The Microstructure of the U.S. Housing Market Evidence from Millions of Bargaining Interactions*

Haaris Mateen[†], Franklin Qian[‡], Ye Zhang[§]

November 9, 2020

Click here for the latest version.

Abstract

We study the microstructure of the U.S. housing market using a novel data set comprising housing search and bargaining behavior for millions of interactions between sellers and buyers. We first establish a number of stylized facts, the most important being a symmetric spread of the sales price around the final listing price in our data. Second, we compare observed behavior with predictions from a large theoretical housing literature. Many predictions on the relationship between sales price, time on the market, listing price and atypicality are borne out in the data. However, existing models do not adequately explain the symmetric spread of the sales price around the final listing price. Using a modeling strategy that treats listing price changes as revisions of expectations about the sales price, we find sellers under-react to information shocks in estimating the sales price. Last, we find that the bargaining outcomes are influenced by previously undocumented buyers' bid characteristics, e.g., financing contingencies and escalation clauses, that signal a buyer's ability to complete or expedite the transaction. This suggests an important role for buyer bid characteristics, which are not explained by existing theories, in affecting bargaining power and surplus allocation in bilateral bargaining in housing transactions.

*We thank Taylor Marr from Redfin for his valuable insights and support in obtaining the data. We also thank Ben Walzer from Redfin for his help with data access. The data we use does not represent 100% of Redfin business or transaction activities, but rather is a sample subset of deals for research purposes. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of Redfin.

[†]Columbia University. hm2692@columbia.edu.

[‡]Stanford University. zqian1@stanford.edu.

[§]Columbia University. yz2865@columbia.edu.

1 Introduction

Housing constitutes an important part of the portfolio of a typical household (Davis and Van Nieuwerburgh (2015)).[1] Houses are typically bought infrequently, and buying a house often involves the use of financing (Chalabi (2014), Neal (2019)). Housing market price and quantity movements, therefore, have an outsize effect on the balance sheet of households. In addition, the housing market is marked by pervasive information frictions and housing stock heterogeneity (Arnott (1987)). Due to these reasons, there have been a large number of empirical studies investigating housing markets in the U.S. The bulk of these studies have concentrated on data related to sales duration, volume variables, financing, and final sales price. In comparison, there have been very few empirical studies investigating the individual stages of the search and bargaining process due to data limitations. The studies that exist in the literature have typically been through survey responses such as Anglin (1997) and Genesove and Han (2011), or have focused on local metropolitan markets such as Merlo and Ortalo-Magné (2004).

This paper utilizes a novel data set to investigate the search and bargaining process in the U.S. housing market. We take advantage of the richness of our data to first document a number of stylized facts and, where possible, compare them with findings in previous studies in the literature. We then evaluate existing theories of house search and bargaining, and compare their predictions with patterns observed in the data. We introduce data that document the occurrence and prevalence of buyer bid characteristics that accompany a buyer's bid and explore its relationship to sales price and house characteristics. Finally, we conduct an exercise that measures the concession of surplus the seller makes in each round of bargaining, thereby giving a sense of the split of bargaining power between the buyer and seller. We also compare bargaining patterns with theories of bargaining in the literature in a manner similar to Merlo and Ortalo-Magné (2004) and Backus et al. (2020).

Our data set presents a number of novel variables on both seller and buyer side across 44 states in the U.S. during 2012–2019. In particular, for each property listing we have data on the bids made by a subset of potential buyer(s), allowing us to track the history of offers and counter-offers up until the final off-market negotiations. We are able to observe the timing of different stages in the house search and bargaining process from the perspective of the seller. On the seller side, we have the listing date as well as the date of each revision of the listing price. On the buyer side,

[1] The home ownership rate in the US has consistently remained above 60% for several decades, according to FRED series ID RHORUSQ156N. Davis and Van Nieuwerburgh (2015) note that housing assets account for as much as 90% early in the lifecycle to 50% in old age, using data from the Survey of Consumer Finances.

we observe the date of the offer made in each round of the bargaining process; when the offer gets accepted; and the final sales date if the house is sold. We also record the reason for a failed bid. About 37% of all observations or 55,258 property listing events, which we call "bargaining events," have multiple buyers, allowing us to investigate competitive bidding behavior in future work.

We establish a set of stylized facts. First, we find an almost equal split between transactions that end with the sales price being above the final listing price and those that end with the sales price being below the final listing price. This motivates the idea that the listing price is a best estimate of the sales price according to the seller. Second, about 10% of all bargaining events (15,898 observations) receive offers in between listing price revisions. In fact, the first listing price revision takes place in the absence of any offer in only 45% of all bargaining events with at least one listing price revision, suggesting that a revision in listing price is triggered not only by the frequency of offers but by the size of offers. Third, the average duration for properties that end up having sales prices less than the final listing price is three times, at 63 days, of the duration for properties that have sales prices more than the final listing prices. These facts point to differences in buyer interest for the property; they also correlate with the atypicality of the property. Finally, a significant number of listing price revisions, about 12% of all observations with revisions, are price *increases*.

Models of housing search and bargaining can be split into two broad questions. The first looks at the macroeconomic implications of housing on household balance sheets. The macrodynamics of housing markets lead to booms and busts in the larger economy leading to interesting business cycle and asset pricing implications, as discussed in Piazzesi and Schneider (2016) and Davis and Van Nieuwerburgh (2015). The second strand of literature, which is what this paper is concerned with, investigates the microstructure of housing markets. The housing market is characterized by many unique features such as high search frictions and market thinness stemming from the pervasive heterogeneity of housing stocks, as well as uncertainty in the search, matching and bargaining process. Han and Strange (2015) offer a good summary of the recent existing literature.

The traditional housing microstructure literature has typically focused on predictions with respect to sales price and time on the market, since this was the first wave of data that was available to establish stylized facts. One sided search models, such as Haurin (1988) and Salant (1991), and random matching models, such as Novy-Marx (2009) and Genesove and Han (2012), are good examples of such theories. We confirm the predictions of these theories with our data, in particular, with respect to sales price, seller duration, and atypicality of housing. These results act as a valuable consistency check of our data since we confirm what are generally regarded as "established facts" in the literature.

With the advent of better data, theoretical models have investigated other aspects of the search and bargaining process, particularly with respect to the role of the listing price (often referred as "asking price"). These papers argue that the listing price is a strategic instrument balancing the trade-off between sales price and duration, as in Yavas and Yang (1995), as an instrument to direct search, as in Chen and Rosenthal (1996a) and Chen and Rosenthal (1996b), as a signal of seller motivation, like in Albrecht et al. (2016), or as a partial commitment device, as in Han and Strange (2016). Our data weighs in on the predictions of these papers.[2] We find evidence affirming the importance of the listing price as an important strategic variable influencing the time on the market, sales price, and interest from buyers. We find that more atypical houses are more likely to sell below the listing price while less atypical houses are more likely to sell above the listing price.

The symmetric spread of the listing price around the final sales price motivates us to consider a theory where the listing price is the best estimate of the seller about the final sales price of the property. A better estimate may possibly reduce the cost of negotiations. We utilize listing price revisions in our data to test, under the assumption that listing price represents an estimate of the final sales price, the under- or over-reaction of sellers to information shocks. We regress the forecast error of the sellers, defined as the difference between the sales price and the final listing price, on the forecast revisions, defined as the difference between the final listing price and the initial listing price, similar to Bordalo et al. (2020). We find evidence for systematic under-reaction to information. Our estimates precisely estimate a forecast error of 4 dollars per 100 dollars of listing price revision. To the best of our knowledge, there is little or no theory explaining how these asking prices are revised in the presence of information shocks.[3] In a future iteration, we intend to incorporate both information shocks and the strategic value of the listing price into a model.

One of the contributions of our paper is presenting systematic data on the presence and prevalence of buyer bid characteristics in the U.S. housing market. A bid from a buyer consists of several characteristics beyond the offer price. For example, roughly 7% of all buyers in our data waive the financing contingency. A financing contingency allows the buyer to not complete purchase of a property due to non-availability of a mortgage.[4] Waiving this contingency is possibly a proxy

[2] Genesove and Mayer (1997) show that the choice of listing price is influenced by the loan-to-value ratios of sellers. This is an important channel but is outside the scope of our study.

[3] Merlo et al. (2015) is a paper that incorporates a structural model investigating the seller's intertemporal problem with respect to listing price revisions.

[4] A recent paper that documents the importance of cash transactions in housing is Han and Hong (2020).

for buyer traits but also a signal to the seller. Similarly, an escalation clause, present in 8% of all bids, tells the seller that the buyer is willing to increase the bid by a certain percentage if there are competing bids. We look at the association of 5 buyer bid characteristics – financing contingency, escalation contingency, inspection contingency, pre-inspection request, and client letter – with the unit sales price of the house and with its atypicality. We find interesting heterogeneous effects. For example, waiver of financing contingencies becomes more likely with higher sales price through a linear relationship; on the other hand, escalation clauses become more likely with increasing unit sales price only up to a point, after which the curve flattens and somewhat falls with a further increase in the unit sales price, suggesting a quadratic relationship. More atypical houses are positively correlated with waivers of financing contingencies and of inspection contingencies. Client letters and escalation clauses are negatively associated with atypicality. These results suggest a non-monetary aspect of buyer bids that may have a signalling value such as displaying the buyer's ability to reduce uncertainty over the completion of the transaction (waiving financing contingency) or the ability to expedite the transaction (waiving inspection contingency), or expressing motivation (through client letter). Exploring the richness of our data in this respect and understanding its connection with its signalling value and bargaining power is something we intend to pursue in a future draft.

We introduce a measure of the concession of surplus that the seller makes to the buyer in each round of bargaining. Our measure has some commonalities with what Backus et al. (2020) use in their paper. Their measure tracks the weights that each of the seller and buyer alternately put on the previous offer made by the counter-party. A crucial difference in this paper, however, is that our measure *always* provides the concession of surplus that the seller makes because the bargaining round in the housing set-up is naturally defined by revisions of the listing price by the seller. In fact, a unilateral revision of listing price by the seller in the absence of buyer offers is a common occurrence in the housing market. Finally, we look at the distribution of bargaining event outcomes according to the predictions of standard bargaining theories in a manner similar to Backus et al. (2020).[5] The products covered by eBay's Best Buy in Backus et al. (2020) are more homogeneous than housing stocks. The sellers and buyers are often repeat participants, and therefore there are fewer information asymmetries associated with buyer/seller reputation and because of buyer/seller experience. The products are also more comparable because they are sold on other platforms or through other selling formats on eBay, such as auctions. The housing market,

[5]See also the discussion in Merlo and Ortalo-Magné (2004).

on the other hand, has extreme heterogeneity – every house is a unique good. This combined with pervasive information asymmetries means that there is no reason predictions from standard bargaining models would hold in the housing market. Indeed, about 16.5% of our observations do display some amount of gradual back-and-forth between the buyer and the seller as also observed in Backus et al. (2020). However, we also find some evidence for more traditional theories that predict immediate agreement, immediate disagreement, delayed agreement, and delayed disagreement.

This paper is organized as follows. Section 2 provides background on the data platform and introduces our novel data set. Section 3 provides an overview of the housing micro-structure models in the literature and collects their predictions with respect to a number of variables of interest. Section 4 tests these models and also introduces the test for under-/over-reaction by the seller when revising the listing price. Section 5 and Section 6 examine several novel detailed buyer characteristics and how they affect the surplus allocation of the bargaining process. Section 7 concludes.

2 Data

The data in this paper combines two sources of information: data from Redfin when they represent potential buyers as agents, and housing data from the Multiple Listing Service (MLS).

2.1 Institutional Background

We obtain housing market data from Redfin, a full-service brokerage that combines the traditional brokerage system of providing in-person agents, with a sophisticated online interface. As an online real estate brokerage, Redfin started using map-based search in 2004. It combined satellite maps with real estate information and helped improve online real estate search systems at that time. This was before the introduction of Google Maps. As Redfin went public in 2017, it became one of the major real estate web portals in the US. Unlike other popular platforms, Redfin makes money when users buy or sell homes through its platform. Redfin hires agents for both the buyer and seller sides. They help guide their clients to bargain strategically throughout the property sales process. Redfin agents experience exceptionally high customer demand because their commissions are below the market-based fees. [6]

[6] Redfin also adopts a Referral Partner Program that relies on a referral network of over 3,100 agents at other brokerages. This program refers a customer to a partner agent, like RE/MAX or Coldwell Banker, representing the customer from the initial meeting through closing. These partner agents are offered when the properties are outside of Redfin's direct service area.

Our data is on the buy-side. If a Redfin customer is interested in buying a house, Redfin provides an agent to the buyer at no expense. This is the typical market structure where the buyer agent is often a sub-agent of the seller agent.[7] Redfin suggests buyers apply for a pre-approval of a mortgage first. Once the lender approves, the buyer is encouraged to book home tours. The tour can be in-person or through video chat. Then the buyer can reach out to an agent to start an offer. Figure 1 depicts the panel seen by a prospective buyer when starting an offer on Redfin. The buyer's agent is hired for free since the seller pays all the real estate commissions.[8] The buyer can even get a refund, which can be applied towards closing costs,[9] if purchasing with a Redfin agent.

The agent can help determine how much to offer and what contingencies to include when making the offer for the buyer. An contingency is an event or condition that must occur before the deal can close. Buyers waive these contingencies as a bonus to the offers. The waiver or inclusion of these contingencies is taken as the buyer bid characteristics for each offer. These characteristics help the buyer stand out from multiple offers received by the seller and persuade the seller to accept this offer.[10]

2.2 Data Description

We first provide an overview of our data. Figure 2 presents a flow chart of the bargaining process between a seller and all potential buyers. A seller puts up her house on the MLS and therefore broadcasts the house listing to several potential buyers who are searching for a property of their preference. We call the price at which the seller puts up her property as the initial listing price. Throughout, the term "buyer" is used to refer to the user interested in potentially buying the item, whatever be the outcome. The buyers can choose to send their offers to the seller through Redfin. The seller can choose to accept the offer *or* to revise the price based on new information received and/or based on the interest in the market. We see every revision of price by the seller on the MLS. We also see the offer price(s) submitted by Redfin on behalf of the buyer. Ultimately, the seller decides to take the property off-market (on a recorded off-market date) and continue negotiations

[7]There is an extensive literature on the role of agents and the MLS in the U.S. housing market. See Han and Strange (2015), Benjamin et al. (2007), Miceli et al. (2007), Zietz and Sirmans (2011) for more details. Our focus is not on this aspect of the market although we control for agent fixed effects in our analysis.

