New Listing Alert: Alternative Theory of Housing Search*

Sonia Gilbukh†

March 21, 2025

Abstract

This paper examines the validity of a stock-flow approach in modeling housing search. Contrary to the commonly used random search model where frictions arise from buyers' inability to sample the entire distribution of houses, in the stock-flow model, all successful matches between buyers and sellers are realized immediately. The current stock of buyers and sellers is therefore mismatched, and both sides are waiting for the new arrivals in the market to find a successful counter party. I show that a parsimonious stock-flow model can match empirical patterns that are difficult to achieve in a random search framework. Using data of 3.3 million listings from one of the largest Multiple Listing Service platforms in the United States, I find that, consistent with the model, 1) sales are much more correlated with new listings than they are with the market inventory, 2) sale hazard rates decline after the initial period, and 3) houses that sell in the initial period are likely to sell at a premium while those that spend a long time on the market sell at a discount. Lastly, I use a sample of listings that were taken off the market at some point during the listing to corroborate the mismatch story of buyer and seller matching.

*First version: April 1, 2021. This version: March 21, 2025. I am grateful to Alina Arefeva, Eric Smith, Lu Han, and Melvin Coles for thoughtful discussions that greatly improved this paper. I thank Stijn Van Nieuwerburgh, Jessie Handbury, Maisy Wong, and Will Strange for their support and guidance. In addition, I thank conference and seminar participants at the New York City real estate conference, Baruch Real Estate, AREUEA National Conference, CEA, Penn State University, Manhattan College, University of Southern California, Haverford College, University of Connecticut, and the Housing Search Conference. Lastly, I thank the Steven L Newman Real Estate Institute for the generous funding support.

†Baruch College, City University of New York. Email: sophia.gilbukh@gmail.com

1 Introduction

Models of search have long been used for understanding unique features of the housing market. Starting with Wheaton (1990), models that incorporate search frictions have been helpful in explaining puzzling excess price volatility (Arefeva (2020) Ngai and Tenreyro (2014), Ngai and Sheedy (2020)), price momentum (Head et al. (2014), Guren (2018)), and understanding liquidity and intermediation (Genesove and Han (2012), Gilbukh and Goldsmith-Pinkham (2024)). In these models, search frictions arise from the inability of buyers to assess the entire distribution of available homes all at once. Buyers sample the inventory of listings at random, one listing at a time, and decide whether to purchase or not. The matching functions determine how many matches are made in each period based on the *total* number of agents on each side of the market.

Although search friction models have become standard, the process of buying and selling a home in the United States has changed significantly over the last twenty years. With the emergence of national aggregator platforms like Zillow.com and Redfin.com, buyers are now well informed about for-sale properties.[1] In most markets, an online search will return a reasonable number of listings that a buyer can examine in one sitting.[2] In addition, recent advances in technology now allow buyers to have virtual 3D tours of the property, helping them narrow down their search further. For the remaining properties, a buyer can schedule in-person visits within days and see several options in one weekend. If a suitable match is not found, they can choose to be notified by the internet platform (or their agent) of new for-sale listings that fit their needs, so that their subsequent search focuses on properties they have not already considered. On the other side of the matching process, a new listing often hosts an open house in the first few weeks of being listed and attracts a sizable number of buyers to view the property all at once. Thus, random search has become a less attractive simplification of the market.

This paper examines an alternative approach to model search in which frictions arise due to mismatch between buyers and sellers. Here, the simultaneous presence of buyers and sellers in the market is due to none of the buyers liking any of the homes currently for sale. All market participants therefore have to wait for new potential counter parties to arrive in order to match successfully. I show that

[1] A survey conducted by National Association of Realtors (National Association of Relators (2023)) finds that all home buyers in 2023 used the Internet to search for a home.

[2] Piazzesi et al. (2020) documents search activity on a competing national platform named Trulia.com in San Francisco and makes a distinction between local buyers and broad searchers. However, they find that broad searchers tend to target particular high inventory areas.

a simple mismatch model can explain several patterns in the data that would be difficult to achieve in a random search framework. In addition, it provides a parsimonious way to model endogenous multilateral matching where in hot markets sellers might have several interested buyers at a time, while in downturns, buyers are particularly unwilling to negotiate with sellers since they have many outside options of for-sale properties.

In the mismatch model, henceforth referred to as a “stock-flow model", buyers have complete information about the market and can see all available homes for sale. However, each buyer has idiosyncratic preferences and only likes a particular house with probability α. All successful matches are realized *immediately*. New buyers and new houses arrive in the market one at a time, adding match opportunities to existing (old) buyers and sellers. A buyer who has been looking for homes for a while would only match with a new listing (otherwise they would have matched up earlier), and similarly a listing that has been sitting on the market would only match with a new buyer. The base trading price is determined via Nash bargaining; however, when a new market participant finds multiple successful counter parties that are willing to trade with them, they negotiate the price and capture all of the surplus from trade. Thus, new listings might find that several old buyers are willing to buy the house and the buyers bid the price up, or a new buyer finds several homes that they like and is able to negotiate the price down. To illustrate the simplicity of the model, I characterize it's equilibrium for a symmetric case where both the stock and the arrival rates are equal between buyers and sellers.

The model has several predictions. First, the number of matches will depend more on the *flow* of new buyers and sellers (turnover), rather than the stock. In other words, the number of matches is strongly correlated with the arrival of new listings and new buyers, even conditional on inventory, because all matches occur with one party being new to the market. Second, the hazard rate of sale falls sharply following the initial period since listing. In contrast with random matching models where there is no distinction between new and old listings, in this model a new listing has an opportunity to match with all of the old buyers in the market with no competition, but an old listing competes for the attention of new buyers (who come in one at a time) with the entire inventory of for-sale listings. Third, price negotiations naturally occur when there are several buyers who are interested in a new listing, while older listings only get attention from one new buyer at a time, and thus do not generate bidding wars. This means that on average, faster-selling homes sell at a price premium while older listings sell at a discount.

I test these predictions using a proprietary data set from one of the largest Multiple Listing Service platforms in the country. The data include all listings that were represented by a real estate agent between 2005 and 2020 in the Mid-Atlantic region that includes Maryland, Delaware, and significant parts of New Jersey, Pennsylvania, Virginia and West Virginia. Using regression analysis, I find that, consistent with the model, sales are better predicted by new listings than they are by the inventory of for-sale properties both in the aggregate series and within narrowly defined sub-markets. Adding new listings into the regression of sales on inventory improves the explanatory power from $R^2 = 0.33$ to $R^2 = 0.82$ in the pooled sample. Each listing increases the probability of a single sale by 3 percent in a given week, while a new listing increases the probability of a sale by as much as 60 percent. Next, I confirm the second prediction and show that hazard rates drop off significantly following the first two weeks of the listing. Notably, the data allows me to define a match at the time when the listing status changes to "in contract" rather than at the closing date. This is an important distinction, because the "in contract" status signifies a match between a buyer and a seller, while the closing date can come at a later date that is agreed upon by both parties.[3] I find that upon entry, each listing has about 18 percent chance of selling in the first two weeks while the average weekly hazard going forward is about 1.8 percent.[4] I confirm this result in a regression to make sure that the aggregate statistics do not pick up a mix of different markets, time periods, or list price strategies. I show that this pattern holds for each year observed in our data, but the hazard rate is "spikiest" in years 2005, 2019, and 2020 (precisely the housing boom years). Finally, I confirm the model prediction of price premiums for new listings and price discounts for listings with longer time on market. I first conduct the hedonic price prediction analysis. A residual from a hedonic regression of sale price is systematically positive and high for listings that sell right away, and low and negative for listings that sell later. New listings sell at about a three percent premium relative to the baseline price, while older listings sell at a discount of around one percent. I confirm this pattern by examining list prices. I show that new listings have a particularly high probability of selling above the asking price in the first two weeks (as high as 35 percent in the second week since listing) while for older listings this happens only about 10 percent of the time.

In addition to matching data patterns on prices and sales timing, a simple stock-flow model features endogenous multilateral matching. This endogenous variation in the number of trading partners for both sellers and buyers leads to cyclical bargaining power in price negotiations. It is an intuitive

[3] Anenberg and Laufer (2017) construct a house price index using the in-contract date rather than the closing date and find that this distinction alone helps explain several puzzles in the housing market, such as serial auto-correlation in prices.

[4] While 18 percent might seem low, in our data only 64 percent of listings sell at all. So among the sold properties, 28 percent sold within the first two weeks.

aspect of the housing market and is often discussed in the media[5]. When there are very few buyers and many homes on the market, a new listing is unlikely to generate interest from multiple buyers, and an old listing will compete with others for the attention of new buyers. On average, the Nash bargaining power for a seller in this market is close to zero. On the other hand, if there are few homes available for sale and many buyers looking, a new listing is likely to attract several buyers who will compete for it in a bidding war. In this market, the Nash bargaining power of a seller is close to one. This cyclical bargaining power can generate an additional amplifying force for pricing cycles.[6] Although it is not possible to test for the bargaining procedures in the listings data, the model predicts that early sales are more likely to end up in a bidding war. Thus, the "spikiness" of the hazard function can be a proxy for the seller's bargaining power. I find that the hazard function is particularly spiky in years 2005, 2019 and in 2020, consistent with these periods being described as "sellers' markets".[7] Endogenous multilateral matching also generates endogenous price dispersion.

