

The Decision to Move House and Aggregate Housing-Market Dynamics*

L. Rachel Ngai[†]

London School of Economics

Kevin D. Sheedy[‡]

London School of Economics

August 2018

Abstract

Using data on house sales and inventories, this paper shows that housing transactions are driven mainly by listings and less so by transaction speed, thus the decision to move house is key to understanding the housing market. The paper builds a model where moving house is essentially an investment in match quality, implying that moving depends on macroeconomic developments and housing-market conditions. The number of transactions has implications for welfare because each transaction reduces the degree of mismatch for homeowners. The quantitative importance of the decision to move house is shown in understanding the U.S. housing market boom during 1995–2003.

JEL CLASSIFICATIONS: D83; E22; R31.

KEYWORDS: housing market; search and matching; endogenous moving; match quality investment; mismatch.

*We are grateful to the editor and four anonymous referees, and Elliot Anenberg, Johannes Boehm, Paul Carrillo, Adam Guren, Allen Head, Yannis Ioannides, Boyan Jovanovic, Steven Laufer, Kurt Mitman, Chris Pissarides, Morten Ravn, Richard Rogerson, Rob Shimer and Ben Williams for helpful comments. We thank Lucy Canham, Thomas Doyle, and Lu Han for providing data. We also thank seminar participants at Atlanta Fed, Bank of Canada, Bank of England, U. Bern, Boston U. Carlos III, U., CREI, EUI, GWU-FRB Real Estate Seminar, U. Helsinki, U. Hong Kong, IIES, U. Louvain, U. Mannheim, Norwegian Business School, São Paulo School of Economics, Southwest U. of Finance and Economics, Stockholm School of Economics, Sveriges Riksbank, Trinity College Dublin, UCL, the Anglo-French-Italian macroeconomics workshop, Conference on Research on Economic Theory and Econometrics, ESSIM, HULM — St. Louis, Midwest Macro Meeting, NBER Summer Institute, Search and Matching Network annual conference, SED, and U. Toronto for their comments. Rachel Ngai acknowledges support from the British Academy Mid-Career Fellowship. This paper was previously circulated under the title “Moving House”.

[†]LSE, CEPR, and CfM. Email: L.Ngai@lse.ac.uk

[‡]LSE and CfM. Email: K.D.Sheedy@lse.ac.uk

1 Introduction

The number of transactions in the housing market quantifies the reallocation of the housing stock among households. The allocation of housing has consequences for efficiency and welfare because homeowners have idiosyncratic and time-varying preferences for particular houses, with the degree of mismatch between houses and their owners' preferences representing misallocation of the housing stock. Each housing-market transaction generates gains from trade by reducing this mismatch. Given that the housing stock constitutes a substantial proportion of an economy's total capital stock, the number of transactions and the resulting degree of housing misallocation is of first-order importance.

In the presence of search frictions in the housing market, how quickly houses come off the market and how quickly they come on to the market jointly account for the number of transactions. How quickly houses come off the market depends on the time-consuming and costly search process of buyers and sellers. On the other hand, how quickly houses come on to the market depends on the decisions of homeowners to put houses up for sale as well as changes in the housing stock.

The first contribution of this paper is to present evidence establishing that the main driver of transactions is changes in how frequently houses come on to the market. Furthermore, U.S. data show that the dominant factor in explaining the number of houses coming on to the market is homeowners' decisions to move house, rather than changes in the housing stock. Following on from this, understanding the decision to move house is crucial for explaining changes in transactions. The second contribution of the paper is to build a model of moving house. The model is based on the idea that moving house constitutes an investment in the quality of the match between a homeowner and a particular house. This investment entails selling a sufficiently mismatched house and buying one that is a better match. Such a model where moving house is endogenous is necessary to think about how the aggregate moving rate might be affected by macroeconomic and policy variables such as interest rates and taxes. Finally, the model is calibrated to the U.S. economy to study the housing-market boom of 1995–2003.

The claim that the moving rate of homeowners is of much greater importance for transactions than the selling rate of houses on the market can be understood with some minimal empirical discipline combined with a basic stock-flow accounting identity. Compared to the average time homeowners spend in a house (more than a decade), the average time taken to sell a house is very short (a few months). This means the average sales rate is around thirty times larger than the average moving rate, and the stock of houses for sale is about thirty times smaller than the stock of occupied houses. An increase in the sales rate with no change in the moving rate would rapidly deplete the stock of houses for sale leaving little overall impact on transactions. On the other hand, an increase in the moving rate adds to the stock of houses for sale, which increases transactions even with no change in the sales rate.

Using data on sales and inventories of existing single-family homes from the National Association of Realtors (NAR), this paper constructs a time series for listing (houses put up for sale) from a stock-flow accounting identity. Listing rates and sales rates can then be calculated using this data

together with an estimate of the housing stock.¹ The resulting listing rates provide estimates of homeowners' moving rates.² A counterfactual exercise derives the number of transactions that would have occurred if only one of the sales rate, moving rate, and housing stock had varied as in the data while the other two were constant at their average values. This exercise confirms the claim that the moving rate is the dominant factor in explaining changes in transactions. It also shows that changes in the housing stock capture less than a third of the rise in transactions. Since changes in the owner-occupied housing stock can occur either through new construction or by houses switching from being rented to being owner occupied, this counterfactual exercise reveals the limited role played by the construction boom and the rise in the homeownership rate in accounting for the boom in transactions.³

To understand what might drive changes in the moving rate, this paper presents a search-and-matching model of endogenous moving. Central to the model is the idea of match quality: the idiosyncratic values homeowners attach to the house they live in. This match quality is a persistent variable subject to occasional idiosyncratic shocks, representing life events such as changing jobs, marriage, divorce, and having children.⁴ These shocks degrade existing match quality, following which homeowners decide whether to move. Eventually, after sufficiently many shocks, current match quality falls below a 'moving threshold' that triggers moving to a new house and a renewal of match quality.

Moving house is a process with large upfront costs that is expected to deliver long-lasting benefits, and hence is sensitive to macroeconomic and policy variables such as interest rates that influence other investment decisions.⁵ These variables affect the threshold for existing match quality that triggers moving. Thus, while idiosyncratic shocks are the dominant factor in moving decisions at the individual level, changes in the moving threshold lead to variation in the aggregate moving rate. In contrast, in a model of exogenous moving, the aggregate moving rate is solely determined by the arrival rate of an idiosyncratic shock that forces moving.

The endogeneity of moving generates new transitional dynamics that are absent from models imposing exogenous moving. Endogenous moving means that those who choose to move are not a random sample of the existing distribution of match quality: they are the homeowners who were only moderately happy with their match quality. Together with the persistence over time of existing match quality, endogenous moving gives rise to a 'cleansing effect'. An aggregate shock that changes the moving threshold leads to variation in the degree of cleansing of lower-quality matches. Since match quality is a persistent variable, more cleansing now leads to less cleansing in the future, which

¹The sales rates of houses on the market are calculated using a measure of inventories that reflects the actual number of houses available for sale, as opposed to the usual approach of taking the number of vacant houses for sale.

²In practice, the measured listing rate may not be the same as the moving rate due to houses being withdrawn from the market and subsequently relisted. This issue is addressed in [section 2.3](#).

³The homeownership rate, as defined in the Census, is the stock of owner-occupied homes divided by the stock of all occupied homes (including those that are rented).

⁴These are the main reasons for moving according to the American Housing Survey. Match quality is modelled as a persistent variable because there are many dimensions to what is considered a desirable house, and a life event will typically affect some but not all of those dimensions.

⁵The implication that interest rates have a negative impact on mobility is consistent with the empirical evidence presented by [Quigley \(1987\)](#) and [Ferreira, Gyourko and Tracy \(2010\)](#).

implies overshooting of the moving rate and transactions.⁶

The importance of the decision to move house is shown through a quantitative application of the model to the boom in the U.S. housing market between 1995 and 2003. The period 1995–2003 in the U.S. housing market is noteworthy as one of rapidly rising activity and prices. Using NAR data, three stylized facts emerge during those years: transactions rise, houses are selling faster, and houses are put up for sale more frequently. The increase in the listing rate is consistent with the independent empirical finding by [Bachmann and Cooper \(2014\)](#) of a substantial rise in the own-to-own moving rate using household-level PSID data. This period featured a number of developments that have implications for moving decisions according to the model, such as the post-1995 surge in productivity growth, the rise of internet-based property search, and cheaper and more easily accessible credit. These developments are represented in the model by changing parameters, and the model is solved analytically to conduct a series of simulations.

The implications of the model are that all three developments unambiguously increase the moving rate. This rise in the moving rate works through a higher moving threshold, which means that homeowners' tolerance of mismatch is reduced. An increase in productivity growth raises income and increases the demand for housing, which raises the marginal benefit of a better match and increases the incentive to move to a house with a higher match quality. The adoption of internet technology reduces search frictions, making it cheaper for homeowners to move in order to invest in a better match. Finally, lower mortgage rates, interpreted as a fall in the rate at which future payoffs are discounted, create an incentive to invest in improving match quality because the capitalized cost of moving is reduced.

Data on the hazard function for moving house are derived from the American Housing Survey (AHS) to pin down the key parameters of the model. The empirical hazard function displays a u-shaped pattern where the aggregate hazard rate for moving first declines and then rises as duration of since the previous move increases. In order to rationalize the downward sloping portion of the hazard function, the model is enriched to allow for heterogeneous homeowners with different arrival rates and sizes of idiosyncratic shocks. Those with a higher arrival rate or a larger idiosyncratic shock size will have a higher moving rate. Given that they will move sooner on average, the aggregate moving rate can be decreasing in duration owing to a compositional effect. The increasing section of the hazard function can be explained owing to the persistence of match quality: an idiosyncratic shock that does not immediately trigger moving nonetheless makes it more likely a subsequent shock will trigger moving.

The parameters of the idiosyncratic shock distributions are chosen to minimize the sum of squared deviations between the empirical and model-implied hazard function. The remaining set of parameters is calibrated to match features of the U.S. housing market. More specifically, the

⁶The persistence of match quality and its implications differentiate the endogenous moving decision here from the endogenous job-separation decision in the labour market studied by [Mortensen and Pissarides \(1994\)](#). There, when an existing match is subject to an idiosyncratic shock, a new match quality is drawn independently of the match quality before the shock, that is, the stochastic process for match quality is 'memory-less'. Here, idiosyncratic shocks degrade match quality, but an initially higher-quality match remains of a higher quality than a lower-quality match hit by the same shock.

cost parameters match data on transaction costs, search costs, and maintenance costs; while the parameters related to the search process match average time-to-sell and the average number of viewings per sale.

The three developments together imply a 42% rise in the moving rate in the short run and a 25% rise in the long run, with the empirical moving rate rising by 34%. There is a considerable amount of overshooting due to the cleansing of the match quality distribution, and this continues for around a decade. The model can thus generate a large and long-lasting increase in transactions of 24%, explaining most of the 27% rise observed in the data. It is not obvious whether the decline in mortgage rate was expected to persist, but even after removing the effects of lower mortgage rates, the first two developments still imply a 18% increase in transactions, accounting for two thirds of the rise in the data. With house prices determined through simple Nash bargaining, the predicted increase in the moving rate and transactions is accompanied by an increase in prices.

There is a large literature (starting from [Wheaton, 1990](#), and followed by many others) that studies frictions in the housing market using a search-and-matching model as done here. This extensive literature is surveyed by [Han and Strange \(2015\)](#).⁷ The novel feature of this paper is in studying moving house and showing its importance in understanding the dynamics of transactions. The empirical finding using NAR data that changes in the number of listings are crucial for understanding the dynamics of transactions is confirmed in a recent paper by [DeFusco, Nathanson and Zwick \(2017\)](#) using U.S. data from CoreLogic.⁸

The focus of the paper is on analysing the dynamics of transactions resulting from homeowners moving from one home to another within the same housing market. This type of transaction is shown to be the main source of fluctuations among all housing transactions by [Anenberg and Bayer \(2013\)](#) using data from Los Angeles between 1988 and 2008. Taking moving as the exogenous outcome of a mismatch shock, they study homeowners' choice of whether to buy first or sell first. The order of transactions is also the focus of [Moen, Nenov and Sniekers \(2015\)](#), who emphasize strategic complementarity in the sequencing decisions of mismatched owners as a source of multiple equilibria that could explain housing cycles. They also make use of two datasets from Copenhagen to construct a time series for the proportion of existing owner-occupiers who buy first or sell first. The present paper complements these two papers by analysing the moving decision itself, abstracting from decisions about whether to buy first or sell first. A key contribution is to allow for an endogenous degree of housing mismatch: instead of imposing an exogenous two-state classification of homeowners as either matched or mismatched, match quality is a continuous variable here.

[Burnside, Eichenbaum and Rebelo \(2016\)](#), [Guren \(2014\)](#), and [Hedlund \(2016\)](#) also explicitly allow for a moving decision.⁹ The last two papers do not quantify changes in transactions, which

⁷See, for example, [Albrecht, Anderson, Smith and Vroman \(2007\)](#), [Anenberg and Bayer \(2013\)](#), [Caplin and Leahy \(2011\)](#), [Coles and Smith \(1998\)](#), [Díaz and Jerez \(2013\)](#), [Krainer \(2001\)](#), [Head, Lloyd-Ellis and Sun \(2014\)](#), [Moen, Nenov and Sniekers \(2015\)](#), [Ngai and Tenreyro \(2014\)](#), [Novy-Marx \(2009\)](#), [Piazzesi and Schneider \(2009\)](#), and [Piazzesi, Schneider and Stroebel \(2015\)](#). See also [Davis and Van Nieuwerburgh \(2015\)](#) for a survey of housing and business cycles, including models without search frictions.

⁸[Díaz and Jerez \(2013\)](#) also show that housing-supply and moving-rate shocks are essential in understanding the observed cyclicality of transactions. Both of these are closely related to changes in the listing rate.

⁹[Karahan and Rhee \(2013\)](#) use a similar framework to [Hedlund \(2016\)](#) and study when a homeowner chooses to

is the focus of the present paper.¹⁰ Burnside, Eichenbaum and Rebelo (2016) build a model where homeowners have different beliefs about the value of owning (relative to renting) and thus have different probabilities of putting their houses up for sale. As in the present paper, their mechanism also implies a positive co-movement between transactions and prices, which is an important feature of the housing market first emphasized by Stein (1995). Their quantitative model generates a change in transactions about 17 times smaller than the change in prices, whereas this ratio is approximately 2 in the baseline results of the present paper, which is much closer to the data for the period 1995–2003.¹¹ While they give a key role to buyers’ speculative motives during the substantial house-price appreciation of 2003–2006, this paper provides a complementary mechanism based on the incentives of homeowners to invest in match quality.¹² For this reason, in the application to the housing boom, the focus here is on the period 1995–2003.¹³

The plan of the paper is as follows. Section 2 uses NAR data to substantiate the claim that changes in the moving rate are the key determinant of housing-market dynamics. Section 3 presents the model of endogenous moving, and section 4.3 derives the equilibrium analytically. Section 5 calibrates the model to study the quantitative importance of the decision to move during the 1995–2003 U.S. housing boom. Section 6 concludes.

2 The importance of changes in the moving rate

This section presents evidence showing that the decision of homeowners to put their houses up for sale is crucial for understanding the dynamics of housing transactions.

2.1 The basic idea

A stock-flow accounting identity is a natural starting point when thinking about any market with search frictions:

$$\dot{U}_t = N_t - S_t, \quad \text{where } N_t = n(K_t - U_t) \text{ and } S_t = sU_t, \quad [2.1]$$

move to another region to look for a job.

¹⁰Hedlund (2016) focuses on the cross-sectional correlations between leverage, listing prices, and time-on-the-market and shows that foreclosures can be caused by housing illiquidity without requiring negative equity. Guren (2014) focuses on how the existence of a concave demand curve can amplify price insensitivity in the housing market, helping to account for the positive autocorrelation of house-price changes (momentum). A recent paper by Anenberg and Laufer (2017) argues that some of the positive autocorrelation of price changes could be due to the house-price index being based on prices at closing dates, rather than contract dates.

¹¹To be precise, Burnside, Eichenbaum and Rebelo (2016) predict a 289% increase in prices and a 17% increase in transactions for the period 1995–2006 (Figure 5 in their paper). Note that they compare their house price prediction to the 144% increase in the Case-Shiller price index for Boston, Los Angeles, and San Francisco. Using data from FHFA and NAR for the U.S. as a whole, the increases in real house prices and transactions are both around 60%.

¹²DeFusco, Nathanson and Zwick (2017) also emphasize the role of investment motives in the dynamics of prices and transactions volume during the housing bubble period. In line with our findings, they show that the increase in transaction volume is mainly due to the increase in the listing rate, which is modelled as a shortening of the holding periods of real estate investors.

¹³As shown in Figure 9a of Burnside, Eichenbaum and Rebelo (2016), the fraction of affirmative responses in the Michigan Survey of Consumers to the question “Is it a good time to buy a house because it is a good investment?” was either falling or flat in the period 1997–2003.

with U_t denoting the number of existing houses for sale, \dot{U}_t the derivative of U_t with respect to time t , N_t the number of existing houses put on the market for sale, and S_t the number of transactions. The stock of all owner-occupied houses (including those for sale) is K_t . The rate at which homeowners decide to move is n , so $n(K_t - U_t)$ is the number of existing houses put on the market for sale. The rate at which houses for sale are sold is s . The variable U_t is defined as existing houses for sale (excluding newly constructed houses for sale) to make it consistent with the data from NAR used below. The growth rate of K_t is denoted by g , where changes in K_t are due to newly constructed houses being bought by owner-occupiers or existing houses switching from being rented to being owner occupied.

Let u_t denote existing houses for sale U_t as a fraction of the total housing stock K_t , which satisfies the following differential equation:

$$\dot{u}_t = n(1 - u_t) - (s + g)u_t. \quad [2.2]$$

Given the flow rates s , n , and g , there is a steady state for the fraction of houses for sale:

$$u = \frac{n}{n + s + g},$$

and transactions S_t^* when u_t is at its steady-state value are:

$$S_t^* = \frac{sn}{s + n + g} K_t. \quad [2.3]$$

Convergence to the steady state for u_t occurs at rate $s + n + g$ (the coefficient of u_t in 2.2), and given that houses sell relatively quickly (a few months on average), s is large enough that transitional dynamics are of limited importance. Therefore, understanding the evolution of transactions over any period of time longer than a few months is mainly a question of understanding how changes in s , n , g , and K_t affect equation (2.3). The total derivative of transactions S_t^* is:

$$\frac{dS_t^*}{S_t^*} = \frac{s + g}{s + n + g} \frac{dn}{n} + \frac{n + g}{s + n + g} \frac{ds}{s} - \frac{g}{s + n + g} \frac{dg}{g} + \frac{dK_t}{K_t}. \quad [2.4]$$

This implies the relative size of the effects on transactions of a one percent change in n and a one percent change in s largely depends on the ratio of the sales rate s to the moving rate n , given that g is rather small empirically. Any change in the level of K_t has a proportional effect on transactions.

The average time taken to sell a house is $1/s$, and the average time homeowners spend living in a house is $1/n$. The impact of a change in the moving rate relative to the same proportional change in the sales rate is therefore related to the ratio of the average time homeowners spend living in a house to the average time to sell. The former is more than a decade and the latter is a few months, suggesting s is around 200%, n is about 7%, and the ratio of the two is approximately 30. Any plausible value of g would be far smaller than $s + n$. Since this means $(n + g)/(s + n + g)$ is very small, huge changes in sales rates would be required to have any significant lasting effect on transactions.

Intuitively, with no change in the moving rate, the stock of houses for sale would be rapidly depleted by faster sales, leaving overall transactions only very slightly higher. On the other hand, since $(s + g)/(s + n + g)$ is close to one, changes in the moving rate can have a large and lasting impact on transactions. This is because even if the moving rate increased slightly, as the stock of potential movers is so large relative to the stock of houses for sale (the ratio $(1 - u)/u$ is equal to $(s + g)/n$), this can have a sustained impact on the number of homeowners who move.

This simple exercise establishes that any attempt to understand sustained changes in transactions through only changes in the sales rate will founder. This leaves two potential explanations: changes in the moving rate, or changes in the housing stock. The next section uses U.S. data to show that changes in the housing stock contribute relatively little to understanding transactions compared to changes in the moving rate.

2.2 Empirical evidence

A quantitative analysis of stocks and flows in the housing market can be performed by using data on transactions and inventories of unsold houses to construct a measure of new listings (the number of houses put up for sale). NAR provides monthly estimates of transactions during each month and inventories of houses for sale at the end of each month for existing homes including single-family homes and condominiums.¹⁴ The NAR data on transactions and inventories are for existing homes, so sales of newly constructed houses are excluded. The focus here is on data for single-family homes, which represent around 90% of total sales of existing homes. Monthly data on transactions and inventories covering the period from January 1989 to June 2013 are first deseasonalized.¹⁵ The data are then converted to quarterly series to smooth out excessive volatility owing to possible measurement error. Quarterly transactions are the sum of the monthly transaction numbers, and quarterly inventories are the level of inventories at the beginning of the first month of a quarter.

Let N_t denote the inflow of houses that come on to the market during quarter t (new listings), and let S_t denote transactions (the outflow from the market) during that quarter. If I_t denotes the beginning-of-quarter t inventory (or end-of-quarter $t - 1$), the stock-flow accounting identity is:

$$N_t = I_{t+1} - I_t + S_t. \tag{2.5}$$

A quarterly listings series N_t is constructed to satisfy the accounting identity above.¹⁶ Assuming inflows N_t and outflows S_t both occur uniformly within a time period, the average number of houses U_t available for sale during quarter t is equal to:

$$U_t = I_t + \frac{N_t}{2} - \frac{S_t}{2} = \frac{I_t + I_{t+1}}{2}. \tag{2.6}$$

¹⁴The methodology and recent data are available at <http://www.realtor.org/research-and-statistics/housing-statistics>.

¹⁵Multiplicative monthly components are removed from the data.

¹⁶It is possible that listings as measured in this way do not perfectly correspond to moving because unsold houses might be withdrawn from sale, or because S_t and I_t include non-owner-occupied houses. The first issue is addressed in [section 2.3](#) and the second in [appendix A.1](#).

Since the time series for inventories I_t 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). This way of measuring of houses for sale is an improvement on the usual measure based on the ‘vacant for sale’ series from the American Housing Survey (AHS) because the most houses for sale are not actually vacant (on average, ‘vacant for sale’ is less than half of inventories). More importantly, there is considerable variation over time in the ratio between ‘vacant for sale’ and inventories. Using the constructed series U_t for houses for sale, the sales rate is $s_t = S_t/U_t$.

The listing rate n_t is defined as the ratio of new listings N_t to the stock of owner-occupied houses not already for sale, that is, $n_t = N_t/(K_t - U_t)$. The housing stock K_t is defined as the stock of single-family homes excluding renter-occupied houses and homes for rent.¹⁷ These data are available only at a biennial frequency, so linear interpolation is used to produce a quarterly series. Figure 1 plots transactions, listings, houses for sale, and the total housing stock in the upper panel, and the sales and listing rates and the fraction of houses for sale in the lower panel. All series are plotted as differences in log points relative to the first quarter of 1989.

The figure shows that both the sales rate and the listing rate rise and fall with transactions. At first sight, the rise in the listing rate between 1993 and 2007 might seem inconsistent with the long-run decline in the U.S. mobility rate. Empirical work by Bachmann and Cooper (2014) has shown that the declining mobility rate is accounted for by a fall in the rent-to-rent moving rate, while they find that the own-to-own moving rate actually rises over this period.¹⁸

A simple counterfactual exercise is used to quantify the importance of the listing rate in understanding changes in transactions. Given that convergence to the steady state in equation (2.3) occurs within a few months, the evolution of transactions over time can be understood through the lens of the following equation:

$$S_t^* = \frac{s_t n_t}{s_t + n_t + g_t} K_t, \tag{2.7}$$

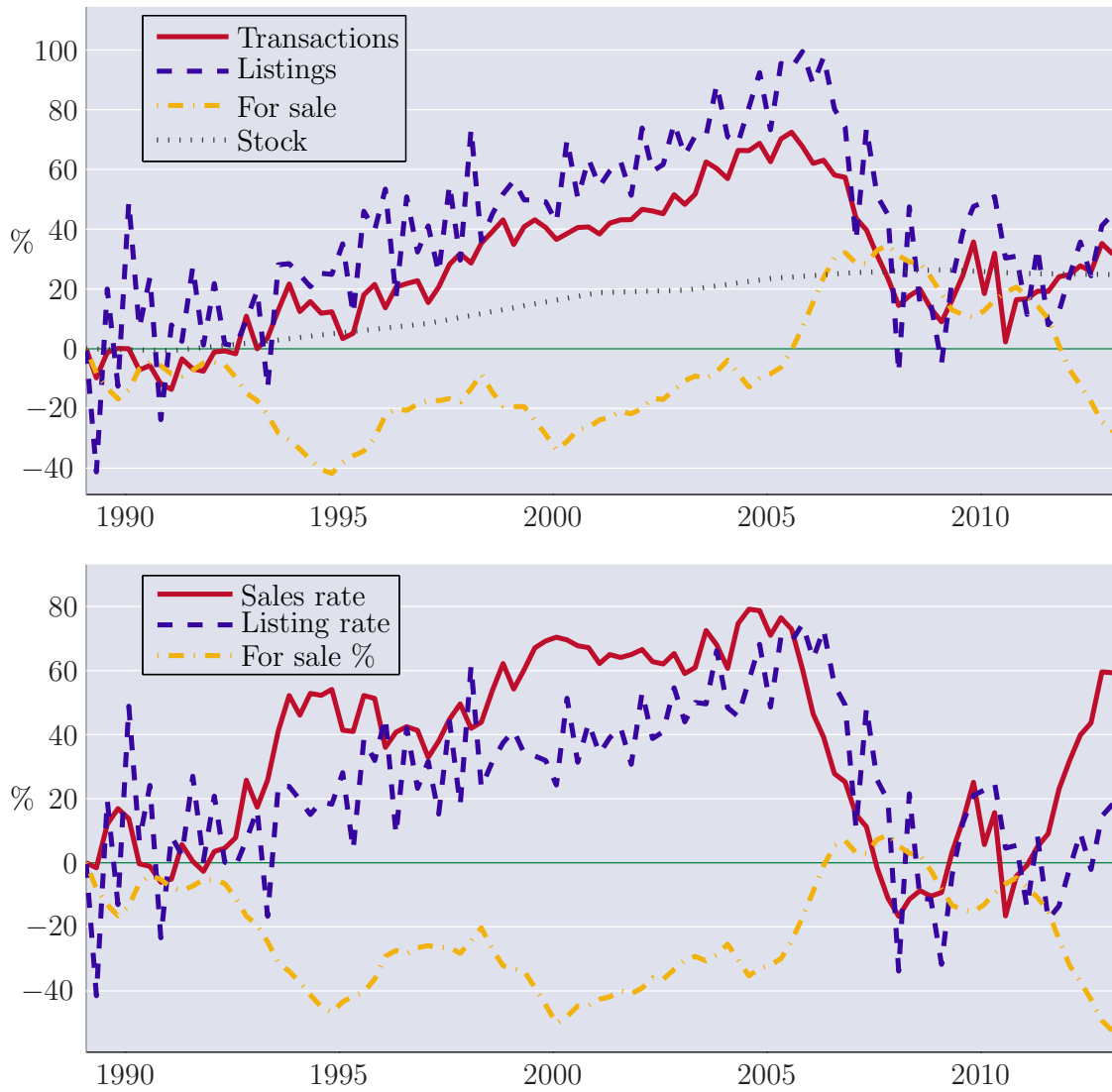
where s_t and n_t are the empirical sales and listing rates and $g_t = (K_t - K_{t-1})/K_{t-1}$ is the growth rate of the housing stock. The variable S_t^* is what the steady-state transactions volume would be at each point in time given the (time-varying) sales and listing rates and the housing stock. The correlation between S_t^* and actual transactions S_t is very high (the correlation coefficient is 0.90), which is not surprising given that convergence to the steady-state fraction of unsold houses is fast.

To see the relative importance of the sales rate, the listing rate, and the housing stock, consider

¹⁷The housing stock is the sum of one-unit structures in the ‘owner occupied’, ‘vacant for sale’, and ‘vacant, but rented or sold’ categories of Table 1A–1 in the AHS. The third category is included because it is likely houses that are sold are the dominant component of that category. In any case, the third category is tiny compared to the first two.

¹⁸Bachmann and Cooper (2014) use an approach similar to many studies of the labour market to compute transition rates between and within the renter-occupied and owner-occupied segments of the housing market using data from the PSID. As shown in their Figure 13, the own-to-own moving rate rises by approximately 30% during the period considered here. Moreover, they point out that the CPS data used by most studies of mobility do not include information on households’ previous tenancy status, so they cannot be used to compute disaggregated moving rates. They show that their aggregate moving rate is comparable to those derived from CPS data, suggesting that the disaggregated moving rates they compute using PSID data are a good representation of U.S. trends.

Figure 1: *Housing-market activity*



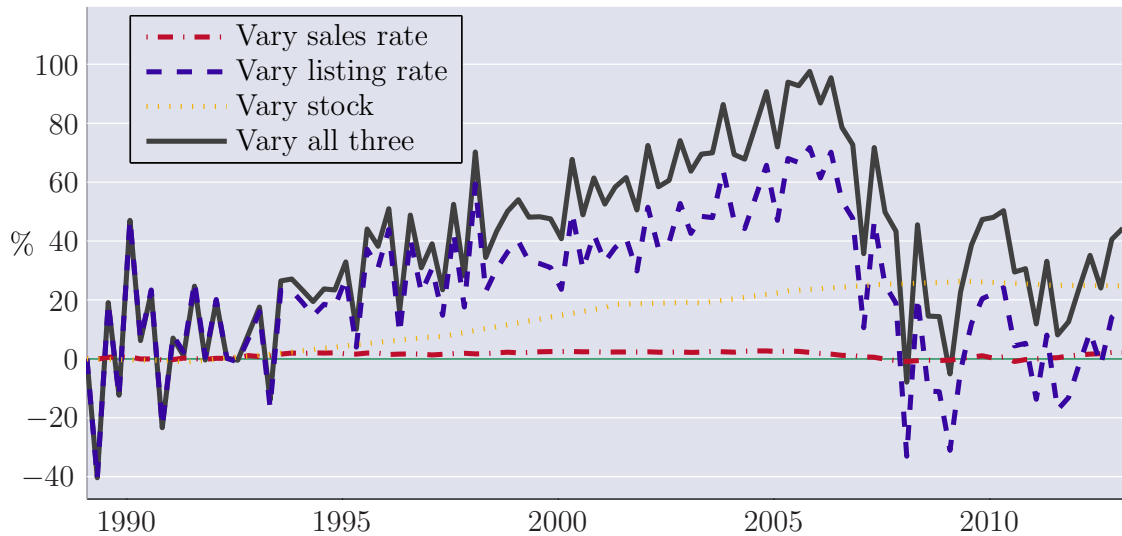
Notes: Series are logarithmic differences from the initial data point. Monthly data (January 1989–June 2013), seasonally adjusted, converted to quarterly series. Definitions are given in [section 2.2](#).
Sources: National Association of Realtors, American Housing Survey.

what (2.7) would be in turn if only one variable at a time (sales rate s_t , listing rate n_t , and stock of houses for sale K_t) is allowed to vary as in the data:

$$S_{s,t}^* = \frac{s_t \bar{n}}{s_t + \bar{n}} \bar{K}, \quad S_{n,t}^* = \frac{\bar{s} n_t}{\bar{s} + n_t} \bar{K}, \quad \text{and} \quad S_{K,t}^* = \frac{\bar{s} \bar{n}}{\bar{s} + \bar{n} + g_t} K_t,$$

where \bar{s} , \bar{n} , and \bar{K} are the average sales rate, listing rate, and stock of houses for sale. The time series of S_t^* , $S_{s,t}^*$, $S_{n,t}^*$, and $S_{K,t}^*$ are plotted in [Figure 2](#) below as log differences. The series $S_{n,t}^*$ (the blue line) closely resembles S_t^* (the grey line), whereas $S_{s,t}^*$ (the red line) misses much of the variation in S_t^* . It is striking, yet consistent with the basic idea set out above, that even large increase in the sales rate (see [Figure 1](#)) cannot account for the large increase in transactions.

