Observational Learning and Information Disclosure in Search Markets

Ziying Fan, Xi Weng, Li-An Zhou and Yiyi Zhou*

May 2021

Abstract

This paper studies the role of observational learning in search markets where buyers do not take the list price as a take-it-or-leave-it offer. Using a unique data of the Beijing housing market, we estimate a structural model in which buyers infer a seller's reservation value from the home's list price, time on market and in-person home viewings and then, they decide whether to view the home and how much to offer. We use the estimated model to quantify the welfare impacts of different information disclosure rules. We find that buyer surplus is reduced and seller surplus is increased on average if only the time-on-market information is disclosed. However, disclosing the home-viewing information in addition to the time-on-market information has very different impacts on individual homes. We find that due to the disclosure of this additional information, buyers are slightly better off while sellers are slightly worse off on average.

*Ziying Fan: Shanghai University of Finance and Economics (e-mail: fan.ziying@mail.shufe.edu.cn); Xi Weng: Peking University (e-mail: wengxi125@gsm.pku.edu.cn); Li-An Zhou: Peking University (e-mail: zhoula@gsm.pku.edu.cn); Yiyi Zhou: Stony Brook University (e-mail: yiyi.zhou@stonybrook.edu). We thank Lianjia for its generosity of providing us with the data. We are grateful for useful comments from Andrew Ching, Ali Hortaçsu, Ting Liu, Steven Stern and participants in seminars and conferences. We also thank Boyang Li for his excellent research assistance. The authors are solely responsible for all errors.

1 Introduction

In many markets, sellers list their product with a list price on the market, prospective buyers must incur a cost to know their private valuations and after that, they make offers to the seller and even bid against each other if multiple buyers are interested in the same product. Examples include houses, used cars, and many consumer-finance products. In such markets, the sale time and sale price of a listed product depend not only upon buyers' preferences and search costs, but also upon the buyers' perception of the seller's reservation value which is usually unknown to buyers. The belief of buyers who arrive later in time can be influenced by the choices of buyers who arrive earlier and the choice of the seller earlier on, which is commonly referred as observational learning in the literature. This paper estimates a structural model of observational learning applied to search markets where potentially multiple buyers bid for the same product. Since observational learning is tightly linked to what information buyers observe when they make their search and bid decisions, an additional goal is to quantify the welfare effect of different information disclosure rules in such search markets.

We study one particular setting, the Beijing housing market, for which we have access to a unique data set that contains detailed information of a large number of home listings at Lianjia, the largest real estate company in Beijing, from August 2015 to July 2016. The main novelty of the data is the record of all list price changes and all home viewings made by prospective buyers between initial listing and sale agreement. The housing market provides an ideal setting for our study. Homes are the largest consumption and investment item for most families. Home buyers usually spend a large amount of time to acquire information before purchase, including browsing brokers' websites, viewing homes in person, hiring professionals to inspect homes, and so on. They are also largely influenced by others' opinion and choices. Moreover, list price does not have full commitment in this market; instead, the transaction price is usually achieved through buyers submitting offers to the seller who decides whether to accept and which offer to accept. In our data, 81% houses are sold below their list price, 15% houses are sold exactly at the list price, and the remaining 4% are sold above the list price. The fact that the transaction price can deviate from the list price has been widely documented in the literature of housing markets, for example, Han and Strange (2016)[1].

One uniqueness of our setting is that during our sample period, Lianjia not only updates the information of time on market for each home listing but also posts the real-time home viewings on its website. So home buyers in our setting observe a home's time on market and the occurring times of all home viewings in the past when they make their own choices of whether to view the home in person and how much to offer later on. The unique feature enables us to quantify the welfare impacts of disclosing the time-on-market information and the home-viewing information, which has important implications for policy makers.

[1] Han and Strange (2016) analyze an American home buyer and seller survey data conducted by the National Association of Realtors. They report that over the 2003-2006 boom, 57% were sold below the list price, 29% were sold at the list price, and the remaining 14% of houses were sold above the list price. Turning to the 2007-2010 bust, the share of below-list sales increased to 74%, while the share of at-list sales reduced to 17%.

Using the unique data on the Beijing housing market, we estimate a structural model that incorporates several key features of housing markets in general. In the model, sellers are passive. Each seller has a reservation value which is the lowest offer that he will accept. A seller's reservation value can be interpreted as the value of his outside option. Prospective buyers are short-lived. Buyers do not know the seller's reservation value and infer it from the home's list price, time on market, and home viewings. After they form a belief, they decide whether to view the home to know their private values by paying a viewing cost. If a buyer is interested in a home, she makes offer to the seller who decides whether to accept. If multiple buyers are interested in the same home, a bidding war may occur. We show that the equilibrium bidding strategy has a cutoff solution: there exists a unique cutoff which is higher than the list price such that buyers with private values lower (higher) than that cutoff choose to reject (accept) the list price and bid their true values if they need to bid. As a result, the equilibrium transaction price can be either higher than, or equal to, or lower than the list price and moreover, the density has a mass point at the list price. This prediction is perfectly in line with the widely documented empirical fact of various housing markets, including the Beijing housing market.

Economists, marketers and policy makers have long sought to understand how information affects individuals' decision making and welfare. Understanding the impacts of different information disclosure rules is crucial for marketers' optimal choice of whether to provide such information and policy makers' optimal choice of whether to regulate the disclosure of such information. With the estimates, we conduct a series of counterfactual exercises to quantify the welfare effect of the disclosure of two important pieces of information that buyers observe in our setting: time on market and home viewings. We consider a no information scenario in which buyers do not observe these two pieces of information and another partial information scenario in which buyers observe the time on market only but not the home viewings. The effect of disclosing the time-on-market information is measured by the differences between the partial information model and the no information model, while the effect of disclosing the home-viewing information in additional to the time-on-market information is measured by the differences between our benchmark model and the partial information model. We find that if only the time-on-market information is disclosed, the transaction price is increased, buyer surplus is reduced and seller surplus is increased on average. Moreover, we find that the additional impacts of disclosing the home-viewing information are ambiguous. For some homes, the transaction price is increased, buyers are worse off while sellers are better off. However, the impacts can be totally opposite for some other homes. On average, buyers are slightly better off while sellers are slightly worse off due to the disclosure of this additional information.

Our results shed important light on intermediaries' information design and information disclosure regulations in the housing market and other search markets with similar features. The role of the time-on-market information has been studied extensively in the literature and has also been the focus of many policies and regulations. Conventional studies (e.g., Taylor (1999)) have emphasized that buyers do not have perfect information about home quality so a property staying long on the market is perceived by buyers as having low quality. Empirically, Tucker et al. (2013)

examine a policy adopted by Massachusetts in 2006 that prohibits home sellers from resetting their properties' time on market through relisting. They explore this exogenous policy change and find that the sale price is $16,000 lower due to the availability of true time-on-market information.

Different from those conventional studies, we emphasize that home buyers do not take the list price as a take-it-or-leave-it offer and they learn of the seller's reservation value. If a home has spent too long on the market, buyers speculate that the seller may have a high reservation value so that previous buyers do not buy it. As a consequence, the disclosure of the time-on-market information would push the sale price up, which is opposite to the prediction of the conventional quality-learning story. In section 2, we provide empirical evidence that home buyers do learn of sellers' reservation value in our setting. We also argue that the information asymmetry on home quality is not a big concern in our setting because almost all homes in Beijing are apartments in high buildings and home quality is largely captured by those observed attributes. Notice that we do not intend to rule out the quality learning story which may capture the main feature of other housing markets where information asymmetry on home quality is more prevalent. We do not incorporate quality learning into our framework mainly because doing that would impose challenges both on our theoretical analysis of buyers' equilibrium strategy and on our empirical identification. Nevertheless, we believe that by highlighting buyers' learning of sellers' reservation value, our paper provides a new angle for studying the role of time on market in housing markets. In this sense, our paper is complementary to the existing literature that has mainly focused on quality learning.

In addition, this paper also examines the implications of disclosing the home-viewing information. To the best of our knowledge, this paper is the first study that examines the welfare effect of disclosing this kind of information in the housing market. Regulating the disclosure of the information related to search activities of previous buyers has received scant attention from the literature and little is known about its welfare impacts. This is particularly meaningful in our setting because Beijing government banned real estate platforms from disclosing this information in February, 2021. While we do not have data covering the time periods just before and after this latest regulation to conduct difference-in-difference types of analysis to evaluate its impact on housing price and other endogenous outcomes, our counterfactual analysis based on our sample (2015-2016) suggests that on average, this ban would slightly increase the transaction price, harm home buyers and benefit home sellers. As a matter of fact, besides China's case, some real estate platforms in western countries (e.g., Redfin) also choose to disclose the number of in-person home tours and most of them disclose the number of online views.

This paper contributes to several different strands of literature. First are the literature of observational learning. Most previous papers in this field have assumed that buyers have imperfect information about product quality and infer it from the choices of previous buyers. This can lead to herd behavior: individuals basing their decisions purely on what others have done. Examples of theoretical studies include Banerjee (1992), Bikhchandani et al. (1992), and Hendricks et al. (2012), while examples of empirical studies include Cai et al. (2009), Zhang (2010), Newberry (2016), and Luca (2016). We contribute to this literature by empirically studying the welfare effects of observational learning in search markets where the list price does not have full commitment. Moreover, by highlighting buyers' learning of sellers' reservation value, our study complements the existing literature that has mainly focused on quality learning.

This paper contributes to the growing literature on the microstructure of housing markets. Han and Strange (2015) provide a comprehensive summary of the recent literature in this field. A large amount of papers have focused on the relationship between time on market and sale price, such as Levitt and Syverson (2008), Hendel et al. (2009), Genesove and Han (2012), and Tucker et al. (2013). Another strand of literature aims to explain the spread of the sale price around the final list price, including Horowitz (1992), Yavas and Yang (1995), Chen and Rosenthal (1996a), Chen and Rosenthal (1996b), Merlo et al. (2015), Albrecht et al. (2016), and Han and Strange (2016). With the novel data, we contribute to this literature by providing new facts of housing markets regarding the dynamic patterns of home viewings. Moreover, we provide a coherent framework for the empirical analysis of housing markets that models buyers' learning, searching, and bidding. The model predicted spread of the sale price and list price is perfectly in line with its empirical distribution.

We are also related to the empirical literature on consumer search, for example, Hortaçsu and Syverson (2004), Hong and Shum (2006), De los Santos et al. (2012), Honka (2014), Moraga-González et al. (2018), Yavorsky et al. (2020) and Allen et al. (2019). Most of these papers have focused on the implication of search frictions on firms' market power where firms choose a posted price that consumers either take it or leave it. One exception is Allen et al. (2019) who introduce price negotiation into the search framework to study the Canadian mortgage market. We contribute to this literature by studying the implications of buyers' observational learning in a search environment.

Lastly, this study contributes to the large body of empirical literature on information disclosure. Most studies in this area have focused on the effect of product price disclosure or quality disclosure on consumer choice and market competition. For example, Schultz (2005), Rossi and Chintagunta (2016), and Luco (2019) study the implications of price disclosure, while Jin and Leslie (2003) and Bollinger et al. (2011) study the effect of quality disclosure. Different from the literature, we study the welfare impacts of information that affect buyers' observational learning.

The rest of the paper is organized as follows. Section 2 presents details on the Beijing housing market and introduces our data set. Section 3 presents the model. Section 4 discusses the estimation strategy and Section 5 describes the empirical results. Section 6 analyzes the welfare impacts of different information disclosure rules. Finally, section 7 concludes.

2 Institutional Background and Data Description

Our primary data are obtained from Lianjia, the largest real estate company in Beijing, China. In this section, we will first provide a brief description of the Beijing housing market, including the roles of real estate agents in home searching, listing, and transactions. Along with that, we will also describe what information real estate companies provide on their website from which prospective home buyers can obtain information before their search and purchase decisions. After that, we will describe the data that we use in the study and provide descriptive evidence outlining the key features of the Beijing housing market that we want to capture. Finally, we will provide suggestive evidence that home buyers do learn of sellers' reservation value.

2.1 Institutional Background

In Beijing, 87.2% of transactions of second-hand residential properties are represented by real estate agents in 2016 which is our sample period. Lianjia was the largest real estate company in Beijing during this time period, not only in terms of its market share of home listings but also in terms of its market share of home transactions. During 2016, Lianjia listed more than 80% of all second-hand residential properties that were on sale in Beijing and participated in 54% of all transactions that occurred during this time. In contrast, the market shares of home transactions of the second, third, and fourth top real estate companies were 13%, 5%, and 4%, respectively.[2]

Noticeably, many home listings are marketed under sole agency agreements. Under this type of agreement, a home is represented by a single real estate company that coordinates all market related activities regarding that home from the initial listing time until the time it is either sold or withdrawn. First, the representing agent lists the home on the company's website by publishing a sheet of home characteristics, photos, and a list price. The list price can be revised at any time if the seller wants to, and most of the time, after consulting with the representing agent.

The vast majority of residential homes in Beijing are apartments in high buildings. Most of those high buildings are built after 1990 and the floor plans are quite similar. It is fair to say that the heterogeneity of home quality after accounting for the observed home attributes, including the local amenities, is relatively small in Beijing, compared to single family homes in western countries. Potential buyers usually start their home search by browsing the websites of real estate companies where they can obtain the information of what homes are on sale and detailed information of those homes. For each listed home, the selling agent provides a detailed description in words along with tons of photos and even virtual tours. If a buyer does not have an Internet access, she can visit a physical store of those real estate companies to get those information. If a buyer is interested in some homes, she can request to view those homes in person accompanied by the agents who work at the same company as the selling agents of those properties.

If a potential buyer is still interested in a home after viewing it, she can make an offer to the seller. If the offer is equal to or higher than the list price, the seller is committed to sell it. This commitment is often written into the contract between the seller and his agent in the form of a clause requiring him to pay the full commission fee to his agent if he rejects any offer at or higher than the list price. If the offer is lower than the list price, the seller can either accept or reject the offer. If the seller rejects it, the potential buyer can either make another offer or walk away. If multiple buyers submit offers at the same time, a bidding war occurs. If an agreement occurs,

[2]Our general understanding of the Beijing housing market is from various industrial reports and news, including Sohu News, see https://m.sohu.com/n/458280155/.

both parties sign a contract and start the administrative procedure of completing the transaction. Throughout the process, the seller and the potential buyers are in touch only through their agents.

If a transaction occurs, the buyer pays her agent a commission fee which is usually a fixed proportion of the transaction price. The commission fee differs across real estate companies and over time. If the buyer has signed a sole agency agreement with a real estate company, she usually gets a discount in the commission fee. In this case, the buyer is represented by the agents of the representing company only. If she is found that she eventually buys a home through agents in other companies, by law she needs to pay a full commission fee to the company that she has signed a sole agency agreement with.