The simple version of the stock-flow model presented in this paper is elegant and tractable, but it has some limitations. The simplicity of the equilibrium solution relies on symmetry of stock and flow between buyers and sellers. While restrictive, I argue that the symmetry assumption is naturally present in all steady-state analysis of housing search. A unique feature of this market is that a seller most often becomes a buyer after selling their house, so it is not unreasonable to assume that there will be the same number of individuals on each side. Furthermore, a symmetric model can be easily extended to a general equilibrium framework by modeling homeownership in addition to search. I propose a general equilibrium extension in Appendix A. Using this framework one can microfound changes in the equilibrium flows of sellers and buyers. For example, starting from 2022, many US households are "locked" into favorable mortgage terms and are reluctant to move. Both partial and general equilibrium setups can be, however, expanded to study asymmetric markets, where the stock of sellers is different from the stock of buyers. While the equilibrium solution will no longer be trivial, numerical methods can be used to compute the equilibrium.

A second limitation is that the stock-flow matching function does not feature constant returns to scale. This makes it difficult to introduce aggregate fluctuations and study, for example, price

[5]For example, one of the most prominent brokerages called Redfin keeps track of how many of their listings faced a bidding war and reports the results periodically in their blog: https://www.redfin.com/news/real-estate-bidding-wars-september-2021/

[6]Cyclicality of prices in random search is captured by the variations in outside option for sellers during hot and cold markets. If there are many/few buyers, a transaction will have a match surplus with a high/low a price commensurate with seller's attractive/unattractive outside option. However, in the set of possible prices, the cyclical Nash bargaining parameter will introduce further amplification of the pricing cycle.

[7]See more evidence in RedFin.com blog posts on bidding wars: https://www.redfin.com/news/bidding-wars/2/

momentum. However, the framework can still be used to study cyclicality. Like in Ngai and Tenreyro (2014), who analyze thick and thin markets by examining two equilibrium steady states, here too, one can look at markets with a large inventory and a small inventory.[8] In Ngai and Tenreyro (2014) the high price in hot markets is driven by buyers expecting a better match quality from every listing they sample. Here, the hot market is generated from a larger pool of counter parties to choose from for each new listing and a buyer, and thus a more likely successful match, higher turnover, and a higher likelihood of a bidding war. This naturally makes entering the market more attractive during the "peak season" for both buyers and sellers, reinforcing higher turnover. Seasonality in prices in this setting would arise naturally from the endogenous multilateral matching process.

In the last section of the paper, I discuss the stock-flow narrative in comparison to the random search one. In the data, older listings sell at a lower price and with lower probability while newer listings sell at higher prices with higher probability. Both can not be achieved in a simple random search model with homogeneous homes. This is because the only mechanism for attracting buyers (and thus increasing sale probability) is by lowering the price. The only way to modify a random search model to match both facts of the data is to include adverse selection. If “more attractive” homes sell for a premium and are more likely to find a match, then the older listings are by selection more likely to be “less attractive" and thus sell at a discount.[9][10] To test whether the stock-flow narrative is relevant in the data beyond the adverse selection story, I turn to examining a common phenomenon in the housing market - relists. Often sellers would pull their house from the market for seemingly no reason and then re-list the property at a later time. The adverse selection story would predict that this has no effect on the subsequent probability of sale, because the underlying property stays the same. The stock-flow model, however, predicts that a returned listing is more likely to sell upon re-entry because it faces an increased pool of new buyers who have entered during the period where the listing was off-market. I find that indeed the hazard rates of re-listed properties increase relative to homes that were never taken off the market. This corroborates the stock-flow story as a likely mechanism in the housing sale process.

[8] A related concept in finance is "the depth of the market", defined as the number of buy and sell orders outstanding for a given asset. This measure is used to quantify liquidity.

[9] Albrecht et al. (2007) introduced heterogeneity on the seller side instead. The paper models increasing impatience in sellers who get a random Poisson shock that increases their cost of holding on to the property. In this setting, prices for older listings might be lower to capture seller’s willingness to sell sooner. However, the model predicts increasing hazard rates, as buyers are more attracted to cheaper (older) listings, while newer (more expensive) listings sell with a lower probability. This is not what I find empirically.

[10] Naturally, with the extent of heterogeneity in housing, there is role for adverse selection in both stock-flow and random search narratives.

This paper contributes to an extensive literature that incorporates search frictions in models of housing to explain the co-movements and aggregate trends of house prices and sales, as well as the interplay between the for-sale and rental market. Han and Strange (2015) gives an overview of this literature where the vast majority of papers rely on the random search framework. The most recent paper examining sources of search and matching frictions in the housing market is Badarinza et al. (2024). This paper uses extensive data on buyer housing search in the U.K. to disentangle the importa of each friction in the buyer and seller matching process. It finds that close to half of the frictions comes from mismatch between the seller and the buyer. However, in this paper, the interpretation of mismatch occurs within the random search framework. In other words, a buyer randomly visits properties and then decides whether or not to buy based on their realized match quality. In contrast, in the stock-flow approach match quality is realized for *all* properties without the intermediate step of random meetings between buyers and sellers.

Although novel in the housing market space, a stock-flow matching function was introduced in the labor market setting by Coles and Smith (1998)[11]. There are a select few papers that incorporate stock-flow approach to study the impact of auctions on the aggregate price dynamics in housing. For example, Smith (2020) embeds the stock-flow matching into an auction model and adds endogenous seller entry to rationalize hot and cold markets.[12] A more recent working paper Eric Smith et al. (2022) explores how aggregate shocks affect prices, sales, and inventory using a simulation of the Smith (2020) model. In contrast to these auction models, this paper proposes a simpler price mechanism (allowing for three possible levels of pricing for each transaction) and lends itself to a tractable way to expand on the stock-flow framework.[13] For example, in Appendix A I offer a tractable way to embed the stock-flow matching into a general equilibrium framework.

The rest of the paper proceeds as follows. Section 2 describes the model and the equilibrium dynamics. In this section, I outline the predictions of the model that contrast the standard search model used in the literature. Section 3 describes the data and tests the predictions of the model. Section 4 discusses useful features and limitations of the stock-flow model. Section 6 concludes.

[11] Coles and Smith (1998) explores job finding rates for people of different unemployment periods, but does not incorporate prices (wages) in the model or the empirical analysis

[12] In Smith (2020) sellers enter when they expect multiple buyers on the market who could potentially enter the bidding war. Once a number of sales have been made, sellers expect that few buyers remain and wait for a period of time to "replenish" the buyer stock before listing their home.

[13] While within an individual transaction, the model only allows for three levels of pricing, the probability of each level being realized is an endogenous variable. This allows for the set of possible average prices in the market to have continuous support

2 Model

This section describes a simple housing search stock flow model, outlines an equilibrium solution for a special case of symmetrical parameter values, and presents testable predictions to be examined in the data.

2.1 Model Setup

There are two agents in the model: sellers who own one stock of an indivisible house for sale, and buyers who are willing to buy one house. Each seller's valuation for the house is normalized to 0. A buyer's valuation for a particular house is a random variable that follows a binomial distribution. With probability α it is 0, and with a complimentary probability 1-α it is π. Time is continuous. Sellers and buyers arrive on the market with Poisson rates s and b respectively by paying the entry costs c_s for the seller and c_b for the buyer. There are no meeting fictions in the market, meaning that if there are gains from trade between a buyer and seller pair, the match occurs immediately. It follows that all the trades happen with exactly one of the counterparties being new to the market. This is because if a buyer and seller were both on the market for some time, the trade would have occurred previously. I also assume that it is a probability zero event that a buyer and a seller show up on the market simultaneously and thus it is impossible for two new agents to trade. Henceforth a buyer or a seller will be called "new" in the instant that they enter the market, and "old" otherwise.

Let S and B be the current stocks of old sellers and old buyers in the market. The first type of matching occurs when a new buyer finds at least one of the S houses attractive. This happens with probability $1-\alpha^S$, a complimentary probability to that of drawing valuation 0 for all S homes. Then $b(1-\alpha^S)$ is the arrival rate of matches made between new buyers and old sellers. Each old seller is expected to have such a match at a rate $\frac{b(1-\alpha^S)}{S}$. Similarly, $s(1-\alpha^B)$ is the arrival rate of the matches between new sellers and old buyers. New sellers expect to match with probability $1-\alpha^B$ and each old buyer matches at a rate $\frac{s(1-\alpha^B)}{B}$.

When a new agent enters the market, there is a chance that there are more than one successful potential matches with the current stock of counterparties. If there are two or more willing buyers for a new seller, the transaction price will be bid up to the buyer's reservation price so that the new seller will extract all of the surplus from the transaction. This happens with probability $1-\alpha^B-B\alpha^{B-1}(1-\alpha)$, or the complimentary probability to either none or exactly one successful match. Similarly, if a new buyer is interested in more than one house, they will bargain the price down to the seller's reservation value and will extract all surplus from the transaction by buying at a lower price. This happens with probability $1-\alpha^S - S\alpha^{S-1}(1-\alpha)$. If, on the other hand, a new agent matches with exactly one counterparty, the price is determined via Nash bargaining with buyers having bargaining power γ.