Figure 2: Counterfactuals for transactions in steady state



Notes: The construction of these series is described in [section 2.2](#). The series are reported as log differences relative to their initial values.

The counterfactuals in [Figure 2](#) assume that the fraction of houses for sale adjusts instantaneously to its new steady-state value when the sales or listing rate changes. The counterfactual exercise can also be done without requiring any variable to have adjusted to its steady-state value. Note that equations (2.5) and (2.6) and the definitions of the sales and listing rates imply the following identity:

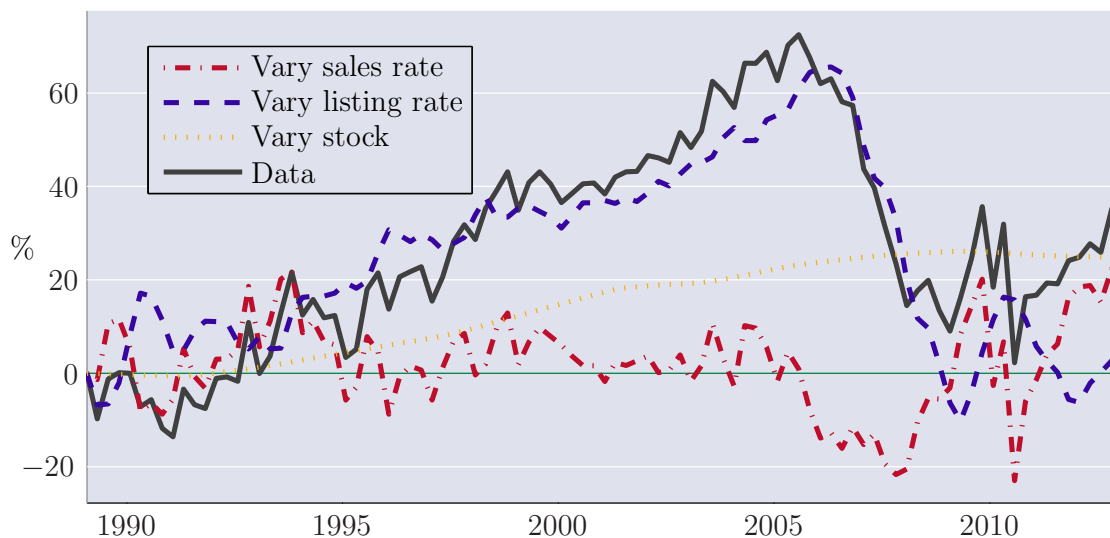
$$I_{t+1} = \frac{\left(1 - \frac{s_t + n_t}{2}\right) I_t + n_t K_t}{1 + \frac{s_t + n_t}{2}}. \quad [2.8]$$

Given time series for the sales rate s_t , the listing rate n_t , and the total housing stock K_t , this equation can be applied recursively to generate a series for inventories I_t starting from an initial value I_1 . The initial value of inventories is assumed to be at its steady state, that is, $I_1 = n_1 K_1 / (s_1 + n_1 + g_1)$, though for subsequent dates t there is no such requirement. With the series for I_t , houses for sale U_t can be calculated using (2.6), which yields a counterfactual series for transactions using $S_t = s_t U_t$. The actual transactions series and the three counterfactuals out of steady state are shown in [Figure 3](#). The findings are largely consistent with those in [Figure 2](#) in that variation in the listing rate is crucial for understanding the actual transactions series.¹⁹

Holding both the sales and listing rates constant, the series $S_{K,t}^*$ (the orange line in both [Figure 2](#) and [Figure 3](#)) shows that changes in the housing stock account for less than a third of the rise in transactions during the boom. This reveals the limited role played by the construction boom and

¹⁹Note the claim that sales rates have relatively little impact on transactions holds everything else constant. It does not imply the correlation between time-on-the-market and transactions is weak. In fact, the correlation between the sales rate and transactions is 0.74, while the correlation between the listing rate and transactions is 0.79. The high correlation between the sales rate and transactions is consistent with the claim because the sales rate is highly correlated with the listing rate (correlation coefficient 0.67).

Figure 3: *Actual and counterfactual transactions out of steady state*



Notes: The series are plotted as log differences relative to their initial values.

the rise in the homeownership rate during the period when transactions were increasing. The homeownership rate is defined as the stock of owner-occupied homes divided by the stock of all occupied homes (including those that are rented). Changes in the housing stock as defined and measured earlier can thus occur either through new construction or houses switching from being rented to being owner occupied. Both the construction boom and the rise in the homeownership rate are therefore captured by the housing stock series, and this accounts for relatively little of the change in transactions.

2.3 Withdrawals and relistings

In the previous section, all outflows from the stock of unsold houses were due to sales. However, some houses are withdrawn from the market without being sold. Since sales and inventories are measured directly in the NAR data, the accounting identity (2.5) will classify withdrawals as ‘negative listings’. Some houses that are withdrawn from the market will subsequently be relisted, and these relistings will also be part of the overall measure of listings derived from the accounting identity (2.5).

Using 1.9 million listings for nine large MSAs from the Multiple Listing Service (MLS) database between 2008 and 2012, [Anenberg and Laufer \(2017\)](#) find that approximately half of all delisted properties were not sold at the point of delisting. However, many of these delisted properties were relisted within six months. [Anenberg and Laufer \(2017\)](#) find that the fraction relisted is around 60%. The analysis of repeat listings in [Carrillo and Williams \(2015\)](#) reveals a similar pattern of delisting and relisting using 1.4 million listings for fourteen MSAs from the MLS database during 2004–2013. Figures 1 and 2 of [Carrillo and Williams \(2015\)](#) report the fraction of properties listed in a particular period that were ‘not sold’ by the first quarter of 2013, that is, where the listing had

either expired or the property was withdrawn from the market. Note that this is a measure of the fraction of withdrawals for a particular cohort of properties that enter the market, rather than a measure of all withdrawals that occur at a specific date. About 50% of these ‘not sold’ properties are subsequently relisted.

If listings N_t is interpreted as the net change in listings then the stock-flow accounting identity (2.1) for U_t remains valid. Net listings N_t can be broken down into the following components:

$$N_t = M_t - W_t + R_t, \quad \text{where } M_t = m(K_t - U_t).$$

Listings associated with a new decision to move house are M_t , and m is the corresponding moving rate. W_t denotes withdrawals of houses from the market without a sale, and R_t denotes relistings of houses that previously failed to sell.

The occurrence of withdrawals implies there is a subset of houses $K_t - U_t$ not currently on the market that have previously failed to sell, but where the owner hopes to make another attempt to sell in the future. The number of such houses is denoted by F_t , which has the following stock-flow accounting identity:

$$\dot{F}_t = W_t - R_t - A_t - mF_t, \quad \text{where } W_t = wU_t, \quad R_t = \rho F_t, \quad \text{and } A_t = \alpha F_t.$$

The variable W_t denotes the inflow due to withdrawals, which are assumed to occur at rate w while a house is on the market. Outflow can occur because of relistings, which occur at rate ρ , or because of reasons that would have triggered a new moving attempt irrespective of the past failed attempt (rate m as before), or finally because the homeowner entirely gives up on selling the house (rate α). The link between the moving rate m and the net listing rate n will depend on the fraction of eventually relisted houses among those that have failed to sell, denoted by β :

$$\beta = \frac{\rho}{\rho + \alpha + m},$$

which is equal to the relisting rate ρ relative to the total outflow rate $\rho + \alpha + m$.

Similarly, the link between m and n will also depend on the fraction of listings that end with a withdrawal rather than a sale, denoted by ϕ :

$$\phi = \frac{w}{s + w}.$$

Figures 1 and 2 of [Carrillo and Williams \(2015\)](#) provide evidence that ϕ tends to be higher when time-to-sell ($1/s$) is high. As can be seen from the formula above, this finding does not reveal whether the withdrawal rate w rises, falls, or remains the same when the sales rate s varies. Even if w were completely constant, a lower sales rate automatically means that houses will linger on the market for longer, making an eventual withdrawal more likely and thus increasing ϕ .

Given β and ϕ , there is a simple link between the net listing rate n and the moving rate m . In

the case where $g = 0$ this is:

$$n = \frac{m}{1 + \frac{\phi(1-\beta)}{1-\phi}}. \quad [2.9]$$

The general formula applicable for any g is derived in [appendix A.2](#), which also shows that as g is very small relative to s and n , the simple formula with $g = 0$ provides a very good approximation. Conditional on the net listing rate n given by (2.9), all the analysis done in [section 2.2](#) still applies. The net listing rate n now depends not only on the moving rate m , but also on withdrawals ϕ and relistings β . However, observe that the only reason withdrawals ϕ matter is because not all houses withdrawn from sale are relisted ($\beta < 1$). If $\beta = 1$ then the net listing rate n equals the moving rate m irrespective of withdrawals and all the earlier conclusions would be identical.

Since the stock-flow accounting identity (2.5) continues to hold with the new interpretation of N_t , the net listing rate n_t can be recovered from the data exactly as before. But equation (2.9) shows that finding the moving rate m_t requires knowing the fractions of withdrawals ϕ_t and relistings β_t . There is no publicly available time series for β_t , either at the national or local level. All that can be done here is to set β_t to the average relisting fraction found in the studies by [Anenberg and Laufer \(2017\)](#) and [Carrillo and Williams \(2015\)](#), which are in the range from 50% to 60%. $\beta = 0.5$ is taken as the baseline value in what follows because any higher value would only reduce the difference withdrawals make to the earlier analysis.

For the withdrawal fraction ϕ_t , [Carrillo and Williams \(2015\)](#) plot time series for five MSAs from 2004 to 2013, and one county for the period 1997–2010, but no national-level data is publicly available. The series are plotted alongside time-to-sell, and the figures suggest that ϕ_t and time-to-sell move in the same direction with an elasticity of approximately 0.6.²⁰ Taking this elasticity and the national-level time-series data on the sales rate from NAR (the inverse of time-to-sell), an estimated national-level time series for ϕ_t is derived. With ϕ_t and β , the time series for the moving rate m_t can be recovered from knowledge of n_t using (2.9).

A counterfactual exercise similar to that from [section 2.2](#) can be performed using the time series for the sales rate s_t , the moving rate m_t , and withdrawals ϕ_t . The first counterfactual varies the sales rate s_t and holds m_t and ϕ_t constant at their average values. The second counterfactual varies the moving rate and holds s_t and ϕ_t constant. The third counterfactual varies withdrawals ϕ_t and holds s_t and m_t constant.²¹ Using the method in [appendix A.2](#) for the case where $g_t \neq 0$, another counterfactual can be calculated where the housing stock K_t varies while s_t , m_t , and ϕ_t are held constant. Each of these counterfactuals implies a hypothetical series for the listing rate n_t , and this can be used to compute the implied level of transactions using the formula in (2.7).

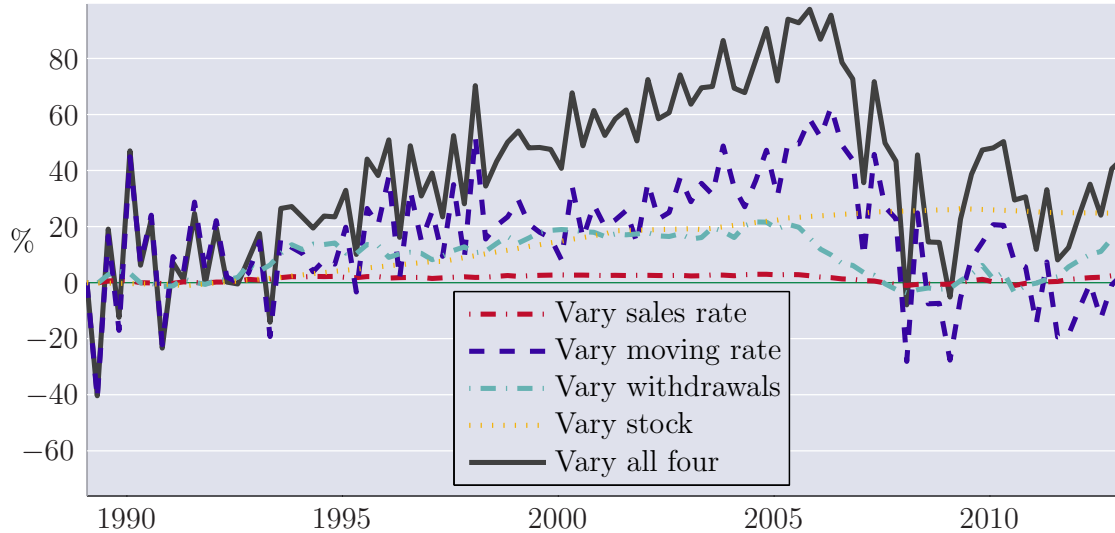
The counterfactuals for steady-state transactions are shown in [Figure 4](#). This generalizes the earlier analysis in [Figure 2](#) by accounting for withdrawals and relistings. As before, the moving rate is very important in explaining the observed variation in transactions. But now some significant

²⁰Ideally, this elasticity would be estimated using a regression, and using data with a broader geographical coverage.

²¹The ϕ_t series is generated from the sales rate series, but after deriving the ϕ_t , the sales rate is held constant in this counterfactual.

portion of the variation in transactions can be explained even when the moving rate is constant. A comparison of the first and third counterfactuals shows that this comes from changes in the fraction of withdrawals, not from changes in the sales rate itself.

Figure 4: *Counterfactual transactions after accounting separately for withdrawals*



Notes: The series are plotted as log differences relative to their initial values.

What lies behind the difference between the listing rate and moving rate counterfactuals in [Figure 2](#) and [Figure 4](#)? The importance of the moving rate in explaining changes in transactions declines because less of the variation in the listing rate is attributed to the moving rate, not because a given change in the moving rate has a smaller effect on transactions. Equation (2.9) shows that the elasticity of n with respect to m is always one, irrespective of the values of ϕ and β , and the elasticity of transactions with respect to n is the same as before.

The evidence in [Carrillo and Williams \(2015\)](#) shows that withdrawals are low when the market is strong and high when the market slows down, in other words, a positive correlation between ϕ_t and time-to-sell. This pattern is built into the analysis above by imposing a positive elasticity of ϕ_t with respect to the time to sell. It is important to note that relistings β_t are also likely to follow the same pattern, that is, relistings would be high when time-to-sell is high. This is because when it takes a long time to sell a property, the listing is either more likely to reach its expiry date, or the seller becomes more likely to withdraw the property so as to avoid the stigma of the property remaining unsold for too long. In both cases, the seller would be expected to relist the property soon after to show the property as new to the market. By fixing relistings β_t to be a constant instead of varying it in line with time-to-sell, [Figure 4](#) provides a lower bound for the importance of the moving rate.

The elasticity of ϕ_t with respect to time-to-sell used in this analysis was based on a small set of data. If the elasticity were higher, the importance of withdrawals would increase and the importance of the moving rate would decline. As explained above, this would be due to claiming that changes in the moving rate had been smaller. However, direct evidence from [Bachmann and Cooper \(2014\)](#)

shows that the own-to-own moving rate increased by 33% during the period 1995–2003. This is close to the increase in the listing rate n of 34% found in the NAR data, suggesting withdrawals ϕ_t could not have fallen by too much in response to the decrease in time-to-sell during that period. That would mean the elasticity of ϕ with respect to time-to-sell cannot be too large.

In summary, this section has shown evidence that understanding changes in moving rates is important in explaining changes in the number of transactions in the housing market.

3 A model of investment in housing match quality

This section presents a search-and-matching model of the housing market that studies both the decision of when to move for an existing homeowner, and the buying and selling decisions of those already in the market. Thus both sales rate and moving rate are endogenous in the model. The model focuses on the market for existing houses.²²

3.1 Houses

The economy contains 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 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 house to buy and occupy.

It is implicit in the model that families moving house might temporarily use the rental market in between selling and buying. However, there is no explicit modelling of the rental market: effectively, this is treated as a distinct segment of the housing market. This view is consistent with [Glaeser and Gyourko \(2007\)](#), who argue that there is little evidence in support of significant arbitrage between the rental and owner-occupied segments of the housing market because owner-occupied homes typically have different characteristics from rental units, as is also the case for homeowners themselves in comparison to renters. More recently, [Bachmann and Cooper \(2014\)](#) calculate gross flows across and within the owner and renter categories using PSID data. They conclude that rental and owner-occupied markets are distinct segments owing to the dominance of moves within the same tenure category. Moreover, between 1970 and 2009, they find that most transactions are associated with own-to-own moves, rather than own-to-rent moves (the former is 2.3 times the latter), suggesting the majority of owners selling their houses are buying another house.²⁴

²²It abstracts from new entry of houses from either new construction or previously rented houses, and it abstracts from the entry of first-time buyers to the market.

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

²⁴Using a different data source (AHS data from 2001), [Wheaton and Lee \(2009\)](#) find 42.6% of house purchases are by existing homeowners, as opposed to renters and newly formed households. To reconcile this with the conclusion drawn from [Bachmann and Cooper's \(2014\)](#) facts that the majority of owners selling their houses are buying another house, note the following observations. First, using [Wheaton and Lee's \(2009\)](#) data, 57.1% of listings of existing

3.2 Search behaviour and the role of match quality

The mere existence of an inventory of houses for sale together with a group of potential buyers indicates the presence of search frictions in the housing market. The model focuses on two types of search frictions. First, it is costly and time-consuming for buyers and sellers to arrange viewings of houses. Second, buyers need to view a property before knowing how desirable it is.

The first search friction is modelled using a meeting function.²⁵ Let u_t denote the measure of houses available for sale and b_t the measure of buyers. At any instant, each buyer and each house can have at most one viewing. The arrival rate of viewings is determined by a constant-return-to-scale meeting function $\mathcal{V}(u_t, b_t)$. For houses, viewings have Poisson arrival rate $\mathcal{V}(u_t, b_t)/u_t$ (time is continuous). For buyers, the corresponding arrival rate is $\mathcal{V}(u_t, b_t)/b_t$. Buyers incur flow search costs F during this process of search. Homeowners incur maintenance costs M irrespective of whether or not a house is up for sale. 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 ($b_t = u_t$).²⁶ The arrival rates of viewings for buyers and sellers are then both equal to the constant $v = \mathcal{V}(u_t, u_t)/u_t$. This arrival rate summarizes all that needs to be known about the frictions in locating houses to view.

Viewings are essential because the desirability of a property is not easily determined simply by knowing objective features such as size. What is desirable is a good match between the idiosyncratic preferences of the buyer and the idiosyncratic characteristics of the house for sale. This match quality only becomes known to a buyer when a house is actually viewed. A measure of the importance of match quality is the average number of viewings per transactions. [Genesove and Han \(2012\)](#) report data on the number of homes visited and the time taken to sell using the ‘Profile of Buyers and Sellers’ surveys from the NAR in the U.S. for various years from 1989 to 2007. In the UK, monthly data on time-to-sell and viewings-per-sale are available from the Hometrack ‘National Housing Survey’ from June 2001 to July 2013.²⁷ The data are shown in [Figure 5](#).

[Figure 5](#) reveals that viewings-per-transaction are far greater than one, indicating there is substantial uncertainty about match quality prior to a viewing. It also shows that variation in time-to-sell is associated with movements in viewings-per-transaction in the same direction, and is not simply due to variation in the time taken to meet buyers, in other words, a meeting function alone is not sufficient. As in [Jovanovic \(1979\)](#), the second search friction is modelled with a stochastic match quality that only becomes known to a buyer when a house is actually viewed.²⁸ When a view-

houses occur through own-to-own transitions, rather than through own-to-rent or owner exit. Second, some own-to-rent and rent-to-own transitions may be extremely short lived as part of what is effectively an own-to-own move, for example, someone who lives temporarily in a rented home while a newly purchased home is under refurbishment.

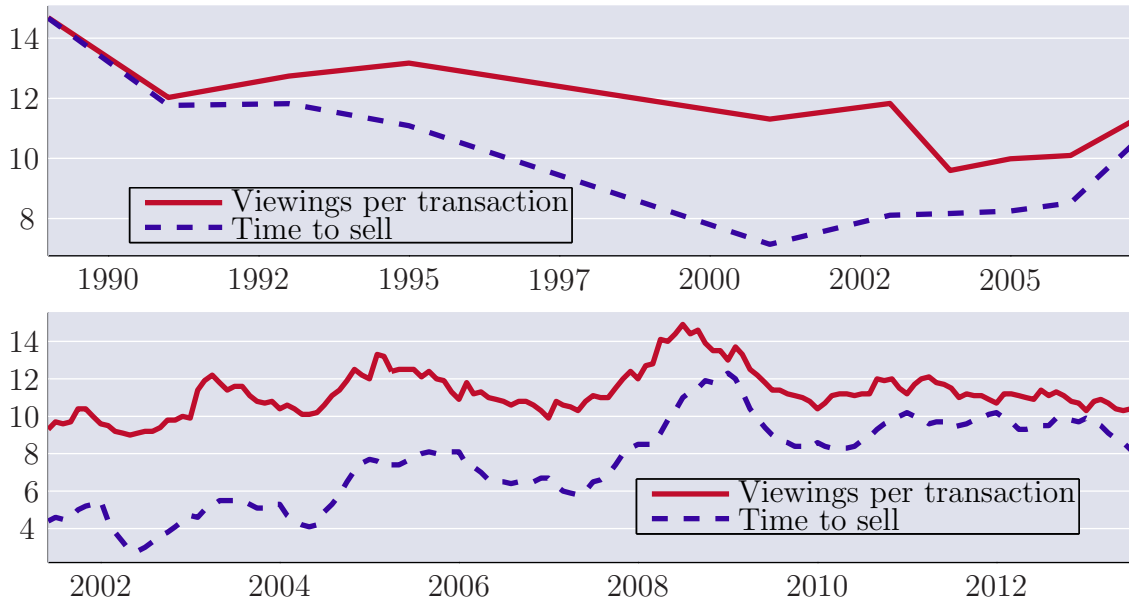
²⁵The term ‘matching function’ is not used here because not all viewings will lead to matches.

²⁶It is important to note that a rise in the homeownership rate does not imply there is an upward trend in market tightness (b_t/u_t). The homeownership rate is defined in terms of the fraction of owner-occupied houses, so a rise in the homeownership rate does not mean that there are more buyers relative to the number of houses.

²⁷Hometrack 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.

²⁸The two search frictions are also present in the labour-market model of [Pissarides \(1985\)](#), and the housing-market

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



Notes: Upper panel, U.S. data, annual frequency (years in sample: 1989, 1991, 1993, 1995, 2001, and 2003–2007); Lower panel, U.K. data, monthly frequency (June 2001–July 2013). Time-to-sell is measured in weeks.

Sources: U.S. data, [Genesove and Han \(2012\)](#); U.K. data, [Hometrack \(www.hometrack.co.uk\)](http://www.hometrack.co.uk).

ing takes place, match quality ϵ is drawn from a distribution with cumulative distribution function $G(\epsilon)$. For analytical tractability, new match quality is assumed to have a Pareto distribution (with minimum value 1, a normalization, and shape parameter λ satisfying $\lambda > 1$):

$$G(\epsilon) = 1 - \epsilon^{-\lambda}. \quad [3.1]$$

When a viewing occurs, the value of ϵ that is drawn becomes common knowledge among the buyer and the seller. The value to a family of occupying a house with match quality ϵ from time t is $H_t(\epsilon)$. By purchasing and occupying this house, the buyer loses the option of continuing to search, with the value of being a buyer denoted by B_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, with the value of owning a house for sale denoted by U_t ('unsatisfied owner'). The buyer and the seller directly bear transaction costs C_b and C_v respectively. If a house with match quality ϵ is sold at a price $p_t(\epsilon)$, the surpluses of the buyer and seller are:

$$\Sigma_{b,t}(\epsilon) = H_t(\epsilon) - p_t(\epsilon) - C_b - B_t, \quad \text{and} \quad \Sigma_{u,t}(\epsilon) = p_t(\epsilon) - C_v - U_t. \quad [3.2]$$

models of [Novy-Marx \(2009\)](#) and [Genesove and Han \(2012\)](#).

The Bellman equations for the buyer and seller value functions are:

$$rB_t = -F + v \int_{y_t} \Sigma_{b,t}(\epsilon) dG(\epsilon) + \dot{B}_t, \quad \text{and} \quad rU_t = -M + v \int_{y_t} \Sigma_{u,t}(\epsilon) dG(\epsilon) + \dot{U}_t, \quad [3.3]$$

where r is the discount rate, \dot{B}_t and \dot{U}_t denote the derivatives of the value functions with respect to time, and y_t is the threshold for transactions defined below. The total surplus resulting from a transaction with match quality ϵ at time t is given by

$$\Sigma_t(\epsilon) = \Sigma_{b,t}(\epsilon) + \Sigma_{u,t}(\epsilon) = H_t(\epsilon) - W_t - C, \quad [3.4]$$

where $W_t = B_t + U_t$ is the combined buyer and seller value function and $C = C_b + C_v$ is the combined transaction cost. As is shown later, $H_t(\epsilon)$ is increasing in ϵ , hence purchases will occur if match quality ϵ is no lower than a threshold y_t , defined by $\Sigma_t(y_t) = 0$. This is the ‘transaction threshold’. 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. The transactions threshold y_t satisfies the following equation:

$$H_t(y_t) = W_t + C. \quad [3.5]$$

Using the Bellman equations for the buyer and the seller in (3.3), the combined value W_t satisfies the Bellman equation:

$$rW_t = -F - M + v \int_{y_t} (H_t(\epsilon) - W_t - C) dG(\epsilon) + \dot{W}_t. \quad [3.6]$$

Intuitively, the first term on the right-hand side captures the flow costs of being a buyer and a seller, while the second term is the combined expected surplus from searching for a house and searching for a buyer.

3.3 Price determination

While the equations (3.5) and (3.6) for the transactions threshold y_t and the value function W_t do not depend on the specific price-setting mechanism, this section briefly discusses price determination under Nash bargaining.

Suppose the seller has bargaining power ω . Using the surplus equations in (3.2) and the Bellman equations in (3.3), the Nash bargaining solution implies the surplus-splitting condition $(1 - \omega)\Sigma_{u,t}(\epsilon) = \omega\Sigma_{b,t}(\epsilon)$, which determines the transaction price $p_t(\epsilon)$. As shown in [appendix A.4](#), the average transactions price P_t is

$$P_t = \frac{\omega}{1 - G(y_t)} \int_{y_t} H_t(\epsilon) dG(\epsilon) + (1 - \omega)C_v - \omega C_b + \frac{\omega F - (1 - \omega)M}{r}. \quad [3.7]$$

The ratio of the seller’s transaction cost C_v to the total transaction cost C is subsequently denoted

by κ , and the model will be parameterized in terms of C and κ rather than C_b and C_v .

3.4 Moving decisions of homeowners

To understand the decision to move, the key variable for homeowners is their current match quality ϵ , which is compared to owners' outside option of search. A homeowner with match quality ϵ receives a utility flow value of $\epsilon\xi$ over time while the house is occupied, where ξ is a variable representing the exogenous economy-wide level of housing demand. Housing demand ξ is common to all homeowners, whereas ϵ is match specific. Moving decisions will lead to an endogenous distribution of match quality across homeowners.

A family's match quality ϵ is a persistent variable subject to occasional 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. The arrival of these shocks follows a Poisson process with arrival rate a . If no shock occurs, match quality remains unchanged. If a shock occurs, match quality ϵ is scaled down from ϵ to $\delta\epsilon$, where δ is a parameter that determines the size of the shocks ($\delta < 1$). Match quality following a shock is still positively related to match quality before the shock because there are many dimensions to what is considered a desirable house, only some of which will be affected by a particular life event. For example, a new job might affect commuting time, but leave other desirable aspects of a particular house unchanged. Note that while the arrival time of the shocks is stochastic, the size of the shock δ is deterministic. If homeowners are ever to consider moving it must be the case that $\delta < 1$.

Following the arrival of an idiosyncratic shock, a homeowner can decide whether or not to move. Those who move become both a buyer and a seller simultaneously. Those who do not experience an idiosyncratic shock face a cost D if they decide to move.²⁹ For tractability, the model is set up so that a homeowner will always choose not to move in the absence of an idiosyncratic shock (formally, this is done by assuming the limiting case of $D \rightarrow \infty$).³⁰

The decision of whether or not to move for those who receive shocks depends on all relevant variables including homeowners' own idiosyncratic match quality, and current and expected future conditions in the housing market. The Bellman equation for $H_t(\epsilon)$ is

$$rH_t(\epsilon) = \epsilon\xi - M + a(\max\{H_t(\delta\epsilon), W_t\} - H_t(\epsilon)) + \dot{H}_t(\epsilon). \quad [3.8]$$

The value function $H_t(\epsilon)$ is increasing in ϵ . Thus, when a shock to match quality is received, a homeowner decides to move if match quality ϵ is now below a 'moving threshold' x_t defined by:

$$H_t(x_t) = W_t. \quad [3.9]$$

²⁹This cost represents the 'inertia' of families to remain in the same house, which is in line with empirical evidence. According to the American Housing Survey and the Survey of English Housing, common reasons for moving include being closer to schools, closer to jobs, or because of marriage or divorce.

³⁰The assumption of a positive D for those who do not receive idiosyncratic shocks has no consequences for the analysis of the steady state of the model. Furthermore, even in the analysis of the model's dynamics, if aggregate shocks are small in relation to the size of transaction costs then the assumption of a positive D has no consequences for those homeowners who have not yet received an idiosyncratic shock.

This equates the value of a marginal homeowner to the outside option of selling an existing house and searching for a new one.

3.5 Stocks and flows

Let n_t denote the rate at which the stock $1 - u_t$ of houses occupied by their owners are listed (put up for sale), and let s_t denote the rate at which the stock u_t of houses for sale are sold. The accounting identity that connects stocks and flows is

$$\dot{u}_t = n_t(1 - u_t) - s_t u_t. \quad [3.10]$$

The listing (inflow) and sales (outflow) rates n_t and s_t are endogenously determined by the moving decisions of individual homeowners and the transactions decisions of individual buyers and sellers.

Given that transactions occur when the match quality ϵ from a viewing exceeds the transactions threshold y_t , by using the Pareto distribution of new match quality (3.1) the sales rate s_t is:

$$s_t = v\pi_t, \quad \text{with } \pi_t = y_t^{-\lambda}, \quad [3.11]$$

where π_t is the proportion of viewings for which match quality is above the transactions threshold y_t . This term captures the second search friction due to buyers' idiosyncratic tastes. The first friction is captured by the viewing rate v .

The moving rate n_t is derived from the distribution of existing match quality among homeowners together with the moving threshold x_t . The evolution over time of the distribution of match quality depends on the idiosyncratic shocks and moving decisions. The derivation of the moving rate is much more complicated than the sales rate. Surviving matches differ along two dimensions: (i) the initial level of match quality, and (ii) the number of shocks received since the match formed. By using the Pareto distribution assumption (3.1) for new match quality, the following moving rate n_t is derived in [appendix A.5](#):³¹

$$n_t = a - \frac{a\delta^\lambda x_t^{-\lambda} v}{1 - u_t} \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau. \quad [3.12]$$

This equation demonstrates that given the moving threshold x_t , the moving rate n_t displays history dependence. The reason is the persistence in the distribution of match quality among existing homeowners.