Like almost all other real estate companies in Beijing, Lianjia updates the information of time on market for each home listing. In addition, it also posted the real-time home viewings on its website before February, 2021 when the Beijing government banned the disclosure of home-viewing information. If any home buyer physically visited a property before February 2021, the time of this visit was posted on the website, see Figure 6. In contrast, most real estate companies in other countries usually publish the information of time on market but not the in-person home viewings. One reason is that in other housing markets, home buyers do not need the accompany of their agents, if they have, when they physically visit a property. They can just stop by and view the property they are interested in when the sellers or seller agents hold open houses. As a result, it is practically challenging to collect this type of information. However, this is not the case in the setting of Beijing housing market, where home buyers have to contact the seller agents through their own agents to arrange for home tours. The real estate companies collect the home-viewing information, with the purpose of learning bout their representing buyers and the market value of their listed properties.

2.2 Data Description

Our sample includes 27,703 apartments that are listed and sold through Lianjia from August 2015 to July 2016. All of those sellers have signed sole agency agreements with Lianjia and therefore, all in-person home viewings are witnessed by Lianjia and are included in our data. We only include apartments that are located in the five core districts of Beijing - Dongcheng, Xicheng, Chaoyang, Haidian, and Fengtai, because the housing market conditions of other districts are quite different from the five core districts.

For each apartment in the sample, we know the date when it is initially listed, the date when it is sold, the transaction price, whether it is mortgaged, its address, number of bedrooms, number of living rooms, number of bathrooms, HOA fees, property size (in square meters), building year, the total floors of the building, and the floor number of the apartment. We also know the initial list price, and all subsequent revisions of the list price. Most importantly, we know the occurring time of all home viewings since the first date of listing until the final transaction.

Table 1: Summary Statistics of Homes in the Sample

Variable	Mean	Std. dev	Min	P25	P50	P75	Max
Weeks on sale	5.439	4.955	1	2	4	8	25
Number of home viewings	11.629	11.348	1	4	8	16	79
Initial list price (Million CNY)	2.989	1.204	1	2.1	2.7	3.6	9.9
Last list price (Million CNY)	2.985	1.208	1	2.1	2.7	3.6	9.9
Transaction price (Million CNY)	2.938	1.196	0.96	2.08	2.68	3.53	9.8
Property age	18.845	9.696	1	12	18	26	63
Property size (square meters)	67.191	23.301	25	52.633	61.2	77	246

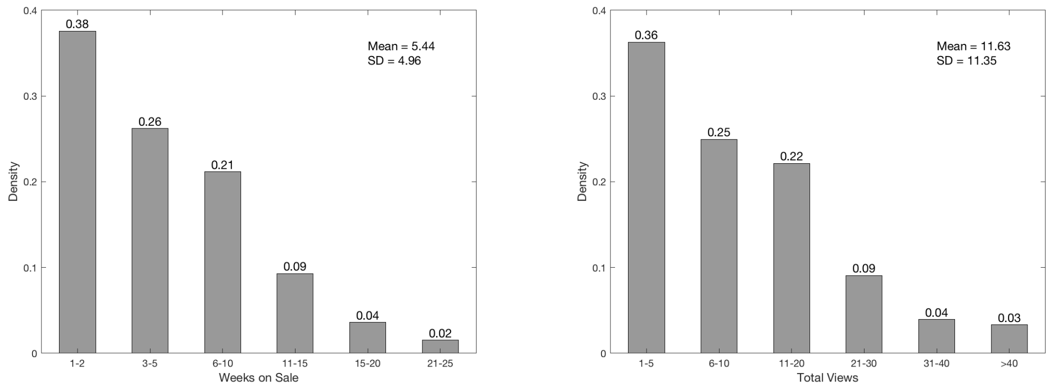
Note: The data includes 27,703 apartments that are represented by Lianjia, listed for sale and sold from August 2015 to July 2016.

Table 1 reports summary statistics of the data. The weeks on sale of an apartment is the number of weeks since the seller initially lists his apartment until he signs a contract with the buyer. On average, home sellers in Beijing spent 5.439 weeks on market with a standard deviation of 4.955. Figure 1a presents the distribution of weeks on sale across all apartments included in our data. 38% apartments were sold within two weeks since being initially listed, 26% were sold between two and five weeks, 21% were sold between five and ten weeks, 9% were sold between ten to fifteen weeks, and the remaining 5% were sold longer than fifteen weeks. According to Genesove and Han (2012), American home sellers spent 7.32 weeks on the market between 1986 to 2007, with a standard deviation of 5.60 weeks. These statistics indicate that Beijing home sellers in our data spent two weeks shorter on market than American home sellers on average.

The number of home viewings is the number of all in-person home tours made by prospective buyers before the home is sold. In our data, the mean is 11.629, the median is 8, and the maximum is 79. Figure 1b presents the distribution of total home viewings. As far as we are aware of, Merlo and Ortalo-Magne (2004) is the only paper that reports the total number of buyer viewings since initial listing. In their sub-sample of 200 properties located in the local market within the Greater London metropolitan area, the average is 9.5, the median is 7, and the maximum is 51. Compared with their statistics, Beijing home sellers accumulated more home viewings from potential buyers than the U.K. home sellers on average, and the heterogeneity in Beijing was also much larger than that in U.K..

The list price ranges from 1 million to 9.9 million CNY, because we drop the apartments whose last list price is either lower than 1 million or higher than 10 million CNY. The average list price is 2.989 million CNY and the average transaction price is 2.938 million CNY. In the next subsection, we will show more details about the list prices and the transaction price. The average property age is around 19 years and the average property size is 67 square meters.

Figure 1: Sale Duration and Number of Home Viewings



(a) Weeks to Sale (b) Number of Home Viewings

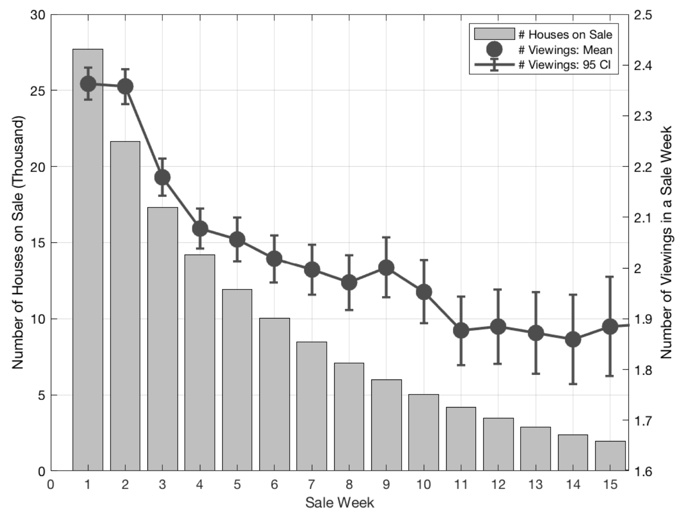
Note: The data includes 27,703 apartments that are represented by Lianjia, listed for sale and sold from August 2015 to July 2016.

In what follows we present several key features of the data. These features will be captured by the model that we present in the next section.

Feature 1: The average number of home viewings declines in the time on market.

The uniqueness of our data is that it includes the information of when buyers view a home, allowing us to examine the dynamics of home viewings. In Figure 2, the horizontal axis is the sale week that is the number of weeks a home has been on sale, the bars indicate the number of homes that are still on sale in each sale week, and the dots indicate the average number of home viewings during each sale week. We do not report the statistics after 15 weeks just because very few homes are left after being on market for 15 weeks and the statistics become very noisy. Unsurprisingly, the number of homes for sale declines in the sale time, because homes exit along the process due to sale. The figure also clearly shows that the average number of home viewings declines in the sale time, starting from around 2.4 in the first week and dropping to 1.9 in the 15th week.

Figure 2: Homes on Sale and Home Viewings by Sale Week

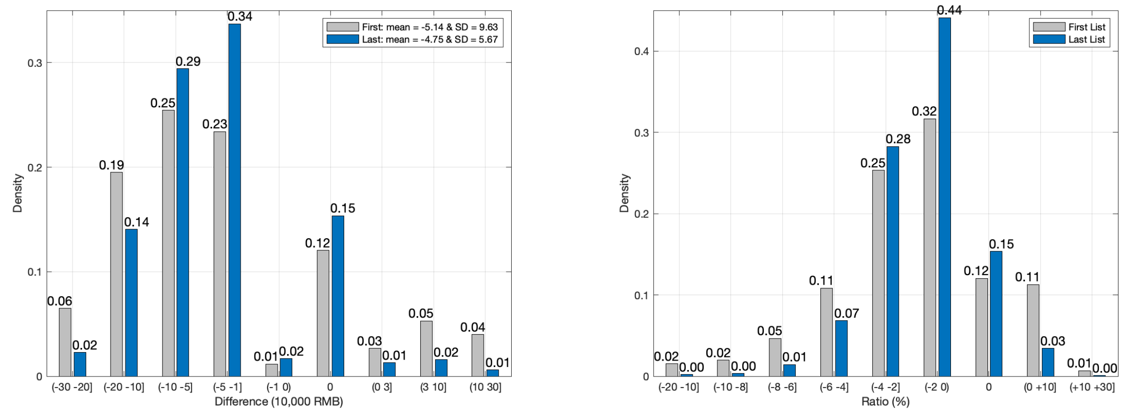


Note: The data includes 27,703 homes that are listed for sale on Lianjia.com and sold from August 2015 to July 2016. The bars indicate the number of homes on sale in each sale week, and the dots indicate the average number of home viewings in each sale week.

Feature 2: The transaction price can be higher than, or equal to, or lower than the list price.

Among the homes included in our data, 81.16% are sold below their last prices price, 15.35% are sold exactly equal to their last list price, and the remaining 3.49% are sold above their last list price. For each home, we compute the the difference between the transaction price and the initial or the last list price. We also compute the ratio of the transaction price over the initial or the last list price. Figure 3 plots the distributions of the differences and the ratios. On average, the final transaction price is 47,525 CNY and 1.63% lower compared with the last list price, and 51,399 CNY and 1.82% lower compared with the initial list price.

Figure 3: Transaction Price v.s. List Prices



(a) Difference

(b) Ratio

Note: The data includes 27,703 homes that are listed for sale on Lianjia.com and sold from August 2015 to July 2016.

Feature 3: Homes' observed attributes explain most of the variations in the transaction prices.

We run a hedonic regression of a home's transaction price on an extensive set of home characteristics that are available in the data set, including 167 neighborhood dummies to indicate the home location, square meters of floor space, number of living rooms, bathrooms, bedrooms, the total floors of the building, the house age, and year-month dummies indicating when the home is initially listed. The R square is 0.89, indicating that the a home's observed attributes provide a very accurate prediction of its actual transaction price.

Feature 4: List prices are highly sticky.

The list prices are highly sticky in the Beijing housing market. Among the 27,703 home sellers in the data, 17,457 sellers (63%) have never revised the list price before sale, 6,966 sellers (25.2%) have revised the list price only once, 2,360 sellers (8.5%) have revised twice, and the remaining 920 sellers (3.3%) have revised more than twice. This feature is not specific to our data. Merlo et al. (2015) also document that the list prices are highly sticky in the UK housing market.[3]

2.3 Suggestive Evidence of Observational Learning

Below we further explore the dynamic pattern of home viewings and the relationship between the home viewings and the transaction, with the purpose of providing some suggestive evidence that home buyers do learn of sellers' reservation value.

[3]In their data, 77% home sellers have never changed the initial list price, 18% have changed once, 4% have changed twice, and the remaining 1% have changed more than two times.

2.3.1 Observational Learning

The declining pattern of the number of home viewings in sale time, shown by Figure 2, can be explained by two different stories. It is possible that home buyers do not have all relevant information about the property or the seller before visiting the home. So they infer these information based on home viewings by previous buyers. The fact that a home has accumulated a large number of viewings is a signal that the home has some flaws or/and the seller has a high reservation value so that previous buyers did not buy it. Consequently, buyers who arrive later in time are less likely to view a home that has accumulated more viewings. This is the observational learning explanation that we have mentioned in the introduction. It is also possible that there is no information asymmetry. Buyers have full information about the home quality and the seller's reservation value. However, homes that have low quality or/and the sellers who have high reservation value are less likely to be sold and hence are left on market longer. As a consequence, the number of home viewings is declining in the sale time. This is the adverse selection explanation.

To distinguish these two stories, we run a regression of a home's viewings newly occurred in a week on the number of viewings that the home has accumulated before that week, the home's list price in that week, controlling for the home fixed effects and a set of calendar week dummies that capture aggregate shocks. Notice that the dependent variable is the viewings occurred during a given sale week and the explanatory variables are all measured at the beginning of that sale week. Moreover, since home fixed effects are included in the specification, the error term captures those time-varying shocks that affect the number of viewings of a home during a given sale week. It is reasonable to assume that those contemporary shocks in a sale week are not correlated with the accumulated viewings before that week.

Table 2 reports the estimation results. We include home fixed effects in all three specifications. We also include the calendar week dummies in Column (II) and (III). In Column (III), we interact the accumulated viewings with dummy variables for different ranges of the accumulated viewings. Across all specifications, the coefficient before the accumulated viewings is negative and significant at 1% significance level, suggesting that a home with a larger amount of accumulated viewings is likely to have fewer home viewings later on. In Column (III), the estimates of the coefficients on the interactions indicate that this negative effect becomes stronger when the number of accumulated viewings is larger. The price coefficient is negative and significant at 1% level in all specifications. It implies that within a home, the number of home viewings in a week declines in its list price. Overall, these results are consistent with the story that observational learning is occurring.