[8]The seller pays the listing and buyer agents' commissions. The typical listing fee is 2.5% - 3%, but Redfin charges 1.5% of the sales price. The listing fee would be 1% and the 0.5% more would be refunded if the seller buys with Redfin within 365 days of the sale. Meanwhile, the buyer covers expenses like settlement fees, lender fees and title insurance.

[9]Closing cost for buyers vary depending on the buyer's loan program, but they typically range from 2% - 5% of the purchase price.

[10]We describe these contingencies in detail in Section 5.

with one or more buyers. We cannot observe off-market negotiations but we record if the buyer(s) represented by different Redfin agents succeeds or fails in buying the house. In case of rejection, we see the recorded reason for rejection. We record the sales price of the house.

We define each bargaining round as beginning when the seller posts a new listing price. Therefore, the first bargaining round starts when the house is first listed. Further rounds begin when the seller revises their seller listing price. Within the duration between a list price and its revision there may or may not be offers submitted by prospective buyers. We define each property listing event, starting from the initial listing to the final sales, as a "bargaining event."

We obtain data from Redfin for the bargaining and sales process of properties listed on Redfin. This consists of 146,675 unique properties listed on Redfin and 147,701 bargaining events for these properties across 44 states from Jan 1, 2012 to Sep 9, 2019.[11] Figure A.2 shows the geographical distribution of these bargaining events. Table A.1 shows the number of bargaining events in each of the 44 states. As mentioned before, we define a bargaining event as the whole duration starting from the property listing to its final sales. The bargaining event includes the initial listing price and date, dates and prices for any listing price revisions, offer prices and offer dates of any interested buyers in our data, the off-market date of the property, and final sales price and date. For a more rigorous discussion of bargaining events with examples, refer to Appendix C.1.

Table 1 provides summary statistics for our data set. The data set contains bid-level Redfin data in the US from 2012 to 2019. For all the analysis in this paper, we deflate prices so that they all in 2010 US dollars. Panel A presents summary statistics of the 147,770 listings what we refer to as Bargaining Events. The average initial listing price is $479,582, and the average sales price is 98% of the initial listing price. The listing duration, i.e. the number of days from the initial listing date to the off-market date, is 30 days on average. Of the listings, 28% have the listing prices revised at some point by the sellers during the listing duration. The average sales price comes down to 97.6% of the initial listing price. There is large variation in the ratio of the final revised listing price to the initial listing price. Some sellers choose to increase the price but on average there is a drop in the listing price. We will investigate this variation in later sections.

Panel B includes descriptive statistics of 107,498 buyers. On average, buyers make offers on 2.04 unique bargaining events, make about 1.05 purchases and are assigned to agents with an average experience of participating in 139 bargaining events. Panel C provides descriptive statistics of the

[11]This data does not represent 100% of Redfin business or transaction activities, but rather is a sample subset of deals for research purposes.

buyers at the bargaining event level. There are 151,336 such buyer-bargaining event pairs. Buyers make 1.04 offers in each bargaining event. We use several dummy variables to measure the bid characteristics of the buyers: among all the buyers, 7.47% have financing contingency waived, 9.4% have inspection contingency waived, 8.02% make the escalation clause, 2.94% make an inspection before making an offer, and 23.2% write a client-letter. The definitions of buyer characteristics are provided in Section 5.

Panel D presents summary statistics of 245,739 rounds, where each round is defined as the period between two listing price revisions. After the listing price has been put up, multiple buyers may make offers for the property listed. The listing price and any offer price submitted constitute one bargaining round. A new bargaining round begins when the seller revises the listing price. Buyer experience in this table is the number of rounds, including the current round, that a buyer has participated in at the time of the current round. Table 1 shows that buyers have an average of 1.55 rounds of experience. Finally, Panel E provides property-level data for our data set. One of the unique aspects of our data set is the presence of walk, transit and bike scores that are provided by Walk Score, a subsidiary of Redfin. More standard housing controls for MLS will be added to this list. We see that our sample consists of larger houses with an average of a little more than 3 bedrooms. Below, we will compare our sample property value with those in national representative surveys like that by the ACS.

Figure 3 provides a distribution of the number of buyers in each event. Our information for the number of buyers is available at each bargaining round level. That is, between two sets of listing price revisions for the same bargaining event, we see the number of buyers who participate in each bargaining round. We construct the number of buyers by taking the maximum of this number across all bargaining rounds in each bargaining event. Therefore, our data for the number of participating buyer for each bargaining event is a weak underestimate of the number of buyers who actually participated. As in previous studies, we see a large number of bargaining events where there is only one recorded buyer. However, there are a non-trivial number of events with more than 10 buyers. The information on the number of buyers is collected in real time by the buyer-side agents.

About 37% (55,258) of all events have recorded additional offers from other interested buyers. Interestingly, 30537 of these or 55% of these observations see the buyer fail in obtaining the house. Failure in buying houses when multiple buyers are submitting offers constitutes 20% of the 71,591 total failed bid observations we record in the data.

Out of the 147,770 bargaining events in our data, 15,898 receive offers *in between* initial and final listing price revisions. At 10% of the sample, they constitute a substantial number of observations where buyers respond to listing prices that have already been put up on the market but the seller then decides to revise the listing price.

Figure 5 provides a sense of the distribution of the original list price, the final list price before the property goes off-market, the offer price and the sales price. Intuitive patterns are observed: the final sales price is shifted to the left of the original list price. Figure 4 shows the distribution of the unit sales price. The unit sales price ranges from \$0.115 to \$336498 and the average is \$279.[12] It is noted that the mode stays around \$130. The size of this data is 70,439, indicating 70,439 of these 147,820 transactions succeed.

Figure 6 compares our data for six states with the median home sales price provided by the the ACS in 2013 and 2018. All prices are deflated to 2010 values. We find that the houses represented by Redfin are consistently high valued compared to ACS houses. There is considerable stability in Redfin prices over this period of 5 years.

Table 2 lists reasons for being rejected. We see that the top reason for rejection is due to multiple offers followed by an unsatisfactory price being submitted. A non-trivial fraction is comprised of late-stage reasons such as a Failed Inspection or because of some other contingencies. In a future iteration, we will investigate the richness of our data in this respect.

2.3 Stylized Facts

We make an effort here to compare our data with previous studies in the literature in the past twenty years. We start with survey data originally from Genesove and Han (2012) and reported in Han and Strange (2016). Table 3 provides a sense for the relationship between the final sales price and the final listing price in our data. In contrast to Han and Strange (2016), we see a nearly 50-50 split between houses that sold below final listing price and those that sold above final listing price. In fact, the spread almost looks as if the final listing price is an estimate on the part of the seller about the ultimate selling price of the house. In contrast to their survey, we report fewer occurrences of multiple buyers.[13]

Table 4 continues the comparison by looking at duration and multiple buyer break-up by the relationship between sales price and the final listing price. We find that there is a significant

[12]The unit sales price is winsorized by first 99th percentile for a better illustration.

[13]The spread of the sales price around the listing price is similar to Case and Shiller (2003). However, while 50% of their data showed properties sold at the listing price, at-listing price sales comprise only about 15% of our data.

difference in the days on the market between houses that sell below the listing price, at the listing price and above the listing price. The duration on the market when the sales price is lower than the final listing price is nearly three times that when the sales price is higher than the final listing price. This suggests that sellers agree to sell below their listing price when they are struggling to find buyers on the market. Interestingly, the number of buyers doesn't increase a lot for cases when the sales price is greater than the listing price. There is no difference in the proportion of sales with multiple buyers either.

Next, we compare the listing price changes over time with findings from Merlo and Ortalo-Magné (2004) who obtained data from four real estate agencies in England between 1995 and 1998. Table 5 shows that the percentage of properties with no listing price revisions is 66% that is somewhat comparable to the 77% figure in Merlo and Ortalo-Magné (2004). The proportion of properties with only one listing price revision is almost the same between the two papers. However, the proportion of two or more listing price revisions is about three times higher in our data.

We also find that the average price change for the first listing price revision is 2.42%, a figure that is less than half of the figure in Merlo and Ortalo-Magné (2004). The time taken to change the listing price is also half of that in their paper at 5.3 weeks (compared to 12). Part of the reason for this reduction is the fact that houses sell much faster two decades from the time Merlo and Ortalo-Magné (2004) did their study. The average time on the market for their study was 11 weeks while it is around 4 weeks for our data. In that respect, the observation that a listing price revision happens when a house that does not sell on the market is plausible. However, a notable difference when considering this story is that the number of houses that receive no offer before a listing price revision in our data is only 45% compared to 92% in Merlo and Ortalo-Magné (2004). This is a significant difference and points to the idea that it is not merely an absence of offers but the *magnitude* of offers that is also causing sellers to delay sale. A similar pattern is obtained for the second price change.

Another important difference with Merlo and Ortalo-Magné (2004) is the presence of a significant number of cases where the final listing price is *higher* than the initial listing price. About 3.6% of all observations see an increase in price changes, which constitutes 12% of all listing price change observations. The idea that sellers choose to revise their prices upwards suggests information shocks that make the sellers revise their expectations for the sales price for their house.

Lastly, the number of observations where the seller takes the house off-market on the day of receiving the offer from the represented buyer is 19,577 which represents 13.2% of all observations.

Although not a one-to-one comparison, because we cannot see if the offer recorded in our data is the first offer made to the buyer, it represents a stark contrast to the 40% of sales that happen with the first offer received in Merlo and Ortalo-Magné (2004). Only 20% of the bargaining events see all buyers submitting their bids on the same day. As mentioned in Han and Strange (2016), these may be coming from markets where there is a designated day for bids to be submitted.

3 Models in the Housing Literature

3.1 A Standard Housing Model

A typical model of the house search and bargaining process has a few basic elements. There are m buyers and n sellers, $m, n \geq 1$. One or both sides may be searching for prospective trades. Both sides are generally assumed to be risk neutral.

Each seller may post a listing price $a_j, j \in \{1, \ldots, n\}$.[14] The buyer(s) responds with an offer price $b_{i,j}$, $i \in \{1, \ldots, m\}$. Generally, the model looks at the surplus generated in the transaction, and the trade is completed if it is positive. With multiple buyers, there may be competitive bidding.

The first set of models we will investigate are search models. All search models involve markets that clear both on price and time.

3.2 Random Matching

In a random matching model, both buyers and sellers search for the right match. Typically, a match is decomposed into a meeting function followed by bargaining. We follow the model given in Genesove and Han (2012) although similar models are present in Albrecht et al. (2007) and Novy-Marx (2009).

To that effect, let $f(m, n)$ be a constant returns to scale meeting function, increasing in both arguments. The meeting function is sufficient to allow us to calculate the probability of a buyer being contacted by a seller, $p_b = f(m, n)/m$, and the analogous probability of a seller being contacted by a buyer, $p_s = f(m, n)/n$.

Noting that market tightness is defined as $\theta = m/n$, it's straightforward to see that $p_s(\theta) = \theta p_b(\theta)$. The match utility between the buyer and seller is assumed to be a draw from $u(x)$, known to both buyers and sellers.

[14] The literature often refers to "listing price" as the "asking price."

A transaction is assumed to go through if the surplus is positive. The probability of the surplus being positive, therefore, is an important parameter of the problem. Most models then use a Nash bargaining solution to describe the price on the market. Since the reservation values are endogenous, as is the market thickness, an intertemporal equation governs the trade-off between continuing search with the opportunity cost of delaying purchase.

A positive demand shock increases market tightness, reducing time on the market, and increasing price. One could check this relationship in the time of a housing boom. Given that our data is mainly through the boom period, we investigate this relationship in the housing boom period from 2012-2019. We should see a positive price-volume correlation.

In booms, there are rapid sales and rapid buying because of liquidity clearing. Sellers have a high opportunity cost of failing to sell during a given period because they need to move. Buyers have an opportunity cost of mismatch as well because they need to continue to search which leads to a drop in housing consumption. Therefore, apart from seeing a positive correlation between final sales price and the number of buyers bidding for each property, we should also see a shorter duration on market for each property.

These models typically do not provide insight into the process of individual seller listings, or on the patterns observed between the listing price and the final sales price.

3.3 One Sided Search

One sided search models simplify the random matching set-up by either fixing the seller, or the buyer. We will concentrate on one sided seller models in this section. Relevant papers include Salant (1991), Haurin (1988), Yavas and Yang (1995), and Anglin (1997).

In a typical one sided seller search model, the seller sets the listing price a. This listing price is the price at which the seller commits to sell the house. The fundamental trade-off is setting a price so that the house sells early while getting the best sales price for the house. We can therefore simplify the intertemporal problem in random matching to only have a trade-off for the seller. The buyer on the other hand simply visits the house and their valuations for the house come from a distribution that is commonly known to all.

There are some simple predictions from this set-up. A lower asking price should predict a lower sales price, a lower time on the market and more bids. Another prediction from one sided seller search models is about the relationship between atypicality, provided by Haurin (1988), and time on the market. The idea is that the more atypical a house, the more time it would stay on the market.

Haurin's heterogeneity is calculated by conducting a hedonic regression of house price on housing characteristics. Let the coefficients be β_i for each characteristic. Let $\bar{x}_i$ be the average value for each characteristic. Then the atypicality is measured by:

$$\sum \beta_i |\bar{x}_i - x_i| \tag{1}$$

and atypicality predicts longer time on the market. Han and Strange (2016) use a slightly different expression in finding the absolute value of the exponential of the difference between the house characteristics and the average characteristics. Another recent application is Haurin et al. (2010).

An important drawback of this class of models is that they cannot explain the fact that the final sales price is typically not equal to the initial asking price. In fact, as we will show, they are more frequently above or below the asking price in our sample.

3.4 Directed Search

The directed search literature focuses on the role of the listing price in the search and bargaining process. The listing price can act as a strategic instrument, as a ceiling on the final sales price, or as a partial commitment device. See Han and Strange (2015) for an extended discussion.

The simple strategic value of the listing price implies that a high listing price should predict a high sales price but a longer time on the market. See, for example Yavas and Yang (1995). A more typical trade-off is that a low listing price directs or encourages search.[15] This is because even though the idiosyncratic match value is determined when the buyer and seller meet, the listing price acts as a credible constraint on the final sales price. This appears in models, for example by Chen and Rosenthal (1996a) and Chen and Rosenthal (1996b). These implications are also borne out by our data.

