha}{1 - \beta(\alpha + (1-\alpha)\delta^{\lambda})}\right)x^{1-\lambda}\right), \qquad [3.17a]$$

$$y = \beta\alpha x + \left(\frac{1 - \beta\alpha}{\tau}\right)C. \qquad [3.17b]$$

The value of x can be found by substituting the linear expression for y into the first equation. Conditional on λ, any solution of equations [3.17a] and [3.17b] must be unique (existence can be established under weak conditions described in appendix A.1). If the thick-market parameter θ were sufficiently large it would be possible to find multiple equilibria for x, y, and λ. Throughout this paper it is assumed that thick-market effects are not so large as to generate multiple steady states. In this case, given equation [3.2], rather than specify parameters Λ and θ to describe the distribution of new match quality and its dependence on the stock of houses for sale, the value of λ (in steady state) will be treated as a parameter. This is computationally simpler and without loss of generality when the equilibrium is unique (because it is always possible to change λ by varying the parameter Λ as well as θ). The resulting steady state will be independent of the thick-market parameter θ, though this parameter will be relevant when there are aggregate shocks that lead to fluctuations in the stock of houses for sale.

Given the solution for x and y, the inflow and outflow rates and the stock of houses for sale can be determined. These flows will be 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.1.

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):

$$V_{\mathrm{s}} = y^{\lambda}. \qquad [3.18]$$

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:

$$T_{\mathrm{s}} = y^{\lambda}\left(\frac{\tau}{\nu}\right). \qquad [3.19]$$

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 home-owners 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):

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

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 threshold y, and the moving threshold 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:

$$T_{\mathrm{a}} = \frac{\left(\frac{\tau}{1-\alpha}\right)+\left(\frac{\delta y}{x}\right)^{\lambda}\left(\frac{\tau}{1-(\alpha+(1-\alpha)\delta^{\lambda})}\right)\left(\frac{(2-\delta^{\lambda})(1-\alpha)}{1-(\alpha+(1-\alpha)\delta^{\lambda})}\right)}{1+\left(\frac{\delta y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-(\alpha+(1-\alpha)\delta^{\lambda})}\right)}. \qquad [3.21]$$

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 $T_{\mathrm{a}} < T_{\mathrm{d}}$.

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

$$P = \kappa C - \beta\left(\frac{\tau}{1-\beta}\right)M + \omega\left(\beta\left(\frac{\tau}{1-\beta}\right)+\frac{1}{\pi}\left(\frac{\tau}{\nu}\right)\right)(x+F), \qquad [3.22]$$

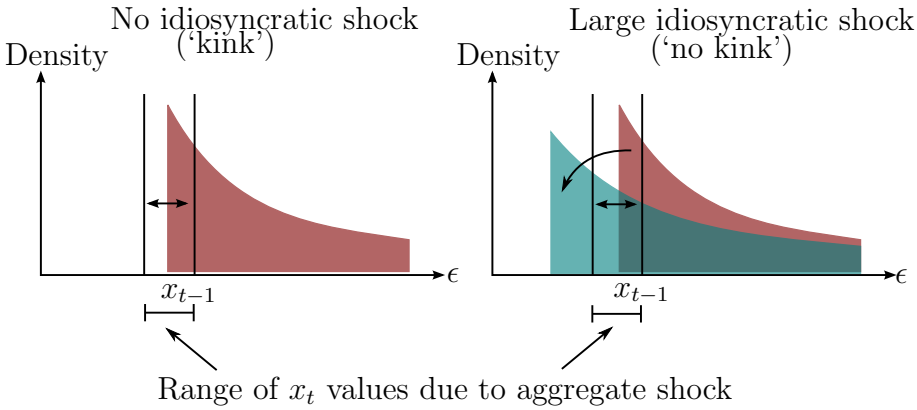
which can be used to calculate the ratio of the transactions, maintenance, and search costs C, M, and F relative to price, denoted by $c \equiv C/P$, $m \equiv M/P$, and $f \equiv F/P$.

3.6 Solution: 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 ($\varsigma = 0$ in [3.12]) described in section 3.5. 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 6. 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). Now there is no differentiability problem. Details of the log linearization are provided in section A.3.

Figure 6: *Differentiability and idiosyncratic shocks*



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 value functions at 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.

3.7 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 [3.15], 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 [3.4]. 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 for two reasons. First, the distribution of match quality for surviving matches changes over time because the threshold for moving can fluctuate. In other words, moving decisions today have consequences for future distributions of match quality. Second, the distribution of new matches can fluctuate because of the thick-market effect.

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.

The thick-market effect also interacts with the cleansing effect in the following way. Anticipating that more listings will occur in the future because of a weaker cleansing effect now, those near the bottom of the distribution of match quality will then be more likely to time their entry into the market to coincide with the increase in listings. The reason is the thick-market effect implies that will be a good time to move. This 'preemption effect' increases cleansing in the future, which then improves the subsequent distribution of match quality. The process described above then goes into reverse, meaning that there is a tendency for transactions to 'bunch' in periods of high activity. Through this process, overshooting can potentially become oscillation.

Finally, there is also a simple direct channel through which endogenous moving affects sales. Even though the sales decision depends on the acceptance threshold and the distribution of new matches, in equilibrium it is affected by moving decisions because of the link between houses for sale and the distribution of new match quality owing to the thick-market effect.

4 Calibration

The model contains a total of 13 parameters $\{a, \delta, \lambda, v, F, M, C, \omega, \kappa, r, \tau, \theta, \rho\}$. 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. This leaves only two free parameters $\{\theta, \rho\}$ to match the fluctuations predicted by the model to those in the data.

Starting with the parameters that have values directly imposed, the length of a discrete time period τ is set to one week ($\tau = 1/52$). 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.

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 moving home.[28] 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.

[28] Weinberg et al. (1981) provide some estimates of moving costs for low-income renters in Phoenix and Pittsburgh.

While the data described above provide information about the ratio of the parameters F, M, and C to the average house price P, since 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 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 (see the discussion of price dispersion in section 3.5). 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).[29]

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 (2012) 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

[29]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.

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 $T_{\rm s} = 6.5/12$. The calibration target for viewings per sale is set to 10 ($V_{\rm s} = 10$) 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 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.[30] 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 ($T_{\rm a} = 11$).

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 ($T_{\rm d} = 12.2$). 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.

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 home-owners 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

[30] 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.

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.

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 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, home-owners 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 home-owners 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.2. The parameters matching the targets and those directly calibrated are all shown in Table 3.

5 Simulations

This section presents the results of simulating the theoretical model with an endogenous moving decision described in section 3 when the economy is subject to aggregate shocks. Given the variation in the inflow rate in the housing market 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 inflows relative to fluctuations in other measures of market activity such as outflows and the stock of houses for sale. Furthermore, does a model with an endogenous inflow decision help in understanding the

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

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

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 for the parameters of the model from section 3 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 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 [3.12] (which scales up the flow value of housing received by all homeowners). The persistence parameter ρ is calibrated to match the persistence of business-cycle fluctuations in GDP as a proxy for the factors that would affect housing demand. The parameter ρ is set so that the half-life of a shock is 3.4 years (this is equivalent to an autoregressive parameter of 0.95 at a quarterly frequency).

With ρ set to represent shocks at business-cycle frequencies, all the remaining parameters of the model (except the thick-market parameter θ) are chosen as described in section 4 (with the modification of this procedure for the exogenous-moving variant of the model described above). Note that the calibration does not use any information from the time-series moments of the data

documented in section 2: the only information used is drawn from the average values of variables, which are matched to the model's equilibrium in the absence of aggregate shocks. The thick-market parameter θ is not identified using information on the average values of aggregate variables, hence a different method must be employed for this parameter. The approach is first to present results in the case of no thick-market effect ($\theta = 0$), and then to experiment with values of θ leading to results that best match the time-series moments of the data.

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 presented in two forms. First, impulse response 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. Second, as a table of relative standard deviations and correlation coefficients 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, a separate variable for time-to-move is not constructed. Finally, for consistency, note that since the model contains no seasonal shocks and the results will not be filtered in any way, the results are not comparable Table 1. Instead, the relative volatilities and correlations in the data are computed with only seasonal adjustment to compare to the results. The summary statistics are presented in Table 4.

Table 4: *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.

5.1 The benchmark case of exogenous moving decisions

The first case considered is the benchmark model with exogenous moving ($\delta = 0$), and initially no thick-market effect is assumed ($\theta = 0$). The implied relative standard deviations and correlation coefficients (given that the economy is hit by aggregate housing demand shocks as in equation [3.12]) are displayed in Table 5 and the impulse response functions in Figure 7. This benchmark model faces many problems in matching the data in Table 4. 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 are also highly negatively correlated with houses for sale in the model, but almost uncorrelated in the data. 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.