The occurrence of idiosyncratic shocks and the moving and buying decisions affect the distribution of match quality across all current homeowners. Summarizing this distribution by its first

³¹The tractability that results from the Pareto distribution assumption comes from the property that a truncated Pareto distribution is also a Pareto distribution with the original shape parameter. Together with the nature of the idiosyncratic shock process, this is what allows the explicit expression (3.12) to be derived. This property of the truncated Pareto distribution is also useful in calculating the expected surplus from searching for a new house taking into account future moving decisions. Since matches receiving idiosyncratic shocks will survive only if $\delta\epsilon > x$, the calculation of the value function involves only an integral starting from x/δ . This integral can be easily obtained with the Pareto distribution $(\underline{\epsilon}, \lambda)$ because its probability density function is only a function of $\epsilon/\underline{\epsilon}$.

moment, let Q_t denote the average of match quality ϵ . The law of motion of Q_t is shown in [appendix A.5](#) to be:

$$\dot{Q}_t = s_t \frac{u_t}{1 - u_t} \left(\frac{\lambda}{\lambda - 1} y_t - Q_t \right) - (a - n_t) \left(Q_t - \frac{\lambda}{\lambda - 1} x_t \right). \quad [3.13]$$

Average match quality Q_t is a state variable owing to the persistence of individual match qualities. The rate of change of match quality is increasing in both the moving and transaction thresholds x_t and y_t , that is, the qualities of new and ongoing marginal matches capture changes in the overall degree of mismatch.

4 The equilibrium of the model

The equilibrium of the model is derived in two stages. First, moving and transaction thresholds x and y are obtained. The transactions threshold y determines the sales rate s , and both moving and transaction thresholds x and y determine the moving rate n . These then determine transactions and listings, and thus the stock of houses for sale. Throughout, the focus is on the case of perfect foresight with respect to the parameters of the model, that is, changes to parameters are treated as permanent and no changes to these parameters are anticipated.³²

4.1 The moving and transactions thresholds

The analysis assumes a case where the idiosyncratic shock is large enough to induce a homeowner with match quality y (a marginal homebuyer) to move, that is, $\delta y < x$. This is true when the parameters of the model satisfy the condition in equation (4.8) below.

When $\delta y < x$, it follows from homeowners' value function (3.8) that the value for a marginal homebuyer satisfies:

$$(r + a)H(y) = \xi y + aW. \quad [4.1]$$

Using equation (3.8) again, the value for a marginal homeowner (in the sense of being indifferent between remaining a homeowner or moving) satisfies:

$$(r + a)H(x) = \xi x + aW. \quad [4.2]$$

These two values are related as follows using the definitions of the transaction and moving thresholds in (3.5) and (3.9):

$$H(y) = H(x) + C. \quad [4.3]$$

³²The model can be solved exactly for both its steady state and the out-of-steady-state transitional dynamics when perfect foresight is assumed. The perfect foresight assumption is more suited to analysing long-lasting shifts in the economy, and would not be appropriate for studying a version of the model augmented with shocks at business-cycle frequencies.

Equations (4.1)–(4.3) together imply that:

$$y - x = \frac{(r + a)C}{\xi}, \quad [4.4]$$

which is the first equilibrium condition linking the moving and transaction thresholds x and y .

A second equilibrium condition connecting x and y is obtained by deriving the combined buyer and seller value W as a function of the moving and transactions thresholds. First, from the definition of the moving threshold x , equations (3.9) and (4.2) together imply that:

$$H(x) = W = \frac{x\xi}{r}. \quad [4.5]$$

Second, the value W can be obtained directly from the flow value equation (3.6) by computing the surplus from a match. The expected surplus from a new match is shown in appendix A.3 to be given by:

$$\int_y (H(\epsilon) - W - C)dG(\epsilon) = \frac{\xi \left(y^{1-\lambda} + \frac{a\delta^\lambda}{r+a(1-\delta^\lambda)} x^{1-\lambda} \right)}{(r + a)(\lambda - 1)}. \quad [4.6]$$

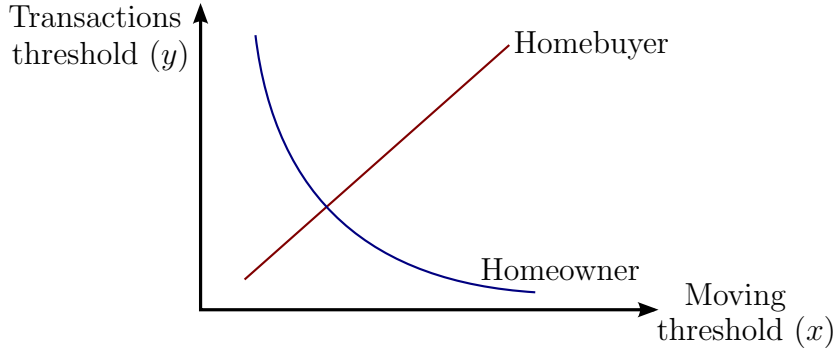
Allowing for moving decisions means that the expected surplus of a new match depends not only on the transactions threshold y but also on the moving threshold x . Combining this equation with (3.6) and (4.5) yields the second equilibrium condition linking x and y :

$$x = \frac{v \left(y^{1-\lambda} + \frac{a\delta^\lambda}{r+a(1-\delta^\lambda)} x^{1-\lambda} \right)}{(r + a)(\lambda - 1)} - \frac{F}{\xi}. \quad [4.7]$$

Together, equations (4.4) and (4.7) can be jointly solved for the thresholds x and y without reference to state variables such as the number of houses for sale or the distribution of existing match quality. Figure 6 depicts the determination of the moving and transaction thresholds as the intersection between an upward-sloping equation (4.4) and a downward-sloping equation (4.7). Intuitively, the upward-sloping line ties the value of a marginal homebuyer to that of a marginal homeowner together with the transaction cost (which is sunk for someone who has decided to become a buyer, but not for an existing homeowner who can choose to stay). This line is referred to as the ‘homebuyer’ curve. The downward-sloping curve ties the value of the marginal homeowner to the expected value of becoming a buyer. This line is referred to as the ‘homeowner’ curve.

In (x, y) space, these two curves pin down the equilibrium values of x and y . If an equilibrium exists, it must be unique. It is shown in appendix A.3 that an equilibrium satisfying the conditions

Figure 6: Determination of the moving (x) and transactions (y) thresholds



Notes: The homebuyer and homeowner curves represent equations (4.4) and (4.7) respectively.

$y > 1$ and $\delta y < x$ exists if and only if:

$$\max \left\{ 1, \frac{(r+a)C}{(1-\delta)\xi} \right\}^{1-\lambda} + \frac{a\delta^\lambda}{r+a(1-\delta^\lambda)} \left(\max \left\{ 1, \frac{(r+a)C}{(1-\delta)\xi} \right\} - \frac{(r+a)C}{\xi} \right)^{1-\lambda} - \frac{(\lambda-1)(r+a)}{v} \left(\max \left\{ 1, \frac{(r+a)C}{(1-\delta)\xi} \right\} - \frac{(r+a)C}{\xi} + \frac{F}{\xi} \right) > 0. \quad [4.8]$$

4.2 Efficiency

The equilibrium of a search-and-matching model is often Pareto inefficient owing to the presence of externalities. For example, a ‘congestion externality’ occurs when individuals do not take account of the extra difficulty faced by others in finding a match when they enter the market. Furthermore, there is typically a ‘hold-up’ problem whereby one party is able to extract surplus from another after sunk costs of search have been incurred, which when anticipated reduces the incentive of the other to enter the market.

It turns out that the equilibrium here is Pareto efficient. The equilibrium is the solution of the following social planner’s problem with objective function

$$\Omega_T = \int_{t=T}^{\infty} e^{-r(t-T)} (\xi(1-u_t)Q_t - Cs_tu_t - Fu_t - M) dt, \quad [4.9]$$

where Q_t , s_t , and u_t are the average match quality, sales rate, and stock of unsold houses resulting from the planner’s choice of the moving and transaction thresholds. It is shown in [appendix A.6](#) that maximizing this objective function subject to the laws of motion derived earlier leads to exactly the same moving and transaction thresholds as in the equilibrium of the model.

Intuitively, the usual congestion externality is not present here because homeowners who decide to move enter the market simultaneously on both sides as a buyer and as a seller. Owing to the constant-returns meeting function, entry has no effect on the likelihood of any other participant in the market meeting a buyer or a seller. Furthermore, the hold-up problem is also absent, again

because homeowners enter on both sides of the market. If participants on one side of the market are able to extract surplus from participants on the other side then homeowners entering the market know that they will face hold up, but will also be able to hold up others. Ex ante, when moving decisions are made, these two effects are expected to cancel out. Private moving and transaction decisions thus result in a socially efficient allocation.

4.3 Steady state

Taking as given the moving and transaction thresholds x and y , there exists a unique steady state for all stocks and flows. This steady state naturally has $\dot{u}_t = 0$, but also the distribution of existing match quality must have converged to its ergodic limit, which in practice requires that both u_t and n_t are constant over time.

First, the transaction threshold y directly pins down the sales rate s using equation (3.11):

$$s = vy^{-\lambda}. \quad [4.10]$$

The steady-state moving rate can be derived from (3.10) and (3.12) using $\dot{u}_t = 0$ and $\dot{n}_t = 0$:

$$n = \frac{a}{1 + \frac{\delta^\lambda}{1-\delta^\lambda} \left(\frac{y}{x}\right)^\lambda}. \quad [4.11]$$

Observe that there is a positive relationship between the moving threshold x and the moving rate n . The transaction threshold y appears in the equation and have a negative effect because it is positively related to the average initial match quality. In the special case when λ is very large (where the distribution $G(\epsilon)$ is degenerate), the effect of thresholds on the moving rate will disappear and only the arrival rate a of the exogenous shock will matter. Intuitively, if everyone has the same match quality then basically everyone who gets a shock has to move otherwise there will be no moving. Put it differently, there is no marginal homeowner. Thus a distribution of match quality is an essential ingredient in endogenising the moving rate.

Given the steady-state sales and moving rates s and n , the steady state for houses for sale follows immediately from (3.10):

$$u = \frac{n}{s + n}. \quad [4.12]$$

The steady-state average transaction price P from (3.7) is:

$$P = C_v - \frac{M}{r} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v} \right) (\xi x + F). \quad [4.13]$$

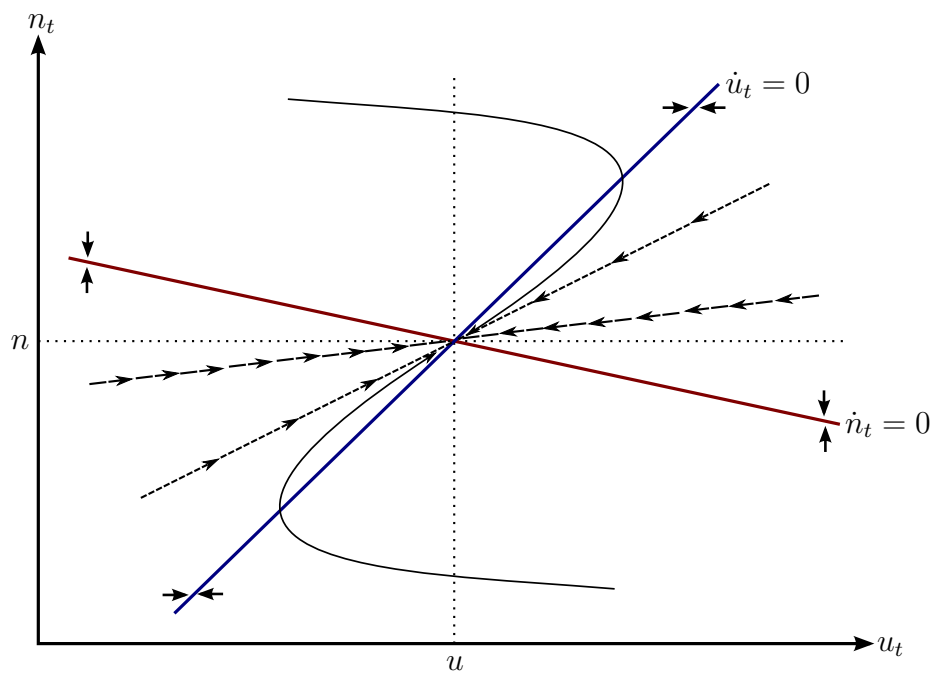
4.4 Transitional dynamics and overshooting

Following a change in the moving and/or transaction thresholds, the housing market will begin a process of convergence to the new steady state for the volume of transactions, the moving rate, and

the stock of houses for sale. There are two facets of these transitional dynamics. First, there is convergence in the stock of houses for sale given the inflow and outflow rates, which is a common feature of most search models. Second, endogenous moving together with persistence in existing match quality gives rise to a novel source of transitional dynamics as the distribution of match quality converges to its ergodic limit.

The second source of transitional dynamics work through the backward looking nature of the aggregate moving rate. In this sense the aggregate moving rate n is also a state variable in addition to the standard state variable u . For any given levels of the moving and transaction thresholds (x, y) , which are the control variables, [appendix A.7](#) shows that the dynamics of the state variables (u, n) are governed by a pair of first order linear differential equations. [Figure 7](#) depicts the phase diagram for these two state variables in the empirically relevant case where the sales rate is large compared to the moving rate.

Figure 7: *Phase diagram of the model*



Notes: The red and blue solid lines depict respectively the points where n_t and u_t are individually stable, and the intersection of the two lines is the steady state. The two dashed lines depict the two eigenvectors of system of differential equations (the eigenvector associated with the dominant eigenvalue is the short dashed line), and the solid curves depict example dynamic paths starting from some initial conditions. These properties are derived analytically in [appendix A.7](#).

Consider a rise in the moving threshold with no change in the transactions threshold. On impact, the moving rate must increase as homeowners are less tolerant of low match quality. Furthermore, since subsequent match quality will be higher on average after this cleansing of the distribution, the initial moving rate will be higher than in the new steady state, which is itself higher than the original steady state. In this sense, there is overshooting of the moving rate. After the overshooting,

the phase diagram confirms that convergence to the steady-state moving rate is monotonic.

There is no immediate change in the stock of houses for sale on impact. However, owing to the higher moving rate, the new steady state level of houses for sale will be higher. The phase diagram shows that houses for sale gradually rises above the new steady state and then falls back towards it. This reflects the gradual clearance of the unsold stock of additional houses that come on to the market. With no change in the sales rate, the dynamics of transactions follow the same overshooting pattern as houses for sale.

Second, consider a fall in the transactions threshold with no change to the moving threshold. In the very short term this creates a boom in low-match quality transactions, but this quickly comes to an end as houses on the market are sold. However, in the future, the new low-quality matches will lead to more moving, including in the new steady state. This replenishes the stock of houses for sale and leads to a recovery in the number of transactions, including in the new steady state. This process means there is overshooting in transactions also for a change in the transactions threshold.

The dynamics described here will be explored quantitatively once the model has been calibrated in [section 5.4](#).

5 The importance of the decision to move house

This section presents an application of the model to housing-market boom between 1995 and 2003. During this time, houses were selling faster, more houses were sold, and at the same time, more houses were put up for sale.³³ The application highlights the role of the endogenous moving decision by contrasting the results of the model to one with exogenous moving. The overshooting dynamics of the endogenous moving model are also quantified.

5.1 The 1995–2003 housing-market boom

[Figure 1](#) shows that during the period 1995–2003, housing transactions rose by 43%, while the sales rate increased by 14% (for consistency with [Figure 1](#), all percentage changes referred to in this section are the log point changes of variables between their 1994 averages and their 2003 averages). Despite the rise in the sales rate, the stock of houses for sale did not fall, and in fact increased by 29%, while the fraction of houses for sale increased by 13%. The time series of listings plays a key role in reconciling the behaviour of transactions, the sales rate, and houses for sale. During this period, the volume of listings rose by 50% while the listing rate increased by 34%. Not only were houses selling faster (the rise in the sales rate), but at the same time homeowners decided to move more frequently. This increase in listings generated a rise in the stock of houses for sale, which also boosted transactions.³⁴

³³It is well known that rapid increases in house prices continued until 2006–2007, while transactions started to fall during 2005. It is likely the 2003–2006 period was affected by factors outside of the model, most prominently, the speculative motive of real-estate investors ([Haughwout, Lee, Tracy and van der Klaauw, 2011](#), [Burnside, Eichenbaum and Rebelo, 2016](#), [DeFusco, Nathanson and Zwick, 2017](#)).

³⁴As mentioned earlier, the 34% increase in the listing rate closely matches the 33% rise in the own-to-own moving rate found by [Bachmann and Cooper \(2014\)](#) using household-level data from the PSID.

It is well known that there was a boom in construction and a rise in the homeownership rate during the period in question. As explained in [section 2.2](#), both factors are captured by the increase in the owner-occupied housing stock series. Equation (2.4) shows that changes in the housing stock have an approximately proportional effect on transactions, which is seen in the 16% rise in the housing stock (the grey dotted line in [Figure 1](#)) and the 16% contribution of the housing stock to transactions (the yellow dotted line in [Figure 2](#)). Thus, net of the contribution from the changing housing stock, transactions rose by 27% and the number of listings increased by 34%. There remains a substantial rise in transactions even after accounting for changes in the owner-occupied housing stock. Finally, note that real house prices increased by 31% during the 1995–2003 period according to the purchase-only price index from the Federal Housing Finance Agency (FHFA) deflated by the personal consumption expenditure (PCE) price index.

There are at least three features of the economic environment during the period 1995–2003 that have implications for moving decisions according to the model and which are consistent with the rise in listings. These are the increase in productivity growth, the adoption of internet technology, and cheaper and more easily accessible credit.

An increase in productivity raises incomes and increases the demand for housing ξ (the reason why higher productivity growth can be interpreted as an increase in the preference parameter ξ is explained in [appendix A.13](#)). The adoption of internet technology reduces search frictions in arranging viewings by distributing information about available houses and their general characteristics more widely among potential buyers. This raises the meeting rate v implied by the meeting function. Finally, lower mortgage rates reduce the opportunity cost of capital and thus lower the discount rate r in the model (the reason why interest rates can be linked to the preference parameter r in the model is also explained in [appendix A.13](#)). Easier access to credit such as a fall in the fixed cost of obtaining a mortgage can be interpreted as a reduction in total transaction costs C (buyers' transaction costs C_b are lower). Another relevant factor that effectively lowered transaction costs C during this period is the Taxpayer Relief Act of 1997, which reduced the tax payable on moving for some existing homeowners (lower transaction costs C_s for sellers).³⁵

Using [Figure 6](#), which depicts equations (4.4) and (4.7), both a rise in ξ and a fall in r imply the two curves shift to the right, as shown in the left panel of [Figure 8](#) below. A rise in v implies the curve representing (4.7) shifts to the right, as shown in the middle panel. Last, a fall in C implies a downward shift of the line representing (4.4), as shown in the right panel. In all cases, the moving threshold x increases, which represents a reduction in the acceptable degree of mismatch for existing homeowners. The increase in x is larger than y , and hence x/y increases, which leads to an unambiguous increase in the moving rate n according to (4.11). The sales rate unambiguously rises when C falls, but the effect is ambiguous in the other cases.

The intuition for these effects is as follows. Take the case of lower mortgage rates r . This increases the present discounted value of flows of housing services, so it increases the incentive to invest in match quality, reducing the tolerance for low current match quality (a higher moving threshold x),

³⁵[Shan \(2011\)](#) and [Heuson and Painter \(2014\)](#) present evidence showing that this policy change led to an increase in mobility and transactions.

resulting in more frequent moves. However, there are two offsetting effects on transactions decisions. On the one hand, buyers are more keen to make a purchase to receive the higher discounted sum of flow values, so they become less picky (a lower transaction threshold y). On the other hand, owing to the reduced tolerance for low quality matches as a homeowner, the expected duration of a match is shortened, which makes buyers more picky (higher y). The intuition is essentially the same for the effects of an increase in housing demand ξ .

In the case of the higher viewing rate v , the effect is to raise the expected surplus from searching. This increases the incentive to search both for existing homeowners and homebuyers, making both more picky (higher moving and transaction thresholds x and y). Finally, a lower transaction cost C shrinks the gap between the least acceptable current match and the least acceptable new match, resulting in more moving and more transactions.

Figure 8: *Comparative statics of moving and transaction thresholds*

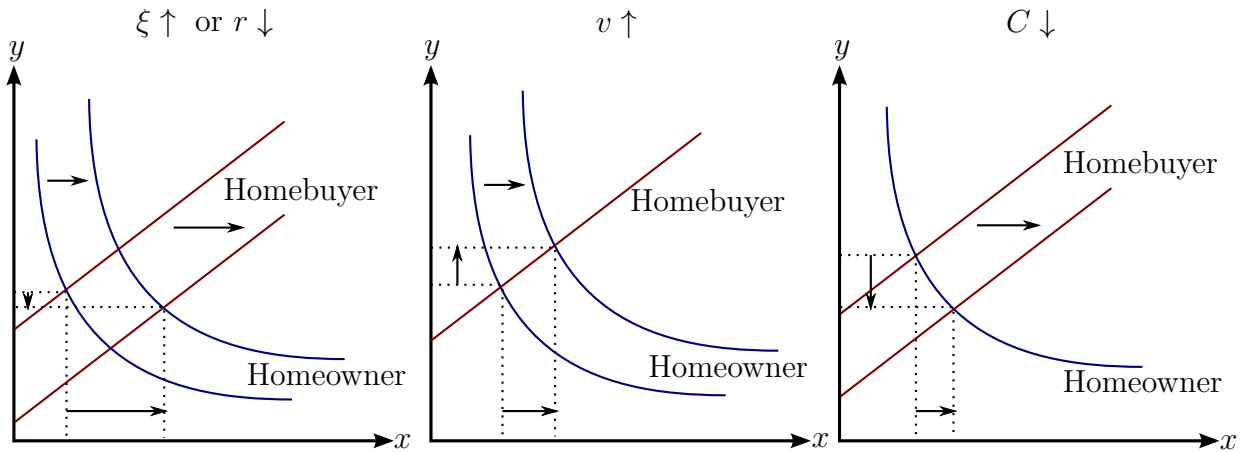


Figure 8 implies that all three factors can lead to an increase in the moving rate and the number of transactions. Quantifying the predictions requires a calibration of the parameters, which is the goal of the next section.

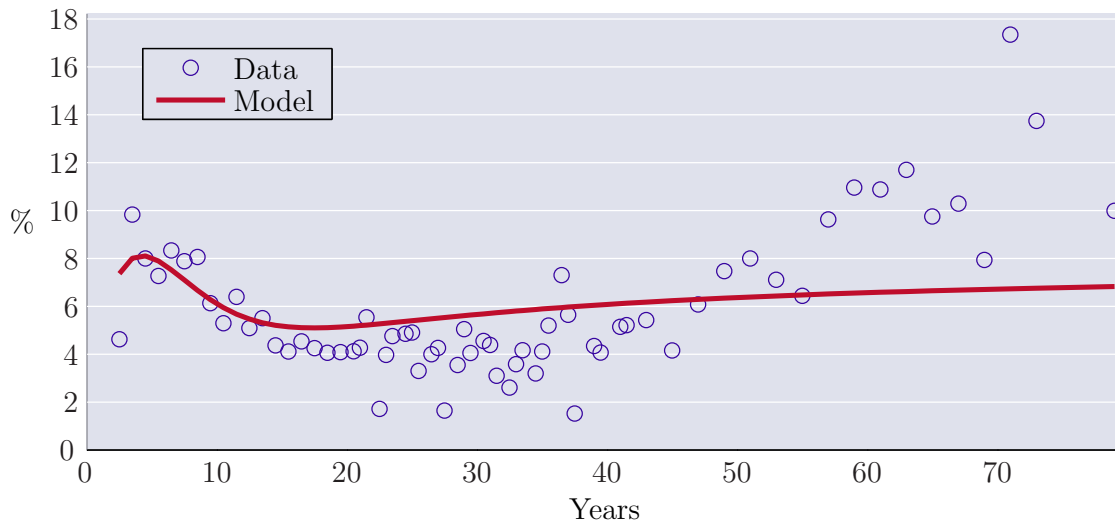
5.2 Quantitative results

The key parameters of the model are calibrated using an estimate of the hazard function for moving house, that is, the rate at which existing homeowners move house as a function of the duration of time they have spent in their current home. The hazard function can be estimated using data available in the American Housing Survey (AHS). Table 3.9 of the AHS reports data on ‘the year householder moved into unit’ for current owner-occupied housing units, given as a frequency distribution over cohorts of homeowners, for example, those that moved in from 1985 to 1989. Each cohort appears in multiple surveys, so the hazard rate can be estimated by comparing the number of remaining homeowners from a cohort in two adjacent survey years. For example, of the cohort that moved in 1985–1989, the number reported in the survey year 2003 was 7.254 million, while in the survey year

2005 it was 6.433 million. This gives an estimate of 0.821 million homeowners out of 7.254 million who move during those two years, which implies an annualized moving rate of 5.8%.

This exercise can be done for every cohort in every pair of survey years. In what follows, all biennial surveys from 1985 to 2015 are used. The hazard function expresses the hazard rate as a function of the duration of stay, and the duration is computed in the data as the difference between the average of the two survey years and the midpoint of the range of years in which a cohort moved in.³⁶ For example, taking the 1985–1989 cohort in the survey years 2003 and 2005, the duration associated with the hazard rate 5.8% is 16.5 years (2004 minus 1987.5). This method yields a collection of duration-hazard rate pairs that are displayed as a scatterplot in Figure 9. Note that where there are multiple hazard rates associated with a given duration, the hazard rates are averaged and appear as a single point in the figure. The data suggest the hazard function displays a ‘u-shaped’ pattern whereby it first declines from around 9% per year to 4% and then rises above 10% as the duration of stay increases.

Figure 9: Empirical and fitted model hazard functions



Notes: The parameters of the model are chosen to minimize the sum of squared deviations between the empirical and model-implied hazard functions.

Source: Authors’ calculations using data from the American Housing Survey.

In order to rationalize the downward-sloping portion of the hazard function, the model of section 3 is enriched to allow for heterogeneity in the arrival rates and sizes of the idiosyncratic shocks faced by families. This extension of the model is described in full in appendix A.8 and the new equations of the model are derived there. In short, after a transaction is complete, a family draws a type $i = 1, \dots, q$, where the types have probabilities $\theta_1, \dots, \theta_q$ ($\sum_{i=1}^q \theta_i = 1$).³⁷ Each type i can have a different distribution of idiosyncratic shocks parameterized by arrival rate a_i and the scaling

³⁶For the cohort ‘1939 or earlier’ the midpoint is taken as 1935.

³⁷This is computationally more tractable than the alternative of families having different types permanently, but the implications of the heterogeneity for the hazard function are similar.

factor δ_i applied to match quality ϵ when a shock is received. The different types can be interpreted as different family compositions, different occupations, or different preferences. The key difference compared to the basic model of [section 3](#) is that families face ex-ante differences in risks to housing match quality, rather than just ex-post differences in the realizations of the idiosyncratic shocks. Such ex-ante differences are likely in light of the most common reasons for moving house reported in the AHS, namely changing jobs, changing schools, marriage, divorce, and having children.

It is shown in [appendix A.8](#) that the model-implied hazard function $h(T)$ is:

$$h(T) = \frac{\sum_{i=1}^q \theta_i a_i e^{-a_i T} \left(1 + \left(\frac{y}{x}\right)^\lambda \left((1 - \delta_i^\lambda) \left(e^{a_i \delta_i^\lambda T} - 1 \right) - \delta_i^\lambda \right) \right)}{\sum_{i=1}^q \theta_i e^{-a_i T} \left(1 + \left(\frac{y}{x}\right)^\lambda \left(e^{a_i \delta_i^\lambda T} - 1 \right) \right)}, \quad [5.1]$$

where $h(T)$ is the average moving hazard rate for families that lived in a house for T years. This depends on the idiosyncratic shock parameters $\{\theta_i\}$, $\{a_i\}$, and $\{\delta_i\}$, the parameter λ describing the distribution of initial match quality, and the moving and transaction thresholds x and y (which are common to all types). The thresholds x and y are determined by a pair of equations that replace [\(4.4\)](#) and [\(4.7\)](#):

$$y - x = \frac{\left(\frac{1}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}} \right) C}{\xi}; \quad [5.2a]$$

$$x = \frac{v \sum_{i=1}^q \frac{\theta_i}{r+a_i} \left(y^{1-\lambda} + \frac{a_i \delta_i^\lambda}{r+a_i(1-\delta_i^\lambda)} x^{1-\lambda} \right)}{\lambda - 1} - \frac{F}{\xi}. \quad [5.2b]$$

These equations reduce to [\(4.4\)](#) and [\(4.7\)](#) in the special case $q = 1$, which represents the original model where the distribution of idiosyncratic shocks is the same for all families.

There are two features of the model that contribute to the shape of the hazard function. First, when $\delta_i > 0$, the ratio of $1 + (y/x)^\lambda ((1 - \delta_i^\lambda)(e^{a_i \delta_i^\lambda T} - 1) - \delta_i^\lambda)$ and $1 + (y/x)^\lambda (e^{a_i \delta_i^\lambda} - 1)$ is increasing in the duration T because $y > x$ and $\delta_i^\lambda > 0$. This reflects the deterioration of match quality in a cohort of homeowners over time as idiosyncratic shocks occur but do not necessarily trigger an immediate move, which makes it more likely a given shock will trigger moving in the future. Second, heterogeneity means that some family types have a greater overall moving rate because of a higher arrival rate or a larger size of idiosyncratic shocks. Given that more of such family types will move sooner and thus leave behind in a cohort more types with lower overall moving rates, there is a tendency for the hazard rate to decline with duration owing to this compositional effect. Mathematically, if $\delta_i = 0$ for all i then the hazard function would be equal to the ratio of $\sum_{i=1}^q \theta_i a_i e^{-a_i T}$ and $\sum_{i=1}^q \theta_i e^{-a_i T}$, which is decreasing in T . For a range of parameters, the aggregate hazard function can display both downward- and upward-sloping sections.

The estimated hazard function provides important information about the parameters of the model. First, the average level of the hazard function reveals how long on average families are

expected to remain in the same home. This statistic is referred to as time-to-move T_n :

$$T_n = \int_{T=0}^{\infty} Th(T) e^{-\int_{\tau=0}^T h(\tau) d\tau} dT = \sum_{i=1}^q \frac{\theta_i}{a_i} \left(1 + \frac{\delta_i^\lambda}{1 - \delta_i^\lambda} \left(\frac{y}{x} \right)^\lambda \right), \quad [5.3]$$

which essentially depends on the average values of the a_i and the δ_i . Second, the extent of negatively sloped sections of the hazard function yields information about the differences in a_i and δ_i across different types of homeowners. Finally, the extent of upward-sloping sections of the hazard function reveals information about the δ_i parameters separately from the a_i . This information can be summarized by the following statistic:

$$\eta = \left. \frac{\partial \log n_t}{\partial \log x_t} \right|_{u_t=u, n_t=n} = \lambda \left(\frac{y}{x} \right)^\lambda \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{1 - \delta_i^\lambda}, \quad [5.4]$$

which is equal to the elasticity of the moving rate n_t with respect to the moving threshold x_t evaluated at the steady state. This statistic essentially depends on the average values of the δ_i . Note that both formulas (5.3) and (5.4) are also applicable to the standard model with a single distribution of idiosyncratic shocks by setting $q = 1$. The next section shows how these insights can be used as part of a strategy to calibrate the model.