There are four types of agents in this model: old and new buyers and sellers. I will look for a stationary equilibrium where the value functions for each type do not change over time and three equilibrium prices are $p_H > p_N > p_L$ corresponding to the new buyer getting all of the surplus, sharing of the surplus via Nash bargaining and new seller getting all of the surplus. Let $V_{at}(\boldsymbol{p};\alpha,\pi,\gamma,\rho,s,b,S,B)$ be the value of agent $a \in s,b$, a seller or a buyer, who is $t \in n,o$ new or old on the market. For simplicity of exposition, I will subsume the parameters of the model $\alpha,\pi,\gamma,\rho,s,b,S,B$. To simplify the equations, define $\gamma_S = S\frac{\alpha^{S-1}(1-\alpha)}{1-\alpha^S}$ and $\gamma_B = B\frac{\alpha^{B-1}(1-\alpha)}{1-\alpha^B}$.

The value functions of each agent type are as follows:

$$
\begin{aligned}
V_{Bn} &= -c_b + (1-\alpha^S)(\pi - \gamma_S p_N - (1-\gamma_S)p_L) + \alpha^S V_{Bo} \\
\rho V_{Bo} &= \frac{s(1-\alpha^B)}{B}(\pi - (1-\gamma_B)p_H - \gamma_B p_N - V_{Bo}) \\
V_{Sn} &= -c_s + (1-\alpha^B)((1-\gamma_B)p_H + \gamma_B p_N) + \alpha^B V_{So} \\
\rho V_{So} &= \frac{b(1-\alpha^S)}{S}(\gamma_S p_N + (1-\gamma_S)p_L - V_{So})
\end{aligned}
$$

A new buyer has to pay the enter fee c_b and has a probability $1-\alpha^S$ of finding a match. If a match is found, they get π value and have to pay one of two prices: with probability γ_S they only like one house and bargain with it's seller for p_N; with a complimentary probability $1-\gamma_S$, they like at least one more house and will hold all of the negotiating power over the sellers. This will result in a low price p_L. An old buyer improves their value only upon purchasing a new listing. This happens with arrival rate $\frac{s(1-\alpha^B)}{B}$. If they buy, they derive utility π. If they are the only bidder for the house, they will be able to get a bargaining price p_N. Otherwise, they will participate in a bidding war and purchase the house at a high price p_H. Value functions for the new and old sellers are symmetrical to those of new and old buyers respectively.

Recall that p_L is the resulting price from negotiations where a seller gives all of their surplus away. It is thus equal to seller's continuation value V_{So}. Similarly, $p_H = \pi - V_{Bo}$, is buyer's indifference price. I can then rewrite the equations as follows:

$$V_{Bn} = -c_b + (1-\alpha^S)(\pi - (1-\gamma_S)V_{So} - \gamma_S p_N) + \alpha^S V_{Bo}$$
$$\rho V_{Bo} = \frac{s(1-\alpha^B)}{B}\gamma_B(\pi - p_N - V_{Bo})$$
$$V_{Sn} = -c_s + (1-\alpha^B)((1-\gamma_B)(\pi - V_{Bo}) + \gamma_B p_N) + \alpha^B V_{So}$$
$$\rho V_{So} = \frac{b(1-\alpha^S)}{S}\gamma_S(p_N - V_{So})$$

I assume that buyers have options to buy in a different sector or rent, so that there is a free entry condition, or $V_{Bn} = 0$. Then the cost of entry for buyers has to be equal to the benefit:

$$c_b = (1-\alpha^S)(\pi - (1-\gamma_S)V_{So} - \gamma_S p_N) + \alpha^S V_{Bo}$$
$$V_{Bo} = \frac{c_b - (1-\alpha^S)(\pi - (1-\gamma_S)V_{So} - \gamma_S p_N)}{\alpha^S}$$

2.2 Equilibrium

In an stationary equilibrium it must be that $b = s$. This is because the outflow of agents from the search market occurs in pairs, once the match happens. If $b \neq s$, then the inflow will not be equal to the outflow for at least one type of agents (sellers or buyers). Equating the inflow and outflow for each type I get:

$$b = s = b(1-\alpha^S) + s(1-\alpha^B)$$

It follows that

$$\alpha^S + \alpha^B = 1$$

I now determine the equilibrium prices and value functions. Let I be the total surplus from a match between a seller and a buyer. Note that the reservation value for a new and an old agent is the same, because if there is no trade both would become old in the next instant.

$$
\begin{aligned}
\gamma I &= \pi - p_N - V_{Bo} \\
(1-\gamma)I &= p_N - V_{So} \\
I &= \underbrace{\pi - p_N - V_{Bo}}_{buyer\ surplus} + \underbrace{p_N - V_{So}}_{seller\ surplus} \\
&= \pi - V_{Bo} - V_{So}
\end{aligned}
$$

Multiplying both sides by ρ and plugging in the value functions I get:

$$
\rho I = \rho\pi - \frac{s(1-\alpha^B)}{B}\gamma_B(\pi - p_N - V_{Bo}) - \frac{b(1-\alpha^S)}{S}\gamma_S(p_N - V_{So})
$$

At this point I simplify the problem by assuming that $B = S$, so there is symmetry between the two agents. Then $\gamma_S = \gamma_B$. In this case,

$$
\begin{aligned}
\rho I &= \rho\pi - \frac{b(1-\alpha^S)}{S}\gamma_S(\pi - V_{Bo} - V_{So}) \\
&= \rho\pi - \frac{b(1-\alpha^S)}{S}\gamma_S I \\
I\left(\rho + \frac{b(1-\alpha^S)}{S}\gamma_S\right) &= \rho\pi \\
I &= \frac{\rho}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\pi
\end{aligned}
$$

I can now solve for the Nash bargaining price p_N. Here I will need the free entry condition of the buyers. This will give me two equations for each surplus share and two unknowns: p_N and V_{So}.

$$
\begin{aligned}
\gamma I &= \pi - p_N - V_{Bo} \\
V_{Bo} &= \pi - p_N - \gamma\frac{\rho}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\pi = \frac{c_b - (1-\alpha^S)(\pi - (1-\gamma_S)V_{So} - \gamma_S p_N)}{\alpha^S}
\end{aligned}
$$

$$
\pi - \gamma\frac{\rho}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\pi - \frac{c_b}{\alpha^S} + \frac{(1-\alpha^S)\pi}{\alpha^S} - \frac{(1-\alpha^S)(1-\gamma_S)V_{So}}{\alpha^S} = \frac{(1-\alpha^S)\gamma_S p_N}{\alpha^S} + p_N
$$

Simplifying yields:

$$p_N = \frac{1}{\alpha^S + (1-\alpha^S)\gamma_S}\left(\pi - \gamma\alpha^S \frac{\rho}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\pi - c_b - (1-\alpha^S)(1-\gamma_S)V_{So}\right)$$

The second surplus equation defines the other relationship between p_N and V_{So}. Plugging in the p_N relationship above and the equation for I, I solve for V_{So} which is equal to the reservation price of the seller, p_L:

$$V_{So} = p_N - (1-\gamma)I = \pi - c_b - \frac{\rho\alpha^S + \rho(1-\gamma)(1-\alpha^S)\gamma_S}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\pi = p_L \tag{1}$$

Next I can plug back the equation for V_{So} to find p_N:

$$p_N = \pi - c_b - \frac{\rho\alpha^S - \rho(1-\gamma)(1 - S\alpha^{S-1}(1-\alpha))}{\rho + b\alpha^{S-1}(1-\alpha)}\pi \tag{2}$$

Now I can solve for V_{Bo}:

$$V_{Bo} = \pi - p_N - \gamma I = c_b - \frac{\rho(1-\alpha^S)(1-\gamma_S + \gamma\gamma_S)}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\pi \tag{3}$$

Which defines the reservation price for buyers at:

$$p_H = \pi - V_{Bo} = \frac{1 + \rho(1-\alpha^S)(1-\gamma_S + \gamma\gamma_S)}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\pi - c_b \tag{4}$$

Finally, I can find the value of the entering seller to be:

$$V_{Sn} = \pi\left(1 + \frac{(1-\alpha^S)(1-\gamma_S) - \alpha^S}{\rho + \frac{b(1-\alpha^S)}{S}\gamma_S}\rho\right) - c_s - c_b \tag{5}$$

This characterizes the equilibrium for the simplified model where $S = B$:

Definition Given the model primitives $\alpha, \pi, \gamma, \rho, s, b, S, B$, that satisfy $S = B$, $b = s$, and $\alpha^B + \alpha^S = 1$, a **stationary equilibrium** is a set of value functions $\{V_{So}, V_{Sn}, V_{Bn}\}$ taking values corresponding to equations 1,5,3; prices $\{p_N, p_L, p_H\}$ corresponding to equations 2, 1, 4; a matching function between buyers and sellers, and a free entry condition of the new buyers $V_{Bn} = 0$. Each period,

$$s(1-\alpha^B) + b(1-\alpha^S) \tag{6}$$

matches are made between sellers and buyers. Conditional on a new listing matching with an older buyer, there is a $\frac{1-\alpha^B - B\alpha^{B-1}(1-\alpha)}{1-\alpha^B}$ chance the price will be p_N, while with the complimentary probability $\frac{B\alpha^{B-1}(1-\alpha)}{1-\alpha^B}$ the price will be bid up to p_H. Conditional on a new buyer matching with an older listing, there is a $\frac{1-\alpha^S - S\alpha^{S-1}(1-\alpha)}{1-\alpha^S}$ chance the price will be p_N, while with the complimentary probability $\frac{S\alpha^{S-1}(1-\alpha)}{1-\alpha^S}$ the price will be discounted to p_L.