Table 5: *Exogenous moving, no thick market effect, shock to housing demand*

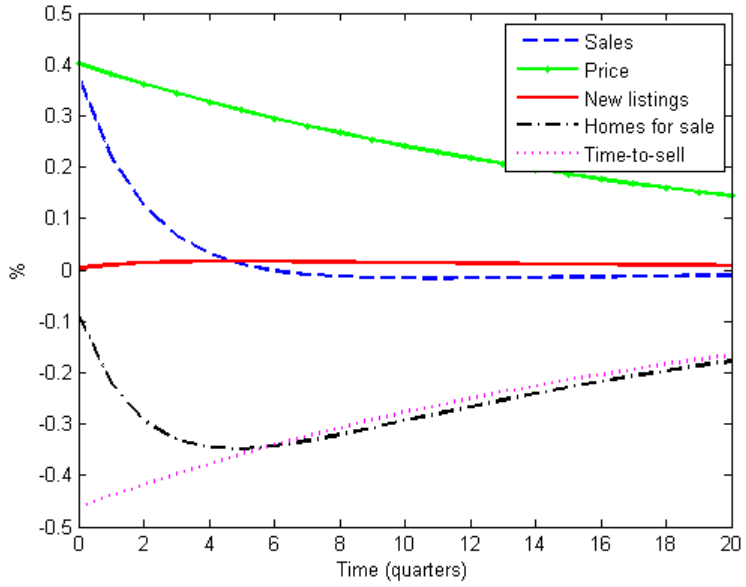
	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	2.780	0.140	2.909	3.207
			Correlation coefficients		
Sales	1				
Price	0.440	1			
New listings	0.126	0.945	1		
Homes for sale	−0.141	−0.951	−0.999	1	
Time-to-sell	−0.440	−1.000	−0.945	0.951	1

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

The problems faced by the model with exogenous moving extend beyond simply the behaviour of new listings. The number of houses for sale is approximately three times more volatile than sales in the model, while these variables have similar volatilities in the data. The reason for this failing can be seen in the impulse response functions in Figure 7. While the shock to the demand for housing pushes up sales, 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). 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 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). Furthermore, price is negatively correlated with the number of houses for sale in the model, while positively correlated in the data.

Figure 7: *Exogenous moving, no thick market effect, shock to housing demand*



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

The correlation of price and time-to-sell is also excessively negative in the model. Some of the problems for the behaviour of prices can be avoided by introducing a thick-market effect into the model. Increasing the thick-market parameter θ initially does not change the qualitative results already discussed. Once θ rises to approximately 0.075, the pattern of results changes qualitatively. Beyond that value of θ, the pattern then remains the same until θ reaches the threshold where multiple equilibria become possible. Table 6 and Figure 8 present the results for the illustrative case of $\theta = 0.075$. Given equation [3.2], $\theta = 0.075$ implies that a doubling of the number of houses for sale would increase the average match quality revealed by a viewing by 7.5%.

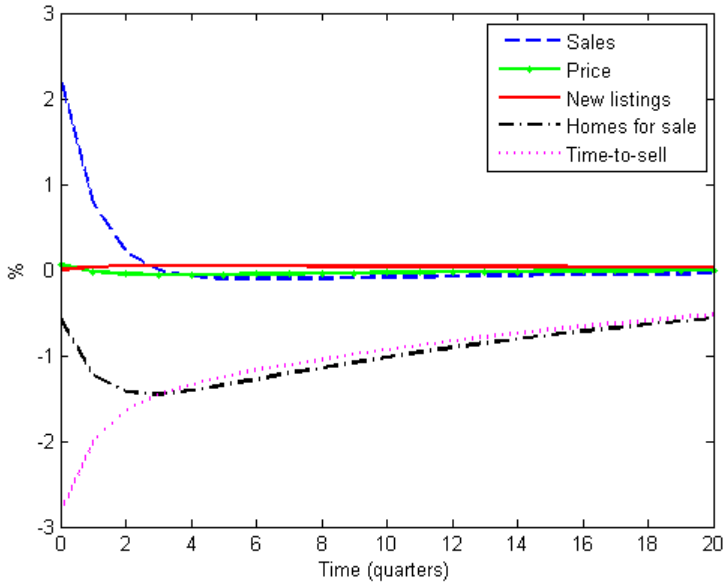
Table 6: *Exogenous moving, thick market effect, shock to housing demand*

	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	0.077	0.091	2.048	2.362
			Correlation coefficients		
Sales	1				
Price	0.327	1			
New listings	−0.024	−0.430	1		
Homes for sale	−0.095	0.668	−0.924	1	
Time-to-sell	−0.505	0.441	−0.791	0.907	1

Notes: Quarterly frequency. Parameters: $\delta = 0$, $\theta = 0.075$.

The impulse response functions in Figure 8 show that adding the thick-market effect to the exogenous-moving model leads to an increase in the magnitude of the fluctuations from a housing demand shock of a given size. However, this need not improve the fit of the model. Relative to sales, the volatility of new listings actually falls, and the volatility of house prices is now also too low. Furthermore, the volatilities of homes for sale and time-to-sell remain much too high (more than double the volatility of sales). In terms of correlations, while the model does better (as might be expected) in matching the positive correlation of prices and houses for sale, its performance deteriorates for other correlations. Listings are predicted to be negatively correlated with all other variables, while in the data have positive correlations with all variables except time-to-sell.

Figure 8: *Exogenous moving, thick market effect, shock to housing demand*



Notes: Quarterly frequency. Parameters: $\delta = 0$, $\theta = 0.075$.

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. Consider first the model of endogenous moving without the thick-market effect ($\theta = 0$). The relative volatilities and correlations are displayed in Table 7 and the impulse response functions in Figure 9. It can be seen that the relative volatilities of the model move closer to what is found in the data: listings become substantially more volatile, though still less volatile than found in the data; houses for sale and time-to-sell are less volatile (relative to sales) than before, but still less volatile than found in the data. Prices are also now less volatile, which goes in the direction of matching the data, but again not by enough. The model with endogenous moving also performs much better in terms of the strongly positive correlation between listings 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 problem is that in the absence of any thick-market effect, there is not a strong incentive for marginal home-owners to take advantage of the option to enter the housing market. Finally, note that the overshooting discussed earlier is masked in this example because the exogenous shock is highly persistent (a small degree of overshooting can be seen in the impulse response of sales, though).

Table 7: *Endogenous moving, no thick market effect, shock to housing demand*

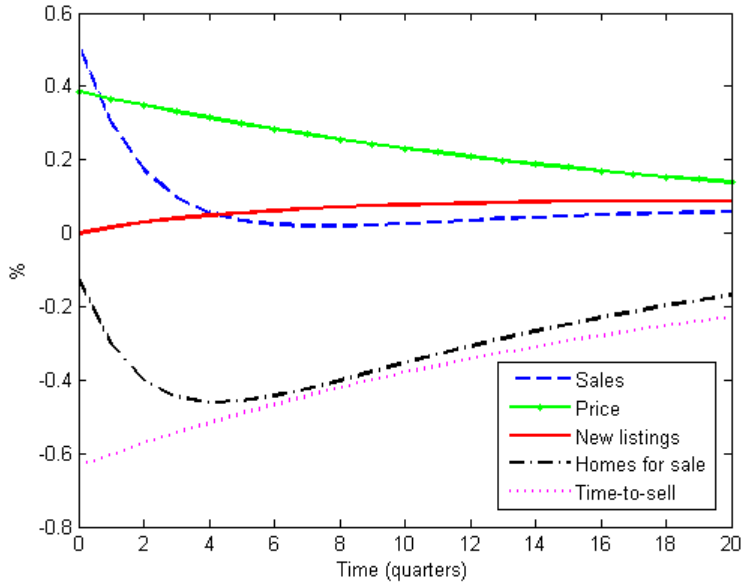
	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	1.655	0.724	2.099	2.721
			Correlation coefficients		
Sales	1				
Price	0.734	1			
New listings	0.557	0.707	1		
Homes for sale	−0.475	−0.946	−0.651	1	
Time-to-sell	−0.734	−1.000	−0.707	0.946	1

Notes: Quarterly frequency. Parameters: $\delta = 0.862$, $\theta = 0$.

Some of these problems, particularly the behaviour of house prices, can be resolved by introducing a thick-market effect. However, it is important that the thick-market effect is not too strong otherwise the performance of the model deteriorates significantly (the relative volatilities of houses for sale and time-to-sell would become much too low, and the correlation between listings and houses for sale would be too high). A value of $\theta = 0.03$ delivers a fairly good overall performance in terms of matching the data. Quantitatively, this is a rather weak thick-market effect: a doubling of the number of houses available for sale would improve the average draws of match quality from viewings by only 3%. Nonetheless, this weak thick-market effect interacts with the endogenous moving decision in an interesting way that changes the dynamics of the model, as can be seen in Figure 10. Following the demand shock, the housing market initially experiences an increase in sales and a drop in listings. As listings fall endogenously, homeowners near the bottom of the distribution of match quality delay their moves, which reduces the degree of 'cleansing' of the distribution. Hence in future time periods, moving becomes more likely. In the presence of a thick-market effect, that time is then a good time to move, so those who are near the bottom of the distribution in future then bring forward their moves ('preemption'). This explains the rebound of listings and the subsequent increase in houses for sale observed in the impulse response functions. Consequently, the distribution of match quality then improves as cleansing increases, so the process then goes into reverse. This explains the oscillations observed in the impulse response function as transactions have a tendency to bunch together, which generates additional volatility in listings and time-to-sell.

The relative volatilities and correlations predicted by the endogenous moving model with a weak

Figure 9: *Endogenous moving, no thick market effect, shock to housing demand*



Notes: Quarterly frequency. Parameters: $\delta = 0.862$, $\theta = 0$.

but positive thick-market effect ($\theta = 0.03$) are shown in Table 8. The volatility of new listings is 90% of the volatility of sales (in the data, listings are 30% more volatile. The volatility of houses for sale is approximately 10% more than sales, not too far from the data where these variables have approximately the same volatility. Time-to-sell is 15% less volatile than sales, which is an improvement on the earlier model, though still some way from the 36% higher volatility found in the data. Taken together, this version of the model gets the relative magnitudes of listings, houses for sale, and time-to-sell much closer to the data than the other versions considered. In terms of correlations, the model predicts the correct sign for all correlations except the negative correlation between listings and time-to-sell found in the data.