5.2.1 Calibration

The parameters of the model to be calibrated are $\{q, \{\theta_i\}, \{a_i\}, \{\delta_i\}, \lambda, v, C, F, M, \omega, \kappa, r, \xi\}$. Some parameters are directly matched to the data, while others are determined indirectly by choosing values that make the predictions of the model consistent with some empirical targets. For the parameter ω , a value of 0.5 is directly imposed to give buyers and sellers equal bargaining power. The (annual) discount rate r is set to 5.7% ($r = 0.057$), which is the 1994 average of the HP-filtered 30-year real mortgage rate series (subtracting PCE inflation from the nominal mortgage rate).

The first empirical target for the calibration is the average time-to-sell, that is, the reciprocal of the average sales rate ($T_s = 1/s$). This is obtained using data from NAR on transactions and inventories (for existing single-family homes), as described in [section 2.2](#).³⁸ Average time-to-sell over the period 1991–2013 is 6.5 months, hence $T_s = 6.5/12$. Previous research on housing markets has used a variety of sources for data on time-to-sell, and there is considerable dispersion in these estimates. This evidence is discussed in detail in [appendix A.10](#). 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 time-to-sell to be more than four months. Since the predictions of the model will be compared to variables constructed from the NAR transactions and inventories data, a measure of time-to-sell consistent with this data is used.

The empirical target for average time-to-sell provides information about the parameter λ . The

³⁸Time-to-sell is highly correlated with the ‘months supply’ variable reported by NAR, which is defined as inventories divided by transactions.

reason is that 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, all else equal.

The second empirical target is for viewings per sale, that is, the reciprocal of the probability of a transaction conditional on a viewing ($V_s = 1/\pi$). The target is set to 10 ($V_s = 10$) on the basis of the evidence on viewings per transaction presented in [section 3.2](#). Together with time-to-sell, this target provides information about the parameter v . This is because v is the reciprocal of the average time between viewings, which is equal to the ratio of average time-to-sell and average viewings per sale (see [equation 3.11](#)), and hence $v = V_s/T_s$.

The parameters C , F , M , 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 $c = C/P$, $f = F/P$, and $m = M/P$ denote the costs C , F , and M 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 c , f , and m are adopted that will determine C , F , and M indirectly. The cost-to-price ratios predicted by the model are derived in [appendix A.12](#), where it is shown that the parameters C , F , and M appear only as ratios to ξ . This is true in all other equations of the model, so the value of ξ can be normalized to 1.

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 costs incurred in buying and selling houses comprise the one-off transactions cost C and the flow costs of search F . For transaction costs, [Quigley \(2002\)](#) estimates the total costs as being in the range 6–12% of price 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 $1/3$.

There are almost no estimates of the flow costs F of searching. The approach taken here is to base an estimate of F on the opportunity cost of the time spent searching. Assuming one house viewing entails the loss of a day’s income, the value of F can be calibrated by adding up the costs 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 $T_s F$ per housing transaction on average. With viewings per sale equal to the average number of viewings made by a buyer, the total search cost should be equated to $V_s Y/365$, where Y denotes average annual income. Thus, the calibration assumes $T_s F = V_s Y/365$, and by dividing both sides by $P T_s$, this implies $f = (1/365)(Y/P)(V_s/T_s)$. Using a house-price to income ratio of 2 as a reasonable average value ([Case and Shiller, 2003](#)) together with the values of $T_s = 6.5/12$ and $V_s = 10$ discussed earlier, 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.³⁹

The five empirical targets above provide information about the five parameters $\{\lambda, v, C, F, M\}$ and are listed in [Table 1](#) below. While each target is intuitively linked to one parameter, the calibration of the parameters to match the targets must be done jointly. Furthermore, the link between these parameters and the targets also depends on the parameters q , $\{\theta_i\}$, $\{a_i\}$, and $\{\delta_i\}$ describing the distributions of idiosyncratic shocks. It turns out that conditional on q , θ_i , a_i , and δ_i^λ , it is possible to find λ , v , C , F , and M to match T_s , V_s , c , f , and m exactly (the method is described in [appendix A.11](#)). This leaves only the idiosyncratic shock parameters to be determined.

Table 1: *Targets used in the calibration*

Target description	Notation	Value
<i>Targets matched exactly</i>		
Average time-to-sell	T_s	6.5/12
Average viewings per sale	V_s	10
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
<i>Derived from the estimated hazard function</i>		
Average time-to-move	T_n	16.6
Elasticity of the moving rate	η	18.7

Notes: The data sources for these empirical targets are discussed in [section 5.2.1](#). The other calibration target is the empirical hazard function shown in [Figure 9](#), and the average time-to-move and elasticity of the moving rate are derived from the model-based hazard function fitted to the data.

The idiosyncratic shock parameters all appear in the expression for the hazard function in (5.1), along with λ and the moving and transactions thresholds x and y , where the latter are endogenous variables. Given q , θ_i , a_i , and δ_i^λ , the procedure to set the other parameters to match T_s , V_s , c , f , and m exactly pins down λ and the thresholds x and y and thus allows the model-implied hazard function to be calculated. It is then possible to choose $\{\theta_i\}$, $\{a_i\}$, and $\{\delta_i\}$ to match the empirical hazard function as closely as possible, conditional on a number of types q . Formally, for a given q , the procedure is to minimize the sum of squared deviations between the points on the empirical hazard function and the model-implied hazard function. While q cannot be chosen in this way because that would result in an infinite-dimensional parameter space, it is possible to start from $q = 1$ and then increase the number of types, stopping when there is only a negligible improvement in the goodness of fit to the empirical hazard function. The details of this procedure are provided in [appendix A.11](#).

Starting from $q = 1$, the minimized sum of squared deviations between the empirical and model-implied hazard functions is 0.0896. By considering a model with two types, the sum of squared

³⁹The calibration of costs ignores any non-pecuniary costs of house search and moving. A sensitivity analysis considering higher costs is performed in [appendix A.15](#).

deviations is reduced to 0.0662, so the hazard function data strongly prefers some heterogeneity. However, adding more than two types has a negligible impact on the fit to the empirical hazard function (there is no change to the minimized sum of squares up to four decimal places) and the estimated proportions of the additional types in the population are negligible (zero up to four decimal places). Therefore, the number of types is set to $q = 2$ in what follows.

The model-implied hazard function that minimizes the sum of squared deviations for $q = 2$ types is shown in [Figure 9](#) alongside the data. The model-implied hazard function has both a downward-sloping section and an upward-sloping section, reflecting the pattern seen in the data. Quantitatively, the model matches the downward-sloping section of the empirical hazard function quite closely. It is less successful in matching the upward-sloping section of the hazard function, which is significantly steeper than the model is able to generate. The reason adding more types does not improve the fit to the empirical hazard function is that more heterogeneity does not help in fitting the upward-sloping section of the hazard function, while a model with two types already does well in matching the downward-sloping part of the hazard function.

The full set of parameters minimizing the sum of squared deviations between the empirical and model-implied hazard functions and exactly matching the calibration targets in [Table 1](#) is shown in the ‘Model I’ column of [Table 2](#) (the other columns are discussed later). For the distributions of idiosyncratic shocks, it turns out that there is only heterogeneity in the arrival rates a_i . 57% of families have a quiet life and only receive shocks to their housing match quality on average after 11.4 years ($a_1 = 0.088$). The other 43% have more eventful lives and receive shocks every 2.5 years on average ($a_2 = 0.404$). For both types, an idiosyncratic shock is estimated to reduce match quality by 11% ($\delta_1 = \delta_2 = 0.892$). The reason why there is no heterogeneity in δ_i is that the restriction $\delta_i y \leq x$ is binding (ensuring a marginal homebuyer would be willing to move after a single idiosyncratic shock, which is used to solve the model). A higher value of δ_i raises the gradient of the hazard function, but at the highest values of δ_i satisfying the restriction, the upward-sloping section of the model-implied hazard function is still less steep than it is in the empirical hazard function. The data thus prefer a model with a more steeply upward-sloping hazard function than the current model is able to generate.

The calibrated parameters imply that the expected time a family lives in a particular house (time-to-move T_n) is 16.6 years. By using [Table 3.9](#) from the AHS, [Ngai and Tenreyro \(2014\)](#) find the median number of years since householders have moved into their current homes is about 10 years.⁴⁰ Time-to-move could also be estimated by taking the reciprocal of the average listing rate n computed from the NAR sales and inventories and AHS housing-stock data in [section 2.2](#), that is, $T_n = 1/n$. This leads to an estimate of 15 years, which is close to the average time-to-move derived from the hazard function.

⁴⁰This statistic is measuring the average years existing households have stayed in the same house, which is conceptually different from the average time-to-move. The equivalent of this number implied by the model is 16.7 years.

Table 2: Calibration of the parameters

Parameter description	Notation	Model I Value	Model II Value	Model III Value
<i>Parameters derived from hazard function</i>				
Number of types	q	2	1	1
Fraction of type-1 families	θ_1	0.57	1	1
Fraction of type-2 families	θ_2	0.43	-	-
Arrival rate of shocks for type 1s	a_1	0.088	0.127	0.060
Arrival rate of shocks for type 2s	a_2	0.404	-	-
Size of shocks for type 1s	δ_1	0.892	0.898	0.0
Size of shocks for type 2s	δ_2	0.892	-	-
<i>Parameters matching other calibration targets</i>				
Steady-state distribution of match quality	λ	15.6	17.0	15.7
Arrival rate of viewings	v	18.5	18.5	18.5
Total transaction cost	C	0.61	0.61	0.64
Flow search costs	F	0.15	0.15	0.16
Flow maintenance costs	M	0.27	0.27	0.29
<i>Directly measured parameters</i>				
Share of transaction cost directly borne by seller	κ	1/3	1/3	1/3
Discount rate	r	0.057	0.057	0.057
<i>Imposed parameters</i>				
Bargaining power of seller	ω	1/2	1/2	1/2
Normalization of homeowner flow value	ξ	1	1	1

Notes: Model I has heterogeneity in the distributions of idiosyncratic shocks; it is calibrated to minimize the sum of squared deviations between the empirical and model-implied hazard functions shown in Figure 9, and to match exactly the first five calibration targets in Table 1. Model II has a single idiosyncratic shock distribution; it is calibrated to match exactly the average time-to-move T_n and elasticity η of the moving rate implied by the hazard function of Model I, and the other calibration targets in Table 1. Model III has a single distribution of idiosyncratic shock and is restricted so that moving is exogenous: the elasticity η of the moving rate must be zero, which requires $\delta = 0$; the other parameters are calibrated to match exactly the other targets in Table 1, including T_n .

5.2.2 Quantitative predictions of the model

Having calibrated the model, it is now possible to examine its quantitative predictions for the housing market following the three developments discussed earlier: the productivity boom, the rise of internet-based property search, and cheaper credit.

A simple measure of productivity growth is the increase in real GDP per person, which grew by a total of 21% between 1995 and 2003 (all percentages refer to changes in log points). This is identical to the growth in the HP-filter trend line. A rise in income naturally leads to an increase in housing demand (the parameter ξ in the model), the extent of the increase also depending on the

income-elasticity of housing, which is assumed to be unity here.⁴¹

The rise of internet-based property search would be expected to improve the efficiency of search as captured by the meeting function (the parameter v in the model). Since $v = V_s/T_s$, data on time-to-sell and viewings-per-sale can be used to infer the rise in v . Using data from [Genesove and Han \(2012\)](#), the rise in V_s/T_s over the period 1995–2003 is 17%, and this is taken as the size of the efficiency gains.

Mortgage rates (30-year conventional) declined from 8.4% in 1994 to 5.8% in 2003, while inflation (PCE) decreased from 2.1% to 2.0%.⁴² To abstract from transitory movements in interest rates, the focus is on the HP-filter trend fitted to the real mortgage rate. This trend real mortgage rate fell from 5.7% to 4.0%, which is a 36% drop. The assumption is that all the changes considered here are expected to be permanent.

The quantitative effects of each of these changes individually and taken together are given in [Table 3](#). As can be seen, the calibrated model shows that macroeconomic variables have a large impact on moving and buying decisions. The model is thus able to explain a significant proportion of the movements seen in the data. As discussed earlier, it is changes in moving rates that are the main driver of transactions in the longer term. The model is consistent with large changes in transactions precisely because it is able to explain large changes in moving rates.

Table 3: *Quantitative results*

Factor	Transactions	Listings	Sales rate	Moving rate	Houses for sale	Prices
Productivity boom	17%	17%	11%	17%	6%	32%
Improved search efficiency	1%	1%	1%	1%	0%	2%
Lower mortgage rates	10%	10%	−7%	11%	17%	36%
Productivity & search	18%	18%	12%	18%	6%	33%
Productivity, search, & mortgages	24%	24%	3%	25%	21%	70%
Data (1995–2003)	27%	34%	14%	34%	13%	31%

Notes: The table shows the long-run steady state effects of the shocks. The combined effects are not the sum of the individual effects because the model is not linear.

The analysis assumes the observed change in the HP-filter trend of the real mortgage rate is permanent, which provides an upper bound for the effects of lower interest rates. It is debatable whether individuals at the time expected the decline in mortgage rates to be permanent (the HP-filter trend is estimated using data from 1971 to 2015, which has the benefit of hindsight). If mortgage rates were expected to be mean reverting, it would not be correct to take the observed change as permanent in the model.⁴³ For this reason, [Table 3](#) separately reports the combined

⁴¹[Harmon \(1988\)](#) finds an income elasticity of housing demand in the range 0.7–1.

⁴²The 10-year Treasury rate declined by a very similar amount, and inflation expectations as measured by the Michigan survey showed almost no change.

⁴³[Glaeser, Gottlieb and Gyourko \(2010\)](#) argue that accounting for mean reversion in interest rates means that

effects with and without the change in mortgage rates, which provides an upper and a lower bound for the combined effects. Excluding the effects coming from mortgage rates, the model accounts for two thirds of the increase in transactions and the sales rate, and half of the increase in the moving rate. The prediction for prices is very close to what is found in the data.⁴⁴

Overall, the model does a fairly good job in explaining a significant fraction of the rise in the moving rate because of incentives for homeowners to invest in improving match quality. The key mechanism is through an increase in the moving threshold x , which leads to a reduction in the degree of mismatch for existing homeowners. The model generates substantial movements in house prices because productivity and mortgage rates have a large direct impact. This is especially true for mortgage rates, where the predicted changes in house prices are far in excess of what was observed, perhaps indicating the decline in mortgage rates was not expected to be permanent.

5.3 The role of endogenous moving

The results in the previous section show that the model does a good job in accounting for the 1995–2003 housing boom. Since it is changes in moving rates that are driving changes in transactions, the aggregate implications of the model depend crucially on the elasticity of the moving rate with respect to the moving threshold. This elasticity is closely related to the proportion of ‘marginal movers’: those homeowners who are close to the threshold for moving as a fraction of all those who move.

The formula for the elasticity of the moving rate (denoted by η) implied by the model is given in equation (5.4), which is derived in [appendix A.9](#). Note that the formula is also applicable to the basic model with a single distribution of idiosyncratic shocks to match quality when $q = 1$. As can be seen, the elasticity for the model with heterogeneity is simply an average of what it would be for the basic model using each distribution of idiosyncratic shocks over all types of homeowners. This suggests that the basic model, if calibrated to match the moving rate elasticity of the heterogeneous model, would have similar aggregate implications. If so, the elasticity η is approximately a sufficient statistic that gauges the importance of endogenous moving.

The calibrated model with heterogeneity implies an elasticity η equal to 18.7. As described in [appendix A.12](#), it is possible to calibrate the basic model with no heterogeneity to match exactly this elasticity, the time-to-move T_n implied by the heterogeneous model (which is equal to 16.6, see equation 5.3), as well as the other calibration targets in [Table 1](#) matched by the heterogeneous model.⁴⁵ Using this approach, the calibrated parameters of the basic model are reported in [Table 2](#) as ‘Model II’.

The quantitative predictions of the basic model calibrated to match the same time-to-move and cheap credit explains only a small proportion of the boom in house prices. [Fuster and Zafar \(2015\)](#) also shows that mortgage rates have small impact on house prices.

⁴⁴As discussed earlier, easier access to credit might also be interpreted as a reduction in buyers’ transaction costs C_b . This is more difficult to quantify. As an illustrative example, [appendix A.14](#) considers the case where C_b is reduced by one third, which alone can account for a substantial increase in both the sales and moving rates (and consequently, transactions too), but its impact on prices is muted.

⁴⁵The common parameters a and δ are chosen to match T_n and η rather than the empirical hazard function.

moving rate elasticity of the heterogeneous model are shown in the second row [Table 4](#), alongside the results for the heterogeneous model in the first row for comparison. These results are based on the changes to productivity and search efficiency derived earlier (see the row ‘Productivity & search’ from [Table 3](#)). The results are very similar, so this exercise confirms that knowing time-to-move T_n and the moving rate elasticity η is largely sufficient for deriving the aggregate implications for the sales rate, moving rate, and transactions.⁴⁶ Thus, for the remainder of the paper, the focus is on the basic model with no heterogeneity.⁴⁷

Table 4: *The role of endogenous moving*

Model	Transactions	Listings	Sales rate	Moving rate	Houses for sale	Prices
Endogenous moving, heterogeneous	18%	18%	12%	18%	6%	33%
Endogenous moving, homogenous	16%	16%	11%	17%	6%	33%
Exogenous moving	0%	0%	−1%	0%	1%	34%
Exogenous moving, shock to a	27%	27%	14%	27%	13%	−4%
Data (1995–2003)	27%	34%	14%	34%	13%	31%

Notes: ‘Endogenous moving (heterogeneous)’ is the row ‘Productivity & search’ from [Table 3](#). ‘Endogenous moving (homogenous)’ is the case of ‘Productivity & search’ under a homogenous model with $q = 1$. ‘Exogenous moving’ is the case of ‘Productivity & search’ when $\eta = 0$. ‘Exogenous moving, shock to a ’ shows the effect of an increase in the moving rate a that matches the long-run increase in transactions in a model with $\eta = 0$.

It is natural to ask what is the value added of a model with endogenous moving compared to a model where moving is exogenous. The general model embeds an exogenous moving model as a special case when $\delta = 0$. In this case, any homeowner will move house after an idiosyncratic shock because match quality will drop to zero. Thus, the moving rate is the same as the exogenous arrival rate a of the idiosyncratic shock. This exogenous moving model can be calibrated to match the same value of time-to-move T_n considered earlier for the endogenous moving model, as well as the other calibration targets in [Table 1](#). The exogenous moving model implies an elasticity of the moving rate η equal to zero, and this value of η is imposed when the model is calibrated. The parameters are shown as ‘Model III’ in [Table 2](#).

[Table 4](#) shows the combined effects of the productivity boom and internet search across the endogenous moving and exogenous moving versions of the model (the change in mortgage rates is excluded because of uncertainty about whether mean reversion was expected). The first two rows were discussed earlier, and attention here focuses on the model without heterogeneity in the second row. The results for the exogenous moving model are shown in the third row. As in the endogenous moving model, house prices rise owing to the direct effect of higher ξ , but note that the sales rate actually falls with exogenous moving. More importantly, the model now fails to account for the rise in transactions in the absence of any change in the moving rate. The exogenous moving model essentially explains none of the change in transactions.

⁴⁶The calibrated basic model will match the overall level but not the shape of the empirical hazard function.

⁴⁷A sensitivity analysis of the results to changes in some of the calibration targets is presented in [appendix A.15](#).

This analysis confirms the basic point that changes in the moving rate are necessary to explain the boom in transactions. However, it would be possible to have a model where the moving rate itself changed exogenously (by varying the arrival rate a of idiosyncratic shocks, which is identical to the moving rate in the exogenous moving model). The results of this exercise are shown in the fourth row of [Table 4](#), where the size of the shock is set to match the observed change in transactions.

By raising the moving rate significantly (closely matching the data), the exogenous moving model can replicate quantities very well. However, the moving rate shock implies a decline in house prices because the effective discount rate $r + a$ increases. The exogenous moving model therefore cannot account for the developments in the housing market during the period 1995–2003 without considering additional factors that would push up house prices.⁴⁸ The endogenous moving model presented in this paper shows that the factors pushing up house prices can also explain the rise in the moving rate and transactions.

5.4 Overshooting dynamics

As discussed in [section 4.4](#), the model predicts overshooting in the moving rate due to the cleansing effect on the distribution of existing match quality. This section quantitatively studies the full dynamic response of the housing market to the three developments discussed in [section 5.1](#). An analytical solution for the model’s out-of-steady-state dynamics is derived in [appendix A.7](#) and this is used here to plot impulse response functions.

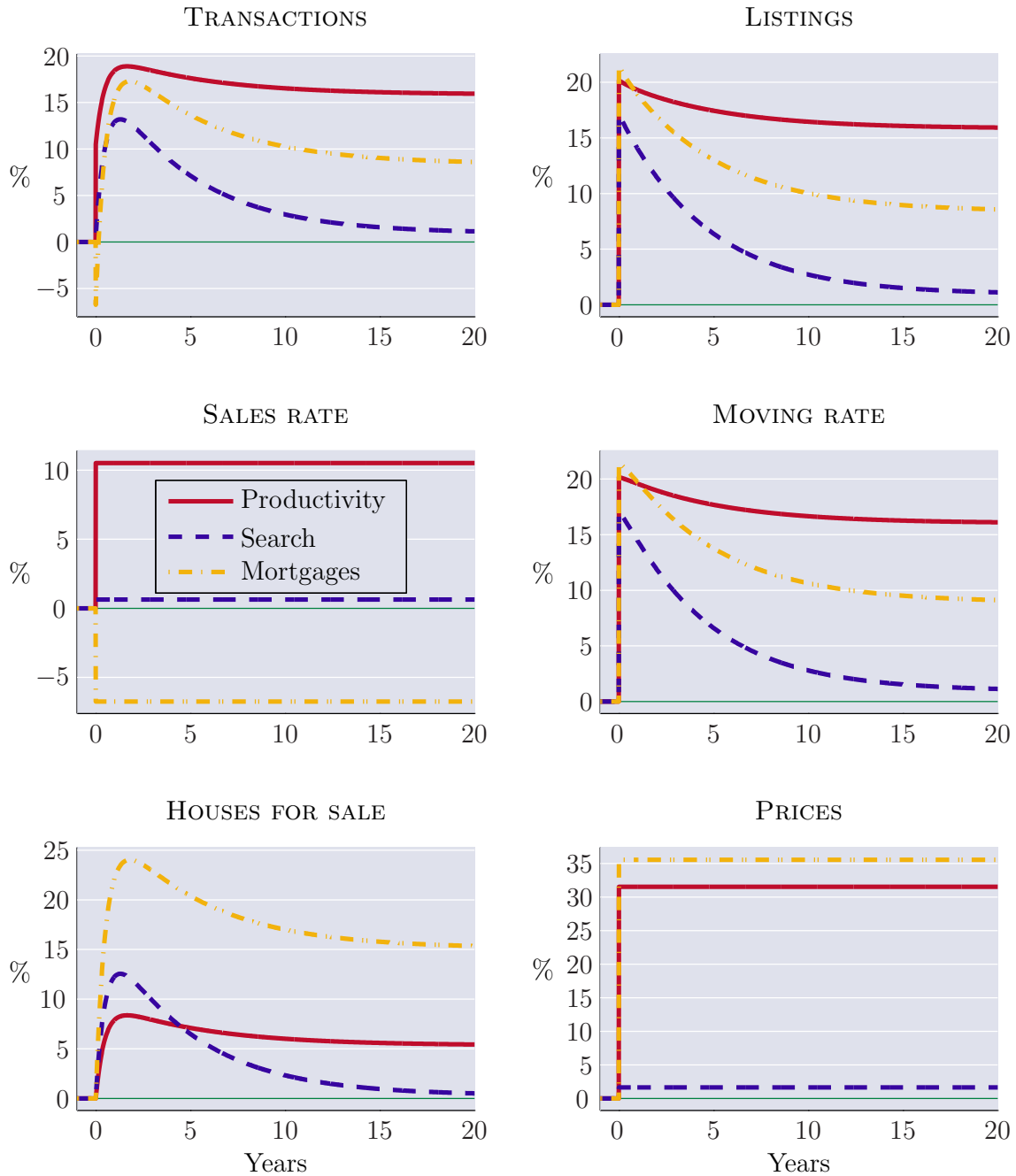
Using the calibrated version of the basic model with no heterogeneity, [Figure 10](#) shows the impulse responses of transactions, listings, the sales rates, the moving rate, houses for sale, and house prices to one-time permanent shocks to productivity, search efficiency, and interest rates. The sizes of these shocks are set to match the changes in the three factors during the 1995–2003 housing boom. The changes in the variables are reported as percentage deviations from their initial values in line with the earlier results shown in [Table 3](#) and [Table 4](#).

For the sales rate and house prices, there are no transitional dynamics. This is because these two variables are purely forward-looking in the model here. All the other variables display the overshooting pattern discussed intuitively in [section 4.4](#) and shown analytically in [appendix A.7](#). These dynamics are due to the persistence of housing match quality, which gives rise to a ‘cleansing effect’ when the moving threshold changes: an increase in moving improves the subsequent distribution of match quality all else equal, which leads to a smaller change in moving in the long run.

The calibrated model shows that overshooting is also quantitatively important, especially following the changes in search efficiency and interest rates. As well as the magnitude of overshooting being quantitatively significant, the duration of the transitional dynamics seen in [Figure 10](#) is also quite prolonged. It typically takes around 4–5 years for about half of the overshooting to dissipate. These dynamics are much more persistent than would be found in a standard exogenous moving

⁴⁸For example, to generate a simultaneous rise in prices and transactions with an exogenous rise in the moving rate, [Díaz and Jerez \(2013\)](#) require positively correlated housing demand and supply shocks, while [Ngai and Tenreyro \(2014\)](#) require a thick-market externality that improves the distribution of new match quality when there are more buyers and houses on the market.

Figure 10: *Dynamic responses to shocks*



Notes: The impulse response functions show the dynamic responses of the variables to unexpected and permanent shocks to productivity, search technology, and mortgage rates as discussed in [section 5.2.2](#).

model where only the stock of houses for sale would be a state variable, not the distribution of housing match quality.

The extent of overshooting for each of the developments individually, for productivity and search together, and for all three developments together is shown in [Table 5](#). This table reports the maximum size of the effects on each variable in the short to medium run, and the ratio of this effect to the ultimate long-run effects shown in earlier tables. As can be seen, the sizes of the transitional

effects of the shocks are now closer to matching the data on the boom in housing transactions 1995–2003.

Table 5: *The quantitative extent of overshooting*

Factor	Transactions	Listings	Sales rate	Moving rate	Houses for sale	Prices
Productivity boom	19%	20%	11%	20%	8%	32%
	[1.2]	[1.3]	[1]	[1.3]	[1.6]	[1]
Improved search efficiency	13%	17%	1%	17%	13%	2%
	[14.1]	[18.1]	[1]	[18.1]	[41.5]	[1]
Lower mortgage rates	17%	21%	−7%	21%	24%	36%
	[2.1]	[2.4]	[1]	[2.4]	[1.6]	[1]
Productivity & search	28%	32%	11%	32%	17%	33%
	[1.7]	[1.9]	[1]	[1.9]	[3.1]	[1]
Productivity, search, & mortgages	36%	42%	3%	42%	34%	69%
	[1.7]	[1.8]	[1]	[1.8]	[1.7]	[1]
Data (1995–2003)	27%	34%	14%	34%	13%	31%

Notes: The table shows the maximum size of each effect during the transition to the new steady state. The ratio of the maximum transitional effect to the ultimate long-run effect is given in brackets below.

6 Conclusions

This paper presents evidence that transaction dynamics are largely explained by changes in the frequency at which houses are put up for sale rather than changes in the length of time taken to sell them. The paper builds a tractable model to analyse moving house where a homeowner’s decision to move is an investment in housing match quality to reduce the degree of mismatch. Since moving house is an investment with upfront costs and potentially long-lasting benefits, the model predicts that the aggregate moving rate depends on macroeconomic variables. The endogeneity of moving means that those who move come from the bottom of the existing match quality distribution, and the non-random selection of movers gives rise to a cleansing effect that leads to overshooting of housing-market variables. The calibrated model can successfully account for some of the key aggregate housing-market dynamics during the 1995–2003 boom.