A recent paper that is of particular interest in the directing role of the listing price is Han and Strange (2016). The authors posit that the listing price acts as a partial commitment device. They present a two-point distribution on match utility. A lower offer price from buyers leads to Nash bargaining. A higher offer price from multiple buyers leads to an auction. In the case of an exactly matching offer by just one buyer, the sale is closed at that price. In this way, the listing price does have bite as a commitment device but it also directs search in the housing market. The trade-off

[15] See Horowitz (1992) for an early example.

of risking a lower sales price, because the listing price as a commitment device if all bidders bid below the listing price, is offset by the possibility of a bidding war that should push up the final price. Listing price is inversely proportional to search activity in booms and also to the atypicality of the house.

Lastly, Albrecht et al. (2016) posit that the listing price represents seller motivation and acts as a signal to prospective buyers on the market. A more motivated seller puts the house up for a lower listing price, compared to a less motivated seller. This leads to more interest from buyers and a shorter time on the market. The model also provides listing prices that are above or below the final sales price.

Interestingly, the authors do not concentrate on revisions of listing price. They posit that all that matters is the existence of a final listing price. This claim runs against the importance of listing price revisions shown in papers like Merlo and Ortalo-Magné (2004). Merlo et al. (2015) is an exception that uses a structural model to understand the patterns observed in Merlo and Ortalo-Magné (2004). The authors provide these results in a non-stationary framework.

Therefore, the directed search literature provides a link between the listing price and sales price and time on the market. More recent models allow for the listing price to not be a ceiling or a full commitment on the part of the seller but predict lower asking price for more atypical houses or during busts. We will test these predictions in our paper.

4 Testing Models

4.1 Seller Duration

The first set of analyses we present concerns the dependence of seller duration on different variables. We define seller duration as the time period between the initial listing date and the final sale date.

Table 6 regresses seller duration on a number of variables that reflect interest from buyers. Column (1) regresses seller duration on buyer duration. We define buyer duration as the time period between the first offer date of any given buyer and the last recorded date till which they are in the bargaining process. This may be the final sales date in case of a successful deal, or the offer-rejection date which may or may not take place before the off-market date. In the latter case, the buyer loses out during final negotiations with the seller. All the regressions in this paper, unless otherwise specified, use year by county by month fixed effects as well as agent fixed effects. The former takes care of time trends, location specific characteristics and seasonality. The latter makes sure the identification comes from comparing property listings represented by the same agent. We also cluster standard errors in all our regressions at the calendar month.

We control for housing characteristics such as property age, property size, number of bedrooms and bathrooms, transport scores and the property type. We find that the seller and buyer duration are positively related to each other, as would be expected.

Column (2) regresses seller duration on the number of buyers. Our data provides the number of buyers in each round of bargaining. We construct number of buyers by taking the maximum of the number of buyers in any bargaining round. This number is a lower bound since we cannot see if the same competing buyers appear in different bargaining rounds. We find that the seller duration reduces with the number of buyers which makes intuitive sense.

We repeat the analysis in Column (3) but split the number of buyers into quartile buckets. The rationale is that the number of buyers variable is both a lower bound and possibly a categorical variable that represents a certain level of buyer interest. We are regressing seller duration on dummies for various quartiles. As expected, the seller duration is highest for the lowest quartile of number of buyers. The duration difference between the 25th percentile and the 75th is of almost 6 days.

Column (4) regresses seller duration on the logarithm of the initial listing price. Various models incorporating the role of listing price predict a longer time on the market with a larger listing price, as discussed in Section 3. The coefficient size and its significance does not change appreciably if we replace the initial listing price with the final listing price.

Finally, Column (5) provides the relationship between seller duration and atypicality. The classic study by Haurin (1988) predicts a longer time on the market for properties that are more atypical. As described in Section 3.2, the atypicality is the sum of absolute deviation of a house's characteristics from its average values. We find that a house that is more atypical takes more days to get sold. A unit increase in atypicality increases the number of days on the market by 15 days, which with an average of 76 days on the market, is almost a 20% increase in the time on the market.

Our data is in a period of housing boom. In a future iteration, we will compare the same results using regional boom-bust episodes to validate results from previous studies such as Glaeser et al. (2014). Random matching models as well as one-sided seller search models both predict seller duration as being inversely proportional to the number of buyers. In addition, matching models predict a lower duration during times of booms.

4.2 Sales Price

Apart from duration, most of the traditional literature has concentrated on predictions with respect to sales price. Table 7 regresses the sales price on the initial listing price and the number of buyers. Column (1) contains the coefficient with respect to the initial listing price and, as expected, we see a positive highly significant coefficient. The fact that it is less than 1 implies that the final sales price is generally negotiated down from the initial listing price. One sided seller search models do predict a positive relationship between the asking price and the final sales price. So do directed search models. Random matching models are generally silent about this aspect of the market structure.

Column (2) regresses the logarithm of sales price on the number of buyers. We find a positive relationship between the sales price and the number of buyers. An increase in one buyer increases the sales price by 0.3%. That is an increase of roughly \$1500. This relationship is interesting because it is predicted by random matching models. Matching models typically predict a positive price-volume correlation. Conditional on a fixed set of properties, this will translate into more buyers and higher sales price during the time of booms. This prediction is not made by one sided seller search models. Column (3) looks at the relationship between sales price and seller duration. We find a negative relationship between the the logarithm of sales price with the seller duration.

4.3 Listing Price

We have already discussed the relationship of seller duration with the listing price in Table 6. There is a strong positive relationship as predicted by several models in the literature. As mentioned in Section 3, a number of recent studies posit that the asking price acts as a signal or some kind of commitment device.

Table 8 regresses the number of buyers on the initial listing price and the final listing price. We find that a higher listing price discourages buyers. Column (1) shows that a 1% increases in the initial listing price almost decreases the number of buyers by 1. Column (2) allows us to compare our results with those in Han and Strange (2016). Their main specification suggested a reduction of the number of buyers by 0.40%. Taking the average number of buyers in their data this suggests a reduction of 0.69 buyers. Our estimate is a reduction of 0.48 buyers. Note that we have not added property tax assessments. We intend to add that in a future iteration of this paper.

Table 9 regresses the listing price on atypicality of the house. We find that a more atypical house reduces the listing price. The coefficient doesn't change if the regressor is the initial listing price or the final listing price. Table 10 conducts a regression of the seller duration on the initial listing price and an interaction term with the atypicality of a house. We do not find an significance on the interaction coefficient though it is negative as suggested by Han and Strange (2016).

4.4 Expectation Formation

A number of recent papers have studied the formation of expectations for a variety of variables, such as inflation and housing construction indices (e.g., Kuchler and Zafar (2019), Bailey et al. (2019), Bordalo et al. (2020), and Gennaioli and Shleifer (2020)). The idea behind these studies is to understand how agents react to new information and revise their expectations. Most studies in the literature use surveys to elicit agents' expectations.

As shown in Table 2, there is a remarkable symmetry of the final sales price around the final listing price. This suggests that the listing price may be an estimate on the behalf of the seller about the correct worth of the house. To the extent that the listing price represents their best estimate of the value of the house, the seller may still commit to the listing price as an upper bound in terms of limited interested either in terms of number of buyers or buyer valuations. The latter point fits a number of theories which posit that the listing price is a commitment device from the seller and lets the buyers know the fair price of sale from the seller's point of view. If we stick to this interpretation, we can do a basic test of how asking price revisions reflect changes in expectations for the sellers about the market clearing price for their house.

The basic model regresses the forecast error, which is the difference between the actual sales price and the final listing price, on the forecast revision, which is the difference between the final and initial listing price.

$$\text{forecast error} = \alpha_0 + \alpha_1 \times (\text{forecast revision}) + \epsilon \tag{2}$$

The left side of the equation stands for the seller's expectation error, which is the difference between the sales price and the final listing price. On the right side of the equation, the difference between final listing price and initial listing price stands for the seller's forecast revision. α_1 represents the coefficient of reaction to information. If the reaction coefficient is larger than 0, sellers under-react to information; if the reaction coefficient is less than 0, sellers over-react. Under the assumption that the listing price is merely a posted price, and under rational expectations, the coefficient in this regression should be equal to zero.

Table 12 shows the results of this regression. Column (1) tells that if there is a revision in the seller's listing price by $10,000 then the estimate will undershoot the actual sales price by $400. The positive reaction coefficient indicates exactly that: the seller's revision between the initial and final listing price systematically underestimates the actual sales price for the house. At this point, we do not mean to imply that this is the actual story behind the choice of listing price. However, it forms an interesting avenue for follow-up research. Columns (2)-(5) repeat the exercise between quartiles of the unit sales price. We see that the under-reaction increases from the lowest quartile to the highest quartile.

For deals in the first quartile, Column (2) indicates that a revision in the listing price of 10,000 dollars leads to a forecast error of 200 dollars. However, for the highest quartile, Column (5) shows that the same revision leads to a positive forecast error that is 3 times that of Column (2), of 600 dollars. Given that the revision in price between the average initial listing price and the final listing price is about 30,000 dollars, Column (1) implies a forecast error of 1,200 dollars.

5 Buyer Characteristics

5.1 Buyer Characteristics Definition

One of the novel features of our data set is the documentation of buyer bid characteristics that go beyond just the offer price. A buyer typically makes an offer that consists of an offer price along with a set of contingencies. These contingencies allow the buyer to step out of their negotiation with the seller depending on the occurrence of certain events. The buyer can also choose to signal their interest and motivation for buying the property by adding a client letter. Table 1 Panel B provides summary statistics with respect to the occurrence of these characteristics. We use several dummy variables as proxies to measure the relative bargaining power of the buyers. It is pertinent to explain the meaning of these characteristics:

1. Financing contingency: The financing contingency allows the buyer to walk out of a deal if the mortgage for the house is not approved by a certain date. If a buyer waives financing contingencies, it may signal the buyer's motivation to buy the property, or it may signal weaker bargaining power of the buyer. It could also mean the buyer has enough cash to pay without loans. About 7% of all bids waive financing contingencies.

2. Inspection contingency: Buyers often want the house to be thoroughly inspected by a licensed inspector. The purchase of a property is tied to the result of this investigation. The contingency

allows buyers to make the deal contingent on the home inspection results. It also gives them the right to negotiate further if there are repair issues. Generally, the inspection happens 5-7 business days after the date of mutual acceptance. A buyer waiving inspection contingencies may signal, as with financing contingencies, extra motivation of the buyer to obtain the property. It may imply that the buyer already has information about the property. It may also signal weaker bargaining power of the buyer. About 9% of all bids waive inspection contingencies, showing that it is a common component of the home buying process.

3. Escalation clause: Escalation clauses are used in competitive markets. It generally indicates that the buyer understands there will be multiple offers for the property. The way it is structured is as follows: the buyer submits a bid b_1 for the house with an "escalator" clause that allows for the bid to go up to b_2 if the seller receives another bid that is greater than b_1. It is clear that the escalation clause allows the buyer to concede more surplus to the seller in the case of competition. An escalation clause is present in 8% of all bids.

4. Pre-inspection: A pre-inspection clause is an ex-ante inspection of the house, as against an ex-post inspection contingency. It allows the buyer to gather more information before submitting the offer. It may also mean that the buyer feels they have enough bargaining power to ask for an inspection even before providing an offer. We see a request for pre-inspection in only 3% of all our bids.

5. Client letter: The buyer may signal their motivation to buy the property by writing a client letter. Anecdotally, this is a way of making an offer more competitive. These letters generally try to approach the seller on a personal level and try to provide the reasons why the buyer is interested in the property and why the buyer should be chosen among other offers. A client letter is another characteristic that acts as a signal from the buyer, and also possibly a proxy for bargaining power. Almost a quarter of all buyer bids in our data is accompanied by a client letter.

Figure 7 presents bin scatter plots of the occurrence of five buyer characteristics observed in listings with respect to their unit sales price. To calculate the unit sales price, we adjust sales prices to 2010 US dollars. The parameter for the bin scatter plot is 25. This divides unit sales prices into 25 bins in ascending order. Each unit sales price corresponds to a successful listing, because only the successful listings have sales prices provided. Then the probability value is the occurrence of one buyer characteristic observed in these groups of bargaining events.

From these plots, we observe some interesting patterns. More dots stay to the left of these plots, where unit sales prices are under $400. This indicates that more listings have unit sales prices below \$400. Panel A shows the probability of financial contingency being waived as unit sales price increases. This relation is obviously positive and the rate of increase in probability slightly grows for unit sales prices below \$300. Panel B shows the probability of inspection contingency being waived with unit sales prices. We see a similar relation to Panel A except one bump at unit sales price around \$320. This bump is more obvious in the later plots. Panel C shows the probability of client letter provided with unit sales prices. The rate of growth in probability is slowing down as unit sales prices reach \$400. Panel D shows the probability of requesting a pre-inspection and Panel E shows the probability of signing an escalation clause. Panel D and E suggest a quadratic relationship. In Panel D, the drastic growth of this probability for listings with unit sales prices in \$200–\$350 indicates that buyers care more about the house qualities as they provide higher offers. But the maximum of this probability is below 5% in absolute terms. This is relatively low compared to those of other plots (i.e. 15% for Panel E, 40% for Panel A and C, 50% for Panel B). Thus the buyers' demand for pre-inspection is not high possibly because waiving the pre-inspection shows their desire for buying the houses. Overall, the probability of waiving contingencies increase with unit sales price. Escalation clauses become more common for higher priced properties and its occurrence remains relatively flat beyond a point. Client letters also become more likely as the unit sales price of the property increases. Panel A, B and C show a linear positive relation and Panel D and E show a concave quadratic relation.

5.2 Do Buyer Characteristics Predict Prices?

Table 13 shows how buyer characteristics are correlated with the logarithm of unit sales price. To calculate the unit sales price, we adjust sales prices to 2010 US dollars. The dependent variable is the logarithm of sales price per square foot in dollars. The independent variables are indicators of buyer characteristics (refer to section 4.1 for details). All regressions include year, agent, county and calendar month fixed effects. Housing control[16] are added. Robust standard errors in parentheses are clustered on the calendar month level.