In Appendix A, I expand the model to a general equilibrium framework by adding homeowners. For a fix housing stock, the homes flow from being well matched with a homeowner to becoming mismatched and listed for sale. This model has a closed form solution and can be useful to study the response of housing markets to aggregate shocks.

2.3 Model Predictions

The search model with embedded stock-flow has several testable predictions that are at odds with a simple random matching model.

Model Prediction 1: The number of total matches increases with turnover of sellers and buyers independent of the current stock of buyers and sellers.

In the model each match necessarily has a new agent involved - either a new buyer or a new seller. This means that the turnover plays a large role in whether matches are made.

Model Prediction 2: Hazard rate sale is highest for new listings and decreases sharply after a short period of time.

Because in the model new listings are exposed to all the current buyers and do not compete with the existing inventory of homes, they have a higher probability of selling. An older listing, in contrast, awaits new buyers and competes for each of them with the entire inventory of homes for sale.

Model Prediction 3: Prices are higher for new listings and lower for old listings.[14]

The model generates three price levels. For new listings the possible prices are p_N and p_H. The premium price p_H is a result or a bidding war for a new listing. If there is exactly one interested buyer

[14] Another prediction is that the fraction of old listings selling at a discount increases with listings' inventory. Similarly, the fraction of new listings selling at a premium increases with the number of buyers searching for a house. It's difficult to show this empirically, since prices are usually evaluated relative to an average price in the market which is endogenous.

the price is negotiated to a Nash bargaining level $p_N < p_H$. For older listings, the two possible price outcomes are p_N and p_L. Either a new buyer likes the house and nothing else and can negotiate to a middle price p_N, or the new buyer likes another house just as much, has a stronger negotiating power and can obtain a discounted price p_L. It follows that the weighted average sale price or new listings is necessarily higher or equal to the average sale price of old listings.

3 Data Analysis

To evaluate the validity of the housing search stock flow model, I use listing-level data from the Bright Multiple Listing Service (MLS) platform, collected by CoreLogic. In this section, I describe the dataset and proceed with empirical tests of each of the three model predictions.

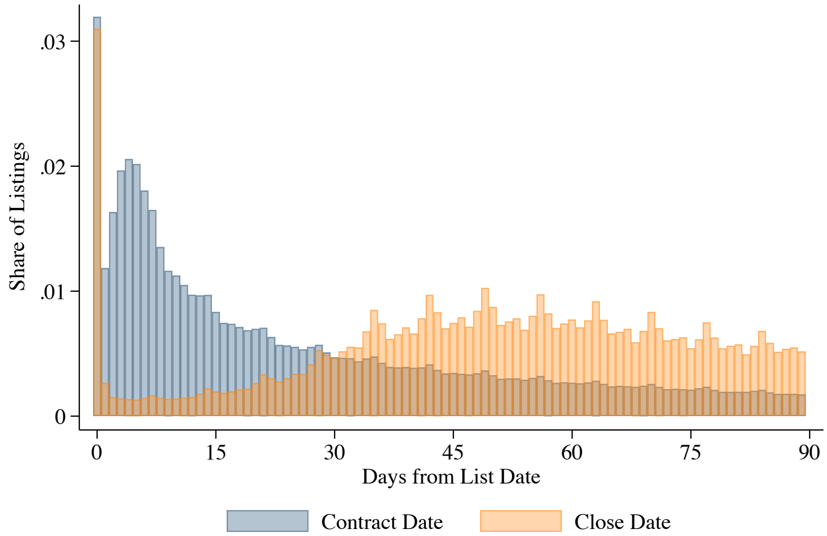
3.1 Data

One particularity of the U.S. housing market is that the vast majority of sellers use a real estate agent to market their home.[15] Bright is one of the biggest MLS platforms in the nation and covers all listings that had real estate agent representation in its geographical coverage. The Mid-Atlantic region that it covers includes Maryland, Delaware, and large parts of New Jersey, Pennsylvania, Virginia, and West Virginia. Although Bright was created in 2017 following a merger of several MLS platforms, the data include listings dating back to the 1990s collected from pre-merger MLS platforms. Since data becomes more populated over time, the analysis is restricted to years 2005 to 2020. Each observation represents a listing and includes very detailed characteristics of the property as well as the status of the listing. Most importantly, these data provide a date when a listing is marked "in contract". This allows me to accurately measure the time a listing spends on the market before meeting a successful buyer, rather than relying on a close date to infer the timing of the match. Typically, buyers enter the contractual agreement as a commitment to buy, barring any unforeseen inspection results or financing constraints.[16] While some buyers waive these contingencies, for most, it means a delay in closing at least until the completion of the inspection and the appraisal. In some cases, the close date gets pushed back even later to accommodate sellers' or buyers' relocation needs. In the data I'm using, the average time between the "in-contract" and "closed" status changes is 50 days. The contrast between two measures is apparent in Figure 1, which plots the distribution of time on the market as measured by days to "in contract" and days to close. In total, the data consists of 2,980,721 listings.

[15] National Association of Relators (2023) found that 89 percent of home sellers worked with a real estate agent.

[16] The price negotiation process often continues on until the sale date, as buyers can ask for concessions related to unexpected appraisal and inspection results

Figure 1: Importance of Contract Timing



Note: This figure plots the distribution of time to sell measured in two different ways. The blue histogram reflects the time passed from the list date until the listing goes into contract. The orange histogram plots the distribution of days between listing and closing.

3.2 Matches and Turnover

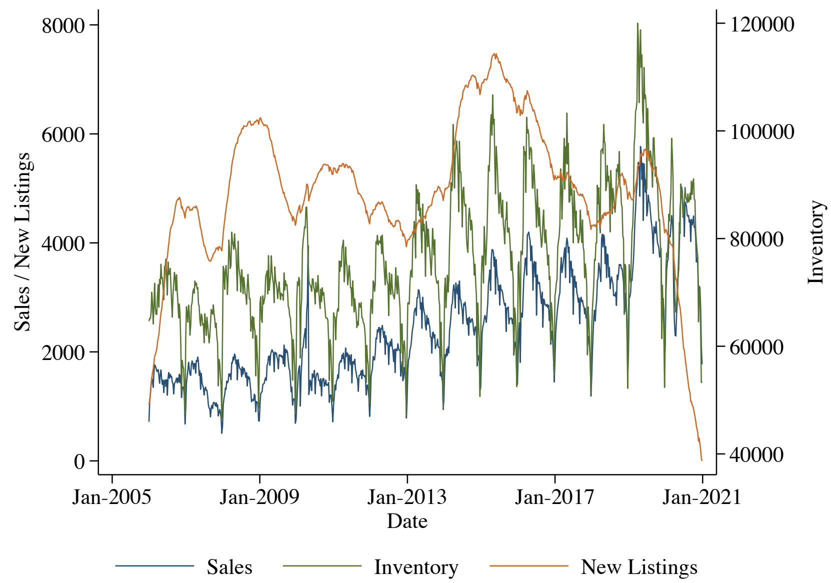
I now examine the first prediction from the stock-flow model that turnover plays a big role in generating matches, even conditional on total inventory of buyers and sellers. Although the data do not have any information about the buyer search process, I do have detailed information on when listings go on the market and when matches are formed. Figure 2 plots the number of sales contracts, new listings, and current inventory of listings over time. The data is aggregated at the weekly level, where the inventory is calculated by the number of listings that have entered the market prior to the corresponding week and have not been sold or removed from the market at the beginning of that week. Sales contracts reflect the number of properties that entered a contract in a given week.[17]

The listings inventory and sales contracts do not seem to co-move. However, sales contracts and new listings clearly follow the same cyclical pattern. Although the plot is suggestive, it shows aggregated data series from many different markets. To formally examine the link between new listings and sales contracts, I next turn to market-level data.

The stock flow model of the housing search assumes that all buyers and sellers are potential counter parties, so when confirming the model predictions, it is important to look at within-market

[17]Note that I only know this information for listings that eventually sold. Failed contracts are not part of the data set. RedFin.com estimates that around 15% of the contracts eventually fail in 2024: https://www.redfin.com/news/home-purchase-cancellations-june-2024/.

Figure 2: Sales Contracts, Inventory, and New Listings



Note: This figure plots the aggregate number of sales contracts, new listings and the current inventory for the entire sample of the Bright MLS data over time. The data is aggregated at the weekly level. Inventory is calculated by the number of listings that have entered the market prior to the corresponding week and have not be sold or take off-them market by the beginning of that week. Sales contracts is the number of listings that enter into a contract with a buyer within those seven days, new listings is the total number of listings that entered the MLS database in the seven day period.

relationships between aggregate variables. This presents a new challenge of defining a market. For this exercise, I define each market by a three-digit zip code and a tercile of square footage within that zip code.[18] I drop observations that are below the 1st and above the 99th percentiles in either square footage or list prices to avoid extreme outliers and entries with typos and exclude markets with fewer than 10 observations. In total, there are 105 unique zip codes and 316 markets. For each market, I aggregate the data weekly and run the following regression:

$$Sales_{m,t} = \beta_1 Inv_{m,t} + \beta_2 NewListings_{m,t} + \alpha_t + \theta_m + \epsilon_{m,t}. \tag{7}$$

Here $Sales_{m,t}$, $Inv_{m,t}$, and $NewListings_{m,t}$ are sales contracts, total inventory of for sale properties, and new listings in week t in market m. In some regression specifications, I include fixed effects for time α_t and markets (θ_m). $\epsilon_{m,t}$ represents the error term.