2.3.2 Learning of Sellers' Reservation Value

Although we have shown some suggestive evidence that buyers do learn from the viewing choices of previous buyers, we do not know whether buyers learn about the home quality or learn about sellers' reservation value. In section 2.2 we have shown that the 89% of the variation in homes' transaction price has been explained by homes' observed attributes. Since the home at-

Table 2: Newly Occurred Home Viewings and Accumulated Home Viewings

	(I)	(II)	(III)
$AccumulatedViewings$	-0.027***	-0.028***	
	(0.001)	(0.001)	
$\times 1(AccumulatedViewings \in [1, 10])$			-0.007***
			(0.003)
$\times 1(AccumulatedViewings \in (10, 20])$			-0.012***
			(0.002)
$\times 1(AccumulatedViewings > 20)$			-0.027***
			(0.001)
$ListPrice$	-1.371***	-2.240***	-2.181***
	(0.140)	(0.144)	(0.146)
Constant	1.464***	1.588***	1.580***
	(0.012)	(0.130)	(0.130)
Home FE	Y	Y	Y
Week dummies	N	Y	Y
R^2	0.07	0.09	0.09
# Home Obs.	27,703	27,703	27,703
# Home-week Obs.	150,887	150,887	150,887

Note: The table displays the results of a regression of the number of home viewings during a sale week on the accumulated home viewings before that sale week, and the list price at the beginning of that sale week. Home fixed effects are included in all specifications. Data includes 150,887 home-week level observations representing 27,703 homes that are listed and sold by Lianjia from August 2015 to July 2016. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

tributes included our hedonic regression are also observed by all prospective buyers before they view homes in person, the fact that the unobserved factors only explains 11% of the variations in the transaction price suggests that the heterogeneity of unobserved home quality is not substantial. Hence, it is fair to say that the information asymmetry regarding home quality, if exists, is not substantial in the Beijing housing market. This is partly because almost all residential properties in Beijing are apartments in high buildings. Buyers can have a good sense of an apartment's quality from the property description (words, photos and even virtual tours) provided by agents before viewing the property in person. Next we provide some extra suggestive evidence that learning of sellers' reservation value is occurring in the housing market.

According to the quality learning story, the observation that a buyer viewed a home but did not buy is a bad signal of the home quality for buyers who arrive later. As a result, buyers are less likely to view a home with more accumulated viewings and they also have lower willingness to pay for a home with with more accumulated viewings. So the quality learning story predicts a negative relationship between a home's transaction price and the number of home viewings. In contrast, if buyers are learning about the seller's reservation value, then the fact that a home has accumulated a large number of non-purchase viewings suggests that the seller has a high reservation value. When buyers perceive that the seller has a high reservation value, they would increase their offers, leading to higher transaction price. Therefore, the transaction price is positively correlated with the number of home viewings.

We run a simple regression of a home's transaction price on a home's total number of home viewings, the total weeks on market, the last list price, controlling for other factors, including home age, home size, home location, and weekly dummies when the transaction occurs. It is worthy pointing out that the coefficients in this regression measure the correlations rather than causal effects because the transaction price and total number of home viewings can be affected by unobserved factors such as home quality, listing strategy, seller motivation, and etc..

Table 3 reports the estimation results. All specifications include home age, home size, home location, and weekly dummies when transaction occurs. Specification (I) and (IV) include the number of home viewings only but not total weeks on sale, Specification (II) includes the total weeks on sale only but not the number of home viewings, and Specification (III) and (V) include both. In the last two specifications, we interact the number of home viewings with dummy variables for different ranges the number of home viewings.

The coefficients before the number of home viewings are all positive and significant, indicating that the transaction price is positively correlated with the number of home viewings and this positive correlation are weaker for homes with a large number of home viewings than those with a small number of home viewings. The coefficient before the total weeks on sale is positive and significant, indicating that the transaction price is also positively correlated with the time on market. In addition, the positive coefficient before the list price suggests that the transaction price is positively correlated with the list price.

We take our finding of a positive correlation as a suggestive evidence that buyers do learn of sellers' reservation price, although we can not say much about the causal effects.

Table 3: Transaction Price and Total Home Viewings

	(I)	(II)	(III)	(IV)	(V)
HomeViews	0.416^{**}		0.337^{*}		
	(0.175)		(0.187)		
$\times 1(HomeViews \in [1,5])$				0.745^{**}	0.703^{**}
				(0.328)	(0.328)
$\times 1(HomeViews \in (5,10])$				0.453^{***}	0.370^{**}
				(0.152)	(0.153)
$\times 1(HomeViews \in (10,20])$				0.160^{**}	0.083
				(0.079)	(0.081)
$\times 1(HomeViews > 20)$				0.067^{**}	0.014
				(0.031)	(0.033)
WeeksOnSale		0.317^{***}	0.358^{***}		0.355^{***}
		(0.067)	(0.074)		(0.074)
LastListPrice	0.889^{***}	0.888^{***}	0.888^{***}	0.889^{***}	0.888^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	-5.490	-5.901	-6.059	-7.345	-7.927
	(7.169)	(7.166)	(7.167)	(7.210)	(7.208)
R^2	0.868	0.868	0.868	0.868	0.868
# Home Obs.	27,703	27,703	27,703	27,703	27,703

Note: The table displays the results of a regression of homes' transaction price on the value of the independent variables when transactions occur. *HomeViews* is the total number of home viewings, *WeeksOnSale* is the total number of weeks on sale, and *LastListPrice* is the last list price. All specifications include home age, home size, home location, and weekly dummies when transaction occurs. Data includes 27,703 homes that are listed and sold by Lianjia from August 2015 to July 2016. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

3 Model

Since our data set only includes properties that are listed and sold via Lianjia, we take the sellers' decision to sell a house and buyers' decision to search a house via this real estate agent company as given. For each home, we examine how prospective buyers update their beliefs about the seller's reservation value based on the history of list prices and home viewings, prospective buyers' decisions of whether to view the home in person, and viewers' decisions of how much to bid for that home. The model we propose incorporates several key features of the home buying process in Beijing illustrated in the previous section.

The next subsections describe the model in detail. We first describe the timing of the events. Then, we describe buyers' utility and their belief of the seller's reservation value. After that, we solve the model and describe the equilibrium strategy. All variables introduced in the model vary at the home level, j, based on observed or unobserved characteristics. To simplify notation we omit the home index j, and will add it back in the next section that describes estimation and identification.

3.1 Timing

A decision period is a week and we use t to denote the number of weeks that a home has been on sale since initial listing. The seller's reservation value is his private information. We model buyers' belief updating of the seller's reservation value and decisions of whether to view the home and if so, how much to bid in each time period t as a three-stage game.

Stage 1: Buyers update their belief about the seller's reservation value. A m_t number of prospective buyers arrive at time t following a process with a probability $\Lambda_t(m_t)$. Buyers know that the generating process of m_t but not the realization of m_t. The list price at time t is revealed to buyers. Buyers observe the history of list prices up to time t and the number of viewers in the past weeks. However, they do not know the seller's reservation value nor their idiosyncratic match values of the house. All prospective buyers form a belief about the seller's reservation value.

Stage 2: Buyers decide on whether to view the home. Each arrived buyer decides whether to view the home by paying a viewing cost. Then, the number of viewers n_t, where $n_t \in \{0, 1, ..., m_t\}$, is realized and revealed to all current buyers and future buyers. After viewing the home, buyers who have viewed the home know their match values. If $n_t = 0$, the home is retained by the seller and will be on sale at time $t + 1$. If $n_t > 0$, we move to the next stage.

Stage 3: Viewers decide on how much to bid. All viewers first decide whether to accept or reject the current list price. If one or more viewers accept the current list price, then the seller commits to sell to the viewer with the highest offer. If no viewer accepts the current list price, then the seller can accept or reject the highest counteroffer. Again, if no sale is made, the home is retained by the seller and will be on sale at time $t + 1$.

3.2 Buyers' Utility

The lifetime utility of a prospective buyer i who arrives at t from buying the home is given by

$$u_{it} = \delta + v_{it} - p_{it},$$

where δ is the home's financial value that all prospective buyers agree on and know without viewing the home in person, v_{it} is buyer i's idiosyncratic match value of the home, and p_{it} is the price that buyer i needs to pay when she buys the home. A buyer who fails to purchase the home receives a payoff of zero.

Buyer i does not know her match value v_{it} before viewing the home in person but knows the distribution with a support of $[\underline{v}, \bar{v}]$. Let $F(\cdot)$ denote the cumulative density function (CDF). To know her exact value, buyer i needs to view the home by paying a viewing cost c_{it} with a CDF of $\Psi(c_{it})$ on $[0, \infty)$. The transaction price p_{it} is not known until an agreement is reached. Later we will describe how the price is determined.

We assume that after consultation with appraisers and the real estate agent, the seller and buyers have a firm assessment of the financial value of the home δ that does not vary over the course of the selling horizon. Following Merlo et al. (2015), we assume that a home's financial value δ is an exponential function of its observed characteristics and an unobserved term:

$$\delta = exp(X\beta + \xi),$$

where X is a vector of observable characteristics of the home (the basis for the traditional hedonic regression prediction of the ultimate transaction price discussed in Section 2.2), and the term ξ captures the unobserved value that are observed by all buyers and sellers but not by econometricians.

Buyers' learning and optimal decisions of viewing and bidding depend critically on the home's common value δ, which in turn depends on a very high dimensional vector of observed home characteristics X as well as unobserved component ξ. To make the estimation computationally feasible, we impose a similar homogeneity assumption as in Merlo et al. (2015) so that we can solve a single problem where the values are defined as differences relative to the home's financial value.[4] In specific, we define the seller's reservation value, buyers' private valuations and bids, list prices and the final transaction price as differences between the actual values and the home's financial value δ. For example, if a seller's normalized reservation value equals to zero, it is equivalent to say that his reservation value equals to the financial value, and if a seller's normalized reservation value is greater than zero, it corresponds to a seller's reservation value that exceeds

[4]In principle, we could treat the estimated hedonic value $exp(X\hat{\beta})$ as a "fixed effect" relevant to a home and solve $J = 27,703$ individual dynamic problems, one for each of the homes in our sample. However, the problem is complicated because of the existence of the unobserved "random effect" ξ. This is a one-dimensional unobserved random variable, and in principle we would need to solve of the J dynamic problems over a grid of possible values of ξ and thereby approximate the optimal viewing and bidding strategies explicitly as a function of all possible values of the unobserved random effect ξ, which would then be "integrated out" in the estimation of the model.

the financial value and so forth. The implicit assumption underlying the homogeneity assumption is that buyers' viewing and bidding behaviors are driven mostly by whether a given seller is perceived to have "a low reservation value" as reflected by the difference of the actual value and the home's financial value. With this homogeneity assumption, we can solve buyers' belief updating, optimal viewing and bidding decisions, and compute the transaction price if sale, where all of those values are normalized by subtracting the actual values with δ.

In our estimation of the model, We further assume that $exp(\xi)$ is a lognormally distributed random variable that is independent of X and $E[exp(\xi)] = 1$. Then, we estimate β via a log-linear regression of the final transaction price on the X characteristics. The underlying assumption of this assumption is that buyers are rational and their beliefs of what the ultimate transaction price are be consistent with the actual values.

3.3 Buyers Learning of the Seller's Reservation Value

Let r_t denote the seller's reservation value at time t *relative to the home's financial value* δ. We do not explicitly model the seller's decision of list prices which is an interesting and complicated problem by itself.[5] We assume that the list price (also expressed as *a difference of the home's financial value* δ) is simply the sum of the seller's reservation value and a non-negative term η that measures the seller's markup over his reservation value when setting the list price:

$$a_t = r_t + \eta. \tag{1}$$

Furthermore, we assume that η follows a distribution with the probability density function (PDF) of h_η on the interval $[0, \bar{\eta}]$, independent of r_t and independent across all sellers.

Buyers do not know the seller's reservation value. Instead, they form a belief about it after observing the list prices up to time t, $\vec{a}_t = [a_1, ..., a_t]$, and the number of viewers in the past periods, $\vec{n}_{t-1} = [n_1, ..., n_{t-1}]$. Let $g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t)$ denote the PDF of buyers' belief at time t. Because of the specific pricing rule that we have assumed, buyers' initial belief at the beginning of the first period is $g_1(r_1 \mid a_1) = h_\eta(a_1 - r_1)$ on the interval $[a_1 - \bar{\eta}, a_1]$.

As we have described in section 2.2, sellers do occasionally change their list price. We do not model sellers' decisions of when to revise list price and if so, by how much. Instead, we assume that the change of list price is caused by the change of the seller's reservation value which is caused by the arrival of some exogenous shocks. As a result, when buyers observe that the current list price is different from the last period, their belief of the seller's reservation value will shift by the same amount as the change of list price.

[5] Lester et al. (2017) present a theoretical model in which sellers compete to attract buyers who need to inspect the goods before purchase by paying an inspection cost and prove that using an asking price, sellers both maximize their revenue and implement the efficient outcome in equilibrium. In Albrecht et al. (2016), sellers with heterogeneous reservation values use asking prices to signal their type, which allows for endogenous market segmentation. Merlo et al. (2015) empirically estimate a dynamic model in which rational and forward-looking home sellers make decisions of which price to list their homes by anticipating the arrival rate of potential buyers and their offers.

3.4 Bidding Stage

Suppose n_t buyers have viewed the home. n_t is a public information when viewers make their bidding decisions, which is a reasonable assumption in our context since the real estate agent lively updates the occurrence time of home viewings on its website.[6] If no buyer has viewed, i.e., $n_t = 0$, the home is retained by the seller and will be on sale at $t+1$. If at least one buyer has viewed the home, buyers who have viewed discover their idiosyncratic match taste. After knowing their private valuations and the number of viewers, the viewers decide whether or not to accept the current list price.

If one or more buyers accept the current list price, then the seller is committed to sell the home. If only one such offer is received, the home is sold to that buyer at the current list price. If two or more such offers are received, the buyers who accept the current list price can bid against each other to buy the home. In this case, those buyers are participating in an ascending auction starting from the current list price.

If no buyer accepts the current list price, then the seller is free to accept or reject the highest counteroffer depending on whether it is higher than his reservation value. A buyer is free to submit a new counteroffer if her previous one is rejected either because it was lower than the seller's reservation value or lower than other buyers' counteroffers. In the case of no sale, the home is retained by the seller and will be listed for sale next period.

The same as in Albrecht et al. (2016), the list price has only limited commitment in our model: a seller is committed to sell if one or more buyers offer to buy at or higher than the list price, but he is free to reject or accept any offer below the list price.

3.4.1 Equilibrium Bidding Strategy

In this subsection, we derive viewers' equilibrium bidding strategy. There are two cases. In the first case, only one buyer has viewed the home, i.e., $n_t = 1$. In the second case, multiple buyers have viewed the home, i.e., $n_t > 1$. The optimal strategy in the first case is straightforward. If the viewer accepts the current list price a_t, she will win the house and pay a_t for sure. If she does not accepts a_t, her optimal strategy is to start from a very low offer and increase it until her offer reaches the seller's reservation value if her private value is higher than that. It is easy to show that in the case of $n_t = 1$, not accepting a_t is always better than accepting it.

Deriving the equilibrium strategy in the case of multiple viewers (i.e, $n_t > 1$) is more complicated. The following proposition characterizes viewers' equilibrium strategy in this case.