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). 4
- ANENBERG, E. AND BAYER, P. (2013), “[Endogenous sources of volatility in housing markets: The joint buyer-seller problem](#)”, working paper 18980, NBER. 4

- ANENBERG, E. AND LAUFER, S. (2017), “A more timely house price index”, *Review of Economics and Statistics*, **99(4)**:722–734 (October). 5, 11, 13, 70
- BACHMANN, R. AND COOPER, D. (2014), “The ins and arounds in the U.S. housing market”, working paper 14–3, Federal Reserve Bank of Boston. 3, 8, 14, 15, 26
- BURNSIDE, C., EICHENBAUM, M. AND REBELO, S. (2016), “Understanding booms and busts in housing markets”, *Journal of Political Economy*, **124(4)**:1088–1147 (August). 4, 5, 26
- CAPLIN, A. AND LEAHY, J. (2011), “Trading frictions and house price dynamics”, *Journal of Money, Credit and Banking*, **43(s2)**:283–303 (October). 4
- CARRILLO, P. E. AND WILLIAMS, B. (2015), “The repeat time-on-the-market index”, working paper 2015-8, Institute for International Economic Policy, George Washington University. 11, 12, 13, 14, 70
- 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. 32
- COLES, M. G. AND SMITH, E. (1998), “Marketplaces and matching”, *International Economic Review*, **39(1)**:239–254 (February). 4
- DAVIS, M. A. AND VAN NIEUWERBURGH, S. (2015), “Housing, finance, and the macroeconomy”, in G. Duranton, J. V. Henderson and W. C. Strange (eds.), *Handbook of Regional and Urban Economics*, vol. 5, North Holland, chap. 12, pp. 753–811. 4
- DEFUSCO, A. A., NATHANSON, C. G. AND ZWICK, E. (2017), “Speculative dynamics of prices and volume”, working paper 23449, NBER. 4, 5, 26
- DÍAZ, A. AND JEREZ, B. (2013), “House prices, sales, and time on the market: A search-theoretic framework”, *International Economic Review*, **54(3)**:837–872 (August). 4, 39, 70
- FERREIRA, F., GYOURKO, J. AND TRACY, J. (2010), “Housing busts and household mobility”, *Journal of Urban Economics*, **68(1)**:34–45 (July). 2
- FUSTER, A. AND ZAFAR, B. (2015), “The sensitivity of housing demand to financing conditions: Evidence from a survey”, Staff report 702, Federal Reserve Bank of New York. 37
- GENESOVE, D. AND HAN, L. (2012), “Search and matching in the housing market”, *Journal of Urban Economics*, **72(1)**:31–45 (July). 16, 17, 36, 69
- GHENT, A. (2012), “Infrequent housing adjustment, limited participation, and monetary policy”, *Journal of Money, Credit and Banking*, **44(5)**:931–955 (August). 32
- GLAESER, E. L., GOTTLIEB, J. D. AND GYOURKO, J. (2010), “Can cheap credit explain the housing boom?”, working paper 16230, NBER. 36
- GLAESER, E. L. AND GYOURKO, J. (2007), “Arbitrage in housing markets”, working paper 13704, NBER. 15
- GUREN, A. M. (2014), “The causes and consequences of house price momentum”, working paper, Boston University. 4, 5
- HAN, L. AND STRANGE, W. C. (2015), “The microstructure of housing markets: Search, bargaining, and brokerage”, in G. Duranton, J. V. Henderson and W. C. Strange (eds.), *Handbook of Regional And Urban Economics*, vol. 5, North Holland, chap. 13, pp. 813–886. 4

- HARMON, O. R. (1988), “The income elasticity of demand for single-family owner-occupied housing: An empirical reconciliation”, *Journal of Urban Economics*, **24(2)**:173–185 (September). 36
- HAUGHWOUT, A., LEE, D., TRACY, J. AND VAN DER KLAAUW, W. (2011), “Real estate investors, the leverage cycle, and the housing market crisis”, Staff report 514, Federal Reserve Bank of New York. 26
- HEAD, A., LLOYD-ELLIS, H. AND SUN, H. (2014), “Search, liquidity, and the dynamics of house prices and construction”, *American Economic Review*, **104(4)**:1172–1210 (April). 4
- HEDLUND, A. (2016), “Illiquidity and its discontents: Trading delays and foreclosures in the housing market”, *Journal of Monetary Economics*, **83**:1–13 (October). 4, 5
- 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). 70
- HEUSON, A. J. AND PAINTER, G. (2014), “The impact of the Taxpayer Relief Act of 1997 on housing turnover in the U.S. single-family residential market”, *Real Estate Economics*, **42(4)**:869–899 (Winter). 27
- JOVANOVIC, B. (1979), “Job matching and the theory of turnover”, *Journal of Political Economy*, **87(5)**:972–990 (October). 16
- KARAHAN, F. AND RHEE, S. (2013), “Geographical reallocation and unemployment during the Great Recession: The role of the housing bust”, Staff report 605, Federal Reserve Bank of New York. 4
- KRAINER, J. (2001), “A theory of liquidity in residential real estate markets”, *Journal of Urban Economics*, **49(1)**:32–53 (January). 4
- MOEN, E., NENOV, P. AND SNIKERS, F. (2015), “Buying first or selling first in housing markets”, discussion paper 10342, CEPR. 4
- 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). 3
- NGAI, L. R. AND TENREYRO, S. (2014), “Hot and cold seasons in the housing market”, *American Economic Review*, **104(12)**:3991–4026 (December). 4, 34, 39
- NOVY-MARX, R. (2009), “Hot and cold markets”, *Real Estate Economics*, **37(1)**:1–22 (Spring). 4, 17
- 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). 4
- PIAZZESI, M., SCHNEIDER, M. AND STROEBEL, J. (2015), “Segmented housing search”, working paper 20823, NBER. 4
- PISSARIDES, C. A. (1985), “Short-run equilibrium dynamics of unemployment, vacancies and real wages”, *American Economic Review*, **75(4)**:676–690 (September). 16
- POTERBA, J. M. (1991), “House price dynamics: The role of tax policy and demography”, *Brookings Papers on Economic Activity*, **1991(2)**:143–203. 32

- QUIGLEY, J. M. (1987), “Interest rate variations, mortgage prepayments and household mobility”, *Review of Economics and Statistics*, **69(4)**:636–643 (November). 2
- (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. 32
- SHAN, H. (2011), “The effect of capital gains taxation on home sales: Evidence from the Taxpayer Relief Act of 1997”, *Journal of Public Economics*, **95(1–2)**:177–188 (February). 27
- 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). 5
- WHEATON, W. C. (1990), “Vacancy, search, and prices in a housing market matching model”, *Journal of Political Economy*, **98(6)**:1270–1292 (December). 4
- WHEATON, W. C. AND LEE, N. J. (2009), “The co-movement of housing sales and housing prices: Empirics and theory”, working paper, MIT. 15

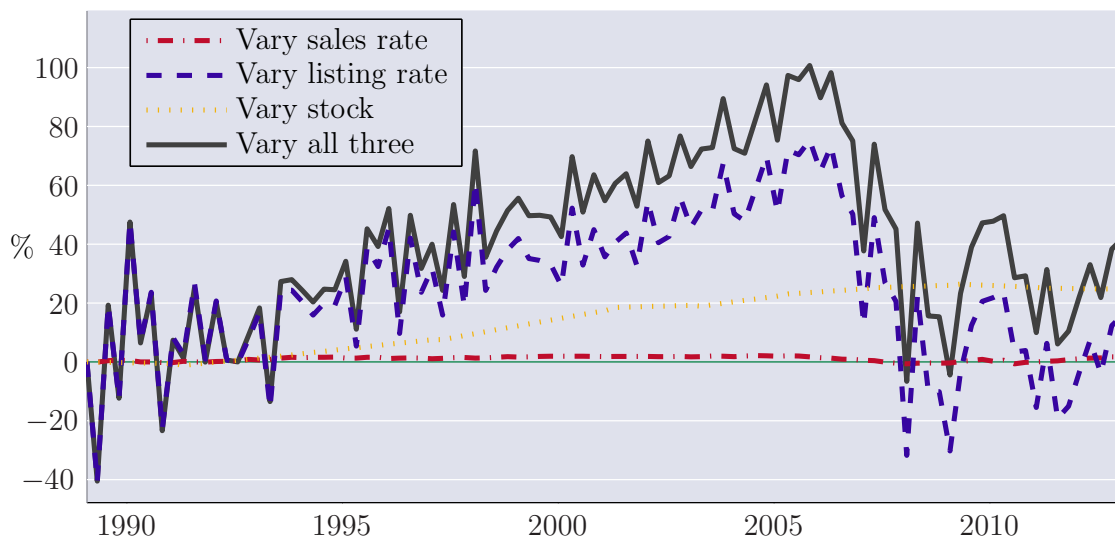
A Appendices

A.1 Measurement of transactions and inventories

The transactions and inventories data from NAR includes all existing single-family homes, so homes that are rented out are counted in this data. An estimate of the fraction of single-family homes that are not rented out can be computed from the AHS. This fraction is around 78% on average over the period 1989–2013. If the fraction were constant over time, the counterfactuals presented in [section 2.2](#) would be unaffected and the only change would be to the average level of the listing rate. The effect on the average level of the listing rate would change the implied average time between moves from 15 years to 19.4 years.

However, the data show some changes over time in the fraction of homes that are not rented out. The fraction rises from 77% to 80% during the boom period, and then falls to 75% by 2013. A simple robustness check is to scale the NAR data on transactions and inventories by the fraction of non-rented homes and then recompute the counterfactuals. The results are shown in [Figure 11](#), which are almost identical to those in [Figure 2](#).

Figure 11: *Actual and counterfactual transactions (adjusted for rented homes)*



Notes: The series are reported as log differences relative to their initial values.

A.2 Withdrawals and relistings

Consider the framework introduced in [section 2.3](#). Suppose that the net listing rate $n_t = N_t/(K_t - U_t)$ is equal to a constant n . The sales rate is s , so $S_t = sU_t$, and the total housing stock grows at rate g , that is, $\dot{K}_t/K_t = g$. This implies that equation (2.1) will hold as in [section 2.1](#) and that the fraction of houses for sale $u_t = U_t/K_t$ will converge to its steady state $u = n/(n + s + g)$.

Let $f_t = F_t/K_t$ denote the stock of houses that have failed to sell but might be relisted relative to the stock of all houses. Using the formula for \dot{F}_t , the implied law of motion for f_t is:

$$\dot{f}_t = wu_t - (\rho + \alpha + m + g)f_t.$$

Given that u_t is equal to its steady-state value u , there is a steady state f for f_t :

$$f = \frac{wu}{\rho + \alpha + m + g}, \tag{A.2.1}$$

and convergence to this steady state takes place at rate $\rho + \alpha + m + g$. Since this rate is strictly greater than the sum of the relisting rate ρ and the rate α at which homeowners give up on attempting a future sale, convergence to the steady state is presumed to be sufficiently fast that f_t is set equal to f in what follows, just as u_t converges quickly enough to u to set $u_t = u$.⁴⁹

Dividing both sides of the accounting identity for net listings by $K_t - U_t$ implies:

$$n = m - w \frac{u_t}{1 - u_t} + \rho \frac{f_t}{1 - u_t}.$$

If u_t and f_t have reached their steady-state values u and f then the net listing rate n is indeed constant as supposed earlier:

$$n = m - w \frac{u}{1 - u} + \rho \frac{f}{1 - u},$$

and noting $u/(1 - u) = n/(s + g)$ and using the expression for f from (A.2.1):

$$n = m - \frac{wn}{s + g} \left(1 - \frac{\rho}{\rho + \alpha + m + g} \right) = m - n \frac{w}{s + g} \left(\frac{\alpha + m + g}{\rho + \alpha + m + g} \right).$$

Divide numerator and denominator of the term in parentheses by $\rho + \alpha + m$, and numerator and denominator of the coefficient by s :

$$n = m - \frac{1}{1 + \frac{g}{s}} \frac{w}{s} \left(\frac{\frac{\alpha + m}{\rho + \alpha + m} + \frac{g}{\rho + \alpha + m}}{1 + \frac{g}{\rho + \alpha + m}} \right) n. \quad [\text{A.2.2}]$$

The formulas for the eventual fraction of withdrawals ϕ and the eventual fraction of relistings β imply that:

$$\frac{\phi}{1 - \phi} = \frac{w}{s}, \quad \text{and} \quad 1 - \beta = \frac{\alpha + m}{\rho + \alpha + m}.$$

These expressions can be substituted into (A.2.2) to deduce:

$$n = m - \left(\frac{\phi}{1 - \phi} \right) \left(\frac{1 - \beta + \frac{g}{\rho + \alpha + m}}{1 + \frac{g}{\rho + \alpha + m}} \right) \left(\frac{1}{1 + \frac{g}{s}} \right) n. \quad [\text{A.2.3}]$$

This is the exact link between the moving rate m and the net listing rate n . In addition to ϕ and β , it depends on the sales rate s , the growth rate g , the relisting rate ρ , and the abandonment rate α .

If the growth rate of the total housing stock g is small in relation to the sales rate s , and the sum of the relisting, abandonment, and moving rates $\rho + \alpha + m$ then the terms g/s , $g/(\rho + \alpha + m)$ are negligible. The equation (A.2.3) linking n and m can then be well approximated by:

$$n \approx m - \frac{\phi(1 - \beta)}{(1 - \phi)} n.$$

Collecting terms in n on one side, this equation simplifies to:

$$n \approx \frac{m}{1 + \frac{\phi(1 - \beta)}{1 - \phi}},$$

which confirms equation (2.9) in section 2.3. Given knowledge of the eventual fraction of withdrawals ϕ and the eventual fraction of relistings β , the moving rate m can be calculated from the net listing rate n , or vice versa, the net listing rate can be calculated from the moving rate.

⁴⁹Confirming this assumption would require an empirical measure of the abandonment rate α .

A.3 Value functions and thresholds

Moving and transaction thresholds

The value functions $H_t(\epsilon)$ and W_t and the thresholds x_t and y_t satisfy the equations (3.5), (3.6), (3.8), and (3.9). No other variables appear in these equations. Given constant parameters, there is a time-invariant solution $H_t(\epsilon) = H(\epsilon)$, $W_t = W$, $x_t = x$, and $y_t = y$. The time-invariant equations are:

$$rH(\epsilon) = \epsilon\xi - M + a(\max\{H(\delta\epsilon), W\} - H(\epsilon)); \quad [\text{A.3.1}]$$

$$H(x) = W; \quad [\text{A.3.2}]$$

$$rW = -F - M + v \int_y (H(\epsilon) - W - C) dG(\epsilon); \quad [\text{A.3.3}]$$

$$H(y) = W + C. \quad [\text{A.3.4}]$$

Attention is restricted to parameters where the solution will satisfy $\delta y < x$.

Evaluating (A.3.1) at $\epsilon = x$, noting that $\delta < 1$ and $H(\epsilon)$ is increasing in ϵ :

$$rH(x) = \xi x - M + a(W - H(x)).$$

Since $H(x) = W$ (equation A.3.2), it follows that:

$$W = H(x) = \frac{\xi x - M}{r}. \quad [\text{A.3.5}]$$

Next, evaluate (A.3.1) at $\epsilon = y$. With the restriction $\delta y < x$, it follows that $H(\delta y) < H(x) = W$, and hence:

$$rH(y) = \xi y - M + a(W - H(y)).$$

Collecting terms in $H(y)$ on one side and substituting the expression for W from (A.3.5):

$$(r + a)H(y) = \xi y - M + \frac{a}{r}(\xi x - M) = \xi(y - x) + \left(1 + \frac{a}{r}\right)(\xi x - M),$$

and thus $H(y)$ is given by:

$$H(y) = \frac{\xi x - M}{r} + \frac{\xi(y - x)}{r + a}. \quad [\text{A.3.6}]$$

Combining the equation above with (A.3.4) and (A.3.5), it can be seen that the thresholds y and x must be related as follows:

$$y - x = \frac{(r + a)C}{\xi}. \quad [\text{A.3.7}]$$

Using the expression for the Pareto distribution function (3.1) and using (A.3.4) to note $H(\epsilon) - W - C = H(\epsilon) - H(y)$, the Bellman equation (A.3.3) can be written as:

$$rW = -F - M + v y^{-\lambda} \int_{\epsilon=y}^{\infty} \frac{\lambda}{y} \left(\frac{\epsilon}{y}\right)^{-(\lambda+1)} (H(\epsilon) - H(y)) d\epsilon, \quad [\text{A.3.8}]$$

which assumes $y > 1$. In solving this equation it is helpful to define the following function $\Psi(z)$ for all $z \leq y$:

$$\Psi(z) \equiv \int_{\epsilon=z}^{\infty} \frac{\lambda}{z} \left(\frac{\epsilon}{z}\right)^{-(\lambda+1)} (H(\epsilon) - H(z)) d\epsilon. \quad [\text{A.3.9}]$$

Since $\delta y < x$ is assumed and $z \leq y$, it follows that $\delta z < x$, and thus $H(\delta z) < H(x) = W$. Equation (A.3.1)

evaluated at $\epsilon = z$ therefore implies:

$$rH(z) = \xi z - M + a(W - H(z)).$$

Subtracting this equation from (A.3.1) evaluated at a general value of ϵ leads to:

$$\begin{aligned} r(H(\epsilon) - H(z)) &= \xi(\epsilon - z) + a(\max\{H(\delta\epsilon), W\} - H(\epsilon)) - a(W - H(z)) \\ &= \xi(\epsilon - z) - a(H(\epsilon) - H(z)) + a\max\{H(\delta\epsilon) - W, 0\}. \end{aligned}$$

Noting that $W = H(x)$ and solving for $H(\epsilon) - H(z)$:

$$H(\epsilon) - H(z) = \frac{\xi}{r+a}(\epsilon - z) + \frac{a}{r+a}\max\{H(\delta\epsilon) - H(x), 0\}. \quad [\text{A.3.10}]$$

The equation above can be substituted into (A.3.9) to deduce:

$$\Psi(z) = \frac{\xi}{r+a} \int_{\epsilon=z}^{\infty} \frac{\lambda}{z} \left(\frac{\epsilon}{z}\right)^{-(\lambda+1)} (\epsilon - z) d\epsilon + \frac{a}{r+a} \int_{\epsilon=z}^{\infty} \frac{\lambda}{z} \left(\frac{\epsilon}{z}\right)^{-(\lambda+1)} \max\{H(\delta\epsilon) - H(x), 0\} d\epsilon. \quad [\text{A.3.11}]$$

First, observe that:

$$\int_{\epsilon=z}^{\infty} \frac{\lambda}{z} \left(\frac{\epsilon}{z}\right)^{-(\lambda+1)} (\epsilon - z) d\epsilon = \frac{z}{\lambda - 1}. \quad [\text{A.3.12}]$$

Next, make the change of variable $\epsilon' = \delta\epsilon$ in the second integral in (A.3.11) to deduce:

$$\begin{aligned} \int_{\epsilon=z}^{\infty} \frac{\lambda}{z} \left(\frac{\epsilon}{z}\right)^{-(\lambda+1)} \max\{H(\delta\epsilon) - H(x), 0\} d\epsilon &= \int_{\epsilon'=\delta z}^{\infty} \frac{\lambda}{\delta z} \left(\frac{\epsilon'}{\delta z}\right)^{-(\lambda+1)} \max\{H(\epsilon') - H(x), 0\} d\epsilon' \\ &= \int_{\epsilon'=\delta z}^x \frac{\lambda}{\delta z} \left(\frac{\epsilon'}{\delta z}\right)^{-(\lambda+1)} 0 d\epsilon' + \int_{\epsilon'=x}^{\infty} \frac{\lambda}{\delta z} \left(\frac{\epsilon'}{\delta z}\right)^{-(\lambda+1)} (H(\epsilon') - H(x)) d\epsilon' \\ &= \left(\frac{\delta z}{x}\right)^{\lambda} \int_{\epsilon=x}^{\infty} \frac{\lambda}{x} \left(\frac{\epsilon}{x}\right)^{-(\lambda+1)} (H(\epsilon) - H(x)) d\epsilon = \left(\frac{\delta z}{x}\right)^{\lambda} \Psi(x), \end{aligned}$$

where the second line uses $\delta z < x$ (as $z \leq y$ and $\delta y < x$) and $H(\epsilon') < H(x)$ for $\epsilon' < x$, and the final line uses the definition (A.3.9). Putting the equation above together with (A.3.11) and (A.3.12) yields the following for all $z \leq y$:

$$\Psi(z) = \frac{\xi z}{(r+a)(\lambda-1)} + \frac{a}{r+a} \left(\frac{\delta z}{x}\right)^{\lambda} \Psi(x). \quad [\text{A.3.13}]$$

Evaluating this expression at $z = x$ (with $x < y$):

$$\Psi(x) = \frac{\xi x}{(r+a)(\lambda-1)} + \frac{a}{r+a} \delta^{\lambda} \Psi(x),$$

and hence $\Psi(x)$ is given by:

$$\Psi(x) = \frac{\xi x}{(r+a(1-\delta^{\lambda}))(\lambda-1)}. \quad [\text{A.3.14}]$$

Next, evaluating (A.3.13) at $z = y$ and using (A.3.14) to substitute for $\Psi(x)$:

$$\Psi(y) = \frac{\xi y}{(r+a)(\lambda-1)} + \frac{a}{r+a} \left(\frac{\delta z}{x}\right)^{\lambda} \left(\frac{\xi x}{(r+a(1-\delta^{\lambda}))(\lambda-1)}\right),$$

and simplifying this equation yields the following expression for $\Psi(y)$:

$$\Psi(y) = \frac{\xi}{(r+a)(\lambda-1)} \left(y + \frac{a\delta^\lambda y^\lambda x^{1-\lambda}}{r+a(1-\delta^\lambda)} \right). \quad [\text{A.3.15}]$$

Using the definition (A.3.9), equation (A.3.8) can be written in terms of $\Psi(y)$:

$$rW = -F - M + vy^{-\lambda}\Psi(y),$$

and substituting from (A.3.5) and (A.3.15) yields:

$$\xi x - M = -F - M + vy^{-\lambda} \left(\frac{\xi}{(r+a)(\lambda-1)} \left(y + \frac{a\delta^\lambda y^\lambda x^{1-\lambda}}{r+a(1-\delta^\lambda)} \right) \right).$$

This equation can be simplified as follows:

$$x + \frac{F}{\xi} = \frac{v}{(\lambda-1)(r+a)} \left(y^{1-\lambda} + \frac{a\delta^\lambda}{r+a(1-\delta^\lambda)} x^{1-\lambda} \right). \quad [\text{A.3.16}]$$

The two equations (A.3.7) and (A.3.16) confirm (4.4) and (4.7) given in the main text. These can be solved for the thresholds x and y .

Existence and uniqueness

By using equation (A.3.7) to replace x with a linear function of y , the equilibrium threshold y is the solution of the equation:

$$\mathcal{I}(y) \equiv \frac{v}{(\lambda-1)(r+a)} \left(y^{1-\lambda} + \frac{a\delta^\lambda}{r+a(1-\delta^\lambda)} \left(y - \frac{(r+a)C}{\xi} \right)^{1-\lambda} \right) - y + \frac{(r+a)C}{\xi} - \frac{F}{\xi} = 0. \quad [\text{A.3.17}]$$

It can be seen immediately (since $\lambda > 1$) that $\mathcal{I}'(y) < 0$, so any solution that exists is unique. A valid solution must satisfy $x > 0$, $y > 1$, and $\delta y < x$. Using equation (A.3.7), the inequality $\delta y < x$ is equivalent to:

$$\delta y < y - \frac{(r+a)C}{\xi},$$

which is in turn equivalent to:

$$y > \frac{(r+a)C}{(1-\delta)\xi}.$$

Thus, to satisfy $y > 1$ and $\delta y < x$, the equilibrium must feature:

$$y > \max \left\{ 1, \frac{(r+a)C}{(1-\delta)\xi} \right\}. \quad [\text{A.3.18}]$$

Observe that $\lim_{y \rightarrow \infty} \mathcal{I}(y) = -\infty$ (using A.3.17 and $\lambda > 1$), so an equilibrium satisfying (A.3.18) exists if and only if:

$$\mathcal{I} \left(\max \left\{ 1, \frac{(r+a)C}{(1-\delta)\xi} \right\} \right) > 0. \quad [\text{A.3.19}]$$

If the condition (A.3.18) is satisfied then by using (A.3.7):

$$x > \max \left\{ 1, \frac{(r+a)C}{(1-\delta)\xi} \right\} - \frac{(r+a)C}{\xi} > \frac{(r+a)C}{(1-\delta)\xi} - \frac{(r+a)C}{\xi} = \frac{\delta(r+a)C}{(1-\delta)\xi} > 0,$$

confirming that $x > 0$ must hold. Therefore, (A.3.19) is necessary and sufficient for the existence of a unique equilibrium satisfying all the required conditions. Using equation (A.3.17), (A.3.19) is equivalent to the condition in (4.8).

Surplus and selling rate

Given x and y , the value functions W , $H(x)$, and $H(y)$ can be obtained from (A.3.5) and (A.3.6). The average surplus can be found by combining (A.3.5) and (A.3.8) to deduce:

$$\int_{\epsilon=y}^{\infty} \frac{\lambda}{y} \left(\frac{\epsilon}{y}\right)^{-(\lambda+1)} (H(\epsilon) - H(y)) d\epsilon = \frac{y^\lambda}{v} (\xi x + F). \quad [\text{A.3.20}]$$

Given x and y , the probability π that a viewing leads to a sale, and the expected number of viewings before a sale $V_s = 1/\pi$ are:

$$\pi = y^{-\lambda}, \quad \text{and} \quad V_s = y^\lambda. \quad [\text{A.3.21}]$$

The selling rate s and the expected time-to-sell $T_s = 1/s$ are given by:

$$s = v y^{-\lambda}, \quad \text{and} \quad T_s = \frac{y^\lambda}{v}. \quad [\text{A.3.22}]$$

A.4 Prices

Nash bargaining

The price $p_t(\epsilon)$ is determined by combining the Nash bargaining solution $\omega \Sigma_{b,t}(\epsilon) = (1 - \omega) \Sigma_{u,t}(\epsilon)$ with the expressions for the buyer and seller surpluses in (3.2):

$$\omega(H_t(\epsilon) - p_t(\epsilon) - C_b - B_t) = (1 - \omega)(p_t(\epsilon) - C_v - U_t),$$

from which it follows that:

$$p_t(\epsilon) = \omega H_t(\epsilon) + (1 - \omega)C_v - \omega C_b + ((1 - \omega)U_t - \omega B_t). \quad [\text{A.4.1}]$$

The surplus-splitting condition implies $\Sigma_{b,t}(\epsilon) = (1 - \omega)\Sigma_t(\epsilon)$ and $\Sigma_{u,t}(\epsilon) = \omega\Sigma_t(\epsilon)$, with $\Sigma_t(\epsilon) = \Sigma_{b,t}(\epsilon) + \Sigma_{u,t}(\epsilon)$ being the total surplus from (4.6). The Bellman equations in (3.3) can thus be written as:

$$rB_t = -F + (1 - \omega)v \int_{y_t} \Sigma_t(\epsilon) dG(\epsilon) + \dot{B}_t, \quad \text{and} \quad rU_t = -M + \omega v \int_{y_t} \Sigma_t(\epsilon) dG(\epsilon) + \dot{U}_t,$$

and a multiple ω of the first equation can be subtracted from a multiple $1 - \omega$ of the second equation to deduce:

$$r((1 - \omega)U_t - \omega B_t) = \omega F - (1 - \omega)M + ((1 - \omega)\dot{U}_t - \omega\dot{B}_t).$$

The stationary solution of this equation is:

$$(1 - \omega)U_t - \omega B_t = \frac{\omega F - (1 - \omega)M}{r},$$

and by substituting this into (A.4.1):

$$p_t(\epsilon) = \omega H_t(\epsilon) + (1 - \omega)C_v - \omega C_b + \frac{\omega F - (1 - \omega)M}{r}. \quad [\text{A.4.2}]$$

Integrating this equation over the distribution of new match quality yields equation (3.7) for the average transaction price.

Average transactions price

In an equilibrium where the moving and transactions thresholds x_t and y_t are constant over time, the value function $H_t(\epsilon)$ is equal to the time-invariant function $H(\epsilon)$. This means that prices $p_t(\epsilon) = p(\epsilon)$ are also time invariant. Using the Pareto distribution function (3.1) and equation (3.7), the average price is:

$$P = \omega \int_y \frac{\lambda}{y} \left(\frac{\epsilon}{y}\right)^{-(\lambda+1)} H(\epsilon) d\epsilon + (1 - \omega)C_v - \omega C_b + \frac{\omega F - (1 - \omega)M}{r}.$$

By using equation (3.5) and (A.3.5), the above can be written as:

$$P = \omega \int_y \frac{\lambda}{y} \left(\frac{\epsilon}{y}\right)^{-(\lambda+1)} (H(\epsilon) - H(y)) d\epsilon + \omega \left(\frac{\xi x - M}{r} + C\right) + \frac{\omega F - (1 - \omega)M}{r} + (1 - \omega)C_v - \omega C_b,$$

and substituting from (A.3.20) yields:

$$P = \omega \frac{y^\lambda}{v} (\xi x + F) + \omega C + \frac{\omega \xi x}{r} + (1 - \omega)C_v - \omega C_b + \frac{\omega F - M}{r}.$$

Simplifying and using $C = C_b + C_v$, the following expression for the average price is obtained:

$$P = C_v - \frac{M}{r} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v}\right) (\xi x + F),$$

which confirms the formula in (4.13). With the definition of $\kappa = C_v/C$, this equation can also be written as:

$$P = \kappa C - \frac{M}{r} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v}\right) (\xi x + F). \tag{A.4.3}$$

A.5 Stocks and flows

The moving rate

The formula (3.12) for the moving rate can also be given in terms of inflows $N_t = n_t(1 - u_t)$, where u_t is the stock of unsold houses:

$$N_t = a(1 - u_t) - a\delta^\lambda x_t^{-\lambda} v \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau. \tag{A.5.1}$$

The first term $a(1 - u_t)$ is the quantity of existing matches that receive a shock (arrival rate a). The second term is the quantity of existing matches that receive a shock now, but decide not to move. The difference between these two numbers (under the assumption that only those who receive a shock make a moving decision) gives inflows N_t .

Now consider the derivation of the second term in (A.5.1). The distribution of existing matches (measure $1 - u_t$) can be partitioned into vintages τ (when matches formed) and the number k of previous shocks that have been received. At time τ , a quantity u_τ of houses were for sale, and viewings arrived at rate v . Viewings were draws of match quality ϵ from a Pareto(1, λ) distribution, and those draws with $\epsilon \geq y_\tau$ formed new matches, truncating the distribution at y_τ . In the interval between τ and t , those matches that have received k shocks now have match quality $\delta^k \epsilon$. Some of these matches will have been destroyed as a result of these shocks, truncating the distribution of surviving match quality. Because the distribution of initial match quality is a Pareto distribution, these truncations also result in Pareto distributions with the same shape parameter λ .

Consider the matches of vintage τ . All of these were originally from a Pareto distribution truncated at $\epsilon \geq y_\tau$. Subsequently, depending on the arrival of idiosyncratic shocks (both timing and number), this

distribution may have been truncated further. Let z denote the last truncation point in terms of the original match quality ϵ (at the time of the viewing). This is $z = y_\tau$ if no shocks have been received, or $z = \delta^{-k} x_T$ if k shocks have been received and the last one occurred at time T when the moving threshold was x_T . Conditional on this last truncation point z , it is shown below that the measure of surviving matches is $z^{-\lambda} v u_\tau$. Furthermore, the original match quality of these surviving matches must be a Pareto(z, λ) distribution.

Now consider the distribution of the number of previous shocks k between τ and t . Given the Poisson arrival rate a , k has a Poisson distribution, so the probability of k is $e^{-a(t-\tau)}(a(t-\tau))^k/k!$. If a shock arrives at time t , matches of current quality greater than x_t survive. If these have received k shocks earlier, this means the truncation threshold in terms of original match quality ϵ is $\epsilon \geq \delta^{-(k+1)} x_t$. Of these matches that have accumulated k earlier shocks, suppose last relevant truncation threshold (in terms of original match quality) was z (this will vary over those matches even with the same number of shocks because the timing might be different), so the distribution of surviving matches in terms of their original match quality is Pareto(z, λ). The probability that these matches then survive the shock at time t is given by $(\delta^{-(k+1)} x_t/z)^{-\lambda}$, and multiplying this by $z^{-\lambda} v u_\tau$ gives the number that survive:

$$(\delta^{-(k+1)} x_t/z)^{-\lambda} z^{-\lambda} v u_\tau = (\delta^\lambda)^{k+1} x_t^{-\lambda} v u_\tau,$$

noting that the terms in z cancel out. This is conditional on z , k , and τ , but since z does not appear above, the distribution of the past truncation thresholds is not needed. Averaging over the distribution of k yields:

$$\begin{aligned} \sum_{k=0}^{\infty} e^{-a(t-\tau)} \frac{(a(t-\tau))^k}{k!} (\delta^\lambda)^{k+1} x_t^{-\lambda} v u_\tau &= \delta^\lambda x_t^{-\lambda} v u_\tau e^{-a(t-\tau)} \sum_{k=0}^{\infty} \frac{(a\delta^\lambda(t-\tau))^k}{k!} \\ &= \delta^\lambda x_t^{-\lambda} v u_\tau e^{-a(t-\tau)} e^{a\delta^\lambda(t-\tau)} = \delta^\lambda x_t^{-\lambda} v u_\tau e^{-a(1-\delta^\lambda)(t-\tau)}, \end{aligned}$$

where the penultimate expression uses the Taylor series expansion of the exponential function $e^z = \sum_{k=0}^{\infty} z^k/k!$ (valid for all z). Next, integrating over all vintages τ before the current time t leads to:

$$\int_{\tau \rightarrow -\infty}^t \delta^\lambda x_t^{-\lambda} e^{-a(1-\delta^\lambda)(t-\tau)} d\tau = \delta^\lambda x_t^{-\lambda} \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} v u_\tau d\tau.$$

Multiplying this by the arrival rate a of the idiosyncratic shocks confirms the second term of the expression for N_t in (A.5.1).

This leaves only the claim that the measure of vintage- τ surviving matches with truncation point z (in terms of the original match quality distribution ϵ) is $z^{-\lambda} v u_\tau$. When these matches first form, they have measure $y_\tau^{-\lambda} v u_\tau$ and a Pareto(y_t, λ) distribution, so the formula is correct if no shocks have occurred and $z = y_\tau$. Now suppose the formula is valid for some z and truncation now occurs at a new point $w > z$ (in terms of original match quality). Since matches surviving truncation at z have distribution Pareto(z, λ), the proportion of these that survive the new truncation is $(w/z)^{-\lambda}$, and so the measure becomes $(w/z)^{-\lambda} z^{-\lambda} v u_\tau = w^{-\lambda} v u_\tau$ (with the term in z cancelling out), which confirms the claim.