Table 1 displays the results. In panel A, I first analyze the aggregate time series plotted in Figure 2 and then, in Panel B, I use a panel of individual markets' listings, sales, and inventories. In columns (1) and (4) I regress the number of sales on the total inventory of listings. In columns (2) and (5) I regress

[18]The results do not differ significantly for coarser definitions of a housing market.

the total sales on the number of new listings. Both variables have a statistically significant coefficient; however, R^2 from the new listing regressions is greater than 0.8, while that of the total inventory is significantly lower (0.33 for the pooled sample and 0.6 for the market panel series). Moreover, in both samples an increase in total inventory by one listing leads to only around 0.03 increase in sales, while a new listing increases the sales number by around 0.6 in a given week. Next, I combine the two explanatory variables in one regression in columns (3) and (6). Here, the coefficient on the total inventory, while still significant, decreases substantially (and even becomes slightly negative), while the coefficient on new listings stays around 0.6 in both samples. In the market-level data, I run additional specifications with controls for the time effects and market effects. I add them sequentially in columns (7) and (8). Controlling for time does not change the coefficients much. Controlling for both market and time effect leads to a smaller coefficient on new listings, however it remains significantly larger than the coefficient on total inventory. Thus, even within a market and conditional on overall time effects, a relative increase in new listings leads to more sales.

Table 1: Sales Contracts, Inventory, and New Listings

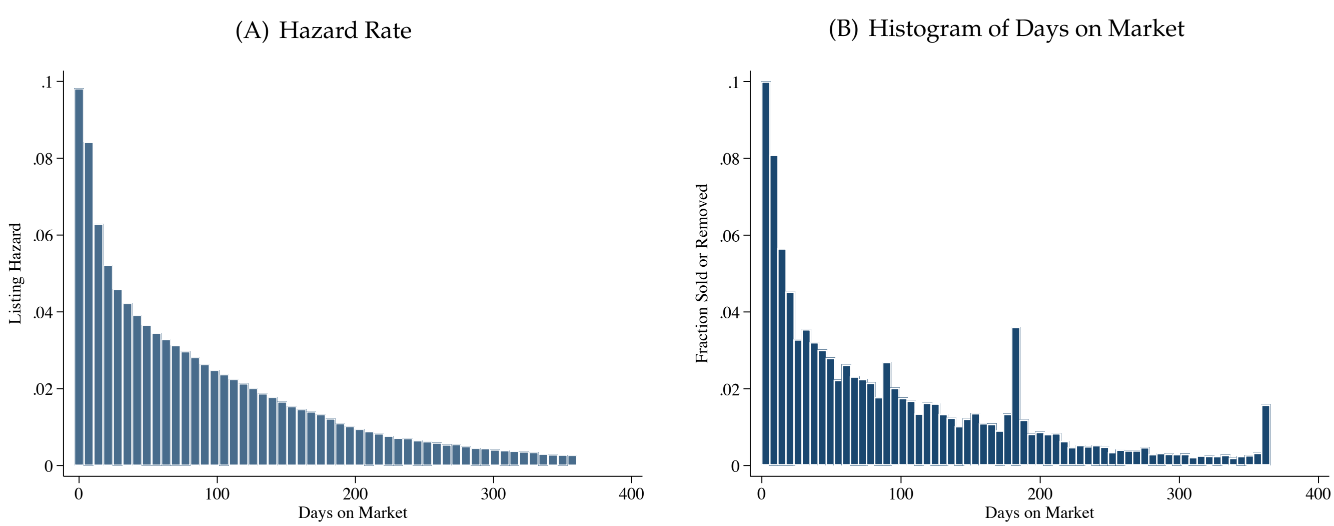
	Panel A: Pooled Data			Panel B: Market Level				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Inventory	0.0288***		0.0088***	0.0346***		-0.0006**	0.0007**	0.0057***
	(0.0014)		(0.0010)	(0.0002)		(0.0003)	(0.0003)	(0.0005)
New Listings		0.6603***	0.5966***		0.5679***	0.5750***	0.5582***	0.3738***
		(0.0151)	(0.0159)		(0.0025)	(0.0043)	(0.0045)	(0.0051)
R^2	0.3296	0.8011	0.8242	0.6024	0.8231	0.8231	0.8515	0.8878
Time Effect	No	No	No	No	No	No	Yes	Yes
Market Effect	-	-	-	No	No	No	No	Yes
N	520	520	520	64108	64108	64108	64108	64108

Note: This table reports coefficients on a set of regressions where the dependent variable is the number of sales contracts, and dependent variables are the total inventory and new listings. Each observation pools data for 7 consecutive days. Sales contracts is the sum of all for-sale contracts within those seven days, new listings is the total number of new listings that entered the MLS database in the 7 day period, and inventory is the number of listings that are unsold by the beginning of the 7 day period and have been on the market for fewer than 365 days. Panel A reports regression results for the pooled sample where an observation is a week, while Panel B uses market level data where an observation is at the market week level. A market is defined by the three digit zip code and the tercile of the for-sale property in terms of square footage within that zip code. In column (7) and (8) I add week fixed effect and column (8) additionally includes market fixed effects.

3.3 Declining Hazard of Sale

The second prediction of the model is a decreasing hazard rate of sale. Figure 3(A) plots the hazard rate of sale over time starting from the first week of the listing. I compute the total inventory of listings for each market duration and report the fraction of them that sell within the following week. About 9.8 percent of new listings sell within a week. The hazard rate declines to 8.4 percent in week two and to 6.2 percent in week three. The more time a listing spends on the market, the lower its chances of selling. In Figure 3(B) plots the histogram of days on market for all listings. Days on market are defined for both sold and removed listings, although typically listings do not get removed much earlier than 180 or 365 days of unsuccessful marketing. I find that about 18 percent of listings sell within two weeks of being listed. Although the hazard rate does not plateau immediately after the initial period, the highest probability of sale is when a listing is brand new. This pattern is consistent with the stock flow model of housing search; however, the aggregate plot can be confounding several factors. The decreasing hazard rate post the initial listing period could be due to selection in the case of buyer preferences being correlated - more attractive properties sell faster, while less attractive take a while to find the right buyer or do not sell at all. Although we can control for unobservable differences, we can see if location or property characteristics play a role in the decreasing hazard pattern. Similarly, the aggregated plot could mask different patterns in many markets. For example, if the hazard rate was flat at different levels for different locations, then putting the data all together could also deliver the declining hazard rate. Finally, this plot ignores the importance of the list prices. Perhaps homeowners have heterogeneity in how flexible they are in adjust their price down and homes that are priced higher stay on the market longer until a willing buyer arrives.

Figure 3: Decreasing Probability of Sale



Note: This figure examines the time on the market for homes listed for sale. A bin corresponds to weeks since the original listing. The first plot shows listing hazard rates - the probability of sale in the corresponding week conditional on being in the market. The second plot is a histogram of days on market for all listings. Included here are listings that do not sell and go off-market.

To formally test for the decreasing hazard rate pattern in the data, I create a listing - week level dataset for weeks since list date. Each listing has 52 observations that can take value 1 is the listing went into contract in the corresponding week, 0 if it has not gone into contract but is still on the market, and "missing" otherwise. I then run the following regression:

$$Hazard_{i,m,t} = \beta FirstWeeks_{i,t} + \delta_1 Weeks_{i,t} + \delta_2 Weeks^2_{i,t} + \delta_3 Weeks^3_{i,t} + \mu \mathbf{W_i} + \alpha_t + \theta_m + \epsilon_{m,t}. \quad (8)$$

Here $Hazard_{i,m,t}$ is an indicator for a contract for listing i in week t since listing in market m; $FirstWeeks_{i,t}$ is an indicator of whether the contract happened in the first two weeks since listing; $Weeks_{i,t}$, $Weeks^2_{i,t}$, and $Weeks^3_{i,t}$ are linear, square and cubic values for the week since listing; $\mathbf{W_i}$ is a vector of listing characteristics; α_t are list month fixed effect and θ_m are fixed effects for markets; finally, $\epsilon_{m,t}$ is the error term. Table 2 shows the results. All specifications include controls for property characteristics. Specification (2) adds the list month effect, specification (3) includes the market effect and specification (4) includes both the list month and the market effect. Even with a third degree polynomial fit of weeks on market predicting the hazard rate, the indicator for whether the contract occurred in the first two weeks has a significant coefficient of over 5 percentage points across all of the specifications.

The coefficient in $FirstWeeks_{i,t}$ can be viewed as the "excess mass" between the true hazard rate of the sale contract and the predicted one. It measures the "spikiness" of the hazard function that is consistent with the stock flow model of the housing search. Through the lens of the model, this excess mass represents the number of latent buyers in the market waiting for new properties to come in. The more old buyers are in the market, the more likely at least one of them will match with a newly listed house. To illustrate the "spikiness" of the hazard function, Figure 4(A) plots the average hazard rate of a sale contract against days on the market together with the cubic spline function that has been fitted to the hazard data for listings contracting 14 or more days after listing. Figure 4(B) then plots the difference between the two functions for different days on market. Unsurprisingly, for listings that go into contract after two weeks of being listed, the difference is negligent by construction. However the fit is particularly bad in the first 14 days, often under-predicting the hazard by over 50 percent.