Proposition 1. *When multiple buyers have viewed a home (i.e., $n_t > 1$), there exists a unique cutoff $\hat{v}_t \in (a_t, \bar{v}]$ such that at equilibrium, (i) buyers with private valuation below $\hat{v}_t$ reject a_t and bids their true value if they need to bid; and (ii) buyers with private valuation higher than $\hat{v}_t$ accept a_t and bid their true*

[6] We can relax this assumption by considering the case when viewers do not know how many other buyers have also viewed the home. In this case, the structure of the equilibrium bidding strategy will be the same as in the baseline model where viewers know n_t when bidding for the home, but the cutoff value will change accordingly.

value if they need to bid. Furthermore, the cutoff $\hat{v}_t = min\{v_t^*, \bar{v}\}$, where v_t^* is determined by:

$$(v_t^* - a_t)\left[F(v_t^*)^{n_t-1} - F(a_t)^{n_t-1}\right] = \int_{a_t-\bar{\eta}}^{a_t} G_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t)F(r_t)^{n_t-1}dr_t + \int_{a_t}^{v_t^*}(v_t^* - z)dF(z)^{n_t-1}. \quad (2)$$

Proposition 1 says that at equilibrium, buyers whose valuations are higher than a cutoff $\hat{v}_t$ accept the list price a_t and buyers whose valuations are below that cutoff reject a_t. In particular, for buyers whose valuations are higher than a_t but lower than $\hat{v}_t$ is to reject a_t in the first place. The main reason is that buyers who accepts a_t need to pay at least a_t in the case of winning; in contrast, if they do not accept a_t, there is a positive probability of paying less than a_t in the case of winning. To prove the strategy described in Proposition 1 is an equilibrium, we need to show that it is optimal for any given buyer to follow this strategy given other viewers following this strategy.

First, it is not hard to show that it is optimal for any buyer with $v_{it} \leq a_t$ not to accept a_t because accepting it implies that she needs to pay at least a_t if she wins. Since buyer i's valuation is not higher than a_t, she should reject a_t in the first place. Then there are two possibilities. In the first case, someone else accepts a_t, so she is out of the game and get a payoff of zero. In the second case, nobody accepts a_t. In this case, she participates an ascending auction with a reserve price equal to the seller's reservation value r_t which is the seller's private information. In equilibrium, the auction ends when the viewer with the second-highest value drops out, so the winner pays an amount equal to the higher one between the second highest value and the seller's reserve price. Essentially, she joins a second-price auction with a reserve price equal to the seller's reservation value r_t which is unknown by bidders. Buyer i's optimal bidding strategy is to bid her true value since bidding true value is the weakly dominate strategy in second-price auctions.

Next we show that the optimal choice for a buyer whose valuation is above a_t but below $\hat{v}_t$ is not to accept a_t. If she accepts a_t, there are two cases. In the first case, someone else also accepts a_t. She essentially needs to participate in a second-price sealed-bid auction starting from a_t. Since others follow the subscribed strategy, it implies that the buyers who accept a_t must have a valuation higher than $\hat{v}_t$ and bid their true values when then need to bid. So in this case, buyer i will not win at all and get a payoff of zero. In the second case, no one else accepts a_t and she is the only one who accepts a_t. She wins and gets a payoff of $v_{it} - a_t$. Given other viewers following the equilibrium strategy described in the proposition, the probability that no other buyers accept a_t is $F(\hat{v}_t)^{n_t-1}$. So, her expected gain of accepting a_t is $(v_{it} - a_t)F(\hat{v}_t)^{n_t-1}$.

In contrast, if she does not accept a_t, there are also two cases. In the first case, someone else accepts a_t and she is out of the game immediately. In the second case, no one accepts a_t. In this case she needs to bid against other viewers. There are two possibilities: (i) if the second highest bid is lower than a_t, then she wins and pays a price equal to the higher one between the second highest value and the seller's reservation value r_t; and (ii) if there are several bidders having valuations above a_t but below $\hat{v}_t$ as well, she wins only if the second highest bid is below v_{it} and she pays that bid. Therefore, her expected gain from rejecting a_t is

$$\begin{aligned}&\int_{a_t-\bar{\eta}}^{a_t}\left[\int_{\underline{v}}^{a_t}(v_{it}-max(z,r_t))dF(z)^{n_t-1}\right]g_t(r_t \mid \vec{n}_{t-1},\vec{a}_t)dr_t+\int_{a_t}^{v_{it}}(v_{it}-z)dF(z)^{n_t-1}\\ = \quad & (v_{it}-a_t)F(a_t)^{n_t-1}+\int_{a_t-\bar{\eta}}^{a_t}G_t(r_t \mid \vec{n}_{t-1},\vec{a}_t)F(r_t)^{n_t-1}dr_t+\int_{a_t}^{v_{it}}(v_{it}-z)dF(z)^{n_t-1}.\end{aligned}$$

The difference between the payoff from accepting a_t and that from not accepting it is

$$(v_{it}-a_t)\left[F(\hat{v}_t)^{n_t-1}-F(a_t)^{n_t-1}\right]-\int_{a_t-\bar{\eta}}^{a_t}G_t(r_t \mid \vec{n}_{t-1},\vec{a}_t)F(r_t)^{n_t-1}dr_t-\int_{a_t}^{v_{it}}(v_{it}-z)dF(z)^{n_t-1}.$$

Notice that when $n_t = 1$, the difference is strictly negative for any finite v_{it}. That is, not accepting a_t is always better than accepting a_t. So the cutoff $\hat{v}_t = \bar{v}$ in the case of $n_t = 1$.

Next consider the case of $n_t > 1$. When $v_{it} = a_t$, the difference is negative, implying that not accepting a_t is better for buyers with valuations equal to a_t. As v_{it} goes to $+\infty$, the difference is strictly positive for any given a_t. Moreover, the difference is strictly increasing in v_{it}. Hence, there must exists a unique $v_t^* > a_t$ such that

$$(v_t^*-a_t)\left[F(v_t^*)^{n_t-1}-F(a_t)^{n_t-1}\right]=\int_{a_t-\bar{\eta}}^{a_t}G_t(r_t \mid \vec{n}_{t-1},\vec{a}_t)F(r_t)^{n_t-1}dr_t+\int_{a_t}^{v_t^*}(v_t^*-z)dF(z)^{n_t-1}.$$

However, v_t^* may be above the upper bound $\bar{v}$. So the cutoff $\hat{v}_t$ is the minimum between the v_t^* and the upper bar $\bar{v}$:

$$\hat{v}_t = min\{v_t^*, \bar{v}\}.$$

To sum up, given the equilibrium strategy of other viewers, rejecting a_t is optimal for any buyer whose valuation is below $\hat{v}_t$ because bidding above it implies that she needs to pay at least a_t if she wins. The argument above also confirms that bidding above a_t is optimal for any buyer whose valuation is above $\hat{v}_t$.

3.4.2 Expected Gain of Equilibrium Strategy

Next we compute the expected gain of a viewer with private valuation v at equilibrium. Let z be the second highest value among all realized match values of buyers who have viewed the home. There are three cases in total.

Case 1: $v \leq a_t$. Her equilibrium strategy is to reject a_t and bid her true value when she needs to bid. So, her expected gain is:

$$\Gamma(v, n_t, a_t, g_t) \quad = \quad \int_{a_t-\bar{\eta}}^{v}\left[\int_{\underline{v}}^{v}(v-max(z,r_t))dF(z)^{n_t-1}\right]g_t(r_t \mid \vec{n}_{t-1},\vec{a}_t)dr_t$$

Case 2: $a_t < v \leq \hat{v}_t$. Her equilibrium strategy is to reject a_t and bid her true value when she needs to bid. So, her expected gain is:

$$\Gamma(v, n_t, a_t, g_t) = \int_{a_t-\bar{\eta}}^{a_t} \left[\int_{\underline{v}}^{a_t} (v - max(z, r_t)) dF(z)^{n_t-1} \right] g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t + \int_{a_t}^{v} (v - z) dF(z)^{n_t-1}$$

Case 3: $v > \hat{v}_t$. Her equilibrium strategy is to accept a_t and bid her true value when she needs to bid. So, her expected gain is:

$$\Gamma(v, n_t, a_t, g_t) = \int_{\underline{v}}^{v} (v - a_t \cdot 1(z \leq \hat{v}_t) - z \cdot 1(z > \hat{v}_t)) dF(z)^{n_t-1}$$

3.4.3 Conditional Transaction Probability and Price Distribution

In this subsection, we calculate the transaction probability and price distribution conditional on transaction, given the number of viewers n_t and all viewers using the equilibrium bidding strategy.

The home is retained only if the valuations of all buyers who have viewed the home fall short of the seller's reservation value. Hence, the probability that the home is not sold at time t conditional on n_t viewers, list price a_t and belief g_t is:

$$\Pr(y_t = 0 \mid n_t, a_t, g_t) = \int_{a_t-\bar{\eta}}^{a_t} [F(r_t)]^{n_t} g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t, \tag{3}$$

and the probability that the home is sold at time t conditional on n_t, a_t, and g_t is:

$$\Pr(y_t = 1 \mid n_t, a_t, g_t) = \int_{a_t-\bar{\eta}}^{a_t} [1 - [F(r_t)]^{n_t}] g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t. \tag{4}$$

Next we derive the conditional distribution of transaction price (*relative to the home's financial value* δ) if an transaction occurs. If only one buyer views the home (i.e., $n_t = 1$), the viewer's optimal bidding strategy is to start from a low offer and increase it until it reaches the seller's reservation value r_t. Hence, the distribution of the transaction price is the same as the distribution of r_t. That is,

$$\Pr(p_t \leq p \mid y_t = 1, n_t = 1, a_t, g_t) = G_t(p \mid \vec{n}_{t-1}, \vec{a}_t). \tag{5}$$

In the case of multiple viewers (i.e., $n_t > 1$), the conditional distribution of transaction price for a given r_t is (Athey and Haile (2002)):

$$\Pr(p_t \leq p \mid y_t = 1, n_t, a_t, g_t, r_t) = \begin{cases} n_t(n_t-1) \int_0^{F(p)} s^{n_t-2}(1-s) ds, & if\ \hat{v}_t \leq p \leq \bar{v} \\ n_t(n_t-1)\Gamma(p), & if\ a_t \leq p \leq \hat{v}_t \\ n_t(n_t-1) \int_0^{F(p)} s^{n_t-2}(F(\hat{v}_t) - s) ds, & if\ r_t \leq p < a_t \\ 0, & if\ p < r_t \end{cases}$$

where

$$\Gamma(p) = \int_0^{F(p)} s^{n_t-2}(F(\hat{v}_t) - s)ds + \int_0^{F(\hat{v}_t)} s^{n_t-2}(1 - F(\hat{v}_t))ds.$$

In the above expression, $n_t(n_t-1)\int_0^{F(p)} s^{n_t-2}(1-s)ds$ is the cumulative distribution function of the $n_t - 1$-th order statistic. This is also the distribution function of transaction price when $\hat{v}_t \leq p \leq \bar{v}$. But when $r_t \leq p < a_t$, we also need the highest valuation to be lower than $\hat{v}_t$. That is why the cumulative distribution function becomes $n_t(n_t-1)\int_0^{F(p)} s^{n_t-2}(F(\hat{v}_t)-s)ds$. The distribution function has a mass point at $p = a_t$, where the mass is $n_t(n_t-1)\int_0^{F(\hat{v}_t)} s^{n_t-2}(1-F(\hat{v}_t))ds$. This arises because as long as the highest valuation is above $\hat{v}_t$ and the second highest valuation is below $\hat{v}_t$, the transaction price will be equal to a_t.

Since r_t is the seller's private information, integrating over r_t yields:

$$\Pr(p_t \leq p \mid y_t = 1, n_t, a_t, g_t) = \begin{cases} n_t(n_t-1)\int_0^{F(p)} s^{n_t-2}(1-s)ds, & if\ \hat{v}_t \leq p \leq \bar{v} \\ n_t(n_t-1)\Gamma(p), & if\ a_t \leq p \leq \hat{v}_t \\ n_t(n_t-1)G_t(p \mid \vec{n}_{t-1}, \vec{a}_t)\int_0^{F(p)} s^{n_t-2}(F(\hat{v}_t)-s)ds, & if\ a_t - \bar{\eta} < p < a_t \\ 0, & if\ p \leq a_t - \bar{\eta} \end{cases} \quad (6)$$

Notice that the distribution of transaction price has two interesting features: 1) the transaction price can be higher, lower or equal to the listed price; and 2) there is a mass point at the listed price as illustrated above.

3.5 Viewing Decision

Now we go back to the second stage when prospective buyers decide whether to view the home and learn their private valuations by paying a viewing cost. When making this decision, they have a belief of the seller's reservation value $g_t(r_t \mid \vec{a}_t, \vec{n}_{t-1})$. They also know their own viewing cost and the distribution of their private valuations, but not the exact value of their private valuations nor the number of viewers who decide to view.

In section 3.4.2 we have derived the expected gain of a viewer with match value v conditional on the number of viewers at equilibrium. However, when buyers making their view decisions, they do not know their match values. So, they their expected gain of viewing the home (before knowing v_{it}) for a given number of viewers n should be integrated over v:

$$EU(n, a_t, g_t) \quad = \quad \int_{\underline{v}}^{\bar{v}} \Gamma(v, n, a_t, g_t)dF(v). \quad (7)$$

The optimal viewing decision of a buyer with viewing cost c_{it} is quite straightforward: view the home if her cost c_{it} is below a cutoff $\hat{c}_t$ and does not view it otherwise, where the cutoff $\hat{c}_t$ depends on buyers' belief of the seller's reservation value, the distribution of match values, and the current list price.

Recall that buyers' viewing cost follows a distribution with CDF of $\Psi(\cdot)$. So, conditional on the cutoff $\hat{c}_t$, each buyer views the home with a probability of $\Psi(\hat{c}_t)$. Hence, the probability of having n viewers at time t conditional on the cutoff $\hat{c}_t$ is:

$$\mathcal{P}_t(n \mid \hat{c}_t) = \sum_{m \geq n} \Lambda_t(m_t = m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Psi(\hat{c}_t)^{n-1}[1 - \Psi(\hat{c}_t)]^{m-n}. \tag{8}$$

The cutoff $\hat{c}_t$ should be equal to the expected gain of viewing the home without knowing the number of viewers:

$$\hat{c}_t = \sum_{n=1}^{\infty} \mathcal{P}_t(n \mid \hat{c}_t) \times EU(n, a_t, g_t), \tag{9}$$

Notice that $\mathcal{P}_t(n \mid \hat{c}_t)$ on the right hand side of equation (9) is a function of the cutoff $\hat{c}_t$ (equation (8)). We can easily show that the solution of equation (9) exists and is unique, because the left hand side of equation (9) is increasing in $\hat{c}_t$, whereas the right hand side is decreasing in $\hat{c}_t$.