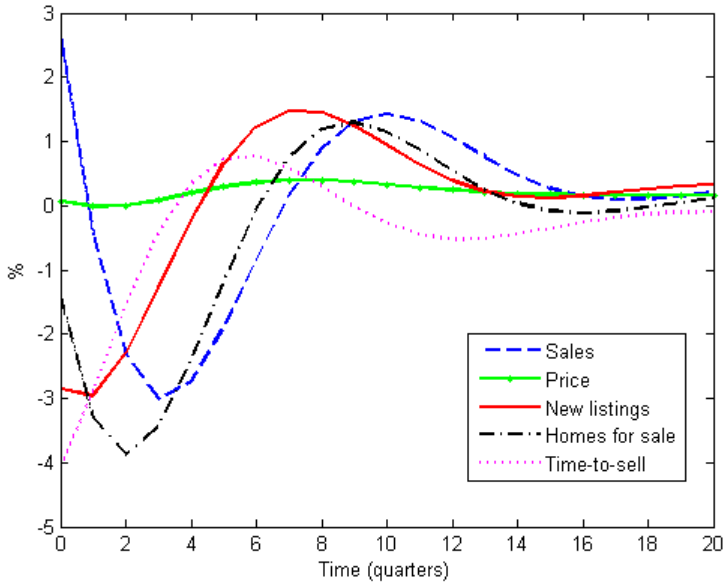
These findings in the theoretical model are consistent with the empirical results of Genesove and Han (2012). Using NAR survey data they find that the short-run responses of time-to-sell and viewings-per-sale to a demand shock are larger than the long-run responses, while the reverse is true for prices. They argue that a search model with a lag in sellers' response could explain these dynamic effects. The search-and-matching model with endogenous moving presented here is able to generate a lag in the sellers' response endogenously, and thus provides a natural mechanism to understand these dynamic effects.

Table 8: *Endogenous moving, thick market effect, shock to housing demand*

	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	0.183	0.893	1.112	0.844
			Correlation coefficients		
Sales	1				
Price	0.179	1			
New listings	0.214	0.462	1		
Homes for sale	0.685	0.123	0.786	1	
Time-to-sell	−0.282	−0.050	0.781	0.505	1

Notes: Quarterly frequency. Parameters: $\delta = 0.862$, $\theta = 0.03$.

Figure 10: *Endogenous moving, thick market effect, shock to housing demand*



Notes: Quarterly frequency. Parameters: $\delta = 0.862$, $\theta = 0.03$.

5.3 Exogenous moving decisions with an aggregate shock to the moving rate

It might be objected that the improved performance of the model with an endogenous moving decision could equally well be obtained by adding an aggregate shock to the moving rate in the exogenous-moving model. This modification would clearly allow that model better to match 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 [3.12]), 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:

$$\log a_t = (1-\rho)\log a + \rho \log a_{t-1} + \eta_t, \quad \text{where } \eta_t \sim \text{i.i.d.}(0, \varsigma^2), \tag{5.1}$$

and where a is the average arrival rate over time. The persistence of this shock (controlled by ρ) is set to reflect the persistence of GDP fluctuations over the business cycle, as was the case for the housing demand shock in [3.12].

Table 9: *Exogenous moving, no thick market effect, shock to moving probability*

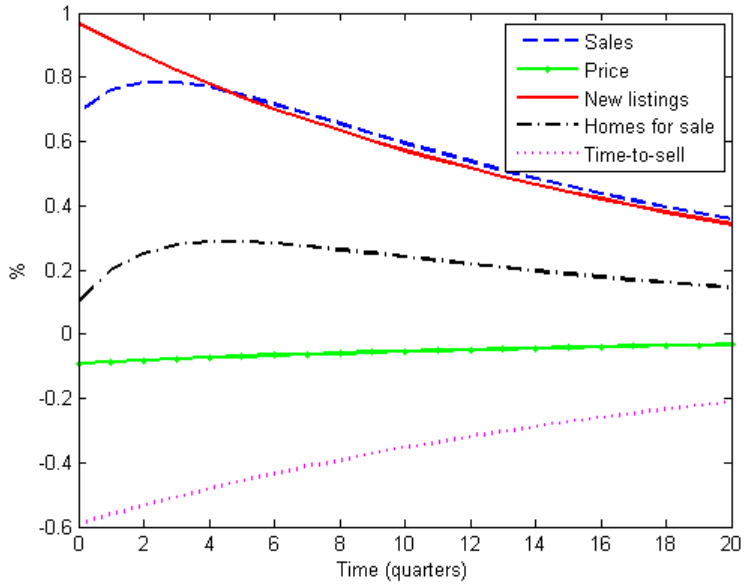
	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	0.097	1.030	0.374	0.635
			Correlation coefficients		
Sales	1				
Price	−0.995	1			
New listings	0.994	−1.000	1		
Homes for sale	0.984	−0.960	0.959	1	
Time-to-sell	−0.995	1.000	−1.000	−0.960	1

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

The first set of results (Table 9 and Figure 11) for the model with this alternative shock assumes there is no thick-market effect ($\theta = 0$). Comparing the empirical evidence in Table 4 to Table 9, 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, the model struggles to match other key features of the data. The number of houses for sale is now too stable relative to the number of 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 there being 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.

Adding a thick-market effect to the model changes the qualitative pattern in the results once θ is above 0.04 approximately. The relative volatility of houses for sale increases to become more

Figure 11: *Exogenous moving, no thick market effect, shock to moving probability*



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

consistent with the data (the relative volatility of listings remains not too far from matching the data). However, the effect on the volatility of time-to-sell relative to sales goes in the other direction, which now has far too little volatility, and the problematic negative correlation of time-to-sell and the number of houses for sale remains. This negative correlation can be eliminated by increasing the strength of the thick-market effect further, which can also reverse some of the fall in the volatility of house prices for sufficiently large θ. Increasing θ by too much leads all variables to be highly positively correlated, though. The results in Table 10 and Figure 12 show the case of $\theta = 0.062$. The relative volatility of houses for sale is about right, but the relative volatility of time-to-sell in the data is approximately seven times larger than what is predicted by the model. Looking at the correlation coefficients, there is too much co-movement among all variables apart from time-to-sell, which co-moves too little with the other variables. Intuitively, as can be seen in Figure 12, the problem is that the thick-market effect makes it more attractive to buy when the exogenous increase in listings occurs. This leads sales to track the increase in houses for sale more closely, which means that the time needed to sell houses changes by little.

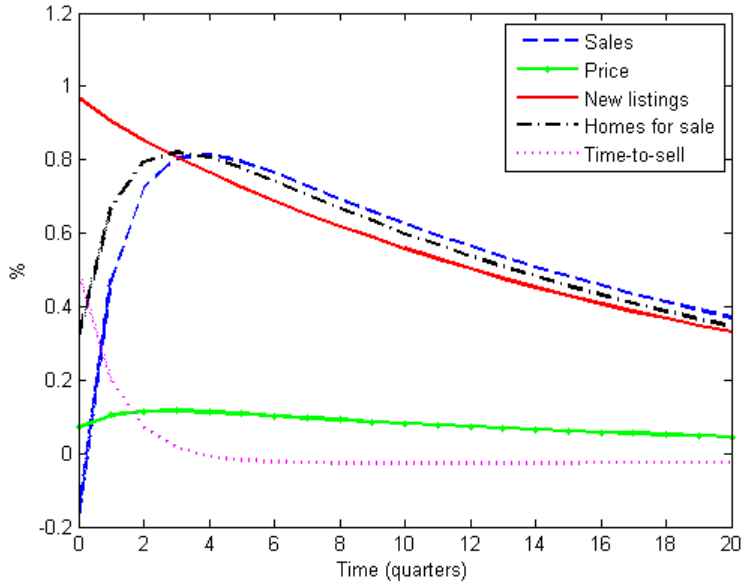
This analysis shows that while a model with exogenous shocks to moving probabilities can generate fluctuations in listings approximately equal in magnitude to fluctuations in sales, such a model does not easily match both the relative volatilities of the number of houses for sale and time-to-sell, and the correlations among these variables. The model with endogenous listings, while it was not able to match all the stylized facts, did perform substantially better in terms of the relative volatilities of houses for sale and time-to-sell, and in terms of most of the correlations.

Table 10: *Exogenous moving, thick market effect, shock to moving probability*

	Sales	Price	New listings	Homes for sale	Time-to-sell
			Relative standard deviations		
Sales	1	0.136	1.018	0.984	0.191
			Correlation coefficients		
Sales	1				
Price	0.961	1			
New listings	0.910	0.984	1		
Homes for sale	0.982	0.996	0.971	1	
Time-to-sell	−0.179	0.100	0.239	0.012	1

Notes: Quarterly frequency. Parameters: $\delta = 0$, $\theta = 0.062$.

Figure 12: *Exogenous moving, thick market effect, shock to moving probability*



Notes: Quarterly frequency. Parameters: $\delta = 0$, $\theta = 0.062$.

6 Conclusions

This paper has documented the role of variation in both outflows (sales) and inflows (new listings) in accounting for fluctuations in the number of houses for sale. Evidence has been presented to show that inflows are arguably just as important as outflows, while existing search models of the housing market have emphasized only the endogeneity of the outflow decision, treating the inflows as exogenous. This paper has calibrated a search-and-matching model with endogenous moving to match key features of the U.S. housing market. 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. It has been argued that the endogenous moving model combined with a moderate thick-market effect performs well in matching the stylized facts. The importance of incorporating an endogenous moving decision is not confined solely to accounting for the volatility of inflows, but also to matching the full set of relative volatilities and correlations.