The distribution of match quality

Now consider the derivation of the law of motion for average match quality Q_t in (3.13). Let total match quality across all families be denoted by \mathcal{E}_t (those not matched have match quality equal to zero), with $\mathcal{E}_t = (1 - u_t)Q_t$ by definition. Total match quality \mathcal{E}_t changes over time as new matches form, when matches are hit by shocks, and when moving decisions are made. With transaction threshold y_t and new match quality drawn from a Pareto(1; λ) distribution, new matches have average quality $(\lambda/(\lambda - 1))y_t$. The contribution to the rate of change of total match quality is that average multiplied by $s_t u_t$. Shocks to existing matches arrive randomly at rate a . If no shock is received then there is no change to match quality and no moving decision. For those who receive a shock, let $\underline{\mathcal{E}}_t$ denote the total match quality of those matches that survive (with matches that dissolve counted as having zero match quality). The contribution

of the shocks and moving decisions to the rate of change of total match quality is to subtract $a(\mathcal{E}_t - \underline{\mathcal{E}}_t)$. The differential equation for \mathcal{E}_t is therefore:

$$\dot{\mathcal{E}}_t = \frac{\lambda}{\lambda - 1} y_t s_t u_t - a(\mathcal{E}_t - \underline{\mathcal{E}}_t). \quad [\text{A.5.2}]$$

Using this formula requires an expression for $\underline{\mathcal{E}}_t$.

Consider the distribution of all matches that formed before time t , survived until time t , and now receive an idiosyncratic shock at time t , but one that is not sufficient to trigger moving. The distribution of surviving matches can be partitioned into vintages τ (when the match formed) and the number of shocks k that have been received previously (not counting the shock at time t). At time τ , a quantity u_τ of houses were for sale, and viewings arrived at rate v . Viewings were draws of match quality ϵ from a $\text{Pareto}(1; \lambda)$ distribution, and those draws with $\epsilon \geq y_\tau$ formed new matches, truncating the distribution at y_τ . Subsequently, a number k of idiosyncratic shocks have occurred, with k having a $\text{Poisson}(a(t - \tau))$ distribution, and these shocks resulting in the distribution of surviving match quality being truncated. With a shock now occurring at time t after k earlier shocks, match quality is now $\delta^{k+1}\epsilon$, and the distribution is truncated at x_t . In terms of the original match quality ϵ , survival requires $\epsilon \geq \delta^{-(k+1)}x_t$.

Consider matches of vintage τ that have previously accumulated k shocks for which the last truncation threshold was z in terms of original match quality (this threshold will depend on when the previous shocks occurred). Since the Pareto distribution is preserved after truncation with the same shape parameter, these matches have a $\text{Pareto}(z; \lambda)$ distribution in terms of their original match quality. It was shown above that the measure of surviving vintage- τ matches with truncation point z is $z^{-\lambda} v u_\tau$ (conditional on z , the number of shocks k is irrelevant, though the number of shocks may be related to the value of z). The measure that remain ($\epsilon \geq \delta^{-(k+1)}x_t$) after moving decisions are made at time t is:

$$\left(\delta^{-(k+1)} x_t / z \right)^{-\lambda} z^{-\lambda} v u_\tau = (\delta^\lambda)^{k+1} x_t^{-\lambda} v u_\tau,$$

noting that the terms in z cancel out. The probability of drawing k shocks in the interval between τ and t is $e^{-a(t-\tau)} (a(t-\tau))^k / k!$, and hence averaging over the distribution of k for vintage- τ matches implies that the surviving measure is:

$$\sum_{k=0}^{\infty} e^{-a(t-\tau)} \frac{(a(t-\tau))^k}{k!} (\delta^{k+1})^\lambda x_t^{-\lambda} v u_\tau = \delta^\lambda x_t^{-\lambda} v u_\tau e^{-a(1-\delta^\lambda)(t-\tau)},$$

which is confirmed by following the same steps as in the derivation of the moving rate above. Integrating these surviving measures over all cohorts:

$$\int_{\tau \rightarrow -\infty}^t \delta^\lambda x_t^{-\lambda} v u_\tau e^{-a(1-\delta^\lambda)(t-\tau)} d\tau = v \delta^\lambda x_t^{-\lambda} \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau,$$

and since the average match quality among the survivors after the shock at time t is $(\lambda/(\lambda - 1))x_t$ for all cohorts, it follows that:

$$\underline{\mathcal{E}}_t = \frac{v \delta^\lambda \lambda}{\lambda - 1} x_t^{1-\lambda} \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau.$$

Substituting this and equation (3.11) into (A.5.2) implies:

$$\dot{\mathcal{E}}_t = \frac{v \lambda}{\lambda - 1} y_t^{1-\lambda} u_t - a \left(\mathcal{E}_t - \frac{v \delta^\lambda \lambda}{\lambda - 1} x_t^{1-\lambda} \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau \right). \quad [\text{A.5.3}]$$

The integral can be eliminated by defining an additional variable Z_t :

$$Z_t = \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau, \quad [\text{A.5.4}]$$

and hence (A.5.3) can be written as follows:

$$\dot{\mathcal{E}}_t = \frac{v\lambda}{\lambda-1}y_t^{1-\lambda}u_t - a\mathcal{E}_t + \frac{av\delta^\lambda\lambda}{\lambda-1}x_t^{1-\lambda}Z_t. \quad [\text{A.5.5}]$$

The evolution of the state variable u_t is determined by combining equations (3.10), (3.11), and (3.12):

$$\dot{u}_t = a(1-u_t) - av\delta^\lambda x_t^{-\lambda} \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau - vy_t^{-\lambda}u_t,$$

where the integral can again be eliminated by writing the equation in terms of the new variable Z_t from (A.5.4):

$$\dot{u}_t = a(1-u_t) - av\delta^\lambda x_t^{-\lambda}Z_t - vy_t^{-\lambda}u_t. \quad [\text{A.5.6}]$$

Differentiating the integral in (A.5.4) shows that Z_t must satisfy the differential equation:

$$\dot{Z}_t = u_t - a(1-\delta^\lambda)Z_t. \quad [\text{A.5.7}]$$

These results can be used to obtain the differential equation for average match quality Q_t in (3.13). Since the definition implies $Q_t = \mathcal{E}_t/(1-u_t)$, it follows that:

$$\dot{Q}_t = \frac{\dot{\mathcal{E}}_t}{1-u_t} + \frac{\mathcal{E}_t\dot{u}_t}{(1-u_t)^2} = \frac{\dot{\mathcal{E}}_t}{1-u_t} + Q_t \frac{\dot{u}_t}{1-u_t}.$$

Substituting from the differential equations (A.5.5) and (A.5.6) leads to:

$$\dot{Q}_t = \left(\frac{v\lambda}{\lambda-1}y_t^{1-\lambda} \frac{u_t}{1-u_t} - aQ_t + \frac{av\delta^\lambda\lambda}{\lambda-1}x_t^{1-\lambda} \frac{Z_t}{1-u_t} \right) + Q_t \left(a - av\delta^\lambda x_t^{-\lambda} \frac{Z_t}{1-u_t} - vy_t^{-\lambda} \frac{u_t}{1-u_t} \right),$$

noting that the terms in Q_t on the right-hand side cancel out, so \dot{Q}_t can be written as:

$$\dot{Q}_t = vy_t^{-\lambda} \left(\frac{\lambda}{\lambda-1}y_t - Q_t \right) \frac{u_t}{1-u_t} - \frac{av\delta^\lambda x_t^{-\lambda}Z_t}{1-u_t} \left(Q_t - \frac{\lambda}{\lambda-1}x_t \right).$$

Comparison with equations (3.11), (3.12), and the definition of Z_t in (A.5.4) confirms the differential equation for Q_t in (3.13).

Steady state

Given the moving rate n and the sales rate s , the steady-state stock of houses for sale is:

$$u = \frac{n}{s+n}. \quad [\text{A.5.8}]$$

The steady-state moving rate n can be derived from the formula (3.12):

$$n = a - a\delta^\lambda x^{-\lambda}v \frac{u}{1-u} \int_{\tau=0}^{\infty} e^{-a(1-\delta^\lambda)\tau} d\tau = a - a\delta^\lambda \left(\frac{y}{x} \right)^\lambda vy^{-\lambda} \frac{n}{s} \frac{1}{a(1-\delta^\lambda)},$$

where the final equality uses $u/(1-u) = n/s$, as implied by (A.5.8). Since $s = vy^{-\lambda}$ according to (A.3.22), the equation above becomes:

$$n = a - \frac{\delta^\lambda}{1-\delta^\lambda} \left(\frac{y}{x} \right)^\lambda n.$$

Solving this equation for n yields:

$$n = \frac{a}{1 + \frac{\delta^\lambda}{1-\delta^\lambda} \left(\frac{y}{x}\right)^\lambda}, \quad [\text{A.5.9}]$$

which confirms the claim in (4.11).

A.6 Efficiency

The social planner's objective function from (4.9) can be written in terms of total match quality \mathcal{E}_t , the transaction threshold y_t , and houses for sale u_t by substituting $\mathcal{E}_t = (1 - u_t)Q_t$ and using equation (3.11):

$$\Omega_T = \int_{t=T}^{\infty} e^{-r(t-T)} \left(\xi \mathcal{E}_t - C v y_t^{-\lambda} u_t - F u_t - M \right) dt. \quad [\text{A.6.1}]$$

This is maximized by choosing x_t , y_t , \mathcal{E}_t , u_t , and Z_t subject to the differential equations for \mathcal{E}_t , u_t , and Z_t in (A.5.5), (A.5.6), and (A.5.7) (the variable Z_t defined in A.5.4 is introduced because the differential equations A.5.5 and A.5.6 are written in terms of Z_t). The problem is solved by introducing the (current-value) Hamiltonian:

$$\begin{aligned} \mathcal{J}_t = & \xi \mathcal{E}_t - C v y_t^{-\lambda} u_t - F u_t - M + \varphi_t \left(\frac{v\lambda}{\lambda-1} y_t^{1-\lambda} u_t - a \mathcal{E}_t + \frac{av\delta^\lambda \lambda}{\lambda-1} x_t^{1-\lambda} Z_t \right) \\ & + \vartheta_t \left(a(1 - u_t) - av\delta^\lambda x_t^{-\lambda} Z_t - v y_t^{-\lambda} u_t \right) + \gamma_t \left(u_t - a(1 - \delta^\lambda) Z_t \right), \quad [\text{A.6.2}] \end{aligned}$$

where φ_t , ϑ_t , and γ_t are the co-state variables associated with \mathcal{E}_t , u_t , and Z_t . The first-order conditions with respect to x_t and y_t are:

$$\frac{\partial \mathcal{J}_t}{\partial x_t} = av\delta^\lambda \lambda x_t^{-\lambda-1} Z_t \vartheta_t - av\delta^\lambda \lambda x_t^{-\lambda} Z_t \varphi_t = 0; \quad [\text{A.6.3a}]$$

$$\frac{\partial \mathcal{J}_t}{\partial y_t} = v\lambda C y_t^{-\lambda-1} u_t - v\lambda y_t^{-\lambda} u_t \varphi_t + v\lambda y_t^{-\lambda-1} u_t \vartheta_t = 0, \quad [\text{A.6.3b}]$$

and the first-order conditions with respect to the state variables \mathcal{E}_t , u_t , and Z_t are:

$$\frac{\partial \mathcal{J}_t}{\partial \mathcal{E}_t} = \xi - a\varphi_t = r\varphi_t - \dot{\varphi}_t; \quad [\text{A.6.3c}]$$

$$\frac{\partial \mathcal{J}_t}{\partial u_t} = -C v y_t^{-\lambda} - F + \frac{v\lambda}{\lambda-1} y_t^{1-\lambda} \varphi_t - (a + v y_t^{-\lambda}) \vartheta_t + \gamma_t = r\vartheta_t - \dot{\vartheta}_t; \quad [\text{A.6.3d}]$$

$$\frac{\partial \mathcal{J}_t}{\partial Z_t} = \frac{av\delta^\lambda \lambda}{\lambda-1} x_t^{1-\lambda} \varphi_t - av\delta^\lambda x_t^{-\lambda} \vartheta_t - a(1 - \delta^\lambda) \gamma_t = r\gamma_t - \dot{\gamma}_t. \quad [\text{A.6.3e}]$$

By cancelling common terms from (A.6.3a), the following link between the moving threshold x_t and the co-states φ_t and ϑ_t can be deduced:

$$\varphi_t = \frac{\vartheta_t}{x_t}. \quad [\text{A.6.4}]$$

Similarly, cancelling common terms from (A.6.3b) implies a link between the transaction threshold y_t and φ_t and ϑ_t :

$$\frac{C}{y_t} + \frac{\vartheta_t}{y_t} = \varphi_t. \quad [\text{A.6.5}]$$

The differential equation for φ_t in (A.6.3c) is:

$$\dot{\varphi}_t = (r + a)\varphi_t - \xi,$$

and since $r + a > 0$, the only solution satisfying the transversality condition is the following constant solution:

$$\varphi_t = \frac{\xi}{r + a}. \quad [\text{A.6.6}]$$

With this solution for φ_t , equation (A.6.5) implies that ϑ_t is proportional to the moving threshold x_t :

$$\vartheta_t = \frac{\xi}{r + a} x_t. \quad [\text{A.6.7}]$$

Eliminating both φ_t and ϑ_t from (A.6.5) by substituting from (A.6.6) and (A.6.7) implies that y_t and x_t must satisfy:

$$y_t - x_t = \frac{(r + a)C}{\xi}. \quad [\text{A.6.8}]$$

Using (A.6.3d) to write a differential equation for ϑ_t and substituting the solution for φ_t from (A.6.6):

$$\dot{\vartheta}_t = (r + a + v y_t^{-\lambda}) \vartheta_t - \gamma_t + F + C v y_t^{-\lambda} - \frac{\xi}{r + a} \frac{v \lambda}{\lambda - 1} y_t^{1-\lambda}. \quad [\text{A.6.9}]$$

Similarly, (A.6.3e) implies a differential equation for γ_t , from which φ_t can be eliminated using (A.6.6):

$$\dot{\gamma}_t = (r + a(1 - \delta^\lambda)) \gamma_t + a v \delta^\lambda x_t^{-\lambda} \left(\frac{\xi}{r + a} x_t \right) - \frac{a v \delta^\lambda \lambda}{\lambda - 1} x_t^{1-\lambda} \left(\frac{\xi}{r + a} \right),$$

which can be simplified as follows:

$$\dot{\gamma}_t = (r + a(1 - \delta^\lambda)) \gamma_t - \frac{\xi}{r + a} \frac{a v \delta^\lambda}{\lambda - 1} x_t^{1-\lambda}. \quad [\text{A.6.10}]$$

It is now shown that there is a solution of the constrained maximization problem where the co-states ϑ_t and γ_t are constant over time. In this case, equations (A.6.7) and (A.6.8) require that x_t and y_t are constant over time and related as follows:

$$y - x = \frac{(r + a)C}{\xi}. \quad [\text{A.6.11}]$$

With $\dot{\vartheta}_t = 0$ and $\dot{\gamma}_t = 0$, (A.6.9) and (A.6.10) imply the following pair of equations:

$$(r + a + v y^{-\lambda}) \vartheta - \gamma + F + C v y^{-\lambda} - \frac{\xi}{r + a} \frac{v \lambda}{\lambda - 1} y^{1-\lambda} = 0; \quad [\text{A.6.12}]$$

$$(r + a(1 - \delta^\lambda)) \gamma - \frac{\xi}{r + a} \frac{a v \delta^\lambda}{\lambda - 1} x^{1-\lambda} = 0. \quad [\text{A.6.13}]$$

Equation (A.6.13) yields the following expression for γ in terms of x :

$$\gamma = \frac{\xi a v \delta^\lambda}{(\lambda - 1)(r + a)(r + a(1 - \delta^\lambda))} x^{1-\lambda},$$

and substituting this and (A.6.7) into (A.6.12) leads to:

$$\frac{\xi(r + a + v y^{-\lambda})}{r + a} x - \frac{\xi a v \delta^\lambda}{(\lambda - 1)(r + a)(r + a(1 - \delta^\lambda))} x^{1-\lambda} + F + C v y^{-\lambda} - \frac{\xi v \lambda}{(\lambda - 1)(r + a)} y^{1-\lambda} = 0.$$

Since $(r + a)C = \xi(y - x)$ according to (A.6.11), multiplying the equation above by $(r + a)$ and substituting

for $(r + a)C$ implies:

$$\xi(r + a + vy^{-\lambda})x - \frac{\xi av\delta^\lambda}{(\lambda - 1)(r + a(1 - \delta^\lambda))}x^{1-\lambda} + (r + a)F + \xi v(y - x)y^{-\lambda} - \frac{\xi v\lambda}{(\lambda - 1)}y^{1-\lambda} = 0.$$

Dividing both sides by ξ and grouping terms in $(r + a)$ on the left-hand side:

$$(r + a) \left(x + \frac{F}{\xi} \right) = \frac{v\lambda}{\lambda - 1}y^{1-\lambda} - vy^{1-\lambda} + \frac{av\delta^\lambda}{(\lambda - 1)(r + a(1 - \delta^\lambda))}x^{1-\lambda},$$

and dividing both sides by $r + a$ and simplifying the terms involving $y^{1-\lambda}$ leads to:

$$x + \frac{F}{\xi} = \frac{v}{(\lambda - 1)(r + a)} \left(y^{1-\lambda} + \frac{a\delta^\lambda}{r + a(1 - \delta^\lambda)}x^{1-\lambda} \right). \quad [\text{A.6.14}]$$

The pair of equations (A.6.11) and (A.6.14) for x and y are identical to the equations (4.4) and (4.7) characterizing the equilibrium values of x and y . The equilibrium is therefore the same as the solution to the social planner's problem, establishing that it is efficient.

A.7 Transitional dynamics and overshooting

Transitional dynamics out of steady state

Equation (3.12) for the moving rate $n_t = N_t/(1 - u_t)$ implies that the quantity of new listings N_t is:

$$N_t = a(1 - u_t) - a\delta^\lambda x_t^{-\lambda} v \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau, \quad [\text{A.7.1}]$$

and equation (3.10) implies the differential equation for the stock of houses for sale u_t is:

$$\dot{u}_t = N_t - S_t, \quad \text{where } S_t = s_t u_t \text{ and } s_t = vy_t^{-\lambda}. \quad [\text{A.7.2}]$$

In the equation above, S_t is the number of transactions and s_t is the sales rate, which is taken from (3.11).

Now suppose that the moving and transactions thresholds x_t and y_t are constant from some date T onwards, that is, $x_t = x$ and $y_t = y$ for all $t \geq T$. Using (A.7.2), the number of transactions S_t and the sales rate s_t are given by:

$$S_t = su_t, \quad \text{and } s_t = s = vy^{-\lambda}. \quad [\text{A.7.3}]$$

Equation (A.7.1) for new listings becomes:

$$N_t = a(1 - u_t) - a\delta^\lambda x^{-\lambda} v \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau, \quad [\text{A.7.4}]$$

where (A.7.3) and (A.7.4) are valid for all $t \geq T$. By taking the derivative of both sides of (A.7.4) with respect to time t :

$$\dot{N}_t = -a\dot{u}_t - a\delta^\lambda x^{-\lambda} v u_t + a(1 - \delta^\lambda) \left(a\delta^\lambda x^{-\lambda} v \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_\tau d\tau \right),$$

and using (A.7.4) to substitute for the integral above an expression involving the current levels of N_t and u_t :

$$\dot{N}_t = -a\dot{u}_t - a\delta^\lambda x^{-\lambda} v u_t + a(1 - \delta^\lambda) (a(1 - u_t) - N_t).$$

This differential equation can be simplified as follows:

$$\dot{N}_t = -a\dot{u}_t - a \left((1 - \delta^\lambda)N_t + \left(a(1 - \delta^\lambda) + \delta^\lambda \left(\frac{y}{x} \right)^\lambda s \right) u_t - a(1 - \delta^\lambda) \right), \quad [\text{A.7.5}]$$

where s is the constant sales rate from (A.7.3). Substituting equation (A.7.3) into (A.7.2):

$$\dot{u}_t = N_t - su_t, \quad [\text{A.7.6a}]$$

and in turn substituting this equation into (A.7.5) and simplifying:

$$\dot{N}_t = -a \left(\left(1 + (1 - \delta^\lambda) \right) N_t + \left(a(1 - \delta^\lambda) + \delta^\lambda \left(\frac{y}{x} \right)^\lambda s - s \right) u_t - a(1 - \delta^\lambda) \right). \quad [\text{A.7.6b}]$$

Equations (A.7.6a) and (A.7.6b) comprise a system of linear differential equations for the stock of houses for sale u_t and new listings N_t .

Now consider a steady state of the system (A.7.6), that is, a solution $u_t = u$ and $N_t = N$ where $\dot{u}_t = 0$ and $\dot{N}_t = 0$ for all t . Equation (A.7.6a) implies $N = su$, and substituting this into (A.7.6b):

$$\left(\left(1 + (1 - \delta^\lambda) \right) s + \left(a(1 - \delta^\lambda) + \delta^\lambda \left(\frac{y}{x} \right)^\lambda s - s \right) \right) u = a(1 - \delta^\lambda),$$

which can be solved for a unique value of u :

$$u = \frac{\frac{a}{1 + \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y}{x} \right)^\lambda}}{s + \frac{a}{1 + \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y}{x} \right)^\lambda}} = \frac{n}{s + n}, \quad \text{where } n = \frac{a}{1 + \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y}{x} \right)^\lambda}. \quad [\text{A.7.7}]$$

This is of course the steady state found in section 4.3, where s is the steady-state sales rate from (3.11) and n is the steady-state moving rate (3.12). Steady-state new listings are $N = su = n(1 - u)$.

Now define the percentage deviations of the variables u_t and N_t from their unique steady-state values u and N :

$$\tilde{u}_t = \frac{u_t - u}{u}, \quad \text{and } \tilde{N}_t = \frac{N_t - N}{N}, \quad \text{or equivalently } u_t = u(1 + \tilde{u}_t), \quad \text{and } N_t = N(1 + \tilde{N}_t), \quad [\text{A.7.8}]$$

and the time derivatives of u_t and N_t and \tilde{u}_t and \tilde{N}_t are related as follows:

$$\dot{\tilde{u}}_t = \frac{\dot{u}_t}{u}, \quad \text{and } \dot{\tilde{N}}_t = \frac{\dot{N}_t}{N}. \quad [\text{A.7.9}]$$

Using (A.7.8), (A.7.9), and $N = su$, the differential equation (A.7.6a) can be written in terms of \tilde{u}_t and \tilde{N}_t as follows:

$$\dot{\tilde{u}}_t = s\tilde{N}_t - s\tilde{u}_t. \quad [\text{A.7.10a}]$$

Likewise, (A.7.8), (A.7.9), and $u/N = 1/s$ imply that the differential equation (A.7.6b) is equivalent to:

$$\dot{\tilde{N}}_t = -a \left(\left(1 + (1 - \delta^\lambda) \right) \tilde{N}_t + \left(\frac{a(1 - \delta^\lambda)}{s} + \delta^\lambda \left(\frac{y}{x} \right)^\lambda - 1 \right) \tilde{u}_t \right). \quad [\text{A.7.10b}]$$

Make the following definition of a variable \tilde{n}_t :

$$\tilde{n}_t = \tilde{N}_t + \frac{n}{s}\tilde{u}_t, \quad \text{and hence } \tilde{N}_t = \tilde{n}_t - \frac{n}{s}\tilde{u}_t \quad \text{and } \dot{\tilde{n}}_t = \dot{\tilde{N}}_t + \frac{n}{s}\dot{\tilde{u}}_t, \quad [\text{A.7.11}]$$

where the final equation follows from taking the time derivative of the definition of \tilde{n}_t . The variable \tilde{n}_t is also approximately the percentage deviation of the moving rate $n_t = N_t/(1 - u_t)$ from its steady-state

value n , but here, \tilde{n}_t is simply taken as the definition of a new variable. The differential equation (A.7.10a) can be written exactly in terms of \tilde{u}_t and \tilde{n}_t by using the second equation in (A.7.11):

$$\dot{\tilde{u}}_t = s\tilde{n}_t - (s+n)\tilde{u}_t. \quad [\text{A.7.12a}]$$

Substituting (A.7.10b) and (A.7.12a) into the third equation from (A.7.11):

$$\begin{aligned} \dot{\tilde{n}}_t &= -a \left(1 + (1 - \delta^\lambda) \right) \left(\tilde{n}_t - \frac{n}{s}\tilde{u}_t \right) - a \left(\frac{a(1 - \delta^\lambda)}{s} + \delta^\lambda \left(\frac{y}{x} \right)^\lambda - 1 \right) \tilde{u}_t + \frac{n}{s} (s\tilde{n}_t - (s+n)\tilde{u}_t) \\ &= - \left((a-n) + a(1 - \delta^\lambda) \right) \tilde{n}_t - \left(\frac{(a-n)}{s} a(1 - \delta^\lambda) + a\delta^\lambda \left(\frac{y}{x} \right)^\lambda - (a-n) - (a-n)\frac{n}{s} \right) \tilde{u}_t. \end{aligned}$$

Rearranging the equation for the steady-state moving rate n in (A.7.7) leads to:

$$a\delta^\lambda \left(\frac{y}{x} \right)^\lambda = \frac{(a-n)}{n} a(1 - \delta^\lambda),$$

and substituting this into the equation for $\dot{\tilde{n}}_t$ above implies that the coefficient of \tilde{u}_t can be simplified:

$$\dot{\tilde{n}}_t = - \left((a-n) + a(1 - \delta^\lambda) \right) \tilde{n}_t - \frac{(a-n)(s+n)}{s} \left(\frac{a(1 - \delta^\lambda)}{n} - 1 \right) \tilde{u}_t. \quad [\text{A.7.12b}]$$

Equations (A.7.12a) and (A.7.12b) form a system of differential equations in the variables \tilde{u}_t and \tilde{n}_t . This system can be written in matrix form as follows:

$$\begin{pmatrix} \dot{\tilde{u}}_t \\ \dot{\tilde{n}}_t \end{pmatrix} = \begin{pmatrix} -(s+n) & s \\ -\frac{n(s+n)\chi_u}{s} & -n\chi_n \end{pmatrix} \begin{pmatrix} \tilde{u}_t \\ \tilde{n}_t \end{pmatrix}, \quad [\text{A.7.13}]$$

where the coefficients χ_u and χ_n are defined by:

$$\chi_u = \frac{a-n}{n} \left(\frac{a(1 - \delta^\lambda)}{n} - 1 \right), \quad \text{and} \quad \chi_n = \frac{a-n}{n} + \frac{a(1 - \delta^\lambda)}{n}. \quad [\text{A.7.14}]$$

The sum of the coefficients χ_u and χ_n is:

$$\chi_u + \chi_n = (1 - \delta^\lambda) \left(\frac{a}{n} \right)^2. \quad [\text{A.7.15}]$$

The expression for n in (A.7.7) implies:

$$\frac{a-n}{n} = \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y}{x} \right)^\lambda, \quad \text{and} \quad \frac{a(1 - \delta^\lambda)}{n} = (1 - \delta^\lambda) + \delta^\lambda \left(\frac{y}{x} \right)^\lambda,$$

and hence the coefficients χ_u and χ_n from (A.7.14) are equal to the following:

$$\chi_u = \frac{\delta^\lambda \left(\frac{\delta y}{x} \right)^\lambda \left(\left(\frac{y}{x} \right)^\lambda - 1 \right)}{1 - \delta^\lambda}, \quad \text{and} \quad \chi_n = \frac{(1 - \delta^\lambda)^2 + (1 - \delta^\lambda) \left(\frac{\delta y}{x} \right)^\lambda + \left(\frac{\delta y}{x} \right)^\lambda}{1 - \delta^\lambda}. \quad [\text{A.7.16}]$$

The coefficient χ_n is strictly positive, as is χ_u because it is always the case that $y > x$. Since $(\delta y/x)^\lambda < 1$ and $(\delta y/x)^\lambda - \delta^\lambda < 1 - \delta^\lambda$, it follows that $\chi_u < 1$. As χ_n is larger than $1 - \delta^\lambda + \delta^\lambda(y/x)^\lambda$, which is greater than 1 because $y > x$, it must be the case that $\chi_n > 1$.

The set of points where $\dot{\tilde{u}}_t = 0$ is given by:

$$-(s+n)\tilde{u}_t + s\tilde{n}_t = 0, \quad \text{and hence} \quad \tilde{n}_t = \frac{s+n}{s}\tilde{u}_t,$$

which is an upward-sloping straight line with gradient $(s+n)/s$ in $(\tilde{u}_t, \tilde{n}_t)$ space. To the left and above,

\tilde{u}_t is increasing over time, and to the right and below, \tilde{n}_t is decreasing. The set of points where $\dot{\tilde{n}}_t = 0$ is given by:

$$-\frac{n(s+n)\chi_u}{s}\tilde{u}_t - n\chi_n\tilde{n}_t = 0, \quad \text{and hence } \tilde{n}_t = -\frac{(s+n)\chi_u}{s\chi_n}.$$

This is a downward-sloping straight line with gradient $-(s+n)\chi_u/s\chi_n$, which is less than the gradient of the $\dot{\tilde{u}}_t = 0$ in absolute value because $\chi_u < \chi_n$. Given that both χ_u and χ_n are positive, \tilde{n}_t is increasing over time to the left and below the line, and decreasing to the right and above.

The characteristic equation for the eigenvalues ζ of the system of differential equations (A.7.13) is:

$$(\zeta + (s+n))(\zeta + n\chi_n) + n(s+n)\chi_u = 0, \quad [\text{A.7.17}]$$

which is a quadratic equation in ζ :

$$\zeta^2 + ((s+n) + n\chi_n)\zeta + n(s+n)(\chi_u + \chi_n) = 0. \quad [\text{A.7.18}]$$

The two eigenvalues ζ_1 and ζ_2 are the roots of this quadratic equation. The eigenvalues are either both real numbers or a conjugate pair of complex numbers. The sum and product of the eigenvalues are:

$$\zeta_1 + \zeta_2 = -((s+n) + n\chi_n), \quad \text{and } \zeta_1\zeta_2 = n(s+n)(\chi_u + \chi_n).$$

Since χ_u and χ_n are both positive, the sum of the eigenvalues is negative and the product is positive. If both are real numbers then both must be negative numbers. If both are complex numbers then the sum is equal to twice the common real component of the eigenvalues, which must therefore be negative. Hence, in all cases, the real parts of all eigenvalues are negative. This establishes that there is convergence to the steady state in the long run starting from any initial conditions.

The condition for the quadratic equation (A.7.18) to have two real roots is:

$$((s+n) + n\chi_n)^2 \geq 4n(s+n)(\chi_u + \chi_n), \quad \text{or equivalently } \left(\frac{s}{n} + 1 + \chi_n\right)^2 \geq 4(\chi_u + \chi_n) \left(\frac{s}{n} + 1\right),$$

where the latter is derived by dividing both sides by the positive number n^2 . By expanding the brackets, this condition can be expressed as a quadratic inequality in the ratio of the sales rate to the moving rate:

$$\left(\frac{s}{n}\right)^2 + 2(1 + \chi_n)\frac{s}{n} + (1 + \chi_n)^2 \geq 4(\chi_u + \chi_n)\frac{s}{n} + 4(\chi_u + \chi_n),$$

which can be simplified as follows:

$$\left(\frac{s}{n}\right)^2 + 2(1 - (\chi_u + \chi_n) - \chi_u)\frac{s}{n} + (1 + \chi_n^2 - 2(\chi_u + \chi_n) - 2\chi_u) \geq 0.$$

It can be verified directly that the inequality can be factorized:

$$\left(\frac{s}{n} + 1 - (\sqrt{\chi_u + \chi_n} - \sqrt{\chi_u})\right) \left(\frac{s}{n} + 1 - (\sqrt{\chi_u + \chi_n} + \sqrt{\chi_u})\right) \geq 0, \quad [\text{A.7.19}]$$

which provides a test in terms of the ratio s/n for whether the eigenvalues of the system of differential equations (A.7.13) are all real (if the test is not satisfied then they are a conjugate pair of complex numbers).