After we solve the cutoff $\hat{c}_t$ from equation (9), we can plug it back to equation (8) to get the conditional probability of viewers conditional on a_t and g_t, denoted as $\mathcal{P}^*_t(n_t \mid a_t, g_t)$.

3.6 Buyers' Belief Updating

Before buyers make their viewing decision, they form their belief about the seller's reservation value using the process described below.

If a home is not sold at time t, buyers who arrive at time $t+1$ will observe the number of viewers up to time t, $\vec{n}_t$. They will use this information to update their belief about the seller's reservation value.

If the list price does not change at time $t+1$, i.e., $a_{t+1} = a_t$, then buyers know that the seller's reservation value does not change. If no buyers views the home at time t (i.e., $n_t = 0$), then buyers' belief at time $t+1$ remains the same as their belief at time t. However, as long as someone views the home home at time t (i.e., $n_t > 0$) but does not buy it, buyers' belief at time $t+1$ is updated using Bayes rule. Therefore, conditional on the number of viewers in the past periods $\vec{n}_t$, buyers' belief of the seller's reservation value becomes:

$$g_{t+1}(r_{t+1} \mid \vec{n}_t, \vec{a}_{t+1}) = \begin{cases} g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t), \; if \; n_t = 0 \\ \frac{[F(r_t)]^{n_t} g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t)}{\int_{a_t - \bar{\eta}}^{a_t} [F(r_t)]^{n_t} g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t}, \; if \; n_t > 0. \end{cases} \tag{10}$$

In the case of $n_t > 0$, the denominator is the probability that the home is not sold at time t given the list price a_t, belief g_t, and the number of viewers n_t.

If the list price changes at time $t+1$, i.e., $a_{t+1} \neq a_t$, then buyers can infer that the seller's reservation value changes. In this case, we first derive the belief of r_{t+1} after observing n_t but keeping the list price unchanged. This density is denoted as $g_{t+1}(r_{t+1} \mid \vec{n}_t, \vec{a}_t)$, which is the same as equation (10). Then, we can derive the belief when observing $a_{t+1} \neq a_t$. Buyers at time $t+1$ form a new belief about r_{t+1} when observing a new list price a_{t+1}. Since we have assumed that

$a_{t+1} = r_{t+1} + \eta$, where η is fixed over time. So we have $a_{t+1} - r_{t+1} = a_t - r_t$. This implies that the density of r_{t+1} is the same as $r_t = r_{t+1} - (a_{t+1} - a_t)$:

$$g_{t+1}(r_{t+1} \mid \vec{n}_t, \vec{a}_{t+1}) = g_{t+1}(r_{t+1} - (a_{t+1} - a_t) \mid \vec{n}_t, \vec{a}_t). \tag{11}$$

4 Estimation

We estimate a parametric version of the model. In this section, we first make several distributional assumptions. Then, we provide an informal discussion of identification. Lastly, we describe our estimation strategy.

4.1 Distributional Assumptions

The model contains four primitives: (i) the arrival process of buyers, $\Lambda_t(m_t = m)$, (ii) buyers' initial belief of the seller's reservation value, $g_1(r_1 \mid a_1)$, (iii) the distribution of buyers' match value, $F(v_{it})$, and (iv) the distribution of viewing cost, $\Psi(c_{it})$. In the estimation, we make the following distributional assumptions.

1. Prospective buyers arrive at time t following a Poisson distribution with an arrival rate λ_t. In particular, we specify λ_t as:
$$\lambda_t = exp(d_t \gamma),$$
where d_t is a vector of factors affecting the arrival rate, including an intercept, weeks on sale, predicted financial value δ and its square. With the Poisson assumption, m prospective buyers will arrive at time t with probability:
$$\Lambda_t(m_t = m) = \frac{\lambda_t^m e^{-\lambda_t}}{m!}.$$

2. We assume that the difference between the list price and the seller's reservation value η follows a uniform distribution on the interval $[0, \bar{\eta}]$. As a result, the initial belief g_1 is a uniform distribution on the interval $[a_1 - \bar{\eta}, a_1]$.

3. We assume that buyers' match values $v_{it} \sim truncated\ N(0, \sigma_v^2)$ on the interval $(-\infty, \bar{v}]$.

4. We assume that the viewing cost follows a Lognormal distribution with parameters μ_c and σ_c. That is, $c_{it} \sim Lognormal(\mu_c, \sigma_c^2)$.[7]

4.2 Identification

The data correspond to a panel data of exogenous variables affecting the arrival rate of prospective buyers, list price, number of viewers, indicator of transaction, and transaction price in the case

[7]The density probability function is $\frac{1}{c_{it}\sigma_c\sqrt{2\pi}} exp\left[-\frac{(ln(c_{it}) - \mu_c}{2\sigma_c^2}\right]$, with mean = $exp(\mu_c + \frac{\sigma_c^2}{2})$, median = $exp(\mu_c)$, and variance = $[exp(\sigma_c^2) - 1] exp(2\mu_c + \sigma_c^2)$.

of transaction. For each home j included in our sample, we observe $D_j = (\{a_{jt}, n_{jt}, y_{jt}, d_{jt}\}_{t=1}^{T_j}, p_{jT_j})$, where T_j is the number of weeks when the home j is sold since being initially listed, a_{jt} is home j's list price in sale week t *relative to its financial value* $\hat{\delta}_j$, n_{jt} is the number of viewers in sale week t, y_{jt} indicates whether the home is sold in sale week t, d_{jt} is a vector of exogenous factors affecting the arrival rate, and p_{jT_j} is the home's transaction price *relative to its financial value* $\hat{\delta}_j$.

The conditional transaction probability (equation (4)) and the conditional transaction price distribution (equations (5)) and (6)) are independent of the buyers' arrival process and the viewing cost distribution. However, the conditional transaction probability and conditional transaction price distribution depend on both match value distribution F and buyer beliefs g_t. So the challenge is how to distinguish between the two sources when discussing the mapping from the reduced form to the primitives of the model.

First notice that $p_t \in [a_t - \bar{\eta}, \bar{v}]$. So the range of transaction price can identify the lower bound of sellers' reservation value r (parameter $\bar{\eta}$) and the upper bound of buyers' match value v (parameter $\bar{v}$). From equation (6), we can see that the distribution of the transaction price when it is above or equal to the list price only depends on the match value distribution F (parameter σ_v). Once the match value distribution is identified, the conditional distribution of the transaction price when it is below the list price can identify the initial belief g_1. In addition, how the transaction probability and price distribution vary in n_t are affected by F and g_t (and hence g_1) in different ways. These moments can also be used to identify F separately from g_1.

Notice that we normalize the mean of v_{it} to be zero because we can not separately identify the mean and the standard deviation. After normalization, a higher σ_v implies a larger probability of getting high match value draws. Holding everything else constant, a realized match value will be more likely exceed the seller's reservation value, resulting in earlier transactions conditional on the number of viewers (equation (4)). Thus, time to sale also identifies σ_v.

After we identify the match value distribution F and initial belief g_1, we can compute the expected gain of viewing for each possible n (equation (7)). From equation (8), we can see that the equilibrium distribution of viewers only depend on the arrival process and the distribution of viewing cost. However, since we only observe the number of viewers n_{jt} but not the number of prospective buyers m_{jt}, we must infer the arrival rate and the distribution of viewing cost from the distribution of n_{jt} only.

How the number of viewers (n_{jt}) varies in those exogenous factors in the arrival rate (d_{jt}), including weeks on sale and financial value, can identify the coefficients before those factors. The intercept in the arrival rate (γ_1) and the parameters associated with the distribution of viewing cost (μ_c and σ_c) are identified by the functional forms. See Appendix for more details. Intuitively, how the number of viewers vary in the list price are affected by the arrival process and the viewing cost distribution in different ways. These moments are used to separately identify the arrival rate intercept and the parameters associated with the viewing cost distribution.

4.3 Likelihood Function

Let θ summarize all other parameters to be estimated: $\theta = [\sigma_v, \mu_c, \sigma_c, \gamma]$, where σ_v is the standard deviation of buyers' match values, μ_c and σ_c are the mean and the standard deviation of buyers' viewing cost, and γ is a vector of parameters associated with the arrival rate of prospective buyers. We estimate them by maximum likelihood. The likelihood contribution of home j is:

$$\begin{aligned} L_j(\theta \mid D_j) &= \prod_{t=1}^{T_j} \Pr(y_{jt} \mid n_{jt}, a_{jt}, g_{jt}) \times \mathcal{P}_t(n_{jt} \mid a_{jt}, g_{jt}) \\ &\quad \times \varphi(p_{jT_j} \mid y_{jT_j} = 1, n_{jT_j}, a_{jT_j}, g_{jT_j}) \end{aligned}$$

where g_{jt} is buyers' belief of seller j's reservation value at time t, given by equation (10) and (11). By construction, $y_{jt} = 0$ for all $t < T_j$ and $y_{jt} = 1$ for $t = T_j$. So $\Pr(y_{jt} \mid n_{jt}, a_{jt}, g_{jt})$ is given by equation (3) for $t < T_j$ and equation (4) for $t = T_j$. $\mathcal{P}_t(n_{jt} \mid a_{jt}, g_{jt})$ is the probability of n_{jt} buyers viewing home j at time t conditional on list price and buyers' belief, given by equation (8)). $\varphi(p_{jt} \mid y_{jt} = 1, n_{jt}, a_{jt}, g_{jt})$ is the price density function conditional on selling, number of viewers, list price, and buyers' belief, which can be derived from its CDF given by equation (5) and (6).

When $n_{jt} = 1$, the corresponding price density function is the same as the buyers' belief of r_{jt} :

$$\varphi(p \mid y_{jt} = 1, n_{jt} = 1, a_{jt}, g_{jt}) = g_{jt}(p \mid \vec{n}_{jt-1}, \vec{a}_{jt}).$$

When $n_{jt} > 1$, there are four cases :

1. The price density for $p \in [a_{jt} - \bar{\eta}, a_{jt})$ is:

$$\begin{aligned} \varphi(p \mid y_{jt} = 1, n_{jt} > 1, a_{jt}, g_{jt}) &= n_{jt}(n_{jt} - 1) G_{jt}(p \mid \vec{n}_{jt-1}, \vec{a}_{jt}) F(p)^{n_{jt}-2} (F(\hat{v}_{jt}) - F(p)) f(p) \\ &\quad + n_{jt}(n_{jt} - 1) g_{jt}(p \mid \vec{n}_{jt-1}, \vec{a}_{jt}) \int_0^{F(p)} s^{n_{jt}-2} (F(\hat{v}_{jt}) - s) ds. \end{aligned}$$

2. There is a mass point at $p = a_{jt}$, given by:

$$\varphi(p \mid y_{jt} = 1, n_{jt} > 1, a_{jt}, g_{jt}) = n_{jt}(n_{jt} - 1) \int_0^{F(\hat{v}_{jt})} s^{n_{jt}-2} (1 - F(\hat{v}_{jt})) ds.$$

3. The price density for $p \in (a_{jt}, \hat{v}_{jt})$ is:

$$\varphi(p \mid y_{jt} = 1, n_{jt} > 1, a_{jt}, g_{jt}) = n_{jt}(n_{jt} - 1) F(p)^{n_{jt}-2} (F(\hat{v}_{jt}) - F(p)) f(p).$$

4. The price density for $p \in [\hat{v}_{jt}, \bar{v}]$ is:

$$\varphi(p \mid y_{jt} = 1, n_{jt} > 1, a_{jt}, g_{jt}) = n_{jt}(n_{jt} - 1) F(p)^{n_{jt}-2} (1 - F(p)) f(p).$$

5 Estimation Results

5.1 Estimates

According to our identification discussion, $\bar{v}$ and $\bar{\eta}$ can be estimated without solving the model. We first compute the upper bound of the transaction price directly from the data and take it as the estimate of $\bar{v}$, which is 0.299 million CNY. This is, buyers' match value is at most 0.299 million CNY higher than a home's financial value. In addition, we compute the upper bound of the difference between the last list price and the transaction price and take it as the estimate of $\bar{\eta}$, which is 0.65 million CNY. This implies that sellers' reservation value is at most 0.65 million CNY lower than their homes' financial value.

The estimates of other parameters of the model are obtained using the maximum likelihood estimation method. Table 4 reports the estimation results. Below we discuss about the economic magnitude of those parameter estimates.

Table 4: Maximum Likelihood Estimation Results

	Estimates	S.E.
Match Value		
Standard deviation: σ_v	3.6052	(0.3153)
Viewing Cost		
Mean: μ_c	-5.1627	(0.0496)
Standard deviation: σ_c	0.0077	(0.0013)
Arrival Rate		
Intercept (γ_1)	0.4145	(0.131)
Weeks on sale (γ_2)	-0.0265	(0.0057)
Predicted financial value (γ_3)	0.2315	(0.0690)
Predicted financial value: square (γ_4)	-0.0156	(0.0081)

Note: The data used for estimation include 27,703 apartments in five core districts of Beijing that are listed and sold by Lianjian from August 2015 to July 2016. Unit: million CNY.

Match value distribution. Recall that the match value is assumed to follow a truncated normal distribution. We have normalized the mean of match value to be zero and the upper bar ($\bar{v}$) has been estimated directly from the empirical upper bound of the transaction price. The estimate of $\bar{v}$ is 0.299 million CNY. Our maximal likelihood estimate of the standard deviation (σ_v) is 3.6052 with a standard error of 0.3153, indicating the estimate is significant at 1% level. As we have explained, σ_v is a direct measure of the potential benefit from searching. Our estimate of σ_v is quite large, suggesting that buyers gain a quite high benefit from physically viewing homes.

In the empirical search literature, most previous studies have normalized σ_v to be one mainly because it cannot be separately identified from the parameters associated with the search cost distribution in their settings (see Moraga-González et al. (2018)). One exception is Yavorsky et al. (2020) who estimate a model of consumer search and purchase of cars. Their estimate of σ_v is 8.16. As we have discussed in Section 4.2, σ_v in our model is identified from the transaction probability and the distribution of the transaction price conditional on the number of viewers which does not depend on the search cost distribution.

Search cost distribution. Recall that the viewing cost is assumed to follow a Lognormal distribution with parameters μ_c and σ_c^2. Based on our estimates, $\hat{\mu}_c = -5.1627$ and $\hat{\sigma}_c = 0.0077$, with both being measured in million CNY. That is, the estimated median is 5,726 CNY, roughly equivalent to $916. This number is roughly half of the average monthly mortgage. How do our results compare to existing estimates of search costs in the literature? Perhaps the closest point of comparison comes from Yavorsky et al. (2020) and Allen et al. (2019) who estimate the distribution of offline search cost. Yavorsky et al. (2020) estimate the average search cost of physically visiting a car dealership in U.S. to be $297 and $446 per mile for rural and urban consumers, respectively. Allen et al. (2019) find that the average cost of searching for multiple mortgage offers is equal to $1,150 (with a median of $784). Although somewhat different, our search-cost estimates are comparable with those found in the literature.