Column 1 tells that buyers who waive their financing contingency will pay 0.05% more for each square foot. Column 2 tells that buyers who waive their inspection contingency will pay 0.046% more for each square foot. Column 3 tells that buyers who have escalation clause will pay 0.035% more for each square foot. Column 4 tells that buyers who didn't pre-inspect the house will pay

[16] Housing controls: property type dummies, property age, property age square, number of bedrooms, number of bathrooms, ratio of number of bedrooms to number of bathrooms, walk score, transit score and bike score.

0.025% for each square foot. Column 5 tells that buyers who have client letters will pay 0.027% more for each square foot. To summarize, Table 13 indicates that the price buyers paid for each square foot will be higher if these dummy variables are 1, which suggests that, at least in a reduced form sense, buyers have less bargaining power with each of these characteristics. In future iterations of this paper, we will look at the joint distribution of these characteristics.

Table 14 regresses atypicality on different buyer characteristics. We find that waivers of financing contingencies and inspection contingencies, Columns (1) and (2), are associated with atypical houses presumably because the idiosyncratic value to the buyer of finding the right match is very high. An escalation clause seems to be associated with more standard housing stock as shown in Column (3). The escalation clause is related to the expectation of a buyer about the number of other buyers who may be competing with them. The coefficient suggests buyers do not expect competing buyers for more atypical houses. Column (4) suggests that ex-ante inspections or pre-inspections happen less often for atypical houses which ties well with the observation that inspection contingencies are more likely to be waived for atypical houses. Finally, Column (5) suggests that client letters are less common for atypical houses.

6 Bargaining and Surplus Measurement

As noted before in Section 2, about 13.2% of all bargaining events involve the property being taken off-market on the same date as the first offer submission. These cases are not preceded by listing price revisions. Around 65% of these off-market offers are accepted. To the extent that the housing market always involves information asymmetries and financing frictions, we can consider these 8.6% (13.2% $\times$ 0.65) of cases to be where there is immediate agreement in the style of Nash (1950), Rubinstein (1982). However, a substantial number of bargaining events do not display such behavior.[17]

Studies that show the possibility of immediate disagreement, like Perry (1986), and of delayed agreement (but ruling out delayed disagreement), like Gul and Sonnenschein (1988) also may explain some of these remaining observations. Immediate disagreement is observed when an offer is rejected without any response from the seller. This happens in 38.8% of total bargaining events. Delayed agreement can be used to rationalize bargaining events where only the seller or only the buyer is revising offers which is observed in another 16.2% of cases. Delayed disagreement, as in Cramton

[17] An earlier discussion of the relationship of the traditional bargaining literature with empirical findings of the microstructure of the housing market is present in Merlo and Ortalo-Magné (2004).

(1992) may also be a possibility. Cramton's model allows the possibility of initial non-serious offers after which one side makes a precipitous concession. The average price change after the first revision is about 2%, which is not large. Also, while the idea of "bully offers" is mentioned in the literature (Han and Strange (2016)), it is not clear how to identify these offers.

In fact, an interesting subset of the observations show back-and-forth between the seller and the buyer. These sequences of offers and counter-offers is reminiscent of the findings in Backus et al. (2020). These gradual concessions from both sides to finally arrive on a sales price is remarkable because, on the face of it, the housing market is marked by informational asymmetries and frictions. About 16.5% of all bargaining events in our data show this pattern.

We generate the "offer price surplus ratio" as a proxy to measure how much surplus is gained by the buyer and how much surplus is gained by the seller during the bargaining rounds. This ratio equals the difference between the revised listing price and the offer price in bargaining round t divided by the difference between the previous listing price in round $t-1$ and offer price in round t. The denominator is the buyer's proposal over how the surplus is to be shared, and the numerator is the concession that is made by the seller.

$$\gamma_t = \frac{\text{revision price}_t - \text{offer price}_t}{\text{revision price}_{t-1} - \text{offer price}_t},\ t = 1, 2...k, \tag{3}$$

where γ_t is the offer price surplus ratio in round t, revision price_t is the price offered by the seller in round t, and offer price_t is the price offered by the seller in round t. We normalize the range of γ_t to from 0 to 1. When $\gamma_t = 0$, the surplus goes to the buyer. When $\gamma_t = 1$, the surplus goes to the seller. Appendix C.2 provides details for the construction of this variable.

This measurement is similar to that used in Backus et al. (2020), which is denoted by γ_t. Their measure tracks the weights that each of the seller and buyer alternately put on the previous offer made by the counter-party. However, a crucial difference is that we measure a bargaining round based on the seller's action. Therefore, each value of γ_t provides a measure of concession made by the seller. We do this because the structure of the bargaining game in the housing market allows the seller to change prices without any further response from the buyer. Figure 8 displays the distribution of offer price surplus ratio in the first four bargaining rounds. The histograms show that the majority of surplus goes to sellers, given that the majority of ratios are larger than 0.5. The histograms also show that the distribution of ratios remains similar within each round, indicating that the part of surplus goes to sellers in each bargaining round is similar.

7 Conclusion

This paper has introduced a novel data set that has allowed us to establish a number of stylized facts about the housing market in the US. There are almost an equal number of transactions that end with the sales price being above the listing price as those that end with the sales price being below the listing price. The average duration for properties with final listing price higher than the sales price is three times that for properties that have a sales price higher than the final listing price. Properties with two or more revisions in listing price form a significant share of all observations. More importantly, a first listing price revision in the absence of any offers constitutes 45% of all first listing price revisions observed in the data. This suggests a revision in listing price may not only be triggered by a lack of offers but also by the magnitude of offer price. The average price change for the first listing price revisions is only about 2.4%. Lastly, a significant number of listing price revisions increase the listing price.

We then proceed to test the predictions of a number of models in the housing literature. Our data set allows these theories to be tested with a data sample that spans 8 years and 44 states. We confirm established facts with respect to the relationship between sales price, seller duration, and atypicality. We also test more recent papers on the link between the listing price and other variables of interest such as atypicality and the number of buyers. We also run expectation error regressions between the forecast error, defined as the difference between the sales price and the final listing price, and the forecast revision, defined as the difference between the final and initial sales price. We are motivated to run this regression owing to the symmetric spread (in frequency terms) of the sales price around the final listing price. We find, following the terminology in the literature, systematic under-reaction on the part of the seller in predicting the final sales price. We intend to investigate the role of informational shocks and changes in expectations in a future iteration of this paper.

We document the presence and prevalence of buyer bid characteristics in the US housing market. We investigate the relationship of these characteristics with sales price and the atypicality of the property. These characteristics suggest a non-monetary aspect of buyer bids that may have a signalling value to them indicating for example, the ability of the buyer to reduce uncertainty in the transaction (by waiving financial contingencies), the ability to raise the offer in the face of competition (with an escalation clause), displaying eagerness to buy the property and to communicate with the seller (through a client letter), the ability to expedite the transaction (by waiving inspection contingencies), or requesting more information (through a pre-inspection request). The richness of these variables and their possible connection with bargaining power offers a tantalizing direction for future research which we intend to pursue.

We use a measure of the concession made by the seller towards surplus sharing the bargaining game and find evidence that suggests the seller wins a larger share of the surplus which is in sync with anecdotal evidence about the housing market being a "seller's market" in recent times. We also investigate the bargaining behavior by comparing occurrence of bargaining outcomes with the theoretical bargaining literature. We find evidence of immediate agreement and disagreement, as well as delayed agreement or disagreement, and finally bargaining events where there is more gradual convergence to the final sales price. Given that the housing market is marked by heterogeneity of housing stock and pervasive informational frictions, traditional theories of bargaining don't necessarily apply. However, the richness of our data allows us to more fully investigate the determinants of the bargaining game that is best applicable between the buyer and seller in different contexts.

There are aspects of our data that we have not tested because they are not simultaneously present in standard models. For example, we have the offer history for the represented buyer. We can track buyers as they bid for multiple houses and we can track multiple buyers bidding for the same property. We have information about the number of buyers as well as listing price revisions. We also have buyer characteristics. We can control for agent fixed effects. To fully utilize the richness of the data, we intend to write a structural model that incorporates these features and allows us to conduct welfare counterfactual experiments.

References

Albrecht, J., A. Anderson, E. Smith, and S. Vroman (2007): "Opportunistic Matching in the Housing Market*," *International Economic Review*, 48, 641–664.

Albrecht, J., P. A. Gautier, and S. Vroman (2016): "Directed Search in the Housing Market," *Review of Economic Dynamics*, 19, 218–231.

Anglin, P. M. (1997): "Determinants of Buyer Search in a Housing Market," *Real Estate Economics*, 25, 567–589.

Arnott, R. (1987): "Chapter 24 Economic Theory and Housing," in *Handbook of Regional and Urban Economics*, Elsevier, vol. 2 of *Urban Economics*, 959–988.

Backus, M., T. Blake, B. Larsen, and S. Tadelis (2020): "Sequential Bargaining in the Field: Evidence from Millions of Online Bargaining Interactions," *The Quarterly Journal of Economics*, 135, 1319–1361.

Bailey, M., E. Dávila, T. Kuchler, and J. Stroebel (2019): "House Price Beliefs And Mortgage Leverage Choice," *The Review of Economic Studies*, 86, 2403–2452.

Benjamin, J. D., P. Chinloy, and D. T. Winkler (2007): "Sorting, Franchising and Real Estate Brokerage Firms," *The Journal of Real Estate Finance and Economics*, 34, 189–206.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110, 2748–2782.

Case, K. E. and R. J. Shiller (2003): "Is There a Bubble in the Housing Market?" *Brookings Papers on Economic Activity*, 34, 299–362.

Chalabi, M. (2014): "How Many Homeowners Have Paid Off Their Mortgages?" .

Chen, Y. and R. W. Rosenthal (1996a): "Asking Prices as Commitment Devices," *International Economic Review*, 37, 129–155.

——— (1996b): "On the Use of Ceiling-Price Commitments by Monopolists," *The RAND Journal of Economics*, 27, 207–220.

Cramton, P. C. (1992): "Strategic Delay in Bargaining with Two-Sided Uncertainty," *The Review of Economic Studies*, 59, 205–225.

Davis, M. A. and S. Van Nieuwerburgh (2015): "Chapter 12 - Housing, Finance, and the Macroeconomy," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Elsevier, vol. 5 of *Handbook of Regional and Urban Economics*, 753–811.

Genesove, D. and L. Han (2011): "Measuring the Thinness of Real Estate Markets," 37.

——— (2012): "Search and Matching in the Housing Market," *Journal of Urban Economics*, 72, 31–45.

Genesove, D. and C. J. Mayer (1997): "Equity and Time to Sale in the Real Estate Market," *The American Economic Review*, 87, 255–269.

Gennaioli, N. and A. Shleifer (2020): *A Crisis of Beliefs*, Princeton University Press.

Glaeser, E. L., J. Gyourko, E. Morales, and C. G. Nathanson (2014): "Housing Dynamics: An Urban Approach," *Journal of Urban Economics*, 81, 45–56.

Gul, F. and H. Sonnenschein (1988): "On Delay in Bargaining with One-Sided Uncertainty," *Econometrica*, 56, 601–611.

Han, L. and S.-H. Hong (2020): "Cash Is King? Evidence from Housing Markets," 54.

Han, L. and W. C. Strange (2015): "Chapter 13 - The Microstructure of Housing Markets: Search, Bargaining, and Brokerage," in *Handbook of Regional and Urban Economics*, ed. by G. Duranton, J. V. Henderson, and W. C. Strange, Elsevier, vol. 5 of *Handbook of Regional and Urban Economics*, 813–886.

——— (2016): "What Is the Role of the Asking Price for a House?" *Journal of Urban Economics*, 93, 115–130.

Haurin, D. (1988): "The Duration of Marketing Time of Residential Housing," *Real Estate Economics*, 16, 396–410.

Haurin, D. R., J. L. Haurin, T. Nadauld, and A. Sanders (2010): "List Prices, Sale Prices and Marketing Time: An Application to U.S. Housing Markets," *Real Estate Economics*, 38, 659–685.

Horowitz, J. L. (1992): "The Role of the List Price in Housing Markets: Theory and an Econometric Model," *Journal of Applied Econometrics*, 7, 115–129.

Kuchler, T. and B. Zafar (2019): "Personal Experiences and Expectations about Aggregate Outcomes," *The Journal of Finance*, 74, 2491–2542.

Merlo, A. and F. Ortalo-Magné (2004): "Bargaining over Residential Real Estate: Evidence from England," *Journal of Urban Economics*, 56, 192–216.

Merlo, A., F. Ortalo-Magné, and J. Rust (2015): "The Home Selling Problem: Theory and Evidence," *International Economic Review*, 56, 457–484.

Miceli, T. J., K. A. Pancak, and C. F. Sirmans (2007): "Is the Compensation Model for Real Estate Brokers Obsolete?" *The Journal of Real Estate Finance and Economics*, 35, 7–22.

Nash, J. F. (1950): "The Bargaining Problem," *Econometrica*, 18, 155–162.

Neal, M. (2019): "Mortgage Debt Has Peaked. Why Has the Share of Homeowners with a Mortgage Fallen to a 13-Year Low?" https://www.urban.org/urban-wire/mortgage-debt-has-peaked-why-has-share-homeowners-mortgage-fallen-13-year-low.

Novy-Marx, R. (2009): "Hot and Cold Markets," *Real Estate Economics*, 37, 1–22.

Perry, M. (1986): "An Example of Price Formation in Bilateral Situations: A Bargaining Model with Incomplete Information," *Econometrica*, 54, 313–321.

Piazzesi, M. and M. Schneider (2016): "Chapter 19 - Housing and Macroeconomics," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and H. Uhlig, Elsevier, vol. 2, 1547–1640.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97–109.

Salant, S. W. (1991): "For Sale by Owner: When to Use a Broker and How to Price the House," *The Journal of Real Estate Finance and Economics*, 4, 157–173.

Yavas, A. and S. Yang (1995): "The Strategic Role of Listing Price in Marketing Real Estate: Theory and Evidence," *Real Estate Economics*, 23, 347–368.

Zietz, E. N. and G. S. Sirmans (2011): "Review Articles: Real Estate Brokerage Research in the New Millennium," *Journal of Real Estate Literature*, 19, 3–40.