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

Aruoba, S. B., Davis, M. A. and Wright, R. (2012), "Homework in monetary economics: Inflation, home production, and the production of homes", working paper 18276, NBER. 4

Caplin, A. and Leahy, J. (2011), "Trading frictions and house price dynamics", *Journal of Money, Credit and Banking*, **43(s2)**:283–303 (October). 3, 6

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. 24

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). 4

Diamond, P. A. (1981), "Mobility costs, frictional unemployment, and efficiency", *Journal of Political Economy*, **89(4)**:798–812 (August). 3

Díaz, A. and Jerez, B. (2012), "House prices, sales, and time on the market: A search-theoretic framework", *International Economic Review*, forthcoming. 3, 4, 5, 6, 14, 25

Fisher, J. D. M. and Gervais, M. (2011), "Why has home ownership fallen among the young?", *International Economic Review*, **52(3)**:883–912 (August). 4

Fujita, S. and Ramey, G. (2009), "The cyclicality of separation and job finding rates", *International Economic Review*, **50(2)**:415–430 (May). 1, 7

Genesove, D. and Han, L. (2012), "Search and matching in the housing market", *Journal of Urban Economics*, **72(1)**:31–45 (July). 12, 13, 14, 25, 34

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). 6, 25

Ghent, A. (2012), "Infrequent housing adjustment, limited participation, and monetary policy", *Journal of Money, Credit and Banking*, **44(5)**:931–955 (August). 24

Glaeser, E. L. and Gyourko, J. (2007), "Arbitrage in housing markets", working paper 13704, NBER. 8

Goodman, J. L., Jr. (1993), "A housing market matching model of the seasonality in geographic mobility", *Journal of Real Estate Research*, **8(1)**:117–137. 3

Head, A., Lloyd-Ellis, H. and Sun, H. (2012), "Search, liquidity and the dynamics of house prices and construction", working paper 1276, Queen's Economics Department. 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). 25

Iacoviello, M. (2005), "House prices, borrowing constraints, and monetary policy in the business cycle", *American Economic Review*, **95(3)**:739–764 (June). 4

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). 6, 25

Merlo, A. and Ortalo-Magné, F. (2004), "Bargaining over residential real estate: evidence from England", *Journal of Urban Economics*, **56(2)**:192–216 (September). 6, 25

Mortensen, D. T. and Pissarides, C. A. (1994), "Job creation and job destruction in the theory of unemployment", *Review of Economic Studies*, **61(3)**:397–415 (July). 2

Ngai, L. R. and Sheedy, K. D. (2014), "Moving house", Tech. rep., London School of Economics. 2, 12

Ngai, L. R. and Tenreyro, S. (2014), "Hot and cold seasons in the housing market", *American Economic Review*, forthcoming. 3

Novy-Marx, R. (2009), "Hot and cold markets", *Real Estate Economics*, **37(1)**:1–22 (Spring). 3, 14

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). 4

Petrongolo, B. and Pissarides, C. A. (2008), "The ins and outs of European unemployment", *American Economic Review*, **98(2)**:256–262 (May). 1, 7, 11

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

Pissarides, C. A. (2000), *Equilibrium Unemployment Theory*, MIT Press, 2nd edn. 4

Poterba, J. M. (1991), "House price dynamics: The role of tax policy and demography", *Brookings Papers on Economic Activity*, **1991(2)**:143–203. 24

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. 24

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). 4

Ungerer, C. (2012), "Monetary policy, hot housing markets and household debt", working paper, London School of Economics. 4, 6

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 Steady state solution

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

Assuming that thick-market effects are not so strong that the steady state is not unique, the approach taken is to take the steady-state match quality distribution as represented by λ as a parameter. Given the value of steady-state houses for sale u that will be found, there is always a value of the parameter Λ that justifies whatever value of λ is chosen. The appropriate value of Λ can be obtained from [3.2]:

$$\Lambda = \left(\frac{\lambda}{\lambda - 1}\right) u^{\theta}.$$

Equation [3.15] implies that $H(x) = W + \tau M$. Since $H(\delta\epsilon) < W + \tau M$, the Bellman equation [3.14] 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:

$$W = \left(\frac{\tau}{1 - \beta}\right)(x - M). \qquad \text{[A.1.1]}$$

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

$$H(x) = \left(\frac{\tau}{1 - \beta}\right)(x - \beta M). \qquad \text{[A.1.2]}$$

Equation [3.4] for the acceptance threshold implies $H(y) = C + \beta W$. Combining this with [A.1.1] yields a formula for $H(y)$:

$$H(y) = C + \beta\left(\frac{\tau}{1 - \beta}\right)(x - M). \qquad \text{[A.1.3a]}$$

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 [3.15] implies $H(x) = W + \tau M$, it follows that $H(\delta y) < W + \tau M$, and hence using the Bellman equation [3.14] with $\epsilon = y$:

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

Substituting the expression for W in [A.1.1] and simplifying:

$$H(y) = \left(\frac{\tau}{1 - \beta\alpha}\right) y + \beta\left(\frac{1 - \alpha}{1 - \beta\alpha}\right)\left(\frac{\tau}{1 - \beta}\right) x - \beta\left(\frac{\tau}{1 - \beta}\right) M. \qquad \text{[A.1.3b]}$$

Equating the two formulas for $H(y)$ in [A.1.3a] and [A.1.3b] leads to:

$$y = \alpha\beta x + \left(\frac{1-\beta\alpha}{\tau}\right) C. \qquad \text{[A.1.4]}$$

At the steady state, the Bellman equation [3.7] becomes:

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

which can be solved for W:

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

Substituting this into equation [A.1.1] yields the following expression for the average surplus Σ:

$$x + F = \left(\frac{\nu}{\tau}\right)\pi\Sigma. \qquad \text{[A.1.5]}$$

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

$$\Psi(z) = \int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} (H(\epsilon) - H(z))\mathrm{d}\epsilon. \qquad \text{[A.1.6]}$$

Since $\Sigma(y) = 0$ and [3.3] states that $\Sigma(\epsilon) = H(y) - \beta W - C$, the surplus can be written as $\Sigma(\epsilon) = H(\epsilon) - H(y)$. Comparing equations [3.6] and [A.1.6]:

$$\Psi(y) = \Sigma = \int_{\epsilon=y}^{\infty} \frac{\lambda}{y}\left(\frac{\epsilon}{y}\right)^{-(1+\lambda)} (H(\epsilon) - H(y))\mathrm{d}\epsilon. \qquad \text{[A.1.7]}$$

At the steady state, the Bellman equation [3.14] 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:

$$H(\epsilon) - H(z) = \tau(\epsilon - z) + \beta\alpha(H(\epsilon) - H(z)) + \beta(1-\alpha)\max\{H(\delta\epsilon) - H(x), 0\}. \qquad \text{[A.1.8]}$$

Substituting this expression into the definition of $\Psi(z)$ from [A.1.6] yields:

$$\Psi(z) = \int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} \tau(\epsilon - z)\mathrm{d}\epsilon + \beta\alpha\Psi(z) + \beta(1-\alpha)\int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} \max\{H(\delta\epsilon) - H(x), 0\}\mathrm{d}\epsilon. \qquad \text{[A.1.9]}$$

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

$$\int_{\epsilon=z}^{\infty} \frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)} (\epsilon - z)\mathrm{d}\epsilon = \frac{1}{\lambda - 1} z. \qquad \text{[A.1.10]}$$

To find the second integral on the right-hand side of [A.1.9], 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.1.10], this equation shows that [A.1.9] implies

$$\Psi(z) = \left(\frac{\tau}{1-\beta\alpha}\right)\left(\frac{z}{\lambda-1}\right) + \beta\Delta\left(\frac{1-\alpha}{1-\beta\alpha}\right)\left(\frac{z}{x}\right)^{\lambda}\Psi(x), \qquad \text{[A.1.11]}$$

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

Evaluating [A.1.11] at $z = x$ (valid since $x < y$) yields an expression for $\Psi(x)$:

$$\Psi(x) = \left(\frac{\tau}{1-\beta\sigma_x}\right)\left(\frac{x}{\lambda-1}\right), \qquad \text{[A.1.12]}$$

where $\sigma_x \equiv \alpha + (1-\alpha)\Delta$ is defined. Then evaluating [A.1.11] at $z = y$ and substituting the expression for $\Psi(x)$ from [A.1.12]:

$$\Psi(y) = \left(\frac{y}{\lambda-1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(1 + \beta\Delta\left(\frac{y}{x}\right)^{\lambda-1}\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\right). \qquad \text{[A.1.13]}$$

Noting that equation [3.5] implies that $\pi = y^{-\lambda}$ in the steady state, by substituting [A.1.13] into [A.1.7] and then [A.1.5]:

$$x + F = \left(\frac{\nu}{\tau}\right)\left(\frac{1}{\lambda-1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(y^{1-\lambda} + \beta\Delta\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)x^{1-\lambda}\right). \qquad \text{[A.1.14]}$$

The steady-state thresholds x and y are the solution of the simultaneous equations [A.1.4] and [A.1.14]. Each equation defines a relationship between x and y. Equation [A.1.4] implies a positive relationship between x and y, while equation [A.1.14] implies a negative relationship between x and y. If a solution exists, it must then be unique. Since [A.1.4] implies x is positive when $y = 0$, and because [A.1.14] 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.1.1], [A.1.2], and [A.1.3a]. The average surplus Δ can be found using [A.1.5]. Using $H(\epsilon) = H(y) + \Sigma(\epsilon)$, the steady-state value of H can be found using $H = H(y) + \Sigma$:

$$H = C + \beta\left(\frac{\tau}{1-\beta}\right)(x - M) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)(x + F), \qquad \text{[A.1.15]}$$

where $\pi = y^{-\lambda}$. The steady-state average price is then found using equation [3.11]:

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

and by substituting the expression for H from [A.1.15]:

$$P = \kappa C - \beta\left(\frac{\tau}{1-\beta}\right) M + \omega\left(\beta\left(\frac{\tau}{1-\beta}\right) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)\right)(x+F). \qquad \text{[A.1.16]}$$

To derive the mean-min ratio $P/p(y)$ measure of price dispersion, start by noting that [3.10] 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.1.3a] 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:

$$\frac{p(y)}{P} = \kappa c - \beta\left(\frac{\tau}{1-\beta}\right) m + \beta\left(\frac{\tau}{1-\beta}\right)\omega\left(\frac{x}{P} + f\right). \qquad \text{[A.1.17]}$$

Dividing both sides of the equation [A.1.16] 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.1.17] and simplifying leads to:

$$\frac{p(y)}{P} = \frac{\beta\left(\frac{\tau}{1-\beta}\right) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)\left(\kappa c - \beta\left(\frac{\tau}{1-\beta}\right) m\right)}{\beta\left(\frac{\tau}{1-\beta}\right) + \frac{1}{\pi}\left(\frac{\tau}{\nu}\right)}. \qquad \text{[A.1.18]}$$

The mean-min ratio is therefore given by:

$$\frac{P}{p(y)} = \frac{1 + \beta\pi\left(\frac{\nu}{\tau}\right)\left(\frac{\tau}{1-\beta}\right)}{\left(\kappa c - \beta\left(\frac{\tau}{1-\beta}\right) m\right) + \beta\pi\left(\frac{\nu}{\tau}\right)\left(\frac{\tau}{1-\beta}\right)}. \qquad \text{[A.1.19]}$$

Turning now to the stocks and the flows, at the steady state, equation [3.16] for the unsold stock of houses becomes:

$$u = 1 - \nu u \sum_{i=1}^{\infty}\left\{ y^{-\lambda}\alpha^i + x^{-\lambda}\sum_{j=0}^{i-1}(1-\alpha)\alpha^j \sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\Delta^{k+1}\right\}, \qquad \text{[A.1.20]}$$

noting the definition $\Delta = \delta^{\lambda}$. The following summation formula is needed to simplify this equation:

$$\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\Delta^k = \sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}$$
$$= (\alpha + (1-\alpha)\Delta)^{i-1-j} = \sigma_x^{i-1-j}, \qquad \text{[A.1.21a]}$$

which is obtained using the binomial theorem and the definition $\sigma_x \equiv \alpha + (1-\alpha)\Delta$. Making use of this

result:

$$\sum_{j=0}^{i-1}(1-\alpha)\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\Delta^{k+1}$$
$$=(1-\alpha)\Delta\sum_{j=0}^{i-1}\alpha^j\left(\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right)=(1-\alpha)\Delta\sum_{j=0}^{i-1}\alpha^j\sigma_x^{i-1-j}$$
$$=(1-\alpha)\Delta\sigma_x^{i-1}\sum_{j=0}^{i-1}\left(\frac{\alpha}{\sigma_x}\right)^j=(1-\alpha)\Delta\sigma_x^{i-1}\left(\frac{1-(\alpha/\sigma_x)^i}{1-(\alpha/\sigma_x)}\right)=\frac{(1-\alpha)\Delta\sigma_x}{\sigma_x-\alpha}\frac{\sigma_x\sigma_x^{i-1}}{\sigma_x^i}(\sigma_x^i-\alpha^i)=\sigma_x^i-\alpha^i, \quad \text{[A.1.21b]}$$

using $\sigma_x-\alpha=(1-\alpha)\Delta$. This equation can then be used to simplify the summations that appear in [A.1.20]:

$$\sum_{i=1}^{\infty}\left\{y^{-\lambda}\alpha^i+x^{-\lambda}\sum_{j=0}^{i-1}(1-\alpha)\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\Delta^{k+1}\right\}$$
$$=y^{-\lambda}\sum_{i=1}^{\infty}\left\{\alpha^i+\left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i-\alpha^i)\right\}=\pi\left(\frac{\alpha}{1-\alpha}+\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\sigma_x}{1-\sigma_x}-\frac{\alpha}{1-\alpha}\right)\right), \quad \text{[A.1.22]}$$

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

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

Substituting this into [A.1.22] and then into [A.1.20] leads to the following equation:

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

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

$$u=\frac{1}{1+\pi\nu\left(\frac{\alpha}{1-\alpha}+\Delta\left(\frac{y}{x}\right)^{\lambda}\frac{1}{1-\sigma_x}\right)}. \quad \text{[A.1.23]}$$

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.1.23]. 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.1.23] leads to the following expression for the inflow probability:

$$n=\frac{1-\alpha}{1+\Delta\left(\frac{y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-\sigma_x}\right)}. \quad \text{[A.1.24]}$$

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:

$$V_{\mathrm{s}} = \sum_{\ell=1}^{\infty} \ell\pi(1-\pi)^{\ell-1} = \frac{1}{\pi}. \qquad \text{[A.1.25]}$$

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:

$$T_{\mathrm{s}} = \left(\frac{\tau}{\nu}\right)\frac{1}{\pi} = \left(\frac{\tau}{\nu}\right) V_{\mathrm{s}}. \qquad \text{[A.1.26]}$$

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 $\Delta = \delta^{\lambda}$. 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 $\sigma_x = \alpha + (1-\alpha)\Delta$. 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:

$$\psi_\ell - \alpha^{\ell} = (\sigma_y - \alpha)\left(\alpha^{\ell-1} + \alpha^{\ell-2}\sigma_x + \cdots + \alpha\sigma_x^{\ell-2} + \sigma_x^{\ell-1}\right). \qquad \text{[A.1.27]}$$

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.1.27] yields the survival function:

$$\psi_\ell = \alpha^{\ell} + \left(\frac{y}{x}\right)^{\lambda}(\sigma_x^{\ell} - \alpha^{\ell}). \qquad \text{[A.1.28]}$$

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:

$$T_\text{d} = \tau\sum_{\ell=1}^{\infty}\ell\zeta_\ell\psi_{\ell-1} = \tau\sum_{\ell=1}^{\infty}\ell(\psi_{\ell-1}-\psi_\ell) = \tau\left(\sum_{\ell=0}^{\infty}(\ell+1)\psi_\ell - \sum_{\ell=0}^{\infty}\ell\psi_\ell\right) = \tau\sum_{\ell=0}^{\infty}\psi_\ell. \qquad \text{[A.1.29]}$$

The formula in [A.1.28] can be used to sum the survival function sequence:

$$\sum_{\ell=0}^{\infty}\psi_\ell = \frac{1}{1-\alpha} + \left(\frac{y}{x}\right)^\lambda\left(\frac{1}{1-\sigma_x} - \frac{1}{1-\alpha}\right) = \left(\frac{1}{1-\alpha}\right)\left(1 + \Delta\left(\frac{y}{x}\right)^\lambda\left(\frac{1-\alpha}{1-\sigma_x}\right)\right), \qquad \text{[A.1.30]}$$

noting that $\sigma_x - \alpha = (1-\alpha)\Delta$. This is substituted into [A.1.29] to obtain the expected duration of a match:

$$T_\text{d} = \left(\frac{\tau}{1-\alpha}\right)\left(1 + \Delta\left(\frac{y}{x}\right)^\lambda\left(\frac{1-\alpha}{1-\sigma_x}\right)\right). \qquad \text{[A.1.31]}$$

Comparing this to equation [A.1.24], the expected duration can also be written in terms of the unconditional inflow probability n:

$$T_\text{d} = \frac{\tau}{n}. \qquad \text{[A.1.32]}$$

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:

$$\phi_\ell = \frac{\psi_\ell}{\sum_{j=0}^{\infty}\psi_j}. \qquad \text{[A.1.33]}$$

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.1.33]. From [A.1.30] 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.1.28] 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.1.33], the equation above implies:

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

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

$$T_{\mathrm{a}} = \frac{\left(\frac{\tau}{1-\alpha}\right) + \Delta\left(\frac{y}{x}\right)^{\lambda}\left(\frac{\tau}{1-\sigma_x}\right)\left(\frac{2-\alpha-\sigma_x}{1-\sigma_x}\right)}{1+\Delta\left(\frac{y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-\sigma_x}\right)}. \qquad [A.1.34]$$

This completes the derivation of the steady state.

A.2 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, and the share κ of the transaction costs borne by the seller. 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. The thick-market parameter θ has no effect on the steady state, so is determined separately.

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.