Arrival rate. The four γ's are the intercept, the coefficients of weeks on sale, the predicted financial value δ, and the square of δ. The estimates indicate that fewer prospective buyers arrive on average if a home has been on sale fora longer time. We also find that the arrival rate is a hump-shaped function of the predicted financial value, suggesting that the arrival rate of medium-valued homes is higher than the low-valued homes and high-valued homes.

5.2 Goodness of Fit

We next provide a number of tests for the goodness of fit of the model. For each home in the data, $j = 1, ..., 27, 703$, we simulate the endogenous outcomes using the estimated model.[8] The endogenous outcomes include the cutoff of the match value that determines whether to accept the current list price, the cutoff of the viewing cost that determines whether to view the home, the equilibrium distribution of the number of viewers , a draw of n_{jt} from the equilibrium distribution, the transaction probability, a draw of y_{jt} from the transaction probability, the transaction price density conditional on viewers if transaction, the expected transaction price conditional on viewers if transaction, and the total weeks to sale.

Table 5 presents summary statistics for the key endogenous outcomes of the model, including the total weeks to sale, total number of home viewings, and the expected transaction price. The top panel summarizes the observed data, while the bottom summarizes the simulated data. Overall, the model is able to match well the distributions of those variables, but the model tends to predict slightly shorter sale time, fewer viewings, and lower transaction price than they are observed in the data.

[8]For home j with the transaction week of T_j in the data, its list price at search week t in the simulation scenario is set to be to the list price at time t observed in the data when $t \leq T_j$ and the last list price observed in the data (the price at the transaction week T_j) when $t > T_j$.

Table 5: Summary Statistics for Simulated and Observed Data

	Mean	SD	P25	P50	P75
Observed					
Weeks to sale	5.439	4.955	2	4	8
Number of home viewings	11.629	11.348	4	8	16
Transaction price	2.938	1.196	2.080	2.680	3.530
Simulated					
Weeks to sale	4.886	5.415	1	3	6
Number of home viewings	10.528	9.341	4	7	14
Transaction price	2.727	1.210	1.854	2.461	3.327

Note: The simulated sample is obtained by simulating the outcomes of 27,703 houses from the estimated model. The transaction price is measured in million CNY.

To understand the source of discrepancy, it is useful to look at the ability of the model to explain the difference between homes with the initial list price lower than the predicted financial value (defined as negative normalized list price, i.e., $a < 0$) and homes with the initial list price higher than the predicted financial value (defined as positive normalized list price, i.e., $a > 0$). The results are reported in Table 6 and Table 7. Like in the data, the model predicts that homes with $a > 0$ stay on the market longer, accumulate more home viewings before sale, and sell at higher price than homes with $a < 0$. However, the model tends to under-predict those outcomes for homes with $a < 0$ while slightly over-predict those outcomes for homes with $a > 0$.

This is because, in the model, the gap between a home's list price and the seller's reservation value, denoted as η, is assumed to be a random term that follows the same distribution across sellers, regardless of the seller's reservation value. Consequently, a seller who sets a low (high) list price has a low (high) reservation value on average, which are expected by rational buyers. A low reservation value implies a high expected gain from viewing the home, leading to more home viewings and shorter sale time. A low reservation value also lowers the transaction price if an transaction occurs, see equations (5) and (6) for the distribution of transaction price. In our model, the mapping from the list price and the reservation value is one dollar to one dollar mapping on average, although the gap η is a random term. In reality, this mapping is likely a one dollar to less than one dollar mapping. Because of our assumption, sellers who charge relatively low list price tend to have lower reservation values in our model than their actual reservation values.

Table 6: Homes with Initial List Price Lower than Predicted Financial Value ($a < 0$)

	Mean	SD	P25	P50	P75
Observed					
Weeks to sale	5.360	4.996	2	3	7
Number of home viewings	11.089	10.973	3	8	15
Transaction price	2.835	1.156	2.010	2.580	3.370
Simulated					
Weeks to sale	3.599	3.642	1	2	4
Number of home viewings	8.286	6.966	4	6	11
Transaction price	2.599	1.156	1.779	2.343	3.132

Note: The simulated sample is obtained by simulating the outcomes of 12,197 houses with initial list price lower than predicted financial value ($a < 0$) from the estimated model. The transaction price is measured in million CNY.

Table 7: Homes with Initial List Price Higher than Predicted Financial Value ($a > 0$)

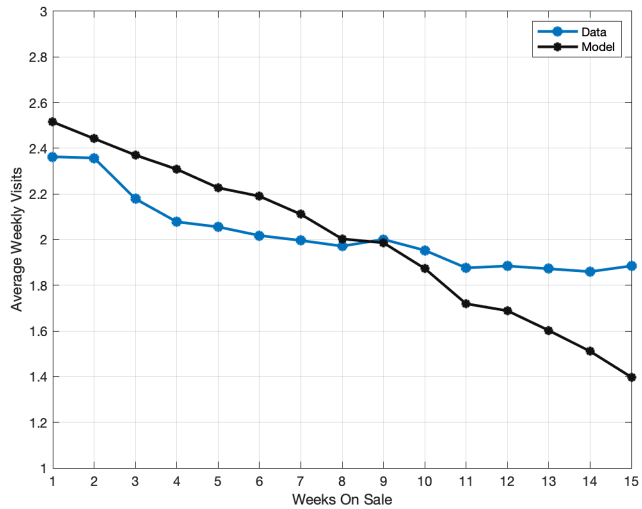
	Mean	SD	P25	P50	P75
Observed					
Weeks to sale	5.501	4.922	2	4	8
Number of home viewings	12.054	11.618	4	8	16
Transaction price	3.018	1.220	2.120	2.760	3.660
Simulated					
Weeks to sale	5.898	6.295	2	4	7
Number of home viewings	12.292	10.519	5	9	17
Transaction price	2.827	1.242	1.923	2.565	3.477

Note: The simulated sample is obtained by simulating the outcomes of 15,506 houses with initial list price higher than predicted financial value ($a > 0$) from the estimated model. The transaction price is measured in million CNY.

Furthermore, we assess the ability of the model to fit the dynamic pattern of home viewings. Figure 4 display the average home viewings by sale week observed in the data and predicted by the model.[9] In general, the model does a good job of predicting the declining pattern of the number of home viewings, although the modeled predicted average values are slightly higher for earlier weeks and slightly lower for later weeks than the data averages.

[9]We do not report the viewings after week 15 just because very few houses are still on the market after those points of time and the sample means are not meaningful anymore.

Figure 4: Observed and Model Predicted Number of Home Viewings by Sale Week



5.3 Buyer Surplus and Seller Surplus

For an individual home j, the expected buyer surplus is defined as the expectation of the net utility of the purchaser (i.e., her match value minus the transaction price) net of the total viewing cost of all viewers:

$$E[BS_j] = \int_{p_j}^{\bar{v}} (v - p_j) dF(v) - \sum_{t=1}^{T_j} n_{jt} \int_0^{\hat{c}_{jt}} c d\Psi(c),$$

where the first term measures the expected gain of the buyer who eventually bought the home j and the second term measures the total viewing cost paid by all buyers who viewed the home j.

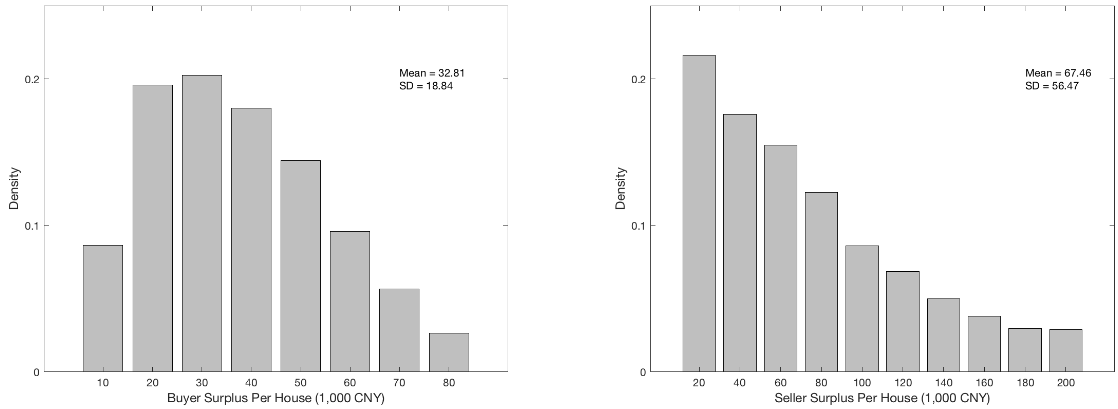
The expected seller surplus of an individual home j is measured by the expectation of the difference between the final transaction price and the seller's reservation value:

$$E[SS_j] = \int_{a_{jT_j} - \bar{\eta}}^{a_{jT_j}} (p_j - r) dG_j(r),$$

where a_{jT_j} is home j's list price when it is sold and $G_j(r)$ is the prior distribution of sellers' reservation value which is assumed to be uniform on the interval $[a_{jT_j} - \bar{\eta}, a_{jT_j}]$.

Using the estimated distributions of match value and viewing cost, we compute the expected buyer surplus and expected seller surplus for each home included in the data. Figure 5a and Figure 5b display the distributions of these two. The mean of the expected buyer surplus is 32,813 CNY and the standard deviation is 18,837 CNY. The mean of the expected seller surplus is 67,460 CNY and the standard deviation is 56,136 CNY.

Figure 5: Expected Buyer Surplus and Expected Seller Surplus



(a) Expected Buyer Surplus

(b) Expected Seller Surplus

6 Information Disclosure and Observational Learning

Observational learning is tightly linked to what information buyers observe when they make viewing and bidding decisions. As we have mentioned in the introduction, one uniqueness of our setting is that the occurrence times of all in-person home viewings are updated on Lianjia's website on a real time basis during our sample period. Therefore, in our model buyers know a home's time on market and the number of home viewings occurred in each of previous weeks when they make their decisions. Based on those information, they update their beliefs about the seller's reservation value. In contrast, home buyers in western countries usually only have the information of the time on market but not the in-person home viewings by previous buyers. Moreover, Lianjia stopped disclosing the information of home viewings in February 2021 to comply with the Beijing government's regulation. In this section, we conduct a series of counterfactual exercises to quantify the effect of these two pieces of information: time on market and home viewings by previous buyers.

6.1 Two Counterfactual Scenarios: NI Case and PI Case

In the first counterfactual scenario, buyers do not have the information of the time on the market nor the home viewings, so buyers do not update their belief about the seller's reservation value over time as long as the list price does not change. Essentially, there is no observational learning at all in this counterfactual scenario and hence we call it no observational learning case (NI). In this NI case, all of the equations in the model do not change except that the updating equation (10) becomes:

$$g_{t+1}(r \mid a_{t+1}) = g_1(r \mid a_{t+1}). \tag{12}$$

In the second counterfactual scenario, buyers observe time on market but not home viewings. In this scenario, buyers update their belief about the seller's reservation value based on the time-on-market information only. We call this case partial information case (PI). In this PI scenario, all equations in our model do not change except for the updating of the seller's reservation value. The updating equation (10) becomes:

$$g_{t+1}(r_{t+1} \mid \vec{a}_{t+1}) = \frac{\sum_n \mathcal{P}_t(n)[F(r_t)]^n g_t(r_t \mid \vec{a}_t)}{\sum_n \mathcal{P}_t(n) \int_{a_t-\bar{\eta}}^{a_t} [F(r_t)]^n g_t(r_t \mid \vec{a}_t) dr_t}. \tag{13}$$

Using the parameter estimates, we simulate the equilibrium outcomes and compute the buyer surplus, seller surplus, and total surplus for each of the 27,703 homes included in the data separately for the NI model and the PI model. The differences between the PI model and the NI model measure the effect of disclosing the time-on-market information only, while the differences between our proposed model and the PI model measure the additional effect of disclosing the viewings information in additional to the time-on-market information. The differences between our proposed model and the NI model measure the total effect of disclosing these two pieces of information. Table 8 reports the simulation results.

6.2 Effects of Disclosing Time On Market and Home Viewings

Effect of Disclosing the Time-on-Market Information Only. Compared with buyers in the NI case who do not know neither time on market nor home viewings, buyers in the PI case who know time on market would update their belief about the seller's reservation value if a home is still on sale. Intuitively, the longer a home has been on sale, the higher the belief is. A higher belief affects the transaction price for two reasons. On the one hand, a higher belief lowers the match value cutoff that determines whether to accept the current list price conditional on the number of viewers (see equation (2)). A higher belief also increases the lower bound of offers that will be accepted by the seller. So, because of this direct effect, the disclosure of the time-on-market information would lead to higher transaction price conditional on the number of viewers. On the other hand, a higher transaction price conditional on the number of viewers implies a lower expected gain from home viewing and hence discourages buyers to view the home in the first place. A fewer number of viewers implies less competition in the bidding stage and hence leads to lower transaction price (see equation (6)). So, because of this indirect effect, the disclosure of the time-on-market information would lead to lower transaction price. At equilibrium, the direct effect dominates the indirect effect. The panel (I) of Table 8 reports the effect of disclosing the time-on-market information, measured by the difference between the PI model and the NI model. The transaction price in PI case is 38,768 CNY higher on average.