Table 1: Descriptive Statistics For the Main Sample

	Mean	Std.Dev.	Min	Median	Max
Panel A: Bargaining-Event-Level					
Property Age (Years)	46	32	1	40	190
Buyers Represented by Redfin	1.02	0.16	1	1	5
Listing Price	479,582	279,646	78,803	410,356	2,734,204
Fraction Revised	0.28	0.45	0	0	1
Revisions	0.53	1.19	0	0	58
Revisions Before First Offer	0.50	1.16	0	0	58
Final Listing Price	449,394	189,420	18,904	410,356	801,808
Final Listing Price/ Initial Listing Price	0.99	0.04	0.64	1	1.43
Duration Until Off-Market (Days)	30	25	0	20	67
Sales Price	488,565	292,562	78,261	413,162	2,750,976
Sales Price/ Initial Listing Price	0.98	0.06	0.63	0.99	1.50
Sales Price Per Square Feet	266.78	168.08	49.31	218.71	1242.88
No. Bargaining Events	147,770				
Panel B: Buyer-Level					
Agent Experience	139	121	1	103	855
No. Attempted Purchases	2.04	1.88	1	1	44
No. Purchases	1.05	0.28	1	1	8
No. Buyers	107,498				
Panel C: Buyer-Event-Level					
No. Offers Per Bargaining Event	1.04	0.20	1	1	6
Financing Contingency Waived	0.07	0.26	0	0	1
Inspection Contingency Waived	0.09	0.29	0	0	1
Escalation Clause	0.08	0.27	0	0	1
Pre-Inspection	0.03	0.17	0	0	1
Client Letter	0.23	0.42	0	0	1
No. Buyer-Event Pairs	151,336				
Panel D: Round-Level					
No. Buyers Represented by Redfin	1.04	0.22	1	1	5
No. rounds a buyer participated	1.55	1.15	1	1	12
No. Rounds	245,693				
Panel E: Property-Level					
Year Built	1974	31	1879	1980	2018
No. Bedrooms	3.36	1.06	1	3	7
No. Bathrooms	2.44	0.86	1	2.50	5.50
Approximate Square Feet	2,100	947.5	574	1,920	5,981
Walk Score	42.71	30.61	0	40	99
Transit Score	38.45	22.60	0	37	100
Bike Score	47.46	24.25	2	45	99
No. Properties	146,675				

Notes: This table presents summary statistics for our main data. The sample consists of buyer and seller interactions from 147,770 housing bargaining events on Redfin's platform across 44 states in the U.S. from January 2012 to September 2019. All variables related to price are adjusted to 2010 dollars. Panel A describes the information for each bargaining event. Panel B describes buyers' characteristics in the whole data. Panel C describes buyers' characteristics within each bargaining event. Panel D provides information for each round of the bargaining event, where a round consists of the seller setting or revising the listing price and buyers making offers. Panel E describes property characteristics. Figure 2 illustrates the bargaining process in details. Table A.2 provides additional details on variable definitions and construction.

Table 2: Rejection Reasons

	Frequency	Percent	Cumulative Frequency
Offer Rejected Due to Multiple Offers	38,954	49.96	49.96
Offer Rejected Due to Price	12,930	16.58	66.54
Offer Rejected Due to Failed Inspection	10,519	13.49	80.03
Offer Rejected Due to Other Contingency	5,419	6.95	86.98
Offer Rejected (Other Reason)	4,853	6.23	93.21
Offer Rejected Due to General Terms	2,830	3.63	96.84
Offer Rejected Due to Failed Financing	2,209	2.83	99.67
Offer Rejected by Lender (Short Sale)	196	0.25	99.92
Invalid: Test Offer/Duplicate	31	0.039	99.96
Pending Offer	18	0.022	99.98
Others	16	0.019	100.00
Total	77,975	100.00	

Notes: This table lists the reasons why an offer is rejected. The sample consists of all 77,975 buyer offers that are rejected. The most frequent rejection reason is "rejected due to multiple offers", which makes up nearly 50% of all rejections. The second most frequent rejection reason is "rejected due to price", which makes up 16.58%. And the third most frequent reasons is "rejected due to failed inspection", which makes up 13.49%.

Table 3: Below-, At-, and Above-Final Listing Sales by Year

Year	Sales/Listing Ratio	Below Listing (%)	At Listing (%)	Above Listing (%)	Mean Price	Sale Volume	Multiple Buyers (%)
2012	1.00	51.50	11.76	36.74	513,421	2,637	3.91
2013	1.03	36.64	11.82	51.54	527,642	11,737	7.29
2014	1.02	44.17	12.62	43.21	490,841	12,547	4.71
2015	1.02	42.46	13.50	44.04	476,334	15,081	4.95
2016	1.02	42.20	14.63	43.17	444,009	18,456	5.24
2017	1.02	38.55	14.50	46.95	436,530	22,591	5.19
2018	1.01	47.23	16.62	36.15	404,646	27,705	3.54
2019	1.00	45.24	18.39	36.37	384,264	36,947	4.05
Total	1.01	43.19	15.28	41.53	436,889	147,701	4.68

Notes: This table presents summary statistics for our main data in terms of listing year, which is the year a property is listed on Redfin. The sample consists of all bargaining events on Redfin's platform across 44 states in the U.S. from January 2012 to September 2019. The first column shows the listing year. The second column shows the ratio of sales price to final listing price. The third, fourth and fifth columns show the share of properties that are sold below the final listing price, at final listing price, and above final listing price (if available). The sixth column shows the average sales price. The seventh column shows the number of properties sold by year. The last column shows the share of properties which are bid by multiple buyers.

Table 4: Summary Statistics for Below-, At-, and Above-Final Listing Sales

	Days on Market		Number of Buyers		% Multiple Buyers	
	Mean	SD	Mean	SD	Mean	SD
Sales Price < Listing Price	62.81	81.28	1.01	0.10	5.33	1.19
Sales Price = Listing Price	39.88	58.97	1.01	0.11	4.93	1.08
Sales Price > Listing Price	18.66	36.37	1.05	0.23	4.99	1.10

Notes: This table presents summary statistics of bargaining events and buyers with respect to the relationship between sales price and the final listing price. The sample consists of all bargaining events. The first and second columns show the mean and standard deviation of how long a property stays on the market (equals the number of days between initial listing date and the off market date). The third and fourth columns show the mean and standard deviation of the number of buyers who bid for the same property (including properties which are bid by only one buyer). The last two columns show the mean and standard deviation of the share of properties that are bid by multiple buyers.

Table 5: Descriptive Summary for Price Change

	Total
Price Change Distribution	
With 0	66.51%
With 1	17.64%
With 2+	15.85%
First Price Change	
Average Price Change	-2.42%
Average Weeks	5.31
With No Offer Yet	45.37%
Second Price Change	
Average Price Change	-2.73%
Average Weeks	5.01
With No Offer Yet	40.97%

Notes: This table presents the summary statistics of the price change in our main data. The sample consists of all bargaining events."Price Change Distribution" shows that the seller revise the price of the property for 0 time, 1 time or more than 2 times. "First Price Change" shows the average price change from the seller's initial listing price to the second listing price of properties (if available), the average weeks between the initial listing date and the second listing date, and the percentage of properties which are not yet bid by buyers after the second listing price is given. "Second Price Change" shows the summary statistics from the second listing price to the third listing price of properties (if available), the average weeks between the second listing date and the third listing date, and the percentage of properties which are not yet bid by buyers after the third listing price is given.

Table 6: Relationship of Seller Duration with Buyer Interest

	Dependent Variable: Seller Duration				
	(1)	(2)	(3)	(4)	(5)
Buyer Duration	1.14***				
	(0.004)				
No. Buyers		-3.47***			
		(0.000)			
No. Buyers < 25-pctile			23.90***		
			(0.000)		
25-pctile < No. Buyers < 50-pctile			18.04***		
			(0.000)		
50-pctile < No. Buyers < 75-pctile			13.67***		
			(0.000)		
Log(Initial List Price)				14.98***	
				(1.43)	
Atypicality					14.83***
					(1.23)
Housing Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Year × County × Month FE	Yes	Yes	Yes	Yes	Yes
Agent FE	Yes	Yes	Yes	Yes	Yes
Control Mean (Duration Seller)	75.75	75.75	75.75	75.75	75.75
Observations	127,620	127,620	127,620	127,620	127,620
R-squared	0.54	0.29	0.31	0.20	0.29

Notes: This table shows how seller duration in days are correlated with the buyer interest. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. The dependent variable is seller duration in days, defined as the number of days between the initial listing date and the final sales date. The independent variables are buyer duration, number of buyers and atypicality of a property. Buyer duration is defined as the number of days between the offer submitted date and the sale date. No.buyers is number of buyers participated in each bargaining event. In column(3), by dividing number of buyers into quartiles, the independent variables are 3 indicators. For example, When No.buyers of one given bargaining event is less than 25-pctile of No.buyers, indicator "No.Buyers < 25-pctile" is set to be 1. Atypicality means the difference between a property's own characteristics and the average characteristics of all the other properties. A precise definition is provided in equation 1. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 7: Relationship of Sales Price with Initial Listing Price

	Dependent Variable: Log(Sales Price)		
	(1)	(2)	(3)
Log(Initial Listing Price)	0.93***		
	(0.003)		
Number of Buyers		0.003***	
		(0.000)	
Seller Duration			-0.0001***
			(0.004)
Housing Characteristics Controls	Yes	Yes	Yes
Year × County × Month FE	Yes	Yes	Yes
Agent FE	Yes	Yes	Yes
R-Squared	0.98	0.78	0.78
Observations	124,917	124,917	124,917

Notes: This table shows how logarithm of sales price is correlated with logarithm of initial listing price. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. We adjust all prices to 2010 real U.S. dollars. The dependent variable is logarithm of sales price. The independent variable are logarithm of initial listing price, number of buyers participated in each bargaining event and seller duration. Seller duration is defined as the number of days between the initial listing date and the final sales date. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 8: Number of Buyers and Initial Listing Price

	Dependent Variable: Number of Buyers	
	(1)	(2)
Log(Initial Listing Price)	-0.83***	
	(0.06)	
Log(Final Listing Price)		-0.48***
		(0.04)
Housing Characteristics Controls	Yes	Yes
Year × County × Month FE	Yes	Yes
Agent FE	Yes	Yes
Control Mean (Number of Buyers)	2.51	2.51
Observations	128,292	124,956
R-Squared	0.28	0.35

Notes: This table shows how number of buyers are correlated with logarithm of initial listing price. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. We adjust all prices to 2010 real U.S dollars. The dependent variable is number of buyers participated in each bargaining event. The independent variables are logarithm of initial listing price and logarithm of final listing price for each property. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 9: Listing Price and Atypicality

	(1)	(2)
	Log(Initial Listing Price)	Log(Final Listing Price)
Atypicality	-0.20***	-0.20***
	(0.02)	(0.02)
Housing Characteristics Controls	Yes	Yes
Year × County × Month FE	Yes	Yes
Agent FE	Yes	Yes
Observations	124,956	124,956
R-squared	0.758	0.766

Notes: This table shows how the atypicality of a property is correlated with logarithm of initial listing price and logarithm of final listing price. We adjust all prices to 2010 real U.S. dollars. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. The dependent variable is logarithm of initial listing price and logarithm of final listing price. The independent variable is the atypicality of a property. Atypicality means the difference between a property's own characteristics and the average characteristics of all the other properties. A precise definition is provided in equation 1. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 10: Seller Duration and listing price atypicality interaction

	Dependent Variable: Seller Duration	
	(1)	(2)
Log(Initial Listing Price)	-0.606***	
	(0.042)	
Log(Initial Listing Price) × Atypicality		-0.008
		(0.006)
Housing Characteristics Controls	Yes	Yes
Year × County × Month FE	Yes	Yes
Agent FE	Yes	Yes
Observations	124,956	124,956
R-squared	0.76	0.77

Notes: This table shows how seller duration is correlated to logarithm of initial listing price and the interaction of logarithm of initial listing price and the atypicality. We adjust all prices to 2010 real U.S. dollars. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. The dependent variable is seller duration in days, defined as the number of days between the initial listing date and the final sales date. The independent variable is logarithm of initial listing price and the interaction of logarithm of initial listing price and atypicality of a property. Atypicality means the difference between a property's own characteristics and the average characteristics of all the other properties. A precise definition is provided in equation 1. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 11: Regression of Sales Type with Atypicality

	(1) Below Final-Listing Sales Indicator	(2) At Final-Listing Sales Indicator	(3) Above Final-Listing Sales Indicator
Atypicality	0.42***	-0.15***	-0.46***
	(0.032)	(0.037)	(0.036)
Housing Characteristics Controls	Yes	Yes	Yes
Year × County × Month FE	Yes	Yes	Yes
Agent FE	Yes	Yes	Yes
Observations	65,418	65,418	65,418

Notes: This table shows how the atypicality of a property is related to different indicators. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. The dependent variables are below final-listing sales indicator, at final-listing sales indicator and above final-listing sales indicator. When sales price is less than final listing price, "below final-listing sales indicator" is set to be 1. When sales price is equal to final listing sales price, "at final-listing sales indicator" is set to be 1. When sales price is larger than final listing price, "above final-listing sales indicator" is set to be 1. The independent variable is the atypicality of a property. Atypicality means the difference between a property's own characteristics and the average characteristics of all the other properties. A precise definition is provided in equation 1. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 12: Regression of Forecast Error on Forecast Revision with respect to Sales Price

	(1) Total	(2) Q1	(3) Q2	(4) Q3	(5) Q4
Forecast Revision	0.04***	0.02**	0.05	0.05***	0.06***
	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)
Housing Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Year × County × Month FE	Yes	Yes	Yes	Yes	Yes
Agent FE	Yes	Yes	Yes	Yes	Yes
Control Mean	4,095	-3,524	-2,160	1,210	18,654
Observations	66,490	13,685	15,071	15,739	16,841
R-squared	0.36	0.36	0.30	0.29	0.39

Notes: This table shows how the forecast revision, which is the difference between final listing price and initial listing price, is correlated with the forecast error, which is the difference between the sales price and the final listing price, for different quantiles of the unit sales price. We adjust all prices to 2010 real U.S. dollars. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. The dependent variable is the difference between sales price and final listing price. The independent variable is the difference between final listing price and initial listing price. Five columns show regressions run in total data set of the unit sales price and quartiles of the unit sales price (Q1, Q2, Q3, Q4) respectively. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 13: Regressions on Unit Sales Price