There are three cases. The first case is where the sales rate relative to the moving rate is above the following threshold:

$$\frac{s}{n} > (\sqrt{\chi_u + \chi_n} + \sqrt{\chi_u})^2 - 1, \quad [\text{A.7.20}]$$

noting that the right-hand side is strictly positive because $\chi_n > 1$. If this holds then the condition (A.7.19) is satisfied and both eigenvalues ζ_1 and ζ_2 are real numbers. Since $\chi_u > 0$, it follows that $\sqrt{\chi_u + \chi_n} > \sqrt{\chi_n}$ and hence (A.7.20) implies $s/n > \chi_n - 1$. This means $s+n > n\phi_n$, and as ζ_1 and ζ_2 are roots of the equation

(A.7.17), it must be the case that the negative eigenvalues satisfy $\zeta_1 > -(s+n)$ and $\zeta_2 > -(s+n)$, recalling that ϕ_u is always positive.

In a model where the moving rate n is exogenous and constant, the only dynamics would come from the differential equation $\dot{\tilde{u}}_t = -(s+n)\tilde{u}_t$ (see A.7.12a), so $s+n$ would be the rate of convergence to the steady state. In the general model, the speed of convergence is determined by the negative of the real components of the eigenvalues. When (A.7.20) is satisfied, it follows that $|\zeta_1| < s+n$ and $|\zeta_2| < s+n$, which means the new dynamics coming from match quality dominate the usual dynamics coming from the evolution of the stock of houses for sale. Convergence is therefore slower than it would be in an exogenous moving model with the same moving rate. As (A.7.20) shows, this case corresponds to the sales rate being sufficiently large, which will make it the empirically relevant one. Using $\sqrt{\chi_u + \chi_n} + \sqrt{\chi_u} < 2\sqrt{\chi_u + \chi_n}$ and the expression for $\chi_u + \chi_n$ in (A.7.15), a sufficient condition for (A.7.20) is:

$$\frac{s}{n} > (1 - \delta^\lambda) \left(\frac{a}{n}\right)^2 - 1,$$

which holds for the calibrated version of the model.

The second case is where the sales rate is relatively low, specifically $s/n < (\sqrt{\chi_u + \chi_n} - \sqrt{\chi_u})^2 - 1$. The condition in (A.7.19) will be satisfied, so both eigenvalues would be negative real numbers. In this case, $s/n < \chi_n - 1$ since $\sqrt{\chi_u} + \sqrt{\chi_n} > \sqrt{\chi_u + \chi_n}$, which means that $s+n < n\phi_n$. Given that the eigenvalues are roots of equation (A.7.17), it follows that $\zeta_1 < -(s+n)$ and $\zeta_2 < -(s+n)$ and hence $|\zeta_1| > s+n$ and $|\zeta_2| > s+n$. Convergence to the steady state is actually faster than an exogenous moving model, and the new dynamics of match quality do not play an important role compared to the usual dynamics coming from the evolution of the stock of houses for sale. This case is not empirically relevant because the required sales rate would need to be too low compared to the moving rate.

The third case is where the dynamics of match quality and the dynamics of the stock of houses for sale are of similar importance, which occurs when s/n lies between $(\sqrt{\chi_u + \chi_n} - \sqrt{\chi_u})^2 - 1$ and $(\sqrt{\chi_u + \chi_n} + \sqrt{\chi_u})^2 - 1$. In this case, condition (A.7.19) does not hold and both eigenvalues ζ_1 and ζ_2 are complex numbers. This case features damped oscillations around the steady state, but is not empirically relevant because the required sales rate is too low compared to the moving rate. In what follows, attention is restricted to the empirically plausible case where the sales rate is sufficiently high that condition (A.7.20) holds.

There are two eigenvectors of the form $(1, \nu)$ associated with the two eigenvalues (as will be seen, the first element can be normalized to 1). The values of ν in the eigenvectors are solutions of the following equations:

$$(-(s+n) - \zeta) + s\nu = 1, \quad \text{and hence } \nu = \frac{(s+n) + \zeta}{s}. \quad [\text{A.7.21}]$$

This equation holds for ν_1 when $\zeta = \zeta_1$ and for ν_2 when $\zeta = \zeta_2$. Without loss of generality, let ζ_1 denote the eigenvalue with the greater absolute value. Consequently, as both are negative, $\zeta_1 < \zeta_2$. Since it has been shown above that $\zeta_1 > -(s+n)$ and $\zeta_2 > -(s+n)$, equation (A.7.21) implies that both ν_1 and ν_2 are positive, with $\nu_1 < \nu_2$. Geometrically in $(\tilde{u}_t, \tilde{n}_t)$ space, both eigenvectors are upward-sloping straight lines, and since both ζ_1 and ζ_2 are negative, their gradient is less than the $\dot{\tilde{u}}_t = 0$ line. The eigenvector associated with the dominant eigenvalue ζ_2 has a steeper gradient than the eigenvector associated with ζ_1 .

Having found the eigenvalues and eigenvectors of the system of differential equations (A.7.13), the solution can be stated as follows:

$$\tilde{u}_t = k_1 e^{\zeta_1(t-T)} + k_2 e^{\zeta_2(t-T)}, \quad \text{and } \tilde{n}_t = k_1 \nu_1 e^{\zeta_1(t-T)} + k_2 \nu_2 e^{\zeta_2(t-T)}, \quad [\text{A.7.22}]$$

where T is the date from which the moving and transaction thresholds will be constant at x and y respectively, and k_1 and k_2 are coefficients to be determined. Since there is convergence to the steady state for any initial conditions, the coefficients k_1 and k_2 are pinned down by knowing the values of \tilde{u}_T and \tilde{n}_T :

$$\tilde{u}_T = k_1 + k_2, \quad \text{and } \tilde{n}_T = k_1 \nu_1 + k_2 \nu_2,$$

and these equations can be solved for k_1 and k_2 :

$$k_1 = \frac{\nu_2 \tilde{u}_T - \tilde{n}_T}{\nu_2 - \nu_1}, \quad \text{and} \quad k_2 = \frac{\tilde{n}_T - \nu_1 \tilde{u}_T}{\nu_2 - \nu_1}.$$

Substituting these expressions into (A.7.22) yields the solution conditional on given initial conditions at date T :

$$\tilde{u}_t = (\nu_2 \tilde{u}_T - \tilde{n}_T) \left(\frac{1}{\nu_2 - \nu_1} \right) e^{\zeta_1(t-T)} + (\tilde{n}_T - \nu_1 \tilde{u}_T) \left(\frac{1}{\nu_2 - \nu_1} \right) e^{\zeta_2 t - T}; \quad [\text{A.7.23a}]$$

$$\tilde{n}_t = (\nu_2 \tilde{u}_T - \tilde{n}_T) \left(\frac{\nu_1}{\nu_2 - \nu_1} \right) e^{\zeta_1(t-T)} + (\tilde{n}_T - \nu_1 \tilde{u}_T) \left(\frac{\nu_2}{\nu_2 - \nu_1} \right) e^{\zeta_2(t-T)}. \quad [\text{A.7.23b}]$$

As T tends to infinity, the vector $(\tilde{u}_t, \tilde{n}_t)$ must approach the origin approximately along the eigenvector associated with the dominant eigenvalue ζ_2 , that is, $(1, \nu_2)$.

Given the initial values of \tilde{u}_T and \tilde{n}_t , (A.7.23) gives an exact solution of the system of differential equations (A.7.13) for variables \tilde{u}_t and \tilde{n}_t defined in (A.7.8) and (A.7.11). Using the definition in (A.7.11), this implies an exact solution for \tilde{N}_t as well. The exact solution for the variables u_t and N_t can then be recovered using the definitions in (A.7.8). Once u_t and N_t are known, the exact solution for the moving rate can be computed using the definition $n_t = N_t/(1 - u_t)$. Finally, the sales rate s_t is simply equal to the constant s , and transactions $S_t = s_t u_t$ can be found given the solution for u_t .

Overshooting

Next, consider how the initial values of \tilde{u}_T and \tilde{n}_T are determined following a change to the moving or transaction thresholds x_t and y_t . Suppose that x_t and y_t were previously constant at x_0 and y_0 , and then move permanently to x and y from date T onwards. The previous sales rate was s_0 , and the steady-state values of the moving rate and houses for sale were n_0 and u_0 . After the change to x and y , there is a new steady-state sales rate s that is reached immediately at date T . Houses for sale is a stock that cannot instantaneously jump, so this variable remains equal to its old steady-state value initially, that is, $u_T = u_0$. The new steady state for u_t is u , and this can be used to compute $\tilde{u}_T = (u_0 - u)/u$.

Using (4.11) and (4.12), an increase in the moving threshold x implies a higher moving rate n and a higher u compared to u_0 , and thus a negative value of \tilde{u}_T . Using (4.10), (4.11), and (4.12), an increase in y implies a lower ratio s/n , which means a higher value of u compared to u_0 , and thus a negative value of \tilde{u}_T . Therefore, either an increase in x or an increase in y implies $\tilde{u}_T < 0$.

Now consider the moving rate at date T when the moving and transactions thresholds change. With houses for sale u_t at its old steady-state value of u_0 for $t < T$, equation (3.12) implies:

$$n_T = a - \frac{a\delta^\lambda x_T^{-\lambda} v_0}{1 - u_T} \int_{\tau \rightarrow -\infty}^t e^{-a(1-\delta^\lambda)(t-\tau)} u_0 d\tau = a - \frac{a\delta^\lambda x^{-\lambda} v_0}{1 - u_0} \frac{u_0}{a(1-\delta^\lambda)} = a - \frac{\delta^\lambda}{1 - \delta^\lambda} \frac{u_0}{1 - u_0} v_0 x^{-\lambda},$$

which uses $x_T = x$ and $u_T = u_0$, and where v_0 denotes the old value of the parameter v (which may change at date T). Noting that $u_0 = n_0/(s_0 + n_0)$, where s_0 and n_0 are the old sales and moving rates, it can be seen that $u_0/(1 - u_0) = n_0/s_0$, where $s_0 = v_0 y_0^{-\lambda}$. Substituting this into the equation above:

$$n_T = a - \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y_0}{x} \right)^\lambda n_0, \quad [\text{A.7.24}]$$

and substituting the expression for the old steady state n_0 from (4.11):

$$n_T = a - \frac{a \left(\frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y_0}{x} \right)^\lambda \right)}{1 + \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y_0}{x_0} \right)^\lambda} = \left(1 + \frac{\delta^\lambda}{1 - \delta^\lambda} y_0^\lambda \left(\frac{1}{x_0^\lambda} - \frac{1}{x^\lambda} \right) \right) n_0. \quad [\text{A.7.25}]$$

This implies that an increase in y has no impact on the initial value of n_T , while an increase in x raises n_T

above n_0 .

Next, the value of n_T is compared to the new steady-state value n . When y increases, equation (4.11) implies the steady-state value of n is lower. Therefore, $n < n_T = n_0$ following an increase in y , and thus $\tilde{n}_T > 0$. Now suppose x increases with no change in y . As explained above, (A.7.25) implies $n_T > n_0$. Combined with (A.7.24):

$$n_T > a - \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y_0}{x}\right)^\lambda n_T, \quad \text{and therefore } n_T > \frac{a}{1 + \frac{\delta^\lambda}{1 - \delta^\lambda} \left(\frac{y_0}{x}\right)^\lambda}.$$

After the increase in x , the right-hand side is equal to the new steady-state moving rate n , so the inequality above implies $n_T > n$. This means the moving rate overshoots its new steady state in the short run. Therefore, following either an increase in x or an increase in y , the initial deviation of the moving rate from its new steady state is such that $\tilde{n}_T > 0$.

In summary, an increase in either x or y leads to $\tilde{u}_T < 0$ and $\tilde{n}_T > 0$, and a decrease in x or y leads to $\tilde{u}_T > 0$ and $\tilde{n}_T < 0$. The transitional dynamics therefore follow the example paths illustrated in Figure 7.

A.8 Model with heterogeneous distributions of idiosyncratic shocks

The search process is the same as in the basic model. u_t denotes the measure of houses for sale, b_t denotes the measure of buyers, and $\mathcal{V}(u_t, b_t)$ denotes the viewing function. The viewing rate for both buyers and sellers is $v = \mathcal{V}(u_t, b_t)/u_t$ given that $u_t = b_t$ in equilibrium. Following a viewing, the buyer draws a match-specific quality ϵ from the Pareto distribution $G(\epsilon)$ in (3.1). If the match quality is sufficiently high then a transaction takes place with the price determined by Nash bargaining.

Value functions, thresholds, and prices

The new aspect of the model with heterogeneity emerges after a transaction occurs. After the buyer has moved in, a type $i \in \{1, \dots, q\}$ is drawn from a distribution with probabilities θ_i , where $\sum_{i=1}^q \theta_i = 1$. The number of types is $q \geq 1$, and the special case $q = 1$ corresponds to the basic version of the model. A type- i homeowner faces idiosyncratic shocks that scale down match quality ϵ by a factor δ_i , with these shocks arriving at rate a_i . The value of occupying a house with match quality ϵ for a type- i homeowner is $H_{i,t}(\epsilon)$. The Bellman equations (the equivalent of 3.8) are:

$$rH_{i,t}(\epsilon) = \epsilon\xi - M + a_i (\max\{H_{i,t}(\delta_i\epsilon), W_t\} - H_{i,t}(\epsilon)) + \dot{H}_{i,t}(\epsilon), \quad \text{for all } i \in \{1, \dots, q\}. \quad [\text{A.8.1}]$$

The moving threshold $x_{i,t}$ for type- i households is the solution of the equation:

$$H_{i,t}(x_{i,t}) = W_t = B_t + U_t, \quad [\text{A.8.2}]$$

for each $i = 1, \dots, q$, which is the equivalent of (3.9), but now each threshold $x_{i,t}$ depends on a type-specific value function $H_{i,t}(\epsilon)$. The value W_t is the sum of values from being a buyer B_t and a seller V_t as in the basic version of the model. The Bellman equations for B_t and V_t are the same as the basic version of the model (the equations in 3.3). The surpluses $\Sigma_{b,t}(\epsilon)$ and $\Sigma_{u,t}(\epsilon)$ are also as given in (3.2), but in the model with heterogeneity, the value function $H_t(\epsilon)$ is a weighted average of the type-specific value functions $H_{i,t}(\epsilon)$:

$$H_t(\epsilon) = \sum_{i=1}^q \theta_i H_{i,t}(\epsilon). \quad [\text{A.8.3}]$$

The total surplus $\Sigma_t(\epsilon)$ is also as given before in (3.4), with $H_t(\epsilon)$ specified by (A.8.3) here. Using (3.4), the equation $\Sigma_t(y_t) = 0$ for the transaction threshold is the same as the basic version of the model, namely equation (3.5), which depends on the $H_t(\epsilon)$ given in (A.8.3). The Bellman equation (3.6) for W_t is also unchanged. Assuming Nash bargaining over transaction prices, the same method in appendix A.4 can be used to show the expression in (3.7) for the average price is unchanged.

Stocks and flows

The stock-flow accounting identity (3.10) for houses for sales u_t is the same as in the basic version of the model. The process by which transactions occur is also the same as in the basic model, so equation (3.11) for the sales rate s_t holds as before. Let $\sigma_{i,t}$ denote the measure of homeowners of type- i at time t . The stock-flow accounting identity for $\omega_{i,t}$ and the link with houses for sale u_t are:

$$\dot{\sigma}_{i,t} = \theta_i S_t - N_{i,t}, \quad \text{and} \quad \sum_{i=1}^q \sigma_{i,t} = 1 - u_t, \quad [\text{A.8.4}]$$

where $N_{i,t}$ denotes the number of type- i homeowners who put their houses up for sale at date t . The same steps used in deriving equation (A.5.1) in appendix A.5 can be applied to show that the $N_{i,t}$ are given by:

$$N_{i,t} = a_i \sigma_{i,t} - \theta_i a_i \delta_i^\lambda x_{i,t}^{-\lambda} v \int_{\tau \rightarrow -\infty}^t e^{-a_i(1-\delta_i^\lambda)(t-\tau)} u_\tau d\tau, \quad [\text{A.8.5}]$$

which holds for $i = 1, \dots, q$. The aggregate number of houses newly put up for sale is $N_t = \sum_{i=1}^q N_{i,t}$, and by using equation (A.8.5), the moving rate $n_t = N_t/(1 - u_t)$ is:

$$n_t = \frac{\sum_{i=1}^q a_i \sigma_{i,t}}{1 - u_t} - \frac{v \sum_{i=1}^q \theta_i a_i \delta_i^\lambda x_{i,t}^{-\lambda} \int_{\tau \rightarrow -\infty}^t e^{-a_i(1-\delta_i^\lambda)(t-\tau)} u_\tau d\tau}{1 - u_t}. \quad [\text{A.8.6}]$$

Equilibrium

As in section 4, if the parameters of the model are constant over time, the equilibrium of the model has constant moving and transaction thresholds, so time subscripts are dropped in what follows.

Using the definition of the moving threshold x_i in (A.8.2) and evaluating the type- i homeowner's value function $H_i(\epsilon)$ at $\epsilon = x_i$:

$$(r + a_i)H_i(x_i) = \xi x_i - M + a_i W, \quad [\text{A.8.7}]$$

which holds for all $i = 1, \dots, q$. Together with (A.8.2) this implies:

$$x_i = \frac{rW + M}{\xi} = x, \quad [\text{A.8.8}]$$

which means that all types of homeowners share a common moving threshold $x_i = x$ in equilibrium. The value W can be written in terms of parameters and the common threshold x :

$$W = \frac{\xi x - M}{r}. \quad [\text{A.8.9}]$$

As in the basic model, idiosyncratic shocks are assumed large enough so that all marginal homebuyers would move upon receiving a shock. This must be true for all types of homeowner, that is, $\delta_i y < x$. It then follows by evaluating the type- i homeowner's value function (A.8.1) at $\epsilon = y$ that:

$$(r + a_i)H_i(y) = \xi y - M + a_i W.$$

Substituting W from (A.8.9) and rearranging yields:

$$(r + a_i)H_i(y) = \xi y - M + a_i \frac{\xi x - M}{r} = \xi(y - x) + \left(\frac{a_i + r}{r} \right) (\xi x - M),$$

and thus the type- i homeowner value function evaluated at the transaction threshold is:

$$H_i(y) = \frac{\xi(y - x)}{r + a_i} + \frac{\xi x - M}{r}.$$

Averaging across all types of homeowners:

$$H(y) = \xi(y - x) \sum_{i=1}^q \frac{\theta_i}{r + a_i} + \frac{\xi x - M}{r}. \quad [\text{A.8.10}]$$

An equation linking the moving and transactions thresholds x and y can be derived from (3.5), (A.8.9), and (A.8.10):

$$\xi(y - x) \sum_{i=1}^q \frac{\theta_i}{r + a_i} + \frac{\xi x - M}{r} = \frac{\xi x - M}{r} + C,$$

which can be rearranged as follows:

$$y - x = \frac{\left(\frac{1}{\sum_{i=1}^q \frac{\theta_i}{r + a_i}} \right) C}{\xi}. \quad [\text{A.8.11}]$$

This reduces to the equilibrium condition (4.4) of the basic model if there is only one type of homeowner ($q = 1$). Another equilibrium condition linking x and y can be obtained by combining equation (3.5) for the transaction threshold with the Bellman equation (3.6) for the value W :

$$rW = -F - M + vy^{-\lambda} \int_{\epsilon=y}^{\infty} \frac{\lambda}{y} \left(\frac{\epsilon}{y} \right)^{-(\lambda+1)} (H(\epsilon) - H(y)) d\epsilon, \quad [\text{A.8.12}]$$

which uses the Pareto distribution of new match quality ϵ from (3.1). Given the expression for $H(\epsilon)$ in (A.8.3), this equation can be rewritten as follows:

$$rW = -F - M + vy^{-\lambda} \sum_{i=1}^q \theta_i \Psi_i(y), \quad \text{where } \Psi_i(z) = \int_{\epsilon=z}^{\infty} \frac{\lambda}{z} \left(\frac{\epsilon}{z} \right)^{-(\lambda+1)} (H_i(\epsilon) - H_i(z)) d\epsilon. \quad [\text{A.8.13}]$$

Expressions for the functions $\Psi_i(z)$ defined above when evaluated at $z = y$ can be found explicitly using the same steps used to derive equation (A.3.15) in appendix A.3:

$$\Psi_i(y) = \frac{\xi}{(\lambda - 1)(r + a_i)} \left(y + \frac{a_i \delta_i^\lambda y^\lambda x^{1-\lambda}}{r + a_i (1 - \delta_i^\lambda)} \right). \quad [\text{A.8.14}]$$

By substituting (A.8.9) and (A.8.14) into the Bellman equation from (A.8.13):

$$\xi x - M = -F - M + vy^{-\lambda} \sum_{i=1}^q \frac{\xi \theta_i}{(\lambda - 1)(r + a_i)} \left(y + \frac{a_i \delta_i^\lambda y^\lambda x^{1-\lambda}}{r + a_i (1 - \delta_i^\lambda)} \right),$$

which simplifies to:

$$x = \frac{v \sum_{i=1}^q \frac{\theta_i}{r + a_i} \left(y^{1-\lambda} + \frac{a_i \delta_i^\lambda}{r + a_i (1 - \delta_i^\lambda)} x^{1-\lambda} \right)}{\lambda - 1} - \frac{F}{\xi}. \quad [\text{A.8.15}]$$

This is the second equilibrium condition involving x and y . It reduces to the equivalent condition (4.7) from the basic model when $q = 1$.

The equilibrium thresholds x and y are solutions of the equations (A.8.11) and (A.8.15). By substituting for x in (A.8.15) using (A.8.11), any equilibrium value of y is a solution of the equation $\mathcal{I}(y) = 0$, where the

function $\mathcal{I}(y)$ is given below:

$$\mathcal{I}(y) = \frac{v}{\lambda - 1} \sum_{i=1}^q \frac{\theta_i}{r + a_i} \left(y^{1-\lambda} + \frac{a_i \delta_i^\lambda}{r + a_i (1 - \delta_i^\lambda)} \left(y - \left(\frac{1}{\sum_{j=1}^q \frac{\theta_j}{r + a_j}} \right) \frac{C}{\xi} \right)^{1-\lambda} \right) - y + \left(\frac{1}{\sum_{i=1}^q \frac{\theta_i}{r + a_i}} \right) \frac{C}{\xi} - \frac{F}{\xi},$$

which is the equivalent of (A.3.17) when $q = 1$. This function is such that $\mathcal{I}'(y) < 0$ and $\lim_{y \rightarrow \infty} \mathcal{I}(y) = -\infty$ because $\lambda > 1$. Since $\mathcal{I}(y)$ is strictly decreasing in y , any solution (if it exists) must be unique. A solution must satisfy $x > 0$, $y > 1$, and $\delta_i y < x$ for all $i = 1, \dots, q$. Generalizing the argument used to derive (A.3.18), the inequalities involving y are equivalent to:

$$y > \max \left\{ 1, \frac{C}{(1 - \max_{i \in \{1, \dots, q\}} \delta_i) \xi} \left(\frac{1}{\sum_{j=1}^q \frac{\theta_j}{r + a_j}} \right) \right\}.$$

Thus, there exists a solution of the equation $\mathcal{I}(y) = 0$ (which is unique) if and only if the function $\mathcal{I}(y)$ is positive when evaluated at the right-hand side of the inequality above. The same argument following (A.3.19) can be used to show that the inequality for y implies $x > 0$, so all the requirements for an equilibrium are satisfied.

Average transaction price

Using the Pareto distribution from (3.1) and the definitions of $C = C_b + C_v$ and $\kappa = C_v/C$, the average transaction price in (3.7) can be written as:

$$P = \omega \int_{\epsilon=y}^{\infty} \frac{\lambda}{y} \left(\frac{\epsilon}{y} \right)^{-(\lambda+1)} (H(\epsilon) - H(y)) d\epsilon + \omega H(y) + (\kappa - \omega)C + \frac{\omega F - (1 - \omega)M}{r}.$$

Using the definition of the transactions threshold from (3.5) and equations (A.8.9) and (A.8.12):

$$P = \omega \frac{(\xi x - M) + F + M}{vy^{-\lambda}} + \omega \left(C + \frac{\xi x - M}{r} \right) + (\kappa - \omega)C + \frac{\omega F - (1 - \omega)M}{r},$$

which can be simplified to:

$$P = \kappa C - \frac{M}{r} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v} \right) (\xi x + F). \quad [\text{A.8.16}]$$

Given a moving threshold x , this is the same equation for P as (4.13) in the basic model.

Average moving rate

Consider a steady state where $u_t = u$, $\sigma_{i,t} = \sigma_i$, and $N_{i,t} = N_i$ for all t . Using the result $x_{i,t} = x_i = x$ from (A.8.8), equation (A.8.5) for the number of houses put up for sale by type- i homeowners becomes:

$$N_i = a_i \sigma_i - \theta_i a_i \delta_i^\lambda x^{-\lambda} v \left(\frac{u}{a_i (1 - \delta_i^\lambda)} \right).$$

In steady state, equation (A.8.4) implies $N_i = \theta_i S = \theta_i s u$, where $S = s u$ is the number of transactions. Substituting into the equation above and dividing both sides by a_i :

$$s u \frac{\theta_i}{a_i} = \sigma_i - x^{-\lambda} v u \theta_i \frac{\delta_i^\lambda}{a_i (1 - \delta_i^\lambda)}.$$

Sum over all $i = 1, \dots, q$ and make use of the link between σ_i and u from (A.8.4):

$$su \sum_{i=1}^q \frac{\theta_i}{a_i} = 1 - u - x^{-\lambda} vu \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{a_i(1 - \delta_i^\lambda)}.$$

Dividing both sides by $1 - u$ and noting that (4.12) implies $u/(1 - u) = n/s$:

$$s \frac{n}{s} \sum_{i=1}^q \frac{\theta_i}{a_i} = 1 - \frac{nx^{-\lambda}v}{s} \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{a_i(1 - \delta_i^\lambda)},$$

and substituting the expression for the sales rate s from (4.10):

$$n \sum_{i=1}^q \frac{\theta_i}{a_i} = 1 - n \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{a_i(1 - \delta_i^\lambda)}.$$

This can be rearranged to give a formula for the steady-state moving rate n :

$$n = \frac{1}{\sum_{i=1}^q \frac{\theta_i}{a_i} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{a_i(1 - \delta_i^\lambda)}}. \quad [\text{A.8.17}]$$

The basic model is a special case of this when $q = 1$.

A.9 The hazard function and the elasticity of the moving rate

The analysis here considers the general model with heterogeneity in the distributions of idiosyncratic shocks across q types of homeowners. The results are applicable to the basic model by considering the special case $q = 1$.

The distribution of time spent in a house

Consider an equilibrium where parameters are expected to remain constant. In this case, the moving and transaction thresholds x and y are constant over time. Let $\psi_i(T)$ denote the survival function for new matches of type- i homeowners, in the sense of the fraction of matches forming at time t that survive until at least $t + T$. Each cohort starts with a match quality distribution $\epsilon \sim \text{Pareto}(y; \lambda)$ at $T = 0$. Now consider some $T > 0$. Moving occurs only if the value of ϵ after shocks have occurred (ϵ') is such that $\epsilon' < x$. Shocks arrive at a Poisson rate a_i , so the number k of shocks that would occur to a match over an interval of time T has a $\text{Poisson}(a_i T)$ distribution, which means the probability that k shocks occur is $e^{-a_i T} (a_i T)^k / k!$. If no shocks occur, $\epsilon' = \epsilon$, so no moving occurs. If $k \geq 1$ shocks have occurred then $\epsilon' = \delta_i^k \epsilon$, where ϵ is the initial draw of match quality. These matches survive only if $\epsilon' \geq x$, that is, $\epsilon \geq x / \delta_i^k$. Since the original values of ϵ are drawn from a Pareto distribution truncated at $\epsilon = y$ with shape parameter λ , this probability is $((x / \delta_i^k) / y)^{-\lambda}$ (this expression is valid for all $k \geq 1$ since $\delta_i y < x$). Therefore, the survival function $\psi_i(T)$ is given by:

$$\begin{aligned} \psi_i(T) &= e^{-a_i T} + \sum_{k=1}^{\infty} e^{-a_i T} \frac{(a_i T)^k}{k!} \left(\frac{x / \delta_i^k}{y}\right)^{-\lambda} \\ &= e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda e^{-a_i T} \sum_{k=1}^{\infty} \frac{(a_i \delta_i^\lambda T)^k}{k!} = e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda e^{-a_i T} e^{a_i \delta_i^\lambda T}, \end{aligned}$$

where the final equality uses the (globally convergent) series expansion of the exponential function. Conditional on each type $i = 1, \dots, q$, the survival function $\psi_i(T)$ is thus:

$$\psi_i(T) = \left(1 - \left(\frac{y}{x}\right)^\lambda\right) e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda e^{-a_i(1 - \delta_i^\lambda)T}.$$

Given the random assignment of types with probabilities $\theta_1, \dots, \theta_q$, the survival function $\psi(T)$ for all members of cohort of new homeowners is:

$$\psi(T) = \sum_{i=1}^q \theta_i \psi_i(T) = \left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \theta_i e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i e^{-a_i(1-\delta_i^\lambda)T}, \quad [\text{A.9.1}]$$

observing that $\psi(0) = 1$.

In the cross-section of homeowners, the stationary distribution of time spent in the home up to now, denoted by $\varpi(T)$, can be derived immediately from the survival function:

$$\varpi(T) = \frac{\psi(T)}{\int_{\tau=0}^{\infty} \psi(\tau) d\tau} = \frac{\left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \theta_i e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i e^{-a_i(1-\delta_i^\lambda)T}}{\left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \frac{\theta_i}{a_i} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \frac{\theta_i}{a_i(1-\delta_i^\lambda)}}. \quad [\text{A.9.2}]$$

For new matches, the distribution $\mu(T)$ of the time T until the next move can also be obtained from the survival function $\psi(T)$ using $\mu(T) = -\psi'(T)$. Hence, by taking the derivative of the survival function $\psi(T)$ from (A.9.1) with respect to the duration T :

$$\mu(T) = -\psi'(T) = \left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \theta_i a_i e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i a_i (1 - \delta_i^\lambda) e^{-a_i(1-\delta_i^\lambda)T}. \quad [\text{A.9.3}]$$

The definition of the hazard function $h(T)$ is the proportional decrease in the survival function for a small change in duration, that is, $h(T) = -\psi'(T)/\psi(T)$. Using (A.9.1) and (A.9.3):

$$h(T) = \frac{\left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \theta_i a_i e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i a_i (1 - \delta_i^\lambda) e^{-a_i(1-\delta_i^\lambda)T}}{\left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \theta_i e^{-a_i T} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i e^{-a_i(1-\delta_i^\lambda)T}}, \quad [\text{A.9.4}]$$

and by simplifying this expression, equation (5.1) is confirmed.