The disclosure of the time-on-market information affects the sale time both in a direct way and

Table 8: Effects of Different Information Disclosure Rules

	Mean	SD	P25	P50	P75
Panel I: Disclosing Time-On-Market Only					
Weeks to sale	1.835	3.312	0	1	5
No. of home viewings	4.480	8.133	0	3	10
Transaction price	38,768	89,684	1,155	39,133	87,404
Buyer surplus	-4,744	10,274	-9,607	-4,432	-53
Seller surplus	35,568	84,273	2,992	37,188	82,196
Total surplus	30,971	75,380	693	31,939	72,490
Panel II: Additional Effect of Disclosing Home-Viewing					
Weeks to sale	-0.110	1.883	0	0	0
No. of home viewings	-0.250	4.330	0	0	0
Transaction price	-3,287	37,004	-11,087	0	3,600
Buyer surplus	411	4,147	-439	0	1,237
Seller surplus	-4,187	38,558	-11,161	0	3,600
Total surplus	-3,777	34,726	-10,155	0	3,170
Panel III: Total Effect of Disclosing Both					
Weeks to sale	1.725	3.263	0	2	5
No. of home viewings	4.230	7.890	0	4	10
Transaction price	35,480	93,671	0	39,440	77,680
Buyer surplus	-4,186	10,811	-9,060	-4,264	0
Seller surplus	31,380	88,127	0	32,256	69,466
Total surplus	27,194	78,612	246	28,289	63,828

Note: The sample includes 27,703 apartments located in five core districts of Beijing. All homes are listed and sold through Lianjia from August 2015 to July 2016. Transaction price, buyer surplus, seller surplus, and total surplus are measured in Chinese Yuan (CNY).

in an indirect way. First, equation (4)) clearly shows that a higher belief directly reduces the sale probability conditional on the number of viewers. So the disclosure of this information leads to longer sale time. Second, a higher belief reduces the number of viewers in the first place and hence extends the sale time in an indirect way. Our simulation results suggest that disclosing the time-on-market information extends the sale time by 1.835 weeks on average. This effect is substantial, considering that the average sale time is 5.439 weeks in the data. Due to the significant extension of sale time, homes accumulate 4.48 more home viewings before sale on average. As a result of a higher transaction price and a larger number of home viewings, the expected buyer surplus per home is 4,744 CNY lower on average. In contrast, the expected seller surplus is 35,568 CNY higher on average. As a whole, the total surplus is 30,971 CNY higher per home on average.

Additional Effect of Disclosing the Home-Viewing Information. The additional effects when the home-viewing information is also disclosed in additional to the time-on-market information are measured by the difference between our proposed model where buyers know and NI model where buyers only know time on market. Buyers in both scenarios update their belief about the seller's reservation value if a home is still on the market. However, the updating rule differs in these two cases. When buyers know the number of home viewings occurred so far, their posterior belief would be based on that information, see equation (10). In contrast, when buyers do not know the home viewings, their posterior belief is an weighted average over all possible number of viewers, see equation (13). Take the case when no one views the home at time t (i.e., $n_t = 0$) as an example. The posterior belief would be the same as the prior belief when buyers know the exact number of viewers, whereas the posterior belief would be higher than the prior belief when buyers do not know the exact number. Intuitively, whether disclosing the home-viewing information would increase buyers' belief is ambiguous, depending on the actual number of viewers is higher or lower than the average number.

The panel (II) of Table 8 reports the additional effects of disclosing the home-viewing information in additional to the time-on-market information. As discussed above, the disclosure of this additional information would have different effects on different homes. For some homes, the transaction price would be higher, buyers are worse off and sellers are better off. However, the effects on other homes are totally opposite: the transaction price would be lower, buyers are better off and sellers are worse off. Our simulation results show that on average, the disclosure of this additional information slightly reduces the sale time and the total number of home viewings. On average, the transaction price is 3,287 CNY lower, the expected buyer surplus per home is 411 CNY higher, the expected seller surplus is 4,187 CNY lower, and the total surplus is 3,777 CNY lower.

Total Effect of Disclosing Both Time-on-Market and Home-Viewing Information. The panel (III) of Table 8 reports the total effects of disclosing both the time-on-market information and the home-viewing information, that is, the difference between our proposed model and and NI model. Compared to the case where buyers do not observe neither time on market nor home viewings, disclosing these two pieces of information leads to 1.725 weeks longer sale time, 4.23 more home viewings, and 35,480 CNY higher transaction price. As a result, the expected buyer surplus is 4,186 CNY lower on average, accounting for 12.76% of the average buyer surplus computed based on our proposed model (32,813 CNY). The expected seller surplus is 31,380 CNY higher on average, which is 55.9% of the average seller surplus suggested by our model (56,136 CNY). The total surplus per home is 27,194 CNY higher on average, a 29% increase.

6.3 Policy Implications

Our study suggests that compared to the case that buyers do not have the time-on-market information, the disclosure of this information increases the final transaction price by 38,768 CNY (1.3% of the average transaction price in the data), lowers buyer surplus by 4,744 CNY and increases seller surplus by 35,568 CNY. These predictions are opposite to most existing studies on the role of the time-on-market information (e.g.,Taylor (1999) and Tucker et al. (2013)). This is because buyers update their beliefs about sellers' reservation value from this information in our model while buyers learn of home quality in those studies. By highlighting the learning of sellers' reservation value, our study provides a new angle for understanding the welfare effect of disclosing the time-on-market information.

Moreover, we have also examined the effects of the disclosure of home-viewing information. According to our study, disclosing this additional information has ambiguous effects on the final transaction price, buyer surplus and seller surplus, compared to the case that buyers only have the time-on-market information. Our simulation results suggest that due to the disclosure of this additional information in our studying market, the transaction price is reduced on average and as a result, buyers are better off and sellers are worse off on average. Notice that on average, the additional effect of disclosing the home-viewing information are in the opposite direction to effects of the disclosure of the time-on-market information. However, this additional information has very different effects on individual homes.

As we are aware of, our study is the first empirical analysis that examines the effects of disclosing the home-viewing information in the housing market. This is particularly important in our studied market because Beijing government recently banned real estate platforms disclosing this information. According to our analysis, this ban would potentially have very different impacts on individual homes. Some homes may get lower transaction price and as a result, buyers would be better off while sellers would be worse off. However, the impacts on some other homes may be totally opposite. The counterfactual exercise based on our sample (2015-2016) suggest that this ban would slightly increase the transaction price, harm home buyers and benefit home sellers on average.

7 Conclusion

In this paper we present a model of buyer learning, search, and bid in search markets and estimate it using a unique data set of the Beijing housing market. The paper makes two main contributions. First, it provides an empirical framework for studying markets in which buyers need to search for best matches and list prices have limited commitment. We demonstrate that at equilibrium, the transaction price can be higher than, or equal to, or lower than the list price and moreover, the distribution of the transaction price has a mass point at the list price. This result is consistent with the empirical evidence that has been widely documented but not sufficiently emphasized in the literature of housing markets.

The second contribution of this paper is that it quantifies the welfare effect of two important pieces of information disclosed to prospective buyers: time on market and home viewings. We find that due to the disclosure of time-on-market information, buyer surplus is lower and seller surplus is higher on average. Moreover, the disclosure of the home-viewing information in addition to the time-on-market information has ambiguous impacts on home buyers and sellers. On average, disclosing this additional information would slightly benefit buyers but harm sellers. Our findings have important implications for intermediaries' design of information disclosure and relevant regulations in this type of search markets.

A few caveats should be noticed. First, although the overall fit of our model is good, it predicts that homes, especially those homes with relatively low list price, have shorter sale time and lower transaction price than they are observed in the data. This difference is directly related to our modeling assumptions: the gap between a home's list price and the seller's reservation value is a random term that follows the same distribution across sellers. This assumption allows us to closely links a home's list price and the seller's reservation value without introducing sellers' strategic behavior into the model. Due to this assumption, however, our model tends to underestimate the reservation values of sellers who set relatively low list prices (relative to their homes' predicted financial values).

Furthermore, we assume that buyers have perfect information about homes' quality. They view homes in person just to resolve their uncertainty about their match values which are assumed to be i.i.d. across buyers. So buyers do not learn about home quality from the choices of other buyers. Instead, we model buyer learning of the seller's reservation value, which we believe is the top reason of learning in our studying market. Nevertheless, ignoring buyers' learning of product quality could bias our results in a non-negligible way. It could over-predict the positive effect on the transaction price when we assess the effect of disclosing the time-on-market information. A natural future research direction is to study the welfare effect of information disclosure by incorporating these two learning channels - learning of product quality and learning of seller's reservation value - into a single framework.

References

Albrecht, James, Pieter A. Gautier, and Susan Vroman, "Directed Search in the Housing Market," *Review of Economic Dynamics*, 2016, *19*, 218–231.

Allen, Jason, Robert Clark, and Jean-François Houde, "Search frictions and market power in negotiated price markets," *Journal of Political Economy*, 2019, *127* (4).

Athey, Susan and Philip Haile, "Identification of Standard Auction Models," *Econometrica*, 2002, *70* (6), 2107–2140.

Banerjee, Abhijit V., "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 1992, *107* (3), 797–817.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades," *Journal of Political Economy*, 1992, *100* (5), 992–1026.

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen, "Calorie Posting in Chain Restaurants," *American Economic Journal: Economic Policy*, 2011, *3* (1), 91–128.

Cai, Hongbin, Yuyu Chen, and Hanming Fang, "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review*, 2009, *99* (3), 964–882.

Chen, Yongmin and W. Robert Rosenthal, "Asking Prices as Commitment Devices," *International Economic Review*, 1996, *37* (1), 129–155.

_ **and** _ , "On the Use of Ceiling-Price Commitments by Monopolists," *The RAND Journal of Economics*, 1996, *27* (2), 207–220.

Genesove, David and Lu Han, "Search and Matching in the Housing Markets," *Journal of Urban Economics*, 2012, *72*, 31–45.

Han, Lu and Will Strange, "The Microstructure of Housing Markets: Search, Bargaining, and Brokerage," *Handbook of Regional and Urban Economics*, 2015, pp. 813–886.

_ **and** _ , "What is the Role of the Asking Price for a House?," *Journal of Urban Economics*, 2016, *93*, 115–130.

Hendel, Igal, Aviv Nevo, and Francois Ortalo-Magne, "The Relative Performance of Real Estate Marketing Platforms: MLS versus FSBOMadison.com," *American Economic Review*, 2009, *99* (5), 1878–98.

Hendricks, Kenneth, Alen Sorensen, and Thomas Wiseman, "Observational Learning and Demand for Search Goods," *American Economic Journal: Microeconomics*, 2012, *2* (1), 1–31.

Hong, Han and Matthew Shum, "Using price distributions to estimate search costs," *The RAND Journal of Economics*, 2006, *37* (2), 257–275.

Honka, Elisabeth, "Quantifying search and switching costs in the US auto insurance industry," *The Rand Journal of Economics*, 2014, *45* (4), 847–884.

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

Hortaçsu, Ali and Chad Syverson, "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds," *The Quarterly Journal of Economics*, 2004, pp. 403–456.

Jin, Ginger Zhe and Phillip Leslie, "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *The Quarterly Journal of Economics*, 2003, *118* (2), 409–451.

Lester, Benjamin, Ludo Visschers, and Ronald Wolthoff, "Competing with Asking Prices," *Theoretical Economics*, 2017, *12*, 731–770.

Levitt, Steve and Chad Syverson, "Market Distortions When Agents are Better Informed: The Value of Information in Real Estate Transactions," *Review of Economics and Statistics*, 2008, *90*, 599–611.

los Santos, Babur De, Ali Hortaçsu, and Matthijs R Wildenbeest, "Testing models of consumer search using data on web browsing and purchasing behavior," *The American Economic Review*, 2012, *102* (6), 2955–2980.

Luca, Michael, "Reviews, Reputation, and Revenue: The Case of Yelp.com," *Harvard Business School Working Paper*, 2016.

Luco, Fernando, "Who Benefits from Information Disclosure? The Case of Retail Gasoline," *American Economic Journal: Microeconomics*, 2019, *11* (2), 277–305.

Merlo, Antonio and Francois Ortalo-Magne, "Bargaining over Residential Real Estate: Evidence from England," *Journal of Urban Economics*, 2004, *56*, 192–216.

_ , _ , and John Rust, "The Home Selling Problem: Theory and Evidence," *International Economic Review*, 2015, *56* (2), 457–484.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R Wildenbeest, "Consumer search and prices in the automobile market," *unpublished manuscript*, 2018.

Newberry, Peter W., "An Empirical Study of Observational Learning," *RAND Journal of Economics*, 2016, *47* (2), 394–432.

Rossi, Federico and K. Pradeep Chintagunta, "Price Transparency and Retail Prices: Evidence from Fuel Price Signs in the Italian Highway System," *Journal of Marketing Research*, 2016, *53* (3), 407–423.

Schultz, Christian, "Transparency on the Consumer Side and Tacit Collusion," *European Economic Reivew*, 2005, *49* (2), 4279–297.

Taylor, Curtis R., "Time-on-the-Market as a Sign of Quality," *The Review of Economic Studies*, 1999, *66* (3), 555–578.

Tucker, Catherine, Juanjuan Zhang, and Ting Zhu, "Days on market and home sales," *RAND Journal of Economics*, 2013, *44* (2), 337–360.

Yavas, Abdullah and Shiawee Yang, "The Strategic Role of Listing Price in Marketing Real Estate: Theory and Evidence," *Real Estate Economics*, 1995, *23* (3), 347–368.

Yavorsky, Dan, Elisabeth Honka, and Keith Chen, "Consumer Search in the U.S. Auto Industry: The Role of Dealership Visits," *Quantitative Marketing and Economics*, 2020, *Forthcoming.*

Zhang, Juanjuan, "The Sound of Silence: Observational Learning in The U.S. Kidney Market," *Marketing Science*, 2010, *29* (2), 315–335.

A Computation Details

A.1 Algebra Details

- The first part of expected gain of a viewer with match value $v \in [a_t, \hat{v}_t]$ from rejecting a_t is:

$$\begin{aligned} & \int_{a_t-\bar{\eta}}^{a_t} \left[\int_{\underline{v}}^{a_t} (v - max(z, r_t)) dF(z)^{n_t-1} \right] g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t \\ = & \int_{a_t-\bar{\eta}}^{a_t} \left[\int_{\underline{v}}^{r_t} (v - r_t) dF(z)^{n_t-1} + \int_{r_t}^{a_t} (v - z) dF(z)^{n_t-1} \right] dG_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) \\ = & (v - a_t) F(a_t)^{n_t-1} + \int_{a_t-\bar{\eta}}^{a_t} G_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) F(r_t)^{n_t-1} dr_t, \end{aligned}$$

  where the last term comes from integration by parts.