	Dependent Variable: Log(Unit Price)				
	(1)	(2)	(3)	(4)	(5)
Financing Contingency Waived	0.05***				
	(0.004)				
Inspection Contingency Waived		0.04***			
		(0.005)			
Escalation Clause			0.03***		
			(0.004)		
Pre-Inspection				0.02***	
				(0.005)	
Client Letter					0.03***
					(0.004)
Housing Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Year × County × Month FE	Yes	Yes	Yes	Yes	Yes
Agent FE	Yes	Yes	Yes	Yes	Yes
Control Mean	208.70	207.40	218.90	219.50	204.60
Observations	66,249	66,249	66,249	66,249	66,249
R-squared	0.82	0.82	0.82	0.82	0.82

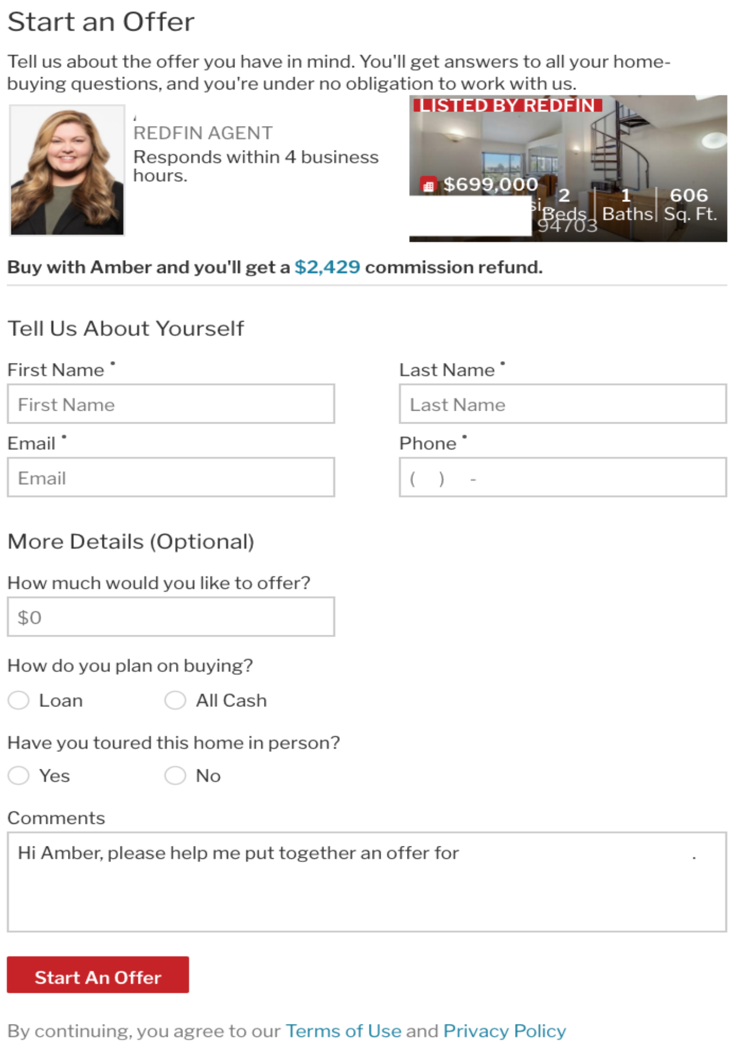
Notes: This table shows how buyer characteristics are correlated with logarithm of the unit sales price. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. We only include bargaining events that resulted in a successful purchase made by a buyer represented by a Redfin agent. The dependent variable is the logarithm of sales price per square feet in dollars (adjusted to 2010 real U.S. dollars). The independent variables are indicators of buyer characteristics, as defined in section 5.1. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 14: Regressions on Atypicality

	Dependent Variable: Atypicality				
	(1)	(2)	(3)	(4)	(5)
Financing Contingency Waived	0.007***				
	(0.002)				
Inspection Contingency Waived		0.004***			
		(0.003)			
Escalation Clause			-0.018***		
			(0.003)		
Pre-Inspection				-0.020***	
				(0.004)	
Client Letter					-0.012***
					(0.002)
Housing Characteristics Controls	Yes	Yes	Yes	Yes	Yes
Year × County × Month FE	Yes	Yes	Yes	Yes	Yes
Agent FE	Yes	Yes	Yes	Yes	Yes
Control Mean	208.70	207.40	218.90	219.50	204.60
Observations	66,265	66,265	66,265	66,265	66,265
R-squared	0.442	0.442	0.442	0.442	0.442

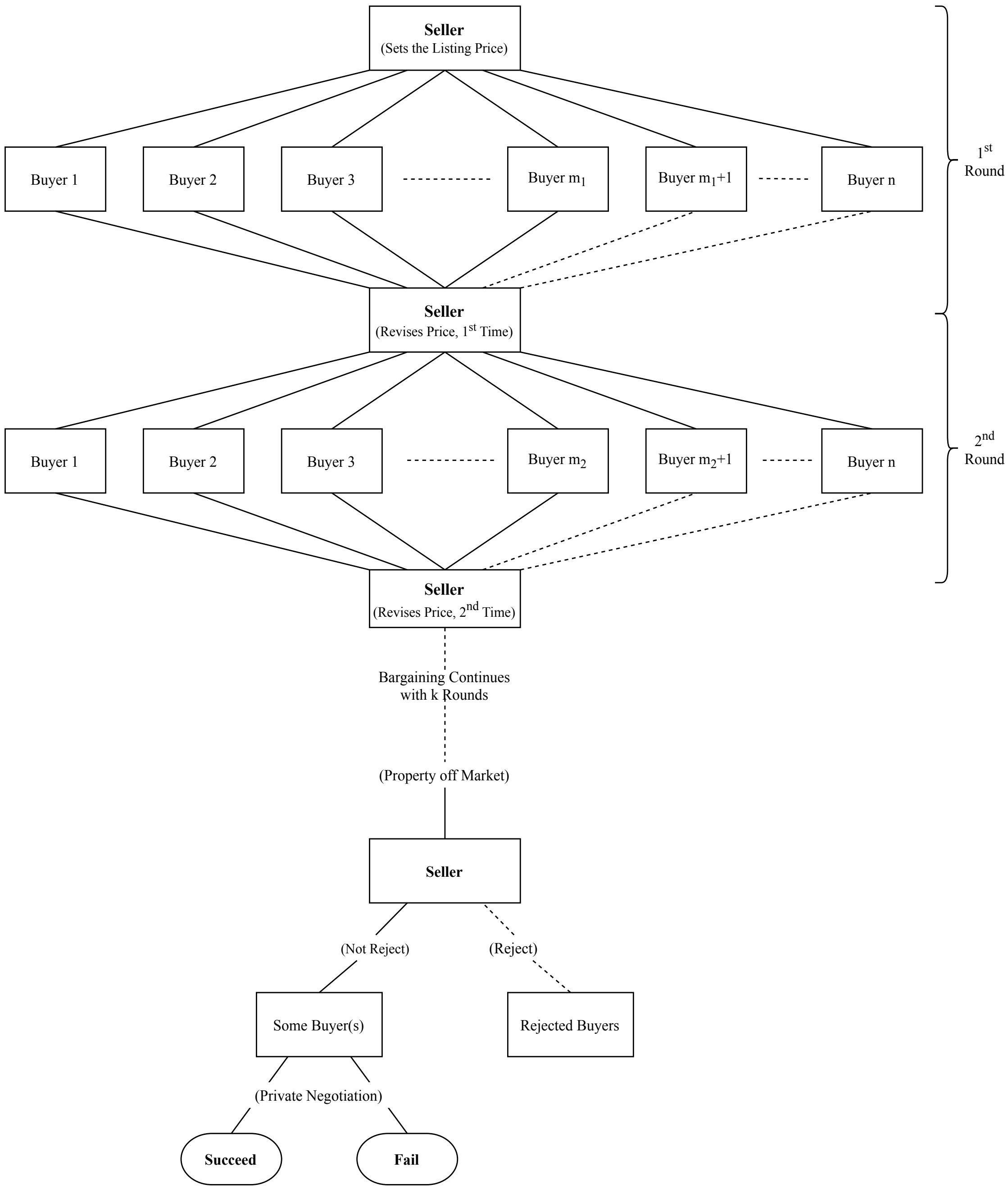
Notes: This table shows how buyer characteristics are correlated with the atypicality of a property. The sample consists of bargaining events happened on Redfin between January 2012 and September 2019 in the U.S. We only include bargaining events that resulted in a successful purchase made by a buyer represented by a Redfin agent. The dependent variable is the atypicality of a property. Atypicality means the difference between a property's own characteristics and the average characteristics of all the other properties. A precise definition is provided in equation 1. The independent variables are indicators of buyer characteristics, as defined in section 5.1. All regressions include year by county by calendar month fixed effects and agent fixed effects. Housing characteristics controls include dummies for property use-code (condominium, single-family, multi-family, etc.), property age, property age squared, number of bedrooms, number of bathrooms, bedrooms-to-bathrooms ratio, walk score, transit score, and bike score. Robust standard errors in parentheses are clustered at the calendar month level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Figure 1: Start an Offer Panel



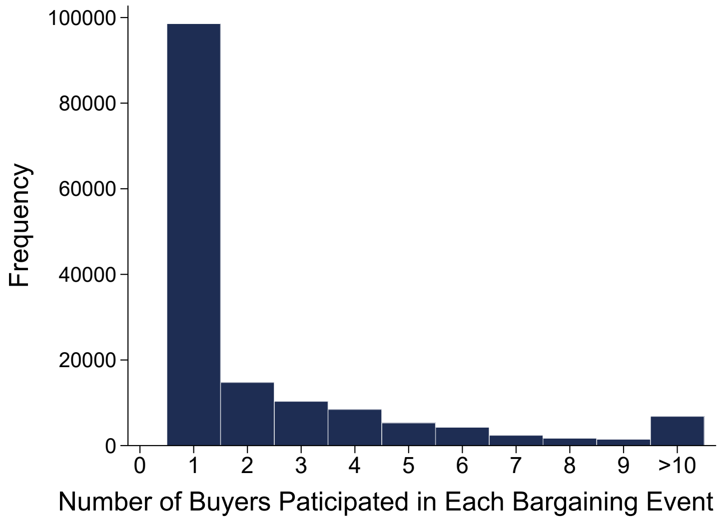
Notes: This figure shows the "Start an Offer" panel on Redfin. This panel is a Redfin page where buyers start to make offers to properties that they are interested in. Once buyers fill in the required information listed on the page and click the red "Start an Offer" button, they are assigned with Redfin agents and are encouraged to declare their needs on this page.

Figure 2: Bargaining Process Flow Chart



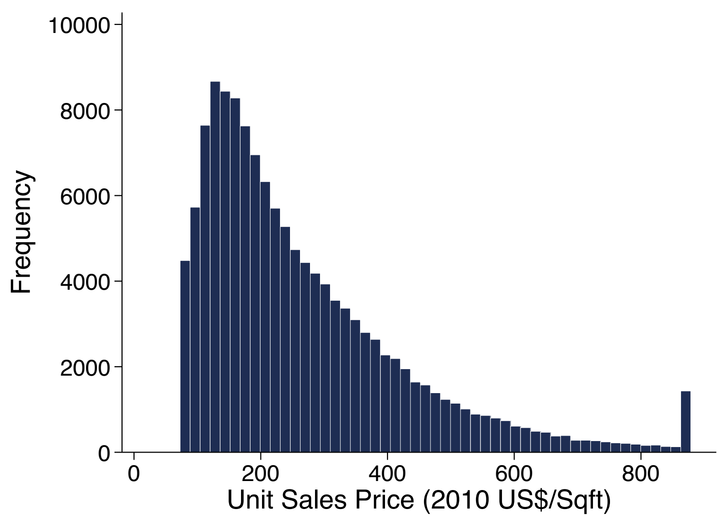
Notes: This figure describes the negotiation process between a seller and multiple buyers. Suppose there are n potential buyers looking for homes on the platform and in total k rounds of bargaining before the property is taken off the market. In the first round of the bargaining process, the seller first lists his or her home on the Redfin platform. m_1 of n buyers reply to the seller with their initial offer prices. In the second round, the seller revises the initial listing price for the first time to provide a counter offer to all n potential buyers. m_2 of these n buyers reply with their new offer prices. The third round of bargaining starts as the seller revises the listing price for the second time. Rounds of bargaining continue until the property is taken off the market by the seller, which could happen with two scenarios: 1) The seller accepts only one offer and the property is automatically taken off the market. 2) The seller takes the property off the market. He or she rejects some buyers and goes through a private negotiation (not available in our data) with the remaining buyers. The deal succeeds if the seller accepts one of the offers from buyers. The deal fails if the seller accepts no offers.

Figure 3: Distribution of the Number of Buyers in Each Bargaining Event



Notes: This figure shows the distribution of the number of buyers in each bargaining event in our data set between January 2012 and September 2019. Number of buyers equals to 1 + the number of additional offers in the bargaining event level. The variable is winsorized at level 1% and 99%.