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:

$$\wp \equiv \Delta\left(\frac{y}{x}\right)^{\lambda}\left(\frac{1-\alpha}{1-\sigma_x}\right). \qquad [A.2.1]$$

Given the calibration target T_{d} and equation [A.1.31], $\wp$ can be found using:

$$\wp = \left(\frac{1-\alpha}{\tau}T_{\mathrm{d}}\right) - 1. \qquad [A.2.2]$$

Now observe using equation [A.2.1] that the expression for T_{a} in [A.1.34] 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:

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

By substituting the expression for $\wp$ from [A.2.2] 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:

$$\sigma_x = 1 - \left(\frac{\tau}{T_\mathrm{d}}\right)\left(\frac{T_\mathrm{d} - \left(\frac{\tau}{1-\alpha}\right)}{T_\mathrm{a} - \left(\frac{\tau}{1-\alpha}\right)}\right). \qquad [A.2.3]$$

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

$$\Delta = \frac{\sigma_x - \alpha}{1-\alpha}. \qquad [A.2.4]$$

With values of α, σ_x, Δ, and $\wp$, equation [A.2.1] implies an expression for $(y/x)^\lambda$:

$$\left(\frac{y}{x}\right)^\lambda = \frac{\wp}{\Delta}\left(\frac{1-\sigma_x}{1-\alpha}\right). \qquad [A.2.5]$$

The next step is to divide both sides of equation [A.1.16] 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.1.26]:

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

which can be solved for x/P as follows:

$$\frac{x}{P} = \frac{1 - \kappa c + \beta\left(\frac{\tau}{1-\beta}\right)m}{\omega\left(\beta\left(\frac{\tau}{1-\beta}\right) + T_\mathrm{s}\right)} - f. \qquad [A.2.6]$$

Making use of the definition of c, divide both sides of equation [A.1.4] by P to obtain:

$$\frac{y}{P} = \alpha\beta\frac{x}{P} + \left(\frac{1-\beta\alpha}{\tau}\right)c, \qquad [A.2.7]$$

and this yields the value of y/P given the value of x/P found from [A.2.6]. 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.2.5], λ is found using the identity:

$$\lambda = \frac{\log\left(\frac{y}{x}\right)^\lambda}{\log\frac{y}{x}}. \qquad [A.2.8]$$

There is one further equation that must hold. Using $\pi = y^{-\lambda}$, equation [A.1.14] can be written as follows:

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

and dividing both sides by P and using the definition of f and the expression for T_s from [A.1.26]

$$\frac{x}{P} + f = \frac{1}{T_\mathrm{s}}\left(\frac{1}{\lambda-1}\right)\left(\frac{\tau}{1-\beta\alpha}\right)\left(\frac{y}{P} + \beta\Delta\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\left(\frac{y}{x}\right)^\lambda\frac{x}{P}\right). \qquad [A.2.9]$$

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.2.9] 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.1.26] and the definition $\nu = 1 - e^{-v\tau}$:

$$v = -\frac{1}{\tau}\log\left(1 - \tau\frac{V_s}{T_s}\right). \qquad [A.2.10]$$

The value of λ is found using equation [A.2.8] via the steps described above. The value of δ can be obtained using the definition $\Delta = \delta^{\lambda}$ and the expression for Δ in [A.2.4]:

$$\delta = \Delta^{\frac{1}{\lambda}}. \qquad [A.2.11]$$

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.1.25] and $\pi = y^{-\lambda}$ it follows that:

$$y = V_s^{\frac{1}{\lambda}}.$$

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.1.14] and using $\pi = y^{-\lambda}$ and [A.1.26], and the values of σ_x and Δ found in [A.2.3] and [A.2.4]:

$$F = \frac{1}{T_s}\left(\frac{1}{\lambda - 1}\right)\left(\frac{\tau}{1 - \beta\alpha}\right)\left(y + \beta\Delta\left(\frac{1-\alpha}{1-\beta\sigma_x}\right)\left(\frac{y}{x}\right)^{\lambda}x\right) - x. \qquad [A.2.12]$$

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.3 Log linearizations

The Bellman equation [3.7] for the combined buyer-seller value W_t has the following log-linear approximate form:

$$\mathsf{W}_t = \beta\mathbb{E}_t\mathsf{W}_{t+\tau} + (1-\beta)\left(\frac{x+F}{x-M}\right)(\nu_t + \pi_t + \Sigma_t), \qquad [A.3.1]$$

where the steady-state values in [A.1.1] and [A.1.5] 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 [3.14] reduces to:

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

which can be iterated forwards to deduce:

$$H_t(\epsilon) = \sum_{\ell=0}^{\infty}(\beta\alpha)^{\ell}\mathbb{E}_t[\tau\epsilon\xi_{t+\tau\ell} - (\beta\alpha)\tau M + \beta(1-\alpha)W_{t+\tau(\ell+1)}]. \qquad [A.3.2]$$

Since $\epsilon = x_t$ satisfies the restriction required to use [A.3.2], equation [3.15] and [A.3.2] 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.1.1]):

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

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.3.1] for W_t into the above leads to:

$$\mathsf{x}_t=\beta\alpha\mathbb{E}_t\mathsf{x}_{t+\tau}-(1-\beta\alpha)\xi_t+(1-\beta\alpha)\left(1+\frac{F}{x}\right)(\nu_t+\pi_t+\Sigma_t),\qquad[A.3.3]$$

which can be written as:

$$\mathbb{E}_t\left[\left(\frac{\mathbb{I}-\beta\alpha\mathbb{F}}{1-\beta\alpha}\right)\mathsf{x}_t\right]=\left(1+\frac{F}{x}\right)(\nu_t+\pi_t+\Sigma_t)-\xi_t.\qquad[A.3.4]$$

It is also the case that $\epsilon=y_t$ meets the requirement ($\epsilon<y+\eta$) needed to apply [A.3.2]. Together with [3.4] this implies:

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

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.3.1] leads to the following equation for y_t:

$$\mathsf{y}_t=\beta\alpha\mathbb{E}_t\mathsf{y}_{t+\tau}-(1-\beta\alpha)\xi_t+\beta\alpha(1-\beta\alpha)\left(1+\frac{F}{x}\right)\left(\frac{x}{y}\right)\mathbb{E}_t[\nu_{t+\tau}+\pi_{t+\tau}+\Sigma_{t+\tau}],\qquad[A.3.5]$$

which can be written as:

$$\mathbb{E}_t\left[\left(\frac{\mathbb{I}-\beta\alpha\mathbb{F}}{1-\beta\alpha}\right)\mathsf{y}_t\right]=\beta\alpha\left(1+\frac{F}{x}\right)\left(\frac{x}{y}\right)\mathbb{E}_t[\mathbb{F}(\boldsymbol{\nu}_t+\boldsymbol{\pi}_t+\boldsymbol{\Sigma}_t)]-\boldsymbol{\xi}_t. \tag{A.3.6}$$

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

$$\Psi_t(z)\equiv\int_{\epsilon=z}^{\infty}\frac{\lambda}{z}\left(\frac{\epsilon}{z}\right)^{-(1+\lambda)}(H_t(\epsilon)-H_t(z))\mathrm{d}\epsilon. \tag{A.3.7}$$

Analogous to the definition of the expected surplus in [3.6], define the expected surplus $\bar{\Sigma}_t$ using the steady-state distribution of ϵ:

$$\bar{\Sigma}_t\equiv\int_{\epsilon=y}^{\infty}\frac{\lambda}{y}\left(\frac{\epsilon}{y}\right)^{-(1+\lambda)}\Sigma_t(\epsilon)\mathrm{d}\epsilon. \tag{A.3.8}$$

Equations [3.3] and [3.4] imply that $\Sigma(\epsilon)=H_t(\epsilon)-H_t(y_t)$, and [A.3.2] 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.3.8] 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.1.5]):

$$\bar{\Sigma}_t=\Psi_t(y)-\nu\pi\left(\frac{y}{x+F}\right)\left(\frac{1}{1-\beta\alpha}\right)\mathsf{y}_t. \tag{A.3.9}$$

By substituting the Bellman equation [3.14] into [A.3.7] for a value of $z<y+\eta$ (which implies $H_{t+\tau}(\delta z)-\tau M<W_{t+\tau}$):

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

For the first term, equation [A.1.10] is used; the second term can be written in terms of $\Psi_{t+\tau}(z)$ using the definition [A.3.7]; and the integral in the third term is analysed by making the change of variable $\epsilon'=\delta\epsilon$:

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

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

$$\begin{aligned}\Psi_t(z)=&\frac{\tau z\xi_t}{\lambda-1}+\beta\alpha\mathbb{E}_t\Psi_{t+\tau}(z)\\&+\beta(1-\alpha)\Delta\left(\frac{z}{x}\right)^{\lambda}\mathbb{E}_t\left[\int_{\epsilon=\delta z}^{\infty}\frac{\lambda}{x}\left(\frac{\epsilon}{x}\right)^{-(1+\lambda)}\max\left\{(H_{t+\tau}(\epsilon)-\tau M-W_{t+\tau},0\right\})\mathrm{d}\epsilon\right],\end{aligned}$$

and using the fact that the value function $H_{t+\tau}(\epsilon)$ is increasing in ϵ and the definition of x_t in [3.15], which

implies $\tau M + W_{t+\tau} = H_{t+\tau}(x_{t+\tau})$:

$$\Psi_t(z) = \frac{\tau z \xi_t}{\lambda - 1} + \beta\alpha \mathbb{E}_t \Psi_{t+\tau}(z) + \beta(1-\alpha)\Delta \left(\frac{z}{x}\right)^{\lambda} \mathbb{E}_t \left[\int_{\epsilon = x_{t+\tau}}^{\infty} \frac{\lambda}{x} \left(\frac{\epsilon}{x}\right)^{-(1+\lambda)} \left(H_{t+\tau}(\epsilon) - H_{t+\tau}(x_{t+\tau})\right) \mathrm{d}\epsilon \right]. \quad \text{[A.3.10]}$$

Note that since $x_{t+\tau} < y + \eta$, equation [A.3.2] 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.3.7], equation [A.3.10] can be written as follows:

$$\begin{aligned}\Psi_t(z) = \frac{\tau z \xi_t}{\lambda - 1} + \beta\alpha \mathbb{E}_t \Psi_{t+\tau}(z) + \beta(1-\alpha)\Delta \left(\frac{z}{x}\right)^{\lambda} \mathbb{E}_t \Psi_{t+\tau}(x) \\ - \beta(1-\alpha)\Delta \left(\frac{z}{x}\right)^{\lambda} \tau \mathbb{E}_t \left[(x_{t+\tau} - x) \sum_{\ell=0}^{\infty} (\beta\alpha)^{\ell} \xi_{t+\tau(\ell+1)} \right] \\ - \beta(1-\alpha)\Delta \left(\frac{z}{x}\right)^{\lambda} \mathbb{E}_t \left[\int_{\epsilon = x}^{x_{t+\tau}} \frac{\lambda}{x} \left(\frac{\epsilon}{x}\right)^{-(1+\lambda)} \left(H_{t+\tau}(\epsilon) - H_t(x_{t+\tau})\right) \mathrm{d}\epsilon \right]. \quad \text{[A.3.11]}\end{aligned}$$