Figure 5(A) repeats the exercise for a select number of years in the data, and Figure 5(B) displays the excess mass calculated as the sum of the difference between the hazard function and the fitted

Table 2: Contract Hazard

	(1)	(2)	(3)	(4)
First Two Weeks	0.05***	0.05***	0.05***	0.05***
	(0.00)	(0.00)	(0.00)	(0.00)
Weeks	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Weeks2	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Weeks3	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Bathrooms	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Bedrooms	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Area	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Area2	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Log(List Price)	0.01***	0.00***	0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.041	0.044	0.043	0.046
Time Effect	No	Yes	No	No
Market Effect	No	No	Yes	Yes
N	70341762	70341762	70341762	70341762

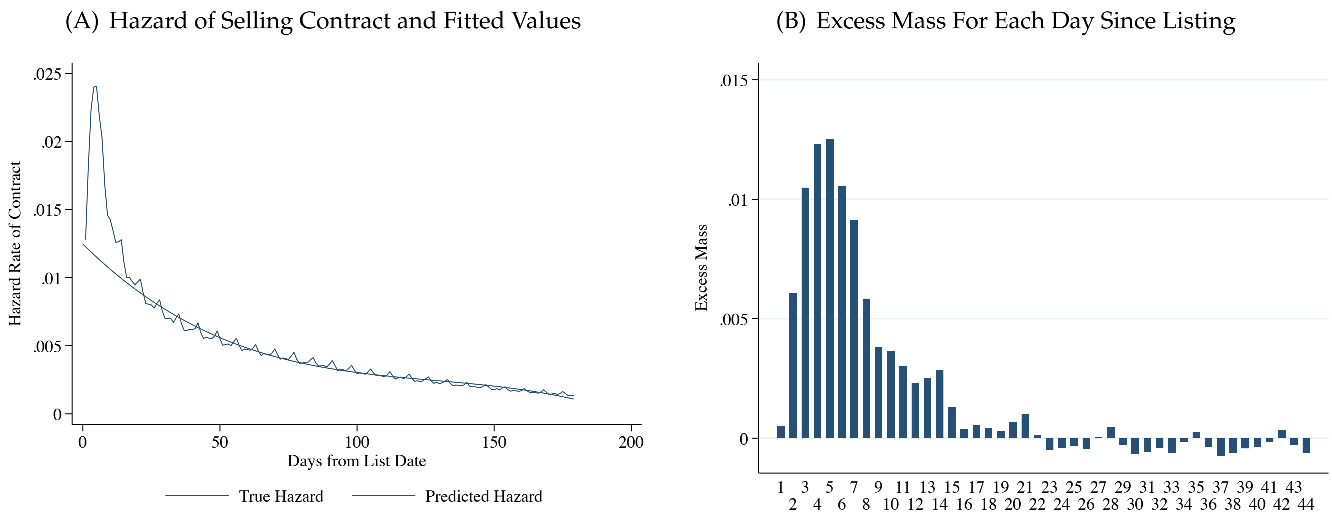
Note: This table show the regression of the hazard rate of sale contract on the timing since listing. The coefficient first a third degree polynomial as well as an indicator of whether the contract happened in the first two weeks of the listing. All specifications include controls for property characteristics. Specification (2) adds the list month effect, specification (3) includes the market effect and specification (4) includes both the list month and the market effect. Market is identified by the three digit zip code and the tercile of square footage within the zip code.

function during the first two weeks since listing. The hazard function of sale contracts is smoother in some years than in others. In the period of the housing crash, there is almost no excess mass. However, years 2020 and 2005 stand out as particularly "spiky". This is supported by the narrative in the media that buyers were difficult to come by in the aftermath of the housing crash, and, on the contrary, there were not enough homes for sale to meet demand during the boom of 2005 and 2020.

3.4 Declining Sales Price over Time

Next, I examine the third prediction of the model that homes sold in the initial period after being listed often sell for a premium. In contrast, properties that sell after being on the market for a while tend to sell at a discount. I confirm this in two ways. First, I compute the residual portion of the price

Figure 4: Fitted Hazard Function: Excess Mass



Note: Figure A plots the average hazard rate of a sale contract against days on the market together with the cubic spline function that has been fitted to the hazard data for listings contracting 14 or more days after listing. Figure B plots the difference between the two functions on each day post listing.

unexplained by observable characteristics and plot this residual against the days on market. Second, I proxy the "bidding war" by whether a house sold above the list price and show the frequency of "above ask" sales for different days on market.

In the residual analysis, the following regression is run on the sample of sold properties:

$$SalePrice_{i,t} = \delta W_{i,t} + \alpha_{i,t} + \epsilon_{i,t}, \tag{9}$$

where $SalePrice_{i,t}$ is the sale price for listing i in time t, $W_{i,t}$ is a vector of property-specific controls: total bathrooms, bedroom, area in squared feet and the square of the area, and $\alpha_{i,t}$ denotes sale month by zip code fixed effect. I predict the sale price for each listing and take the residual between the predicted price and the true sale price. I then scale the residual by the sale price to get the relative discount or premium. Figure 6(A) plots scaled residual described above across days on market. Newly listed homes that sell right away have on average a 2.8 percent premium over the predicted sale price. While listings that spend more time on the market on average sell at a discount.

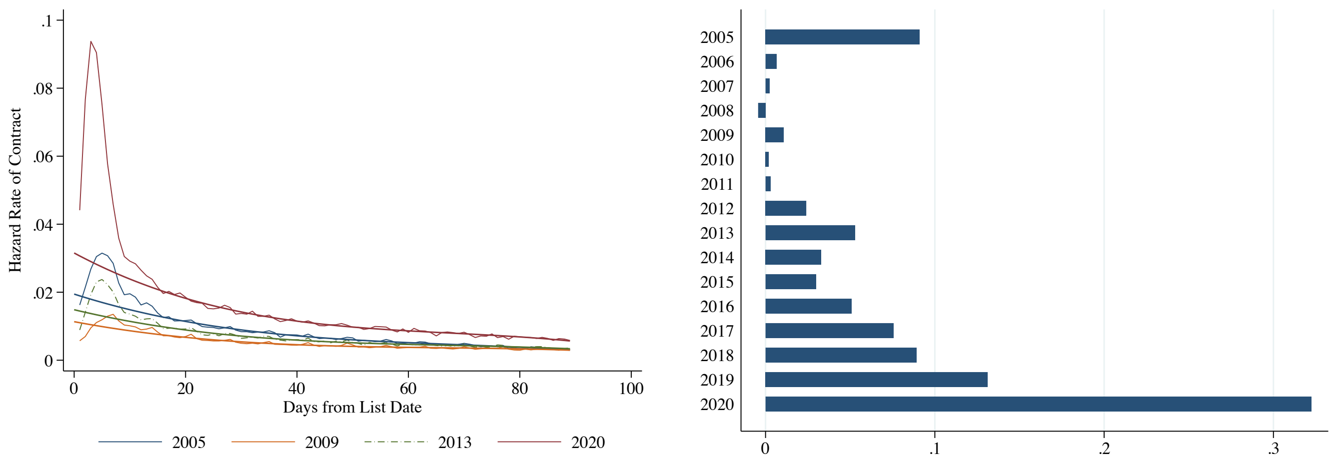
My model does not incorporate the role of list prices. However, the price escalation that occurs when multiple buyers are interested in a listing could be thought of as a bidding war, and I can identify some bidding wars by looking at whether the house sold above the list price[19] Figure 6(B) plots the

[19]This is a subset of occasions where there are multiple willing buyers, because it includes only those bidding wars where buyers are willing to pay *above* the list price.

Figure 5: Fitted Hazard Function: Excess Mass by Year

(A) Hazard of Selling Contract and Fitted Values

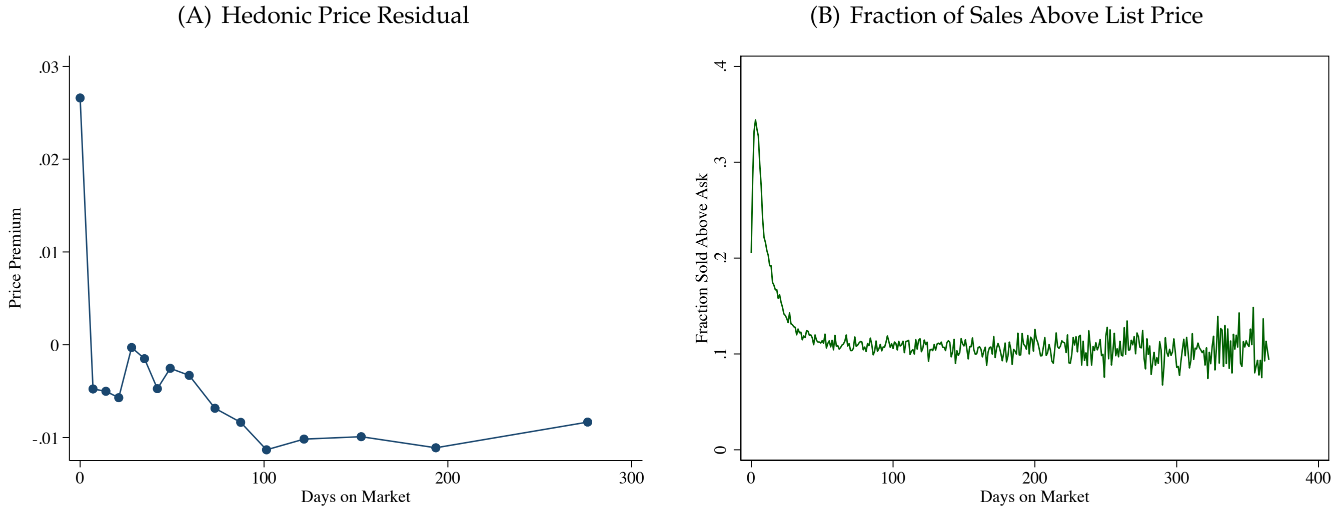
(B) Excess Mass in the First 14 Days



Note: Figure A plots the average hazard rate of a sale contract against days on the market together with the cubic spline function that has been fitted to the hazard data for listings contracting 14 or more days after listing. Each plot represents the hazard function for a particular year and the fitted function represents a function that has been fitted to the data of that year. Figure B plots the sum of the excess mass (the difference between the two functions) for the first 14 days since listing.