The expected time T_n between moves is the expected value of the probability distribution $\mu(T)$:

$$T_n = \int_{T=0}^{\infty} T \mu(T) dT = \left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \frac{\theta_i}{a_i} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \frac{\theta_i}{a_i(1-\delta_i^\lambda)}, \quad [\text{A.9.5}]$$

which can be simplified to derive the expression for T_n in (5.3). It is also the case that T_n is equal to the reciprocal of the average moving rate n , which is in turn equal to $\varpi(0)$. The former can be seen by comparing (A.8.17) and (A.9.5), and the latter by integration by parts, the definition of $\varpi(T)$ from (A.9.2), and $\psi(0) = 1$:

$$\frac{1}{\varpi(0)} = \frac{\int_{\tau=0}^{\infty} \psi(\tau) d\tau}{\psi(0)} = \lim_{T \rightarrow \infty} [\tau \psi(\tau)]_{\tau=0}^T + \int_{T=0}^{\infty} T (-\psi'(T)) dT = \int_{T=0}^{\infty} T \mu(T) dT = T_n = \frac{1}{n}.$$

Finally, n is the average hazard rate $h(T)$ with respect to the cross-sectional distribution of homeowner durations $\varpi(T)$.

$$\begin{aligned} \int_{T=0}^{\infty} h(T) \varpi(T) dT &= \int_{T=0}^{\infty} \left(-\frac{\psi'(T)}{\psi(T)}\right) \left(\frac{\psi(T)}{\int_{\tau=0}^{\infty} \psi(\tau) d\tau}\right) dT = \frac{-\int_{T=0}^{\infty} \psi'(T) dT}{\int_{\tau=0}^{\infty} \psi(\tau) d\tau} \\ &= \frac{\psi(0)}{\int_{\tau=0}^{\infty} \psi(\tau) d\tau} = \varpi(0) = n. \end{aligned}$$

Note that T_n is not generally equal to the average of the cross-sectional distribution of durations at a point

in time, denoted by T_h :

$$T_h = \int_{T=0}^{\infty} T \varpi(T) dT = \frac{\left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \frac{\theta_i}{a_i^\lambda} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \frac{\theta_i}{a_i^2(1-\delta_i^\lambda)^2}}{\left(1 - \left(\frac{y}{x}\right)^\lambda\right) \sum_{i=1}^q \frac{\theta_i}{a_i} + \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \frac{\theta_i}{a_i(1-\delta_i^\lambda)}}.$$

The elasticity of the moving rate

Imposing a common moving threshold $x_t = x_{i,t}$ for all $i = 1, \dots, q$ (as shown in [appendix A.8](#)) and differentiating the moving rate n_t from (A.8.6) with respect to x_t :

$$\frac{\partial n_t}{\partial x_t} = \frac{\lambda x_t^{-\lambda-1} v \sum_{i=1}^q \theta_i a_i \delta_i^\lambda \int_{\tau \rightarrow -\infty}^t e^{-a_i(1-\delta_i^\lambda)(t-\tau)} u_\tau d\tau}{1 - u_t}.$$

Let η denote the elasticity of the moving rate n_t with respect to x_t evaluated in the steady state. The partial derivative above implies:

$$\eta = \left. \frac{\partial \log n_t}{\partial \log x_t} \right|_{u_t=u, n_t=n} = \frac{\lambda x^{-\lambda} v u \sum_{i=1}^q \theta_i a_i \delta_i^\lambda \int_{\tau \rightarrow -\infty}^t e^{-a_i(1-\delta_i^\lambda)(t-\tau)} d\tau}{n(1-u)}. \quad [\text{A.9.6}]$$

Note that the integrals appearing in the expression above are:

$$\int_{\tau \rightarrow -\infty}^t e^{-a_i(1-\delta_i^\lambda)(t-\tau)} d\tau = \frac{1}{a_i(1-\delta_i^\lambda)},$$

and by substituting these into (A.9.6):

$$\eta = \frac{\lambda x^{-\lambda} v \sum_{i=1}^q \theta_i \frac{a_i \delta_i^\lambda}{a_i(1-\delta_i^\lambda)}}{n(1-u)} = \lambda x^{-\lambda} v \frac{u}{n(1-u)} \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{1-\delta_i^\lambda}.$$

Equation (4.12) implies $u/(1-u) = n/s$, and by substituting this and the expression for s from (4.10) into the above:

$$\eta = \lambda \frac{v x^{-\lambda} n}{n v y^{-\lambda}} \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{1-\delta_i^\lambda} = \lambda \left(\frac{y}{x}\right)^\lambda \sum_{i=1}^q \theta_i \frac{\delta_i^\lambda}{1-\delta_i^\lambda}. \quad [\text{A.9.7}]$$

This confirms the formula for η given in (5.4).

A.10 Estimates of time-to-sell

This section provides further discussion of alternative estimates of time-to-sell. 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.

These numbers are significantly smaller than the 6 months estimate of time-to-sell derived from the NAR data on sales and inventories. 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.

To understand the impact withdrawals can have on estimates of time-to-sell, suppose houses on the market have sales rate s and withdrawal rate w as in [section 2.3](#). Let \tilde{T}_s denote the average time taken to sell among those houses that are successfully sold, measuring the time on the market from the start of the

most recent listing. Let T_s denote the average of the total time spent on the market by houses that are successfully sold, ignoring the times between listings when houses are off the market. The two measures of time-to-sell are:

$$\tilde{T}_s = \frac{1}{s+w}, \quad T_s = \frac{1}{s}, \quad \text{and hence } T_s = \frac{\tilde{T}_s}{1-\phi}.$$

The final equation gives the relationship between the two measures in terms of the fraction ϕ of houses eventually withdrawn from sale. Estimates of time-to-sell based on survey data are typically measuring \tilde{T}_s . On the other hand, the NAR data provide an estimate of the sales rate s , and taking the reciprocal yields a measure of T_s .

The studies by [Anenberg and Laufer \(2017\)](#) and [Carrillo and Williams \(2015\)](#) discussed in [section 2.3](#) suggest that the fraction of properties eventually withdrawn from sale lies between 50% and 60%. Using these numbers, the formula above suggests that estimates of T_s should be around 2 to 2.5 times higher than estimates of \tilde{T}_s . This simple observation goes a long way in reconciling the magnitudes of the different estimates. [Carrillo and Williams \(2015\)](#) also show directly that controlling for withdrawals substantially increases the estimated value of average time-to-sell. Similarly, 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).

An alternative approach to estimating time-to-sell unaffected by withdrawals is to look at the average duration for which a home is vacant using data from the American Housing Survey. In the years 2001–2005, the mean duration of 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. Another approach that avoids the problem of withdrawals is to look at the average time taken to sell newly built houses. [Díaz and Jerez \(2013\)](#) use the Census Bureau ‘New Residential Sales’ report to find that the median number of months taken to sell a newly built house is 5.9 (for the period 1991–2012). This is only slightly shorter than the average of the time-to-sell number constructed using NAR data on existing single-family homes, but there is reason to believe that newly built homes should sell faster than existing homes owing to greater advertising expenditure and differences in the target groups of buyers.

A.11 Calibration of the model with heterogeneity

This section shows how the parameters of the model with heterogeneity can be set to give the best fit to the empirical hazard function for moving house, as well as matching other empirical targets. There are parameters q , $\{\theta_i\}$, $\{a_i\}$, and $\{\delta_i\}$ that describe the distributions of idiosyncratic shocks faced by homeowners, and parameters λ , v , C , F , M , κ , ω , and r that are related to other aspects of the model.

The number of types q determines the dimension of the parameter space, and as discussed in [section 5.2.1](#), this can be chosen to be large enough to give a sufficiently good fit to the empirical hazard function. Here, q is taken as given. Three of the other parameters (κ , ω , and r) are also set directly. The other parameters are chosen to minimize the sum of squared deviations between the empirical hazard function and the hazard function $h(T)$ implied by the model, and to match five empirical targets exactly: time-to-sell T_s , viewings per sale V_s , the transaction cost to price ratio c , the flow search cost to price ratio f , and the flow maintenance cost to price ratio m .

The calibration procedure has two stages. First, a numerical search over parameters θ_i , a_i , and δ_i^λ to find the solution to:

$$\min_{\substack{\{\theta_i\}_{i=1}^q, \{a_i\}_{i=1}^q, \{\delta_i^\lambda\}_{i=1}^q \\ \text{s.t. } \sum_{i=1}^q \theta_i = 1}} \sum_T (\hat{h}(T) - h(T))^2, \quad [\text{A.11.1}]$$

where T denotes a duration for which data on the hazard function is available, $\hat{h}(T)$ is the estimated hazard rate described in [section 5.2](#), and $h(T)$ is the model-implied hazard function given in [\(5.1\)](#).⁵⁰ There are

⁵⁰The numerical method used is to draw many initial conditions at random from the parameter space, and then

$3q - 1$ independent parameters given that $\sum_{i=1}^q \theta_i = 1$, and the parameters must satisfy the restrictions $0 < \theta_i \leq 1$, $a_i > 0$, and $0 \leq \delta_i^\lambda < 1$. The procedure specifies δ_i^λ rather than δ_i because this turns out to be more convenient, and the admissible range of δ_i^λ is the same as that of δ_i .

To compute the model-implied hazard function $h(T)$ using (5.1) it is necessary to know the value of the parameter λ and the values of the endogenous variables x and y in addition to $\{\theta_i\}$, $\{a_i\}$, and $\{\delta_i^\lambda\}$. Conditional on the parameters given at the first stage, the values of λ , x , and y are pinned down by the requirement of matching the other five empirical targets: T_s , V_s , c , f , and m . This is the second stage of the calibration procedure.

By dividing the cost parameters C , F , and M (search, transactions, and maintenance) by the average transaction price P from equation (A.8.16), the model's predictions for the targets c , f , and m are:

$$c = \frac{\frac{C}{\xi}}{\kappa \frac{C}{\xi} - \frac{M}{r\xi} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v} \right) \left(x + \frac{F}{\xi} \right)}; \quad [\text{A.11.2a}]$$

$$f = \frac{\frac{F}{\xi}}{\kappa \frac{C}{\xi} - \frac{M}{r\xi} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v} \right) \left(x + \frac{F}{\xi} \right)}; \quad [\text{A.11.2b}]$$

$$m = \frac{\frac{M}{\xi}}{\kappa \frac{C}{\xi} - \frac{M}{r\xi} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v} \right) \left(x + \frac{F}{\xi} \right)}. \quad [\text{A.11.2c}]$$

Note that the model contains one other parameter ξ in addition to those listed earlier, but in all equations determining observables, ξ enters only as a ratio to other parameters. The parameter ξ is therefore normalized to $\xi = 1$. Now take equation (A.8.16) for the average price and divide both sides by P :

$$\kappa c - \frac{m}{r} + \omega \left(\frac{1}{r} + \frac{y^\lambda}{v} \right) \left(\frac{x}{P} + f \right) = 1.$$

Using the expression for T_s in (A.3.22), the equation above can be solved for x/P as follows:

$$\frac{x}{P} = \frac{1 - \kappa c + \frac{m}{r}}{\omega \left(\frac{1}{r} + T_s \right)} - f. \quad [\text{A.11.3}]$$

Now take the linear equation (A.8.11) involving the thresholds x and y and divide both sides by P (recalling that $\xi = 1$):

$$\frac{y}{P} = \frac{x}{P} + \frac{c}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}} = \frac{1 - \kappa c + \frac{m}{r}}{\omega \left(\frac{1}{r} + T_s \right)} + \frac{c}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}} - f, \quad [\text{A.11.4}]$$

and then dividing both sides by x/P and using (A.11.3):

$$\frac{y}{x} = \frac{y/P}{x/P} = 1 + \frac{\frac{c}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}}}{\frac{1 - \kappa c + \frac{m}{r}}{\omega \left(\frac{1}{r} + T_s \right)} - f}. \quad [\text{A.11.5}]$$

Dividing both sides of the second equation (A.8.15) for the thresholds x and y by $P \sum_{i=1}^q \theta_i / (r + a_i)$ and rearranging leads to:

$$\frac{\sum_{i=1}^q \frac{\theta_i}{r+a_i} \left(1 + \frac{a_i \delta_i^\lambda}{r+a_i(1-\delta_i^\lambda)} \left(\frac{y}{x} \right)^{\lambda-1} \right)}{(\lambda - 1) \sum_{i=1}^q \frac{\theta_i}{r+a_i}} = \frac{y^\lambda \left(\frac{x}{P} + f \right)}{v \frac{y}{P} \sum_{i=1}^q \frac{\theta_i}{r+a_i}}.$$

perform a search for the minimum using a simplex algorithm starting from each initial condition, before finally choosing the parameter vector with the smallest sum of squared deviations from among all the searches.

Using $T_s = y^\lambda/v$ from (A.3.22) together with (A.11.3), (A.11.4), and (A.11.5), this equation can be written as follows:

$$\mathcal{C}(\lambda) = \frac{1 + \aleph_\delta \aleph_{yx}^{\lambda-1}}{\lambda - 1} - \aleph = \frac{1 + \aleph_\delta e^{(\log \aleph_{yx})(\lambda-1)}}{\lambda - 1} - \aleph = 0, \quad [\text{A.11.6}]$$

where the coefficients \aleph , \aleph_δ , and \aleph_{yx} can be derived from the calibration targets:

$$\aleph = \frac{\frac{(1-\kappa c + \frac{m}{r})T_s}{\omega(\frac{1}{r} + T_s)}}{c + \left(\frac{1-\kappa c + \frac{m}{r}}{\omega(\frac{1}{r} + T_s)} - f\right) \sum_{i=1}^q \frac{\theta_i}{r+a_i}}, \quad \aleph_\delta = \frac{\sum_{i=1}^q \frac{\theta_i a_i \delta_i^\lambda}{r+a_i(1-\delta_i^\lambda)}}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}}, \quad \text{and} \quad \aleph_{yx} = 1 + \frac{\frac{c}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}}}{\frac{1-\kappa c + \frac{m}{r}}{\omega(\frac{1}{r} + T_s)} - f}. \quad [\text{A.11.7}]$$

Observe that the function $\mathcal{C}(\lambda)$ becomes an arbitrarily large positive number as λ tends to 1, and since $\aleph_{yx} > 1$, it is also the case that $\mathcal{C}(\lambda)$ eventually becomes arbitrarily large as λ increases. Note that the derivative of $\mathcal{C}(\lambda)$ is:

$$\mathcal{C}'(\lambda) = \frac{((\lambda - 1)^2 - 1) \aleph_\delta e^{(\log \aleph_{yx})(\lambda-1)} - 1}{(\lambda - 1)^2}.$$

The denominator of this expression is always positive given that $\lambda > 1$. The sign of the numerator depends only on $\aleph_\delta((\lambda - 1)^2 - 1) - e^{-(\log \aleph_{yx})(\lambda-1)}$, which is strictly increasing in λ for all $\lambda > 1$. Since $\mathcal{C}'(1) < 0$, it follows that the function $\mathcal{C}(\lambda)$ is initially decreasing in λ and subsequently increasing in λ after passing a threshold value of λ . For any $\lambda > 1$, it must be the case that $e^{(\log \aleph_{yx})(\lambda-1)} > 1 + (\log \aleph_{yx})(\lambda - 1)$ because $\log \aleph_{yx} > 0$. This inequality implies the function $\mathcal{C}(\lambda)$ from (A.11.6) has the following lower bound:

$$\mathcal{C}(\lambda) > \frac{1 + \aleph_\delta}{\lambda - 1} - (\aleph - \aleph_\delta \log \aleph_{yx}). \quad [\text{A.11.8}]$$

As a solution for λ requires $\mathcal{C}(\lambda) = 0$, a necessary condition for a solution to exist is $\aleph > \aleph_\delta \log \aleph_{yx}$. When this condition is satisfied, the inequality above implies a lower bound $\underline{\lambda}$ for a solution (if one exists):

$$\lambda > \underline{\lambda}, \quad \text{where} \quad \underline{\lambda} = 1 + \frac{1 + \aleph_\delta}{\aleph - \aleph_\delta \log \aleph_{yx}},$$

which follows because the bound (A.11.8) on $\mathcal{C}(\lambda)$ is decreasing in λ . The parameter λ must also satisfy an upper bound. It is required that $\delta_i y < x$ for all $i = 1, \dots, q$. This is equivalent to $(y/x)^\lambda \delta_i^\lambda < 1$ for all i , and hence $\aleph_{yx}^\lambda \max \delta_i^\lambda < 1$ since $y/x = \aleph_{yx}$. By taking logarithms of both sides, this implies an upper bound for λ :

$$\lambda < \bar{\lambda}, \quad \text{where} \quad \bar{\lambda} = \frac{-\log \max \delta_i^\lambda}{\log \aleph_{yx}}, \quad [\text{A.11.9}]$$

with $\aleph_{yx} > 1$ taken from (A.11.7). Given the properties of the $\mathcal{C}(\lambda)$ function established above, the necessary and sufficient conditions for the existence of a unique solution $\lambda > 1$ to the equation $\mathcal{C}(\lambda) = 0$ with $1 < \lambda < \bar{\lambda}$ are that $\underline{\lambda} > 1$ and $\mathcal{C}(\bar{\lambda}) < 0$. When these conditions are met, the solution for λ can be found by searching the interval $(\underline{\lambda}, \bar{\lambda})$ because $\mathcal{C}(\underline{\lambda}) > 0$ and $\mathcal{C}(\bar{\lambda}) < 0$.

With the solution of (A.11.6) for λ , the transactions threshold y can be obtained from viewings per sale V_s :

$$y = V_s^{\frac{1}{\lambda}}, \quad [\text{A.11.10}]$$

and the moving threshold x can be derived from the above along with equation (A.11.5):

$$x = \frac{V_s^{\frac{1}{\lambda}}}{1 + \frac{\frac{c}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}}}{\frac{1-\kappa c + \frac{m}{r}}{\omega(\frac{1}{r} + T_s)} - f}}. \quad [\text{A.11.11}]$$

With λ , x , and y , the hazard function $h(T)$ can be computed using the formula in (5.1) given the values of $\{\theta_i\}$, $\{a_i\}$, and $\{\delta_i^\lambda\}$. This allows the sum of squared deviations (A.11.1) to be computed, and hence the calibration procedure can be implemented as described above.

Once θ_i , a_i , and δ_i^λ have been chosen to minimize (A.11.1), the parameters $\{\delta_i\}$ can be obtained from $\{\delta_i^\lambda\}$ using the value of λ that solves (A.11.6) and $\delta_i = (\delta_i^\lambda)^{1/\lambda}$. The remaining parameters v , C , F , and M can be obtained as follows. Using (A.3.21) and (A.3.22), the ratio of viewings per sale V_s and time to sell T_s determines the meeting rate v :

$$v = \frac{V_s}{T_s}. \quad [\text{A.11.12}]$$

Combining equations (A.11.4) and (A.11.10) leads to the following expression for P :

$$P = \frac{V_s^{\frac{1}{\lambda}}}{\frac{1-\kappa c + \frac{m}{r}}{\omega(\frac{1}{r} + T_s)} + \frac{c}{\sum_{i=1}^q \frac{\theta_i}{r+a_i}} - f}, \quad [\text{A.11.13}]$$

and this can be used to obtain the parameters C , F , and M using $C = cP$, $F = fP$, and $M = mP$.

A.12 Calibration of the basic model

This section shows how the 10 parameters a , δ , λ , v , C , F , M , κ , ω , and r can be determined in the basic version of the model (with no heterogeneity in idiosyncratic shock distributions, that is, $q = 1$). When $q = 1$, the general expressions for time-to-move T_n and the elasticity of the moving rate η from (5.3) and (5.4) reduce to:

$$T_n = \frac{1 + \frac{\delta^\lambda}{1-\delta^\lambda} \left(\frac{y}{x}\right)^\lambda}{a}; \quad [\text{A.12.1}]$$

$$\eta = \lambda \frac{\delta^\lambda}{1-\delta^\lambda} \left(\frac{y}{x}\right)^\lambda. \quad [\text{A.12.2}]$$

Three of the parameters (κ , ω , and r) are set directly. The other seven are obtained indirectly from five calibration targets: time-to-sell T_s , viewings per sale V_s , the transaction cost to price ratio c , the flow search cost to price ratio f , and the flow maintenance cost to price ratio m , together with two other targets derived from information contained in the hazard function, namely time-to-move T_n and the steady-state elasticity η of the moving rate with respect to the moving threshold.

The price equation (A.4.3) is identical to (A.8.16) in the model with heterogeneity, so the expressions in (A.11.2a)–(A.11.2c) for the ratios of costs (search, transactions, and maintenance) to the average price are also valid here. The model contains one other parameter ξ , but as in appendix A.11, in all equations determining observables, ξ enters only as a ratio to other parameters, hence it can be normalized to $\xi = 1$.

The calibration method begins by setting κ , ω , and r directly. Next, consider a guess for T_δ , the expected time until an idiosyncratic shock occurs. This conjecture determines the parameter a using:

$$a = \frac{1}{T_\delta}. \quad [\text{A.12.3}]$$

The admissible range for T_δ is $0 < T_\delta < T_n$.

Using equation (A.12.1) for time-to-move T_n and the expressions for η and T_δ from equations (A.12.2) and (A.12.3):

$$T_n = \left(1 + \frac{\eta}{\lambda}\right) T_\delta. \quad [\text{A.12.4}]$$

Rearranging equation (A.12.4) and using the calibration targets T_n and η and the conjecture for T_δ :

$$\lambda = \frac{\eta T_\delta}{T_n - T_\delta}. \quad [\text{A.12.5}]$$

This yields the value of the parameter λ .

Since the equation for the average price is the same as the model with heterogeneity, the same steps used to derive the expression for x/P in (A.11.3) are valid here. Now take the linear equation (A.3.7) involving the thresholds x and y and divide both sides by P , recalling that $\xi = 1$:

$$\frac{y}{P} = \frac{x}{P} + (r+a)c = \frac{1 - \kappa c + \frac{m}{r}}{\omega \left(\frac{1}{r} + T_s\right)} + (r+a)c - f, \quad [\text{A.12.6}]$$

which uses the formula for x/P from (A.11.3). Dividing both sides by x/P and using (A.11.3) again:

$$\frac{y}{x} = \frac{y/P}{x/P} = 1 + \frac{(r+a)c}{\frac{1 - \kappa c + \frac{m}{r}}{\omega \left(\frac{1}{r} + T_s\right)} - f}. \quad [\text{A.12.7}]$$

This gives the ratio y/x implied by the calibration targets. Equation (A.12.2) for η can be rearranged to obtain an expression for δ^λ :

$$\delta^\lambda = \frac{\eta}{\eta + \lambda \left(\frac{y}{x}\right)^\lambda},$$

and hence the value of the parameter δ is:

$$\delta = \left(\frac{\eta}{\eta + \lambda \left(\frac{y}{x}\right)^\lambda} \right)^{\frac{1}{\lambda}}. \quad [\text{A.12.8}]$$

Given λ , the transactions threshold y must satisfy equation (A.11.10) in terms of viewings per sale V_s given that (A.3.21) holds for the model with and without heterogeneity. An expression for the moving threshold x can be derived using (A.11.10) and (A.12.7):

$$x = \frac{V_s^{\frac{1}{\lambda}}}{1 + \frac{(r+a)c}{\frac{1 - \kappa c + \frac{m}{r}}{\omega \left(\frac{1}{r} + T_s\right)} - f}}. \quad [\text{A.12.9}]$$

The parameter v must satisfy (A.11.12) in terms of viewings per sale V_s and time to sell T_s given that equations (A.3.21) and (A.3.22) are the same with and without heterogeneity. Combining equations (A.12.6) and (A.11.10) leads to the following expression for P :

$$P = \frac{V_s^{\frac{1}{\lambda}}}{\frac{1 - \kappa c + \frac{m}{r}}{\omega \left(\frac{1}{r} + T_s\right)} + (r+a)c - f},$$

and this can be used to obtain the parameters C , F , and M using $C = cP$, $F = fP$, and $M = mP$. Finally,

equation (A.3.16) must also hold, which requires:

$$x + \frac{F}{\xi} = \frac{v}{(\lambda - 1)(r + a)} \left(y^{1-\lambda} + \frac{a\delta^\lambda}{r + a(1 - \delta^\lambda)} x^{1-\lambda} \right),$$

and this is used to verify the initial conjecture for T_δ .

A.13 Productivity and interest rates

Suppose that a family's flow utility is $C_t^{1-\theta} \mathcal{H}_t^\theta$, where C_t denotes consumption and \mathcal{H}_t denotes housing, and where θ indicates the importance of housing in the utility function ($0 < \theta < 1$). This adds non-housing goods to the model and replaces the flow utility $\xi\epsilon$ assumed earlier. The form of the flow utility function assumes complementarity between consumption and housing services. The housing variable \mathcal{H}_t that enters the utility function is equal to the match quality ϵ of a family with its house, and the evolution of this variable in response to idiosyncratic shocks and moving and transaction decisions is the same as before. The discount rate for future utility flows is the rate of pure time preference ρ . The lifetime utility function from time T onwards is therefore:

$$\mathcal{U}_T = \int_{t=T}^{\infty} e^{-\rho(t-T)} C_t^{1-\theta} \mathcal{H}_t^\theta dt. \quad [\text{A.13.1}]$$

Suppose there are complete financial markets for securities with consumption payoffs contingent on any state of the world, and suppose all families receive the same real income (with no aggregate risk) and initially all have equal financial wealth. Note that only state-contingent consumption, not housing services, can be traded in these markets. With complete financial markets there is full consumption insurance of idiosyncratic risk coming from shocks to match quality and the uncertainties in the search process, implying that the marginal utility of consumption must be equalized across all families. The marginal utility of consumption is $z_t^{-\theta}$, where $z_t = C_t/\mathcal{H}_t$ is the ratio of consumption to housing match quality. If r is the real interest rate (in terms of consumption goods) then maximization of utility (A.13.1) subject to the lifetime budget constraint requires that the following consumption Euler equation holds:

$$\theta \frac{\dot{z}_t}{z_t} = r - \rho. \quad [\text{A.13.2}]$$

In equilibrium, the sum of consumption C_t across all families must be equal to aggregate real income Y_t , which is assumed to be an exogenous endowment growing at rate g over time. Given equalization of $z_t = C_t/\mathcal{H}_t$ across all families at a point in time and given a stationary distribution of match quality $\mathcal{H}_t = \epsilon$ across all families, it follows that all families have a value of z_t proportional to aggregate real income Y_t at all times:

$$z_t = \varkappa Y_t, \quad \text{where } \varkappa = \frac{1}{(1-u)Q}. \quad [\text{A.13.3}]$$

The constant \varkappa is the reciprocal of total match quality $(1-u)Q$ in steady state (noting that unsatisfied owners receive no housing utility flows). Substituting this into the consumption Euler equation (A.13.2) implies that the equilibrium real interest rate is:

$$r = \rho + \theta g. \quad [\text{A.13.4}]$$

Since (A.13.3) implies $C_t = \varkappa Y_t \mathcal{H}_t$, it follows that lifetime utility (A.13.1) can be expressed as follows:

$$\mathcal{U}_T = \varkappa^{1-\theta} \int_{t=T}^{\infty} e^{-\rho(t-T)} Y_t^{1-\theta} \mathcal{H}_t dt.$$

With Y_t growing at rate g , income at time t can be written as $Y_t = e^{g(t-T)}Y_T$. By substituting this into the lifetime utility function and using the expression for the real interest rate r in (A.13.4):

$$\mathcal{U}_T = \varkappa^{1-\theta} Y_T^{1-\theta} \int_{t=T}^{\infty} e^{-(r-g)(t-T)} \mathcal{H}_t dt. \quad [\text{A.13.5}]$$

Lifetime utility is therefore a discounted sum of match quality $\mathcal{H}_t = \epsilon$. The coefficient of match quality (this is the parameter ξ in the main text) is increasing in the current level of real income, and the discount rate (denoted r in the main text) is the difference between the market interest rate and the growth rate of real income. This provides a justification for interpreting a rise in real incomes as an increase in ξ and a fall in the market interest rate as a lower discount rate.

A.14 The effects of easier credit

Table 6 reports the results in the case of a one-third reduction in buyers' transaction costs C_b . The effects are shown in isolation and in combination with the other factors.

Table 6: *Improvement in credit conditions*

Factor	Transactions	Listings	Sales rate	Moving rate	Houses for sale	Prices
C_b	20%	20%	16%	20%	4%	2%
C_b , productivity, & search	31%	31%	22%	31%	9%	35%
All factors	35%	35%	12%	36%	23%	71%
Data (1995–2003)	27%	34%	14%	34%	13%	31%

A.15 Sensitivity analysis

Table 7 below conducts a sensitivity analysis in respect of some of the calibration targets to check the robustness of the results and to identify the key mechanisms at work in the model.

The first exercise is to explore the relative importance of the two search frictions discussed in section 3.2 that are found in the model. The first friction relates to the time taken to find suitable houses to view. The second friction relates to houses having a range of possible match qualities with different buyers that only become known to a buyer once a house is viewed. The sensitivity analysis considers a reduction in the first friction, a reduction in the second friction, and a reduction of both frictions together.

Lowering the first friction is equivalent to increasing the viewing rate v , while lowering the second friction is equivalent to increasing the shape parameter λ of the Pareto distribution of match quality. To increase v while keeping λ constant requires holding the time-to-sell T_s constant. This can only be done by increasing viewings per sale V_s since $v = T_s/V_s$. To increase λ while keeping v constant essentially means decreasing average viewing per sale V_s and lowering time-to-sell T_s in proportion to the reduction in V_s . The effects of reducing the two frictions separately by an order of magnitude are reported in the 'low first friction' and 'low second friction' rows. The row labelled 'both negligible' makes both frictions quantitatively irrelevant by reducing viewings per sale V_s to one and time-to-sell T_s to one day. These three rows clearly demonstrate the importance of the housing mismatch and the need to inspect houses before purchasing.

The next exercise is to vary the size of transaction costs C . This has a large effect on the results, with stronger effects of endogenous moving found when transaction costs are high relative to house prices. To understand this, note that in the special case of zero transactions costs, the model has the surprising feature that its steady-state equilibrium is isomorphic to an exogenous moving model with the parameter a

redefined as $a(1 - \delta^\lambda)$. The logic behind this is that equation (4.4) implies $y = x$ when $C = 0$. From (4.11), this means that $n = a(1 - \delta^\lambda)$, so the moving rate is independent of the equilibrium moving and transactions thresholds. Hence, only those parameters directly related to the shocks received by homeowners affect the moving rate. The equilibrium value of y is determined by replacing x by y in equation (4.7) and simplifying to:

$$\frac{vy^{1-\lambda}}{(r + a(1 - \delta^\lambda))(\lambda - 1)} = y + \frac{F}{\xi}.$$

This has the same form as (4.7) when $\delta = 0$, that is, when moving is exogenous, so all steady-state predictions of the two models would be the same if $C = 0$.

As can be seen in Table 7, the size of the flow cost of search F has a much smaller impact on the results than transaction costs C . Finally, the extent of the seller's bargaining power ω does make a difference to the results, with higher seller's bargaining power increasing the strength of the results. At first glance, this is surprising because in the model, bargaining power should affect only prices, not quantities. However, changes in bargaining power require changes in the other parameters to continue to match the calibration targets. As can be seen from equation (4.13), an increase in ω raises average transaction prices, which requires an increase in transaction costs C to match the calibration target for $c = C/P$. Following the discussion above, it is the required increase in C after raising ω that has a large impact on the results for quantities.

Table 7: *Sensitivity analysis*

Target to vary	Transactions	Listings	Sales rate	Moving rate	Houses for sale	Prices
<i>Frictions in the search process</i>						
Low 1 st fric.: $T_s = 6.5/12$, $V_s = 100$	16%	16%	11%	17%	6%	33%
Low 2 nd fric.: $T_s = 6.5/120$, $V_s = 1$	10%	10%	4%	10%	6%	32%
Both negligible: $T_s = 1/365$, $V_s = 1$	1%	1%	-1%	1%	2%	32%
<i>Transaction costs</i>						
Low: $c = 0.05$	9%	9%	5%	9%	4%	33%
High: $c = 0.2$	30%	30%	21%	30%	9%	34%
<i>Flow costs of search</i>						
Low: $f = 0.0125$	16%	16%	11%	16%	4%	35%
High: $f = 0.05$	19%	19%	10%	19%	9%	29%
<i>Bargaining power of the seller</i>						
Low: $\omega = 0.25$	8%	8%	5%	8%	3%	34%
High: $\omega = 0.75$	25%	25%	16%	25%	9%	32%
Baseline	16%	16%	11%	17%	6%	33%
Data (1995–2003)	27%	34%	14%	34%	13%	31%

Notes: The table shows the long-run steady-state effects of the changes to productivity and internet search in the basic version of the model.