Log linearizing this equation at $z = y$ (noting that $\Sigma = \Psi(y)$, and using the expression for Σ in [A.1.5] and that for $\Psi(x)$ in [A.1.12]):

$$\begin{aligned}\Psi_t(y) = \left(\frac{\nu\pi}{\lambda - 1}\right) \left(\frac{y}{x + F}\right) \xi_t + \beta\alpha \mathbb{E}_t \Psi_{t+\tau}(y) + \left(\frac{\nu\pi}{\lambda - 1}\right) \left(\frac{x}{x + F}\right) \left(\frac{y}{x}\right)^{\lambda} \left(\frac{\beta(1-\alpha)\Delta}{1 - \beta\sigma_x}\right) \mathbb{E}_t \Psi_{t+\tau}(x) \\ - \nu\pi \left(\frac{x}{x + F}\right) \left(\frac{y}{x}\right)^{\lambda} \left(\frac{\beta(1-\alpha)\Delta}{1 - \beta\alpha}\right) \mathbb{E}_t \mathsf{x}_{t+\tau}. \quad \text{[A.3.12]}\end{aligned}$$

Log linearizing equation [A.3.11] now at $z = x$:

$$\Psi_t(x) = (1 - \beta\sigma_x)\xi_t + \beta\sigma_x \mathbb{E}_t \Psi_{t+\tau}(x) - (\lambda - 1)(1 - \beta\sigma_x) \left(\frac{\beta(1-\alpha)\Delta}{1 - \beta\alpha}\right) \mathbb{E}_t \mathsf{x}_{t+\tau}, \quad \text{[A.3.13]}$$

which can be written as:

$$\mathbb{E}_t \left[\left(\frac{\mathbb{I} - \beta\sigma_x \mathbb{F}}{1 - \beta\sigma_x}\right) \Psi_t(x) \right] = \xi_t - (\lambda - 1) \left(\frac{\beta(1-\alpha)\Delta}{1 - \beta\alpha}\right) \mathbb{E}_t[\mathbb{F}\mathsf{x}_t]. \quad \text{[A.3.14]}$$

By substituting equation [A.3.6] into [A.3.9]:

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

and then substituting from [A.3.12]:

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

For the next step, substitute from [A.3.14]

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

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

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

Finally, substitute equation [A.3.4] into the final term to obtain:

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

and by simplifying this expression:

$$\mathbb{E}_t\left[(\mathbb{I}-\beta\alpha\mathbb{F})(\mathbb{I}-\beta\sigma_x\mathbb{F})\bar{\Sigma}_t\right]+\nu\pi\mathbb{E}_t\left[\left\{\alpha(\mathbb{I}-\beta\sigma_x\mathbb{F})+(1-\alpha)\Delta\left(\frac{y}{x}\right)^{\lambda}\right\}\beta\mathbb{F}(\nu_t+\pi_t+\Sigma_t)\right] =$$
$$\nu\pi\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{x}{x+F}\right)\mathbb{E}_t\left[\left\{\left(\frac{y}{x}\right)(\mathbb{I}-\beta\sigma_x\mathbb{F})+\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta\beta\mathbb{F}\right\}\xi_t\right]. \quad [A.3.15]$$

Equation [3.5] for the acceptance probability can be log linearized as follows:

$$\pi_t = -\lambda\mathsf{y}_t-\lambda(\log y)\lambda_t. \quad [A.3.16]$$

The log linearization of the law of motion [3.16] for houses for sale is:

$$
\begin{aligned}
\mathsf{u}_t &= \lambda\nu y^{-\lambda}\sum_{i=1}^{\infty}\alpha^i \mathsf{y}_{t-i} \\
&+ \lambda\nu x^{-\lambda}\left\{\sum_{i=1}^{\infty}\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\mathsf{x}_{t-j}\right\} \\
&+ \lambda\nu y^{-\lambda}\sum_{i=1}^{\infty}\left\{(\log y)\alpha^i + \left(\frac{y}{x}\right)^{\lambda}\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\ldots\right. \\
&\left.\times\left((\log x)-(k+1)(\log\delta)\right)\right\}\lambda_{t-i} \\
&- \nu\sum_{i=1}^{\infty}\left\{\alpha^i y^{-\lambda} + \sum_{j=0}^{i-1}(1-\alpha)\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\left(\frac{x}{\delta^{k+1}}\right)^{-\lambda}\right\}\mathsf{u}_{t-i} \\
&- \nu\sum_{i=1}^{\infty}\left\{\alpha^i y^{-\lambda} + \sum_{j=0}^{i-1}(1-\alpha)\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\left(\frac{x}{\delta^{k+1}}\right)^{-\lambda}\right\}\mathsf{v}_{t-i}. \quad [A.3.17]
\end{aligned}
$$

The terms in u_t in this expression can be written in the following simpler form

$$
\begin{aligned}
&\nu\sum_{i=1}^{\infty}\left\{\alpha^i y^{-\lambda} + \sum_{j=0}^{i-1}(1-\alpha)\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\left(\frac{x}{\delta^{k+1}}\right)^{-\lambda}\right\}\mathsf{u}_{t-i} \\
&= \nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^{\lambda}\left(\sum_{j=0}^{i-1}(1-\alpha)\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\Delta^{k+1}\right)\right\}\mathsf{u}_{t-i} \\
&= \nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i - \alpha^i)\right\}\mathsf{u}_{t-i}, \quad [A.3.18]
\end{aligned}
$$

which uses $\pi = y^{-\lambda}$, the definitions $\Delta = \delta^{\lambda}$ and $\sigma_x = \alpha + (1-\alpha)\Delta$, and the binominal theorem. Identical steps establish that:

$$
\begin{aligned}
&\nu\sum_{i=1}^{\infty}\left\{\alpha^i y^{-\lambda} + \sum_{j=0}^{i-1}(1-\alpha)\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}(1-\alpha)^k\alpha^{i-1-j-k}\left(\frac{x}{\delta^{k+1}}\right)^{-\lambda}\right\}\mathsf{v}_{t-i} \\
&= \nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i + \left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i - \alpha^i)\right\}\mathsf{v}_{t-i}. \quad [A.3.19]
\end{aligned}
$$

The terms in x_t from [A.3.17] can be simplified by a change in the order of summation:

$$\lambda\nu x^{-\lambda}\left\{\sum_{i=1}^{\infty}\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\mathsf{x}_{t-j}\right\}$$
$$=\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\sum_{j=0}^{\infty}\left\{\sum_{i=j+1}^{\infty}(1-\alpha)\Delta\alpha^j\left(\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right)\right\}\mathsf{x}_{t-j}$$
$$=\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\sum_{j=0}^{\infty}\left\{\sum_{i=j+1}^{\infty}(1-\alpha)\Delta\alpha^j\sigma_x^{i-1-j}\right\}\mathsf{x}_{t-j}=\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\sum_{j=0}^{\infty}\left\{(1-\alpha)\Delta\alpha^j\sum_{i=j+1}^{\infty}\sigma_x^{i-(j+1)}\right\}\mathsf{x}_{t-j}$$
$$=\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\Delta\left(\frac{1-\alpha}{1-\sigma_x}\right)\sum_{j=0}^{\infty}\alpha^j\mathsf{x}_{t-j}=\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\Delta\left(\frac{1-\alpha}{1-\sigma_x}\right)\sum_{i=0}^{\infty}\alpha^i\mathsf{x}_{t-i},\quad[\text{A.3.20}]$$

which again makes use of the binomial theorem. Finally, the terms in λ_t that appear in [A.3.17] can be written as follows:

$$\lambda\nu y^{-\lambda}\sum_{i=1}^{\infty}\left\{(\log y)\alpha^i+\left(\frac{y}{x}\right)^{\lambda}\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\dots\right.$$
$$\left.\times((\log x)-(k+1)(\log\delta))\right\}\lambda_{t-i}=\lambda\nu\pi(\log y)\sum_{i=1}^{\infty}\alpha^i\lambda_{t-i}$$
$$+\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}(\log x)\sum_{i=1}^{\infty}\left\{\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right\}\lambda_{t-i}$$
$$-\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}(\log\delta)\sum_{i=1}^{\infty}\left\{\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right.$$
$$\left.+\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}k\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right\}\lambda_{t-i}$$
$$=\lambda\nu\pi(\log y)\sum_{i=1}^{\infty}\alpha^i\lambda_{t-i}+\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}(\log x)\sum_{i=1}^{\infty}\left\{\sigma_x^i-\alpha^i\right\}\lambda_{t-i}$$
$$-\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}(\log\delta)\sum_{i=1}^{\infty}\left\{\sigma_x^i-\alpha^i+\sum_{j=0}^{i-1}(1-\alpha)\sigma_x\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!k}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right\}\lambda_{t-i}.$$
$$[\text{A.3.21}]$$