fraction of sold listings that sell above list price. In the beginning period, as much as a third of all listings sells above list price. This fraction plateaus at about 10 percent for listings sold in their third month or later. Consequently, consistent with the model, the price premiums are likely to occur in the beginning of the listing's life and less likely the more time the listing spends on the market.

Figure 6: Prices and Days on Market



Note: The first plot shows the unexplained portion of a sale price for listings with different timing of the sale. I construct this residual by first regressing the sale price on observable property characteristics and taking the difference between the true sale price and that predicted by the hedonic regression. I then scale the residual to reflect the percentage of the sale price. The second plot shows the fraction of listings that sell above the list price against the timing of the sale.

4 Discussion

4.1 Multilateral Matching

In the last section, I have shown that a parsimonious stock flow model of the housing market is consistent with unique empirical patterns we observe in the data. Another attractive property of the model is that it features endogenous multilateral matching on both the seller and the buyer sides and effectively a time-varying Nash bargaining power that each side has. In this simple model, this can generate amplifications in the house price cycle. In a booming market, there are typically more buyers than sellers, and bidding wars are more common. The seller's negotiation power is close to one, so house prices experience an additional boost. Conversely, when the market is slow, sellers have to compete for buyers and buyers can negotiate the price down, exacerbating the price slump. Although MLS data does not include any information on the buyer and seller negotiation process, I can still proxy a "hot" market with the "spikiness" of the sale contract hazard function. This is because if disproportionately many homes sell in the first few weeks of listing, it means that there are many latent buyers waiting for new homes to enter the market. This, in turn, makes it more likely that more than one buyer finds the property attractive. Figure 5(B) shows that the "spikiest" years were 2005, 2019, and 2020. This is consistent with statistics reported by the media on bidding wars.[20]

Another interesting feature of the model is that multilateral matches are more likely to occur in thick markets where there are more buyers and sellers on each side. This is because each new market participant can draw from more potential trading partners. As a result, in peak season, when there are more buyers and sellers participating in the market, the prices could be amplified by a more likely bidding war. In contrast, in the off-season, a thinner market is less likely to feature bidding wars. Thus the seasonal cycle of house prices might be merely a reflection of bargaining power between buyers and sellers.

Finally, due to multilateral matching, the stock-flow features within period price dispersion. There are three possible levels of prices in each period, but the probability of each occurring is an endogenous outcome.

4.2 Limitations of the Stock-Flow Model

This paper presents a simple model of stock-flow with symmetric stock and flow rates between sellers and buyers. Without symmetry, the model solution is not as trivial and might require numerical methods. However, I argue that the symmetry is not an unusual assumption in the housing models. Since sellers and buyers exit the market together as a pair, and since a seller becomes a subsequent buyer and vice versa, in equilibrium, symmetry is quite an intuitive assumption. To illustrate this, Appendix A shows the general equilibrium set up for the stock-flow model. Here, another type of agent is introduced, a homeowner. This model is also simple to solve and can be used for a market analysis that incorporates the general equilibrium effects of home ownership, time on the market, and prices.

Another limitation of the model is that the matching function is not homothetic. This makes it difficult to use it in models of aggregate fluctuations. However, this limitation can be overcome using the simulation analysis. One puzzle in real estate is house price momentum, which by definition requires aggregate shifts in prices and involves analysis away from the steady-state equilibrium. Eric Smith et al. (2022) relies on simulations to show that stock-flow models can deliver house price momentum observed in the U.S. data.

[20]https://www.redfin.com/news/bidding-wars/

5 Adverse Selection

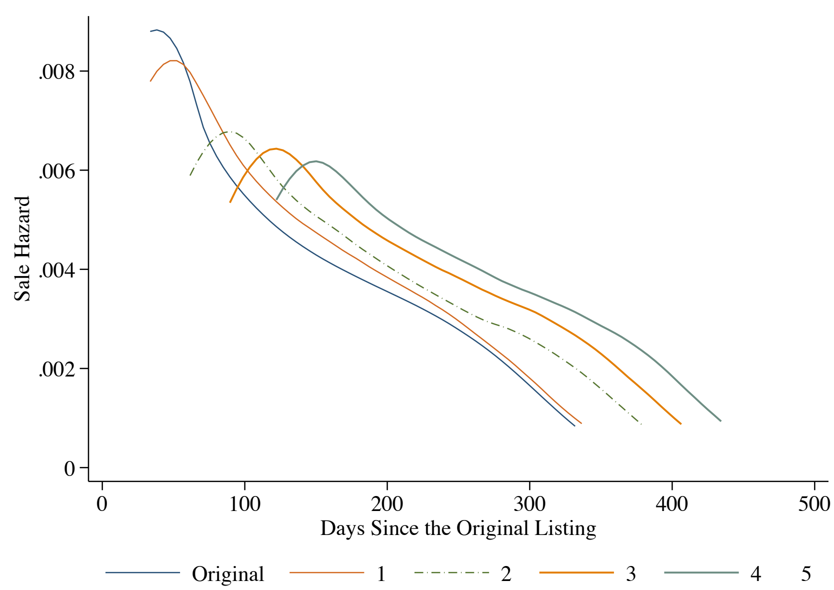
All three predictions of the stock flow model tested in this paper are at odds with models that do not have heterogeneity among buyers or sellers. If there is no distinction between old and new agents, then there can not be differences in expected pricing or match probability between them. By definition, whether a listing is new or old is irrelevant. However, random matching models have been adopted to incorporate heterogeneity and match some of the empirical patterns that I document. Albrecht et al. (2007), for example, introduces impatience that sets in on sellers with a Poisson arrival rate and makes them discount their listings to attract more buyers and increase sale probability. This leads to lower expected prices as the time on market goes up. A directed search price mechanism is at play: cheaper listings attract more buyers and sell faster and expensive listings see fewer visits and are less likely to sell. However, while capturing the pricing patterns, the model delivers *increasing* hazard rates of sale, since it is precisely the discounted homes that have been on the market for a while that sell faster than the newly listed homes. In the data, on the other hand, it is the more expensive (new) listings that sell fastest while cheaper (old) listings linger on the market.

To my knowledge, there is no random matching search model that provides both the price and the hazard rate predictions that I consider. The only way to get both in a random search model is to introduce heterogeneity on the property side. If more attractive listings sell faster and at a premium, then the pool of older listings inevitably has more lemons. If buyers cannot tell the quality of the house ex ante, visiting an attractive listing is more likely to result in a sale and at a higher price, leading to adverse selection and exactly the patterns we confirm in the data. The goal of this paper is not to discredit random search models, but rather to show that a parsimonious approach of the stock-flow framework can achieve empirical data patterns that are difficult to incorporate in a random search framework. However, a random search model with adverse selection tells a very different story than a mismatching stock flow model. If adverse selection fully explains the empirical facts, then perhaps modeling mismatch is not a helpful exercise.

One phenomenon in house selling can help distinguish between a stock-flow mechanism and a random search with adverse selection - relists. A seller often pulls their listing from the market for a period of time only to re-list it afterward with the same information. It used to be a prominent strategy to "reset" the "days on market" field, which buyers often interpreted as a signal of quality. However, platforms like Zillow now display a full history of the listing so the buyers would be able distinguish easily between a re-list and an original listing. In a random search with adverse selection, relisting will therefore not increase the chances of sale as compared to homes with equivalent days on the market. However, in a stock flow model, a seller might experience a boost in the hazard rate that arises from several buyers coming to the market during the time a listing delisted. For those buyers who entered during the time period after a de-list and before a re-list, this listing is effectively "new". Since there could be several of these buyers, a re-list might even generate a bidding war.

I test whether relists experience a hazard rate increase by comparing probability of sale for re-lists and listings that never de-listed given the same period of effective days on market. I fit a Cox proportional hazard model on hazard rates for different vintages of re-lists. In each, I control for the calendar time of the listing. Figure 7 illustrates the findings. The blue line represents a hazard rate as a function of days on market for listings that have not yet de-listed. Subsequent lines represent listings that were de-listed in the first, second, third, and fourth month post the original listing date, respectively. The relist strategy in the initial listing period has no effect on the hazard rate (these are likely because of a typo in the MLS entry). However, re-listing strategy in the second - fifth months leads to increasing hazard rates which points to the stock-flow mechanism in favor of the adverse selection story.