- The expected gain when knowing (v, n_t) is given by:
  - When $v \le a_t$,

$$\begin{aligned} \Gamma(v, n_t, a_t, g_t) &= \int_{a_t-\bar{\eta}}^{v} \left[\int_{\underline{v}}^{v} (v - max(z, r_t)) dF(z)^{n_t-1} \right] g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t \\ &= \int_{a_t-\bar{\eta}}^{v} F(r_t)^{n_t-1} G_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t \end{aligned}$$

  - When $a_t < v \le \hat{v}_t$,

$$\begin{aligned} \Gamma(v, n_t, a_t, g_t) &= \int_{a_t-\bar{\eta}}^{a_t} \left[\int_{\underline{v}}^{a_t} (v - max(z, r_t)) dF(z)^{n_t-1} \right] g_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t \\ &\quad + \int_{a_t}^{v} (v - z) dF(z)^{n_t-1} \\ &= \int_{a_t}^{v} F(z)^{n_t-1} dz + \int_{a_t-\bar{\eta}}^{a_t} G_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) F(r_t)^{n_t-1} dr_t. \end{aligned}$$

– When $v > \hat{v}_t$,

$$\begin{aligned}
\Gamma(v, n_t, a_t, g_t) &= \int_{\underline{v}}^{v} (v - a_t \cdot 1(z \le \hat{v}_t) - z \cdot 1(z > \hat{v}_t)) dF(z)^{n_t - 1} \\
&= (\hat{v}_t - a_t) F(\hat{v}_t)^{n_t - 1} + \int_{\hat{v}_t}^{v} F(z)^{n_t - 1} dz
\end{aligned}$$

Hence, the expected gain for each n_t when v is not known, equation (7) is given by:

$$\begin{aligned}
EU(n_t, a_t, g_t) &= \int_{\underline{v}}^{a_t} \left[\int_{a_t - \bar{\eta}}^{v} F(r)^{n_t - 1} G_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) dr_t \right] dF(v) \\
&+ \int_{a_t}^{\hat{v}_t(n_t)} \left[\int_{a_t}^{v} F(z)^{n_t - 1} dz \right] dF(v) \\
&+ (F(\hat{v}_t(n_t)) - F(a_t)) \int_{a_t - \bar{\eta}}^{a_t} G_t(r_t \mid \vec{n}_{t-1}, \vec{a}_t) F(r_t)^{n_t - 1} dr_t \\
&+ (\hat{v}_t(n_t) - a_t) F(\hat{v}_t(n_t))^{n_t - 1} (1 - F(\hat{v}_t(n_t))) \\
&+ \int_{\hat{v}_t(n_t)}^{\bar{v}} \left[\int_{\hat{v}_t(n_t)}^{v} F(z)^{n_t - 1} dz \right] dF(v)
\end{aligned}$$

A.2 Comparative Statics

We omit the subscript t for notation simplicity. The equilibrium probability of n viewers is determined by the equation (8) and (9) in the text:

$$\mathcal{P}(n \mid \hat{c}) = \sum_{m \ge n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Psi(\hat{c})^{n-1} [1 - \Psi(\hat{c})]^{m-n}$$

$$\hat{c} = \sum_{n=1}^{\infty} \mathcal{P}(n \mid \hat{c}) \times EU(n, a, g_t)$$

Hence, $\mathcal{P}(n)$ is determined by the following equation:

$$\mathcal{P}(n) = \sum_{m \ge n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi\Big(\frac{\sum_{n'=1}^{\infty} \mathcal{P}(n') EU(n')}{\sigma_c} - \frac{\mu_c}{\sigma_c} \Big)^{n-1} \Big[1 - \Phi\Big(\frac{\sum_{n'=1}^{\infty} \mathcal{P}(n') EU(n')}{\sigma_c} - \frac{\mu_c}{\sigma_c} \Big) \Big]^{m-n}.$$

We can derive the following comparative statics. We compute the derivative of the $\hat{c}$ which equals to the expected gain from viewing at equilibrium, i.e., $\hat{c} \equiv \sum_{n=1}^{\infty} EU(n)\mathcal{P}(n)$, w.r.t. the parameters, μ_c, σ_c, and λ.

1. The derivative of $\hat{c}$ w.r.t. μ_c.

$$
\begin{aligned}
\frac{\partial \mathcal{P}(n)}{\partial \mu_c} &= \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Big[(n-1)\Phi^{n-2}\frac{\partial \Phi}{\partial \mu_c}(1-\Phi)^{m-n} - \Phi^{n-1}(m-n)(1-\Phi)^{m-n-1}\frac{\partial \Phi}{\partial \mu_c}\Big] \\
&= \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Big[(n-1)(1-\Phi) - (m-n)\Phi\Big]\Phi^{n-2}(1-\Phi)^{m-n-1}\frac{\partial \Phi}{\partial \mu_c} \\
&= \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1) - (m-1)\Phi\Big]\phi\frac{1}{\sigma_c}\Big[\sum_{n'=1}^{\infty} EU(n')\frac{\partial \mathcal{P}(n')}{\partial \mu_c} - 1\Big]
\end{aligned}
$$

$$
\begin{aligned}
\sum_{n=1}^{\infty} EU(n)\frac{\partial \mathcal{P}(n)}{\partial \mu_c} &= \sum_{n=1}^{\infty} EU(n) \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1} \\
&\quad \times\Big[(n-1) - (m-1)\Phi\Big]\phi\frac{1}{\sigma_c}\Big[\sum_{n'=1}^{\infty} EU(n')\frac{\partial \mathcal{P}(n')}{\partial \mu_c} - 1\Big]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \hat{c}}{\partial \mu_c} &\equiv \sum_{n=1}^{\infty} EU(n)\frac{\partial \mathcal{P}(n)}{\partial \mu_c} \\
&= \frac{\sum_{n=1}^{\infty} EU(n) \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1) - (m-1)\Phi\Big]\phi}{\sum_{n=1}^{\infty} EU(n) \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1) - (m-1)\Phi\Big]\phi - \sigma_c}
\end{aligned}
$$

2. The derivative of $\hat{c}$ w.r.t. σ_c.

$$
\begin{aligned}
\frac{\partial \mathcal{P}(n)}{\partial \sigma_c} &= \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Big[(n-1)\Phi^{n-2}\frac{\partial \Phi}{\partial \sigma_c}(1-\Phi)^{m-n} - \Phi^{n-1}(m-n)(1-\Phi)^{m-n-1}\frac{\partial \Phi}{\partial \sigma_c}\Big] \\
&= \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Big[(n-1)(1-\Phi) - (m-n)\Phi\Big]\Phi^{n-2}(1-\Phi)^{m-n-1}\frac{\partial \Phi}{\partial \sigma_c} \\
&= \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1) - (m-1)\Phi\Big] \\
&\quad \times\phi\frac{1}{\sigma_c}\Big[\sum_{n'=1}^{\infty} EU(n')\frac{\partial \mathcal{P}(n')}{\partial \sigma_c} - \frac{\sum_{n'=1}^{\infty} \mathcal{P}(n')EU(n') - \mu_c}{\sigma_c}\Big]
\end{aligned}
$$

$$
\begin{aligned}
\sum_{n=1}^{\infty} EU(n)\frac{\partial \mathcal{P}(n)}{\partial \sigma_c} &= \sum_{n=1}^{\infty} EU(n) \sum_{m\geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1) - (m-1)\Phi\Big] \\
&\quad \times\phi\frac{1}{\sigma_c}\Big[\sum_{n'=1}^{\infty} EU(n')\frac{\partial \mathcal{P}(n')}{\partial \sigma_c} - \frac{\sum_{n'=1}^{\infty} \mathcal{P}(n')EU(n') - \mu_c}{\sigma_c}\Big]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \hat{c}}{\partial \sigma_c} &\equiv \sum_{n=1}^{\infty} EU(n) \frac{\partial \mathcal{P}(n)}{\partial \sigma_c} \\
&= \frac{\sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1} \Big[(n-1)-(m-1)\Phi\Big] \phi \frac{\sum_{n'=1}^{\infty} \mathcal{P}(n') EU(n') - \mu_c}{\sigma_c}}{\sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1} \Big[(n-1)-(m-1)\Phi\Big] \phi - \sigma_c}
\end{aligned}
$$

3. The derivative of $\hat{c}$ w.r.t. λ.

$$
\begin{aligned}
\frac{\partial \mathcal{P}(n)}{\partial \lambda} &= \sum_{m \geq n} \frac{\partial \Lambda(m)}{\partial \lambda} \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-1}(1-\Phi)^{m-n} \\
&\quad + \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Big[(n-1)(1-\Phi)-(m-n)\Phi\Big] \Phi^{n-2}(1-\Phi)^{m-n-1} \frac{\partial \Phi}{\partial \lambda} \\
&= \sum_{m \geq n} \Big[\frac{m\lambda^{m-1}e^{-\lambda}}{m!} - \frac{\lambda^m e^{-\lambda}}{m!}\Big] \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-1}(1-\Phi)^{m-n} \\
&\quad + \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1} \Big[(n-1)-(m-1)\Phi\Big] \phi \frac{1}{\sigma_c} \Big[\sum_{n'=1}^{\infty} EU(n') \frac{\partial \mathcal{P}(n')}{\partial \lambda}\Big]
\end{aligned}
$$

$$
\begin{aligned}
\sum_{n=1}^{\infty} EU(n) \frac{\partial \mathcal{P}(n)}{\partial \lambda} &= \sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \frac{m\lambda^{m-1}e^{-\lambda}}{m!} \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-1}(1-\Phi)^{m-n} \\
&\quad - \sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \frac{\lambda^m e^{-\lambda}}{m!} \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-1}(1-\Phi)^{m-n} \\
&\quad + \sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1} \\
&\quad \times \Big[(n-1)-(m-1)\Phi\Big] \phi \frac{1}{\sigma_c} \Big[\sum_{n'=1}^{\infty} EU(n') \frac{\partial \mathcal{P}(n')}{\partial \lambda}\Big]
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \hat{c}}{\partial \lambda} &\equiv \sum_{n=1}^{\infty} EU(n) \frac{\partial \mathcal{P}(n)}{\partial \lambda} \\
&= \frac{\sum_{n=1}^{\infty} EU(n)\mathcal{P}(n) - \sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \frac{m\lambda^{m-1}e^{-\lambda}}{m!} \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-1}(1-\Phi)^{m-n}\sigma_c}{\sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1)-(m-1)\Phi\Big]\phi - \sigma_c} \\
&= \frac{\hat{c} - \sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \frac{m}{\lambda}\Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-1}(1-\Phi)^{m-n}\sigma_c}{\sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1)-(m-1)\Phi\Big]\phi - \sigma_c}
\end{aligned}
$$

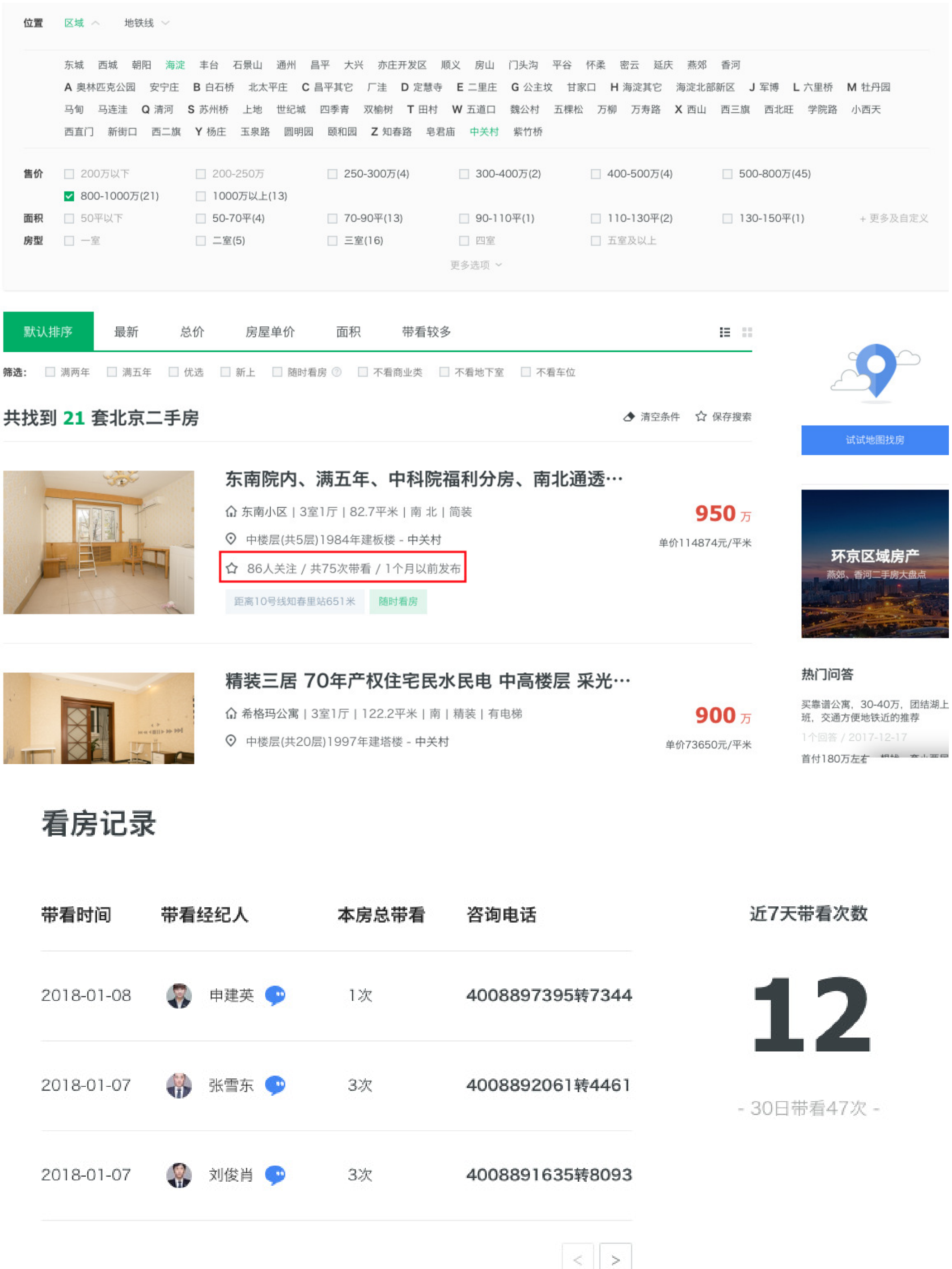
Therefore,

$$
\frac{\frac{\partial \hat{c}}{\partial \sigma_c}}{\frac{\partial \hat{c}}{\partial \mu_c}} = \frac{\sum_{n'=1}^{\infty} \mathcal{P}(n')EU(n') - \mu_c}{\sigma_c} = \frac{\hat{c} - \mu_c}{\sigma_c}
$$

$$
\frac{\frac{\partial \hat{c}}{\partial \lambda}}{\frac{\partial \hat{c}}{\partial \mu_c}} = \frac{\hat{c} - \sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \frac{m}{\lambda}\Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-1}(1-\Phi)^{m-n}\sigma_c}{\sum_{n=1}^{\infty} EU(n) \sum_{m \geq n} \Lambda(m) \begin{pmatrix} n-1 \\ m-1 \end{pmatrix} \Phi^{n-2}(1-\Phi)^{m-n-1}\Big[(n-1)-(m-1)\Phi\Big]\phi}
$$

B Additional Figures

Figure 6: Real-Time Home Viewings Disclosed on Lianjia's Website



Note: The top graph displays the home listings on Lianjia's website. The bottom graph displays the home viewing records of a given home listing.