Further analysis of these terms will make use of the following summation formula:

$$\sum_{k=0}^{i-1-j}\frac{(i-1-j)!k}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}=\sum_{k=1}^{i-1-j}\frac{(i-1-j)!}{(k-1)!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}$$
$$=(i-1-j)\sum_{l=0}^{i-2-j}\frac{(i-2-j)!}{l!(i-2-j-l)!}((1-\alpha)\Delta)^{l+1}\alpha^{i-2-j-l}$$
$$=(i-1-j)((1-\alpha)\Delta)\sum_{l=0}^{i-2-j}\frac{(i-2-j)!}{l!(i-2-j-l)!}((1-\alpha)\Delta)^l\alpha^{i-2-j-l}$$
$$=(1-\alpha)\Delta(i-1-j)(\alpha+(1-\alpha)\Delta)^{i-2-j}=(1-\alpha)\Delta(i-1-j)\sigma_x^{i-2-j},$$

which is derived using the change of variable $l = k - 1$. By making use of this formula:

$$\begin{aligned}
&\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\left(\sum_{k=0}^{i-1-j}\frac{(i-1-j)!k}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right)\\
&=(1-\alpha)\Delta\sum_{j=0}^{i-1}\alpha^j(1-\alpha)\Delta(i-1-j)\sigma_x^{i-2-j}=((1-\alpha)\Delta)^2\alpha^{i-2}\sum_{j=0}^{i-1}(i-1-j)\left(\frac{\sigma_x}{\alpha}\right)^{i-2-j}\\
&=((1-\alpha)\Delta)^2\alpha^{i-2}\sum_{l=1}^{i-1}l\left(\frac{\sigma_x}{\alpha}\right)^{l-1}=((1-\alpha)\Delta)^2\alpha^{i-2}\left\{\frac{1-\left(\frac{\sigma_x}{\alpha}\right)^{i-1}}{\left(1-\frac{\sigma_x}{\alpha}\right)^2}-\frac{(i-1)\left(\frac{\sigma_x}{\alpha}\right)^{i-1}}{\left(1-\frac{\sigma_x}{\alpha}\right)}\right\}\\
&=\frac{((1-\alpha)\Delta)^2}{(\alpha-\sigma_x)^2}\alpha^2\alpha^{i-2}\left\{1-\left(\frac{\sigma_x}{\alpha}\right)^{i-1}-(i-1)\left(1-\frac{\sigma_x}{\alpha}\right)\left(\frac{\sigma_x}{\alpha}\right)^{i-1}\right\}\\
&=\frac{((1-\alpha)\Delta)^2}{(\sigma_x-\alpha)^2}\alpha^i\left\{\frac{\alpha^i-\alpha\sigma_x^{i-1}-(i-1)(\alpha-\sigma_x)\sigma_x^{i-1}}{\alpha^i}\right\}\\
&\qquad=\frac{((1-\alpha)\Delta)^2}{((1-\alpha)\Delta)^2}\frac{\alpha^i}{\alpha^i}\left(\alpha^i-\alpha\sigma_x^{i-1}+(i-1)(\sigma_x-\alpha)\sigma_x^{i-1}\right)=\alpha^i-\alpha\sigma_x^{i-1}+(i-1)(\sigma_x-\alpha)\sigma_x^{i-1},
\end{aligned}$$

which uses the change of variable $l = i - 1 - j$ and the general summation formulae:

$$\sum_{l=1}^{n}la^{l-1}=\frac{1-a^n}{(1-a)^2}-\frac{na^n}{1-a}.$$

It follows that

$$\begin{aligned}
&(\sigma_x^i-\alpha^i)+\left(\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!k}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\right)\\
&=(\Delta^i-\alpha^i)+(\alpha^i-\alpha\sigma_x^{i-1}+(i-1)(\sigma_x-\alpha)\sigma_x^{i-1})\\
&\qquad=(\sigma_x-\alpha)\sigma_x^{i-1}+(i-1)(\sigma_x-\alpha)\sigma_x^{i-1}=(\sigma_x-\alpha)i\sigma_x^{i-1}=(1-\alpha)\Delta i\sigma_x^{i-1}.
\end{aligned}$$

Combining these results yields an expression for the $\tilde{\lambda}_t$ terms in the linearization:

$$\begin{aligned}
&\lambda\nu y^{-\lambda}\sum_{i=1}^{\infty}\left\{(\log y)\alpha^i+\left(\frac{y}{x}\right)^{\lambda}\sum_{j=0}^{i-1}(1-\alpha)\Delta\alpha^j\sum_{k=0}^{i-1-j}\frac{(i-1-j)!}{k!(i-1-j-k)!}((1-\alpha)\Delta)^k\alpha^{i-1-j-k}\dots\right.\\
&\left.\times\left((\log x)-(k+1)(\log\delta)\right)\right\}\lambda_{t-i}\\
&\qquad=\lambda\nu\pi\sum_{i=1}^{\infty}\left\{(\log y)\alpha^i+(\log x)\left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i-\alpha^i)-(\log\delta)\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta i\sigma_x^{i-1}\right\}\lambda_{t-i}. \quad [A.3.22]
\end{aligned}$$

Putting together all the terms:

$$\begin{aligned}
\mathsf{u}_t=\lambda\nu\pi\sum_{i=1}^{\infty}\alpha^i\mathsf{y}_{t-i}+\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\Delta\left(\frac{1-\alpha}{1-\sigma_x}\right)\sum_{i=0}^{\infty}\alpha^i\mathsf{x}_{t-i}-\nu\pi\sum_{i=1}^{\infty}\left\{\alpha^i+\left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i-\alpha^i)\right\}\mathsf{u}_{t-i}\\
+\lambda\nu\pi\sum_{i=1}^{\infty}\left\{(\log y)\alpha^i+(\log x)\left(\frac{y}{x}\right)^{\lambda}(\sigma_x^i-\alpha^i)-(\log\delta)\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta i\sigma_x^{i-1}\right\}\lambda_{t-i}.
\end{aligned}$$

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

$$(\mathbb{I}-\alpha\mathbb{L})\sum_{i=1}^{\infty}\alpha^{i}\mathsf{y}_{t-i}=\alpha\mathbb{L}\mathsf{y}_{t};$$
$$(\mathbb{I}-\alpha\mathbb{L})\sum_{i=0}^{\infty}\alpha^{i}\mathsf{x}_{t-i}=\mathsf{x}_{t};$$
$$(\mathbb{I}-\sigma_{x}\mathbb{L})\sum_{i=1}^{\infty}\sigma_{x}^{i}\mathsf{u}_{t-i}=\sigma_{x}\mathbb{L}\mathsf{u}_{t};$$
$$(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}\sum_{i=1}^{\infty}i\sigma_{x}^{i-1}\lambda_{t-i}=\mathbb{L}\lambda_{t}.$$

Multiplying both sides of [A.3.22] by $(\mathbb{I}-\alpha\mathbb{L})(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}$:

$$\begin{aligned}(\mathbb{I}-\alpha\mathbb{L})(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}\mathsf{u}_{t}&=\lambda\nu\pi(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}\alpha\mathbb{L}\mathsf{y}_{t}+\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\Delta\left(\frac{1-\alpha}{1-\sigma_{x}}\right)(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}\mathsf{x}_{t}\\&-\nu\pi(\mathbb{I}-\sigma_{x}\mathbb{L})\left((\mathbb{I}-\sigma_{x}\mathbb{L})\alpha\mathbb{L}+\left(\frac{y}{x}\right)^{\lambda}((\mathbb{I}-\alpha\mathbb{L})\sigma_{x}\mathbb{L}-(\mathbb{I}-\sigma_{x}\mathbb{L})\alpha\mathbb{L})\right)\mathsf{u}_{t}\\&+\lambda\nu\pi\Bigg((\log y)(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}\alpha\mathbb{L}+(\log x)\left(\frac{y}{x}\right)^{\lambda}(\mathbb{I}-\sigma_{x}\mathbb{L})\left((\mathbb{I}-\alpha\mathbb{L})\sigma_{x}\mathbb{L}-(\mathbb{I}-\sigma_{x}\mathbb{L})\alpha\mathbb{L}\right)\dots\\&\qquad-(\log\delta)\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta(\mathbb{I}-\alpha\mathbb{L})\mathbb{L}\Bigg)\lambda_{t}.\end{aligned}$$

Making some simplifications to this expression:

$$\begin{aligned}&(\mathbb{I}-\sigma_{x}\mathbb{L})\left\{(\mathbb{I}-\alpha\mathbb{L})(\mathbb{I}-\sigma_{x}\mathbb{L})+\nu\pi\left(\alpha(\mathbb{I}-\sigma_{x}\mathbb{L})+\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta\mathbb{I}\right)\right\}\mathsf{u}_{t}\\&=\lambda\nu\pi\alpha(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}\mathbb{L}\mathsf{y}_{t}+\lambda\nu\pi\left(\frac{y}{x}\right)^{\lambda}\Delta\left(\frac{1-\alpha}{1-\sigma_{x}}\right)(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}\mathsf{x}_{t}\\&+\lambda\nu\pi\left\{(\log y)\alpha(\mathbb{I}-\sigma_{x}\mathbb{L})^{2}+(\log x)\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta(\mathbb{I}-\sigma_{x}\mathbb{L})-(\log\delta)\left(\frac{y}{x}\right)^{\lambda}(1-\alpha)\Delta(\mathbb{I}-\alpha\mathbb{L})\right\}\mathbb{L}\lambda_{t}.\end{aligned}$$

Equation [3.2] stating the thick-market assumption can be written in log-linear terms as follows:

$$\lambda_{t}=-\theta(\lambda-1)\mathsf{u}_{t}. \qquad [A.3.23]$$

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

$$\mathsf{s}_{t}=\nu_{t}+\pi_{t}. \qquad [A.3.24]$$

Total sales are defined as $S_t=s_tu_t$, or in log-linear terms:

$$\mathsf{S}_{t}=\mathsf{s}_{t}+\mathsf{u}_{t}. \qquad [A.3.25]$$