Figure 7: Hazard Rates for Re-lists



Note: This figure plots hazard rates for re-lists and listings that never de-listed given the same period of effective days on market.

6 Conclusion

This paper proposes an alternative approach to modeling search frictions in the housing literature. Here, the frictions come from mismatch between buyers and sellers, rather than the buyer's inability to sample all the houses at once. All successful matches realize immediately, so the current stock of houses for sale and buyers who are looking to buy are mismatched. Buyers and sellers have to wait for new counter parties to enter the market in order to form a successful match. I present a parsimonious stock-flow model with endogenous pricing: If there is only one potentially successful match with a seller and a buyer, then the surplus is divided equally between the two parties. However, when more than two buyers are interested in a seller, the price is bid up to the buyers' indifference point. Similarly, if a new buyer finds two or more homes attractive, they negotiate the price down and capture all of the match surplus.

Although simple, this model can deliver several patterns consistent with the data: 1) new listings are highly predictive of sales regardless of the total inventory of for-sale homes; 2) for-sale properties experience declining hazard rates with the first few weeks being the most likely time of sale; 3) listings that sell in the first few weeks are likely to sell above the average price, while older listings are more likely to sell at a lower price. In addition, endogenous multilateral matching in the stock-flow model generates amplified price cycles and can rationalize pronounced seasonality in the housing markets.

Finally, the stock-flow narrative of a mismatch between buyers and sellers is supported by my analysis of the relisted properties. I show that properties that were withdrawn from the market at some stage during their listing period experience a more likely sale post re-listing relative to properties that spent the same amount of time on the market. This is consistent with a new pool of potential buyers that accumulated while the listing was not available.

References

Albrecht, James, Axel Anderson, Eric Smith, and Susan Vroman, "Opportunistic matching in the housing market," *International Economic Review*, 2007, *48* (2), 641–664.

Anenberg, Elliot and Steven Laufer, "A more timely house price index," *Review of Economics and Statistics*, 2017, *99* (4), 722–734.

Arefeva, Alina, "How auctions amplify house-price fluctuations," *Available at SSRN 2980095*, 2020.

Badarinza, Cristian, Vimal Balasubramaniam, and Tarun Ramadorai, "In Search of the Matching Function in the Housing Market," *Available at SSRN 4594519*, 2024.

Coles, Melvyn G and Eric Smith, "Marketplaces and matching," *International Economic Review*, 1998, pp. 239–254.

Genesove, David and Lu Han, "Search and matching in the housing market," *Journal of Urban Economics*, 2012, *72* (1), 31–45.

Gilbukh, Sonia and Paul Goldsmith-Pinkham, "Heterogeneous Real Estate Agents and the Housing Cycle," *The Review of Financial Studies*, 08 2024, *37* (11), 3431–3489.

Guren, Adam M, "House price momentum and strategic complementarity," *Journal of Political Economy*, 2018, *126* (3), 1172–1218.

Han, Lu and William C. Strange, "Chapter 13 - The Microstructure of Housing Markets: Search, Bargaining, and Brokerage," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5 of *Handbook of Regional and Urban Economics*, Elsevier, 2015, pp. 813–886.

Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun, "Search, liquidity, and the dynamics of house prices and construction," *The American Economic Review*, 2014, *104* (4), 1172–1210.

National Association of Relators, "2023 Profile of Home Buyers and Sellers," Technical Report, National Association of Relators 2023.

Ngai, L Rachel and Kevin D Sheedy, "The decision to move house and aggregate housing-market dynamics," *Journal of the European Economic Association*, 2020, *18* (5), 2487–2531.

_ **and Silvana Tenreyro**, "Hot and cold seasons in the housing market," *American Economic Review*, 2014, *104* (12), 3991–4026.

Piazzesi, Monika, Martin Schneider, and Johannes Stroebel, "Segmented housing search," *American Economic Review*, 2020, *110* (3), 720–59.

Smith, Eric, "High and Low Activity Spells in Housing Markets," *Review of Economic Dynamics*, April 2020, *36*, 1–28.

Smith, Eric Eric, Zoe Xie, and Lei Fang, "The Short and the Long of it: Stock-Flow Matching in the US Housing Market," *Available at SSRN 10035*, 2022.

Wheaton, William C, "Vacancy, search, and prices in a housing market matching model," *Journal of political Economy*, 1990, *98* (6), 1270–1292.

A General Equilibrium Stock-Flow Model

In this section, I describe a general equilibrium framework that embeds the stock-flow housing search model. Time is continuous. There are three types of agents in the market: homeowners, home sellers, and home buyers. Homeowners own exactly one house and derive flow utility π from it. At any point in time a homeowner can experience a separation shock that will lead them to not value the home, changing their flow utility from π to 0. The arrival of the shock follows a Poisson process with arrival rate ϕ. Once a homeowner receives this shock, they enter the housing search market as a home seller. Home sellers have exactly one property to sell and value it at 0. They pay flow cost c_s to participate in a search market and try to match with a home buyer. Upon selling their house they immediately become a home buyer, continuing to participate in the housing search market on the other side. Home buyers pay flow cost c_b to participate in the market and are looking to buy a house. However they have idiosyncratic preferences over properties. With probability $1 - \alpha$, a home buyer likes any given house, and, if they buy it, will derive flow utility π from it as a homeowner. With a complimentary probability α, the buyer will not like the house and would not derive any utility from owning it.

When a new seller arrives on the market, all the buyers currently in search of a house will attempt to match with them. In this model all successful matches will realize immediately. The probability of a match is then $1 - \alpha^B$. That is, a match will form if at least one buyer will like the property, which will happen unless none of the buyers like it. If there is no match, then the sellers joins the "stock" of existing (old) sellers S. If a match is made, the buyer becomes a homeowner and the seller instantly becomes a new buyer.

When a new buyer enters the market, they can instantly see all the houses that are currently available for sale from the current stock of S sellers. They will purchase a house and become a homeowner if they like at least one of these homes. This happens with probability $1 - \alpha^S$. If not, they will join the stock of old buyers and will have to wait for new sellers as potential matches.

The flow arrival of matches at any given point is $s(1 - \alpha^B) + b(1 - \alpha^S)$. In other words, all matches are made either between a new seller and an old buyer or with an old seller and a new buyer. This is because the Poisson process precludes a new buyer and a new seller entering the market at the exact same time to meet. And if two "old" counter parties had a viable match, it would have been realized earlier.

I will now describe the prices. If a buyer finds more than one house that they like (this can happen only with a new buyer, because out of the few houses at least one is on the market by an old seller, so if the buyer was also old they would have matched earlier), they can negotiate the price down to the reservation value of the sellers. Similarly, if a seller finds more than one interested buyer (this, again, can only happen with a new seller), they will negotiate the price up to buyer's reservation value. In other words, the house will generate a bidding war.

Given the housing stock H, match probability α, the flow utility from a good match π, and the arrival of the separation shock ϕ, one can characterize a stationary equilibrium of this model as follows. First, for the inflows and outflows of each state (homeowner, home seller, home buyer) have to be equal. If s is the inflow of sellers and b is the inflow of buyers, then it must be that:

$$s = b = s(1 - \alpha^B) + b(1 - \alpha^S) \tag{10}$$

In a symmetric equilibrium where $S = B$ it must be that $S = B = log(1/2)/log(\alpha)$. Then the number of homeowners must be $H - log(1/2)/log(\alpha)$ and the inflow of sellers and buyer is $s = \phi\,(H - log(1/2)/log(\alpha)) = b$.

The prices p_L, p_N, p_H, value functions $\{V_{So}, V_{Bo}, V_{Sn}, V_{Bn}, V_H\}$ and the trade surplus I can be solved in a system of equations:

$$
\begin{aligned}
V_{Bn} &= -c_b + (1-\alpha^S)(\pi - (1-\gamma_S)p_L - \gamma_S p_N) + \alpha^S V_{Bo} \\
\rho V_{Bo} &= \frac{s(1-\alpha^B)}{B}(V_H - (1-\gamma_B)p_H - \gamma_B p_N - V_{Bo}) \\
V_{Sn} &= -c_s + (1-\alpha^B)((1-\gamma_B)p_H + \gamma_B p_N + V_{Bn}) + \alpha^B V_{So} \\
\rho V_{So} &= \frac{b(1-\alpha^S)}{S}((1-\gamma_S)p_L + \gamma_S p_N + V_{Bn} - V_{So}) \\
\rho V_H &= \pi + \phi(V_{Sn} - V_H) \\
p_L &= V_{So} \\
p_H &= V_H - V_{Bo} \\
I &= V_H - p_N - V_{Bo} + p_N - V_{So} \\
\gamma I &= V_H - p_N - V_{Bo} \\
(1-\gamma)I &= p_N - V_{So}
\end{aligned}
$$
[21]

Where $\gamma_S = S\frac{\alpha^{S-1}(1-\alpha)}{1-\alpha^S}$ and $\gamma_B = B\frac{\alpha^{B-1}(1-\alpha)}{1-\alpha^B}$.

While the equations can be messy, there exists a closed form solution to this system, making this a useful framework for studying the housing market.