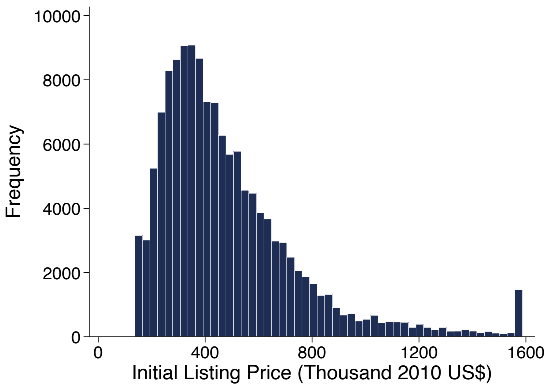
Figure 4: Distribution of Unit Sales Price



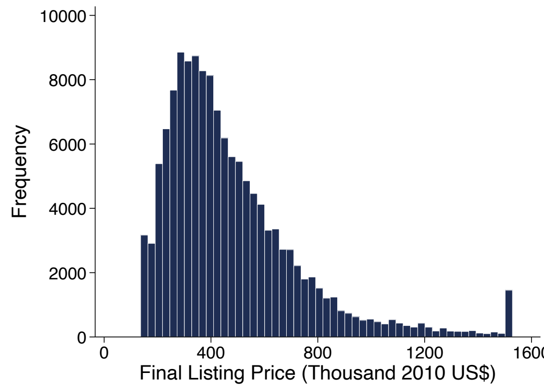
Notes: This figure shows the distribution of unit sales price of U.S. housing properties listed on Redfin between January 2012 and September 2019. Unit sales price is obtained from dividing the property's sales price by the size of the property. Price unit is in 2010 US$/Sqft. The variable is winsorized at level 1% and 99%

Figure 5: Distribution of Prices

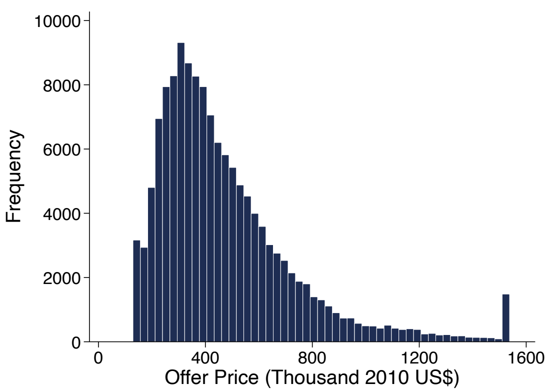
(a) Distribution of Initial Listing Prices



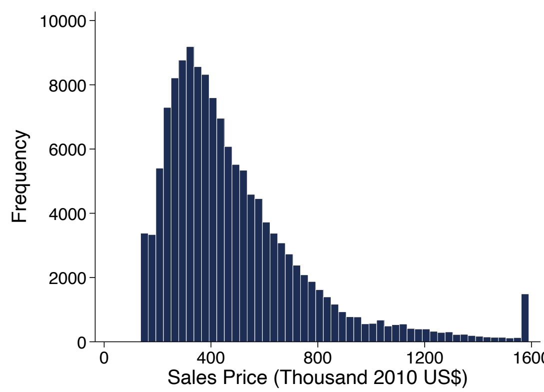
(b) Distribution of Final Listing Prices



(c) Distribution of Offer Prices



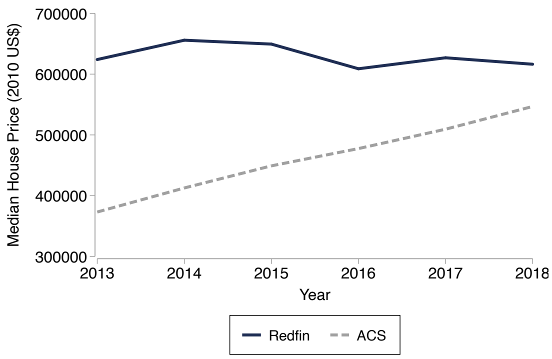
(d) Distribution of Sales Prices



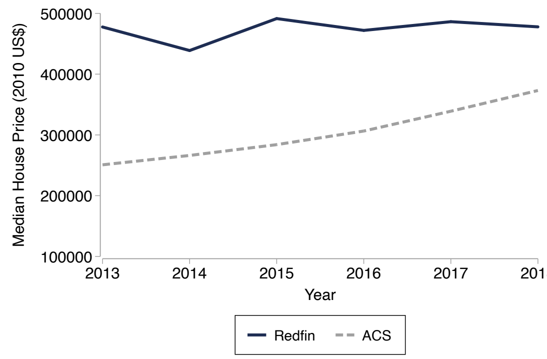
Notes: This figure shows the distributions of initial listing prices, final listing prices, offer prices and sales prices. Sellers make listing prices and buyers make offer prices. Observations are from U.S. housing properties listed on the Redfin platform between January 2012 and September 2019. Price unit is in thousand 2010 US$. All variables are winsorized at level 1% and 99%.

Figure 6: Time Series of Median Sales Price by State

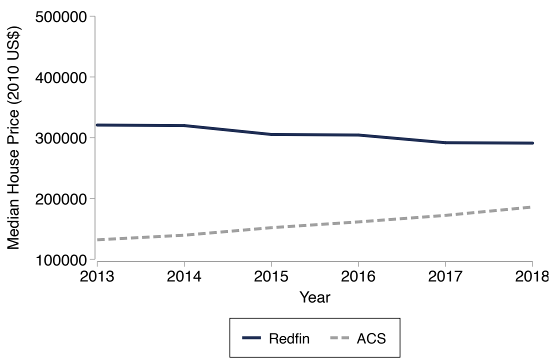
(a) California



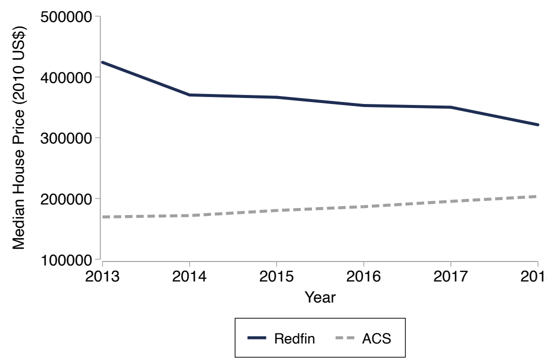
(b) Washington



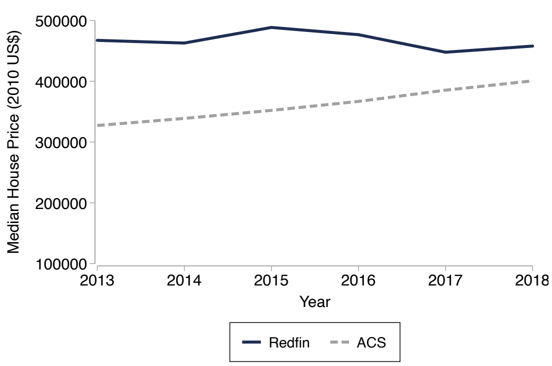
(c) Texas



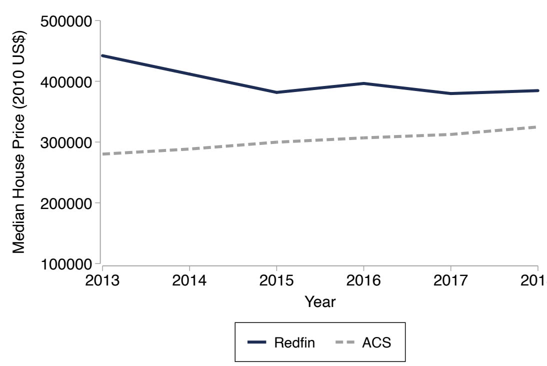
(d) Illinois



(e) Massachusetts



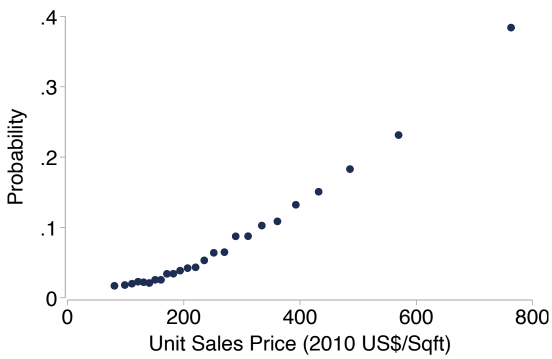
(f) Maryland



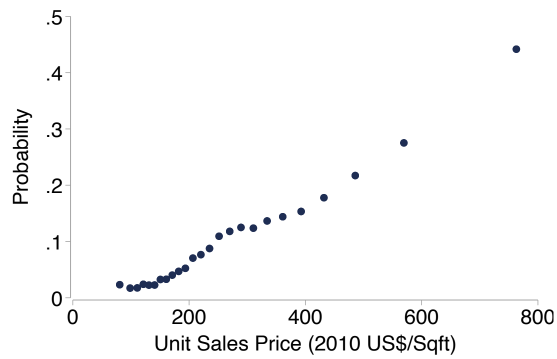
Notes: This figure compares the median state-level house prices from 2013 to 2018 of our Redfin data to that of the American Community Survey (ACS). We start from 2013 because property listings before 2013 only account for 0.7% of our data. The most recent ACS data available is for 2018. The solid lines are median final sales price of Redfin listings in each of the states, while the dashed lines are median home values from ACS in each state. The 6 states, including California (Panel a), Washington (Panel b), Texas (Panel c), Illinois (Panel d), Massachusetts (Panel e) and Maryland (Panel f), together account for 67% of all the property listings in our data. Prices are adjusted to 2010 real U.S. dollars.

Figure 7: Buyer Characteristics by Unit Sales Price

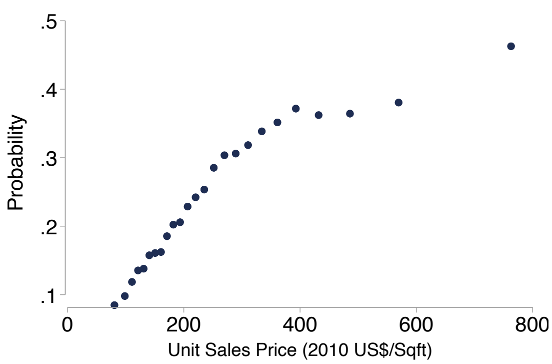
(a) Financial Contingency Waived



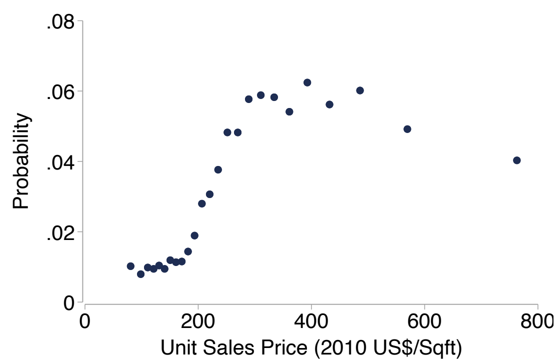
(b) Inspection Contingency Waived



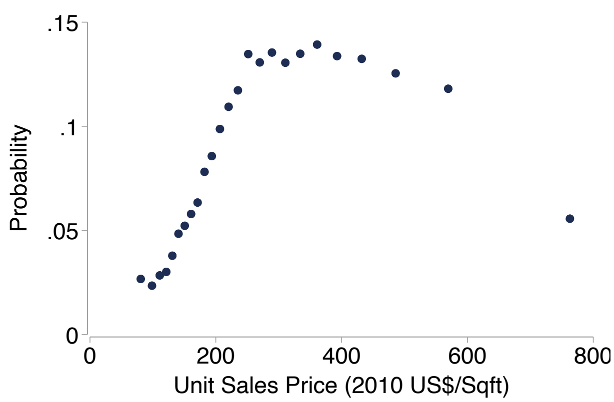
(c) Client Letter Attached



(d) Pre-Inspection Requested

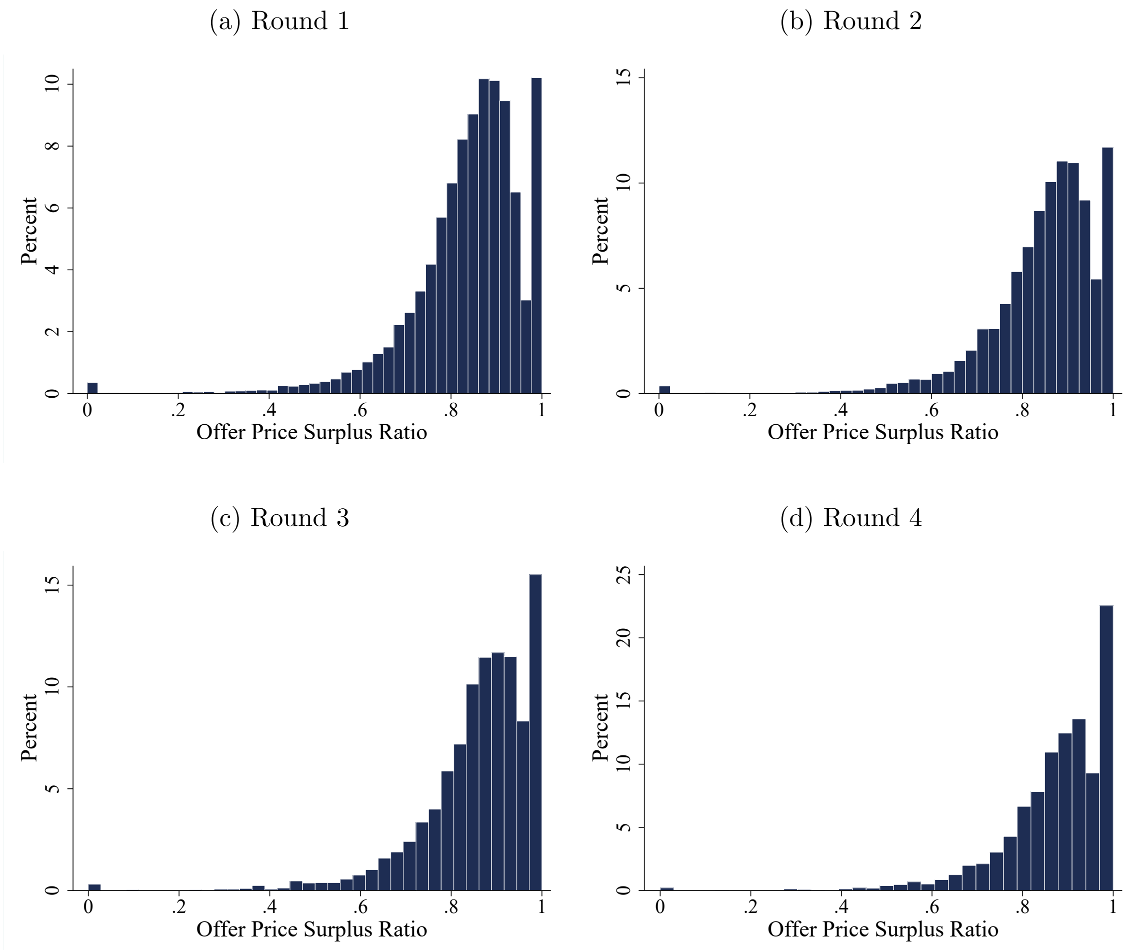


(e) Escalation Clause Signed



Notes: This figure displays the difference in buyer characteristics across 25 quantiles of unit sales prices of listed properties. For each panel, the horizontal axis represents the unit property sales price, computed from dividing sales price (adjusted to 2010 real U.S. dollars) by approximate square feet of each property. We split the whole sample into 25 quantiles by their unit sales price. The vertical axis shows the average probability of buyer characteristics observed in each quantile of the property sales prices. Five buyer characteristics include waiving financial contingencies (Panel a), waiving inspection contingencies (Panel b), attaching client letters (Panel c), requesting pre-inspections (Panel d), and signing escalation clauses (Panel e). A.2 gives a more detailed definition of each buyer characteristic. Unit sales prices are winsorized at level 1% and 99%.

Figure 8: Distribution of Offer Price Surplus Ratio



Notes: This figure shows the distribution of the offer price surplus ratio in bargaining rounds 1 to 4, Panels (a)–(d) respectively. "Offer price surplus ratio in bargaining round t" is defined as the difference between the seller revision price and the buyer offer price in round t, divided by the difference between the seller revision price in round $t-1$ and the buyer offer price in round t. The denominator is the total surplus to be shared by both the seller and the buyer. The nominator is the part of the surplus that goes to the seller.

Appendices

A Additional Tables

Table A.1: Geographical Distribution of Bargaining Events (2012–2019)

State	Count	State	Count	State	Count	State	Count
California	39,165	Florida	3,417	South Carolina	719	Connecticut	187
Washington	18,560	New Jersey	3,071	Utah	719	Kentucky	181
Texas	11,112	Pennsylvania	2,770	Wisconsin	693	Idaho	170
Illinois	11,077	North Carolina	1,889	New Hampshire	640	Nebraska	150
Massachusetts	9,740	Georgia	1,710	Rhode Island	451	Kansas	104
Maryland	8,931	New York	1,560	Hawaii	375	Alabama	100
Virginia	8,347	Minnesota	1,117	Oklahoma	341	Maine	82
Oregon	4,751	Michigan	1,039	Indiana	312	Delaware	67
Colorado	3,620	Ohio	907	Missouri	285	Arkansas	18
Arizona	3,611	Nevada	901	Louisiana	251	Iowa	7
District of Columbia	3,504	Tennessee	846	New Mexico	203	West Virginia	1

Notes: This table shows the geographical distribution of 147,701 property bargaining events across 44 states in U.S. which happened between 2012 and 2019 in our data.

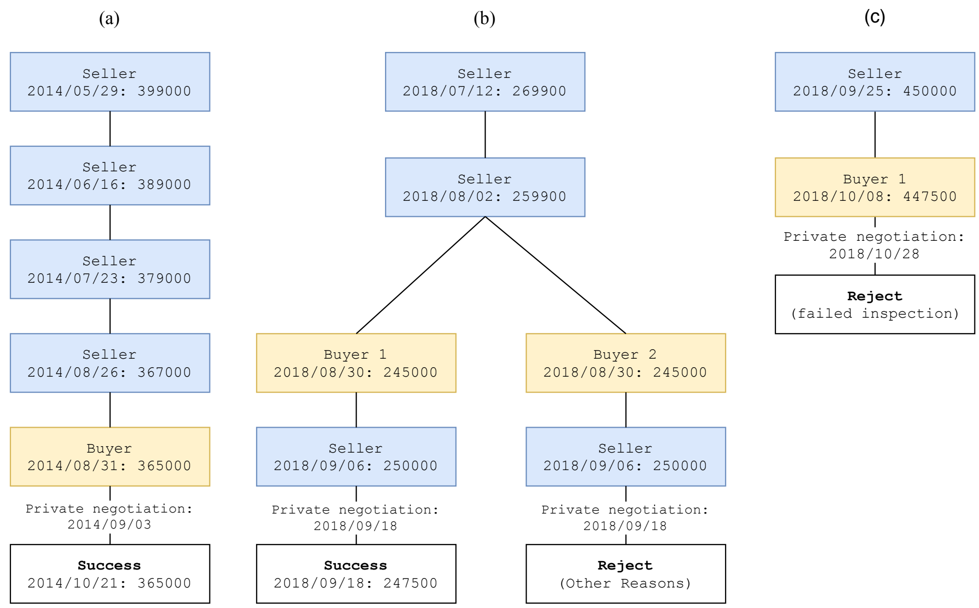
Table A.2: Dictionary of Variables in Table 1

Panel	Variable	Definition
A	Property Age (Years)	Number of years since the property was built
A	Buyers Represented by Redfin	Number of buyers represented by Redfin
A	Listing Price	Initial listing price of the property made by the seller
A	Fraction Revised	Fraction of bargaining events that the seller has revised the initial listing price
A	Revisions	Number of times which the seller revises the listing price
A	Revisions Before First Offer	Number of times which the seller revises the listing price before a buyer makes the first offer to the property
A	Final Listing Price	The listing price after the seller's final revision
A	Final Listing Price/ Initial Listing Price	The ratio of the final listing price to the initial listing price
A	Duration Until Off-Market (Days)	Number of days when the property is listed on Redfin
A	Sales Price	Sales price of the property
A	Sales Price/ Initial Listing Price	The ratio of sales price to initial listing price
A	Sales Price Per Square Feet	Sales price of the property per square feet in 2010$
A	No. Bargaining Events	Number of bargaining events in our data. A bargaining event starts when a seller lists the property on Redfin, ends when the deal is closed (i.e. a success or a rejection to a buyer's offer)
B	Agent Experience	Number of bargaining events which the Redfin agent of a buyer has participated before
B	No. Attempted Purchases	Number of bargaining events that a buyer has participated by making an offer
B	No. Purchases	Number of properties which a buyer has purchased successfully on Redfin
B	No. Buyers	Number of buyers who have made at least one offer in the whole data
C	No. Offers Per Bargaining Event	Number of offers a participating buyer has made in a given bargaining event;
C		participating means the buyer has made at least one offer in a given bargaining event
C	Financing Contingency Waived	Whether a buyer has waived the financial contingency in a given bargaining event A financial contingency allows the buyer to terminate the deal if the mortgage loan is not approved
C	Inspection Contingency Waived	Whether a buyer has waived the inspection contingency in a given bargaining events A inspection contingency entitles a buyer to inspect the property
C	Escalation Clause	Whether a buyer has an escalation clause in a given bargaining events An escalation clause is a clause in a contract which entitles a buyer to change the offer price if another buyer makes a higher offer price
C	Pre-Inspection	Whether a buyer has pre-inspected the property in a given bargaining event A pre-inspection is a thorough examination of conditions of a home by a licensed inspector
C	Client Letter	Whether a buyer has submitted a client letter to the seller in a given bargaining event A client letter connects the buyer to the seller on a personal level, and may increase the buyer's chance of successfully purchasing the property
C	No. Buyer-Event Pairs	Number of pairs of buyer and bargaining event that the buyer has participated in
D	No. Buyers Represented by Redfin	Number of buyers represented by Redfin in a given round
D	No. rounds a buyer participated	Number of rounds a buyer has participated in, including the current round, up to a given bargaining event
D	No. Rounds	Number of bargaining rounds in our data. A round starts when a seller sets or revises the listing price and buyer(s) make offers, ends when the seller revises the listing price again
E	Year Built	Year when the property was built
E	No. Bedrooms	Number of bedrooms in the property
E	No. Bathrooms	Number of bathrooms in the property
E	Approximate Square Feet	Approximate size of the property in square feet
E	Walk Score	Walk score measures the how walkable a commute from the property is on a scale from 0 to 100 A 0-49 score indicates that commute is car dependent A 50-89 score indicates that commute is somewhat walkable. A 90-100 score indicates that it's "Walker's Paradise" Detailed explanation is provided on Redfin's website at https://www.redfin.com/how-walk-score-works
E	Transit Score	Transit Score is a measure of how well a property is served by public transit on a scale from 0 to 100 A 0-49 score indicates that the property is only served by a few public transit options A 50-89 score indicates that the property is served by good or excellent public transportation options A 90-100 score indicates that the property is served by world-class public transportation
E	Bike Score	Bike score measures whether the property is serviced by infrastructure suitable for biking on a scale from 0 to 100 A 0-49 score indicates that the property is serviced by minimal bike infrastructure A 50-89 score indicates that the property is serviced by some convenient bike infrastructure A 90-100 score indicates that the property is "Biker's Paradise"
E	No. Properties	Number of unique properties in our data

Notes: This table presents the definition for all variables in table 1, which provides summary statistics for our main data set. Panel A, B, C, D and E are at the bargaining event level, buyer level, buyer-event level, round level, and property level, respectively.

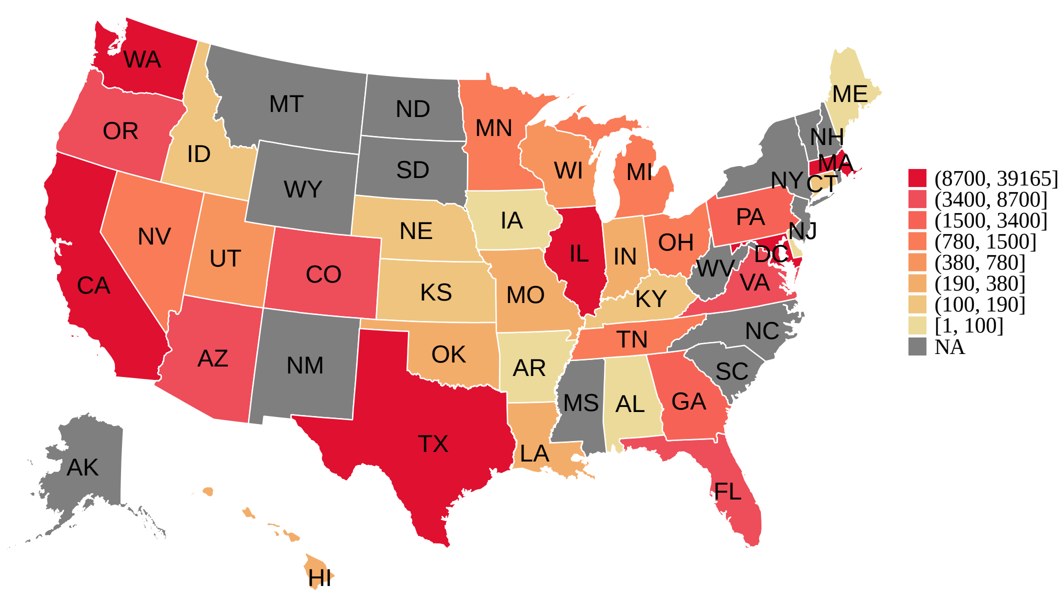
B Additional Figures

Figure A.1: Illustration of Bargaining Events



Notes: This figure illustrates three different bargaining events.

Figure A.2: Geographical Distribution of Bargaining Events (2012–2019)



Notes: This map shows the geographical distribution of 147,701 property bargaining events across 44 states in U.S. that happened between January 2012 and September 2019 in our data. Table A.1 shows the total number of events happened in each state. States colored in gray indicate that data in that state is unavailable.

C Data Construction

C.1 Illustration of Bargaining Events

Figure A.1 (a) shows the situation which one seller bargains with one buyer for multiple rounds. The seller lists the property on the Redfin platform on 2014/05/29 with an initial listing price of \$399,000. Then, on 2014/06/16, the seller revises the price to \$389,000. That constitutes the second round of bargaining. On 2014/07/23, the seller further revises the price as \$379,000. According to our definition, that is the third round of bargaining. On 2014/08/26, the seller revises the price again as \$367,000. On 2014/08/31, one buyer makes an offer with a price of \$365,000. That completes the fourth round of bargaining. On 2014/09/03, the seller takes the property off the Redfin market, and the deal moves to private negotiation stage. After private negotiation, the seller successfully sells the property at \$365,000. This process demonstrates the first situation of a complete event tree.

Figure A.1 (b) shows the situation which one seller bargains with multiple buyers for multiple rounds. The seller lists the property on the Redfin platform on 2018/07/12 with an initial listing price of \$269,900. Then, on 2018/08/02, the seller revises the price to \$259,900. That constitutes the second round of bargaining. On 2018/08/30, both buyer 1 and buyer 2 make respective offers with a price of \$245,000. That completes the second round of bargaining. On 2018/09/06, the seller revises the listing price to \$250,000 to bargain with the two buyers. That constitutes the third round of bargaining. On 2018/09/18, the seller takes the property off the Redfin market, and the deal moves to private negotiation stage. After private negotiation, the seller successfully sells the property at \$247,500 to buyer 1, and rejects buyer 2's offer due to other reason. This process demonstrates the second situation of a complete event tree.

Figure A.1 (c) shows the situation which one seller bargains with one buyer for only one round. The seller lists the property on the Redfin platform on 2019/09/25 with an initial listing price of \$450,000. Then, on 2018/10/08, one buyer makes an offer with a price of \$447,500. That completes the first round of bargaining. On 2018/10/28, the seller takes the property off the Redfin market, and the deal moves to private negotiation stage. After private negotiation, the seller rejects the buyer due to failed inspection. This process demonstrates the third situation of a complete event tree.

C.2 Construction of Surplus Measurement

To make sure that "offer price surplus ratios" are between 0 and 1, we use the following procedures to readjust the ratios:

1. Adjust $\gamma_t = 0$ if revision price$_t$ < offer price$_t$ and revision price$_{t-1}$ > offer price$_t$. In this case the ratio will be less than 0, and the surplus goes to buyer since the seller accept the revision price in bargaining round t, which is less than the offer price in the same round.

2. Adjust $\gamma_t = 0$ if revision price$_t$ < offer price$_t$ and revision price$_{t-1}$ < offer price$_t$ and $\gamma_t > 1$. In this case the surplus goes to buyer for the same reason as step 1.

3. Adjust $\gamma_t = 0$ if revision price$_{t-1}$ = offer price$_t$ and revision price$_t$ < offer price$_t$. In this case the ratio is meaningless, and the surplus goes to buyer for the same reason as step 1.

4. Adjust $\gamma_t = 0.5$ if revision price$_{t-1}$ = offer price$_t$ and revision price$_t$ = offer price$_t$. In this case the ratio is also meaningless, but the surplus is shared by both seller and buyer.

5. Adjust $\gamma_t = 1$ if revision price$_t$ > offer price$_t$ and revision price$_{t-1}$ < offer price$_t$. In this case the ratio will be less than 0, and the surplus goes to seller since the seller accept the revision price in bargaining round t, which is higher than the offer price in the same round.

6. Adjust $\gamma_t = 1$ if revision price$_t$ > offer price$_t$ and revision price$_{t-1}$ > offer price$_t$ and $\gamma_t > 1$. In this case the surplus goes to seller for the same reason as step 5.

7. Adjust $\gamma_t = 1$ if revision price$_t$ > offer price$_t$ and revision price$_{t-1}$ = offer price$_t$. In this case the ratio is meaningless, but the surplus goes to seller for the same reason as step 5.