Dynamic Search in a Non-Stationary Search Environment: An Application to the Beijing Housing Market*

Ying Fan[†]
University of Michigan
CEPR and NBER

Yiyi Zhou[‡]
Stony Brook University

Ziying Fan[§]
Shanghai University of
Finance and Economics

June 8, 2024

Abstract

This paper investigates how dynamics in the search environment affect consumers' search and purchase decisions. We develop a dynamic search model with a non-stationary search environment. We estimate the model using data on consumers' search duration, search intensity in each period, and purchase decision in the Beijing housing market. We find that the dynamics in the search environment have larger effects on consumers' search and purchase decisions and welfare than traditional search frictions such as search costs. Moreover, we find that a static search model would yield unrealistically high estimates of search costs.

JEL: C23, E31, D8, L8
Keywords: dynamic search, dynamic demand, Beijing housing market

1 Introduction

In many markets, consumers need to search before making a purchase decision. The extent to which a consumer searches thus determines the choice set for the purchase. Consequently, search frictions hinder search and limit the choice set. Examples abound in various sectors: the automobile market, the travel accommodation market, the real estate market, the mortgage market, and many more.

*We thank Andrew Joung and Takahiro Moriya for their research assistance.
[†]Department of Economics, University of Michigan, Ann Arbor, MI 48109; yingfan@umich.edu.
[‡]Department of Economics, Stony Brook University, Stony Brook, NY 11794; yiyi.zhou@stonybrook.edu.
[§]School of Public Finance and Administration, Shanghai University of Finance and Economics, Shanghai 200433, China; fan.ziying@mail.shufe.edu.cn.

Furthermore, in many scenarios, the search environment changes over time as prices change, new products enter the market, or existing products exit the market. For example, consumers searching for a car during the COVID-19 pandemic faced rapid price increases and volatile dealer inventories. Homebuyers in 2022–2023 saw significant increases in mortgage rates. Consumers planning vacations during peak seasons may experience rapid changes in the price and availability of accommodations and travel services. The real estate market is also dynamic, with property availability and prices changing over time.

In this paper, we study how dynamics in the search environment affect consumers' search and purchase decisions. Consider a scenario in which prices increase rapidly. On the one hand, by searching longer, a consumer can learn about more products, expand her choice set, and find a more suitable product. On the other hand, if she searches longer, her preferred product in her previously searched set will become more expensive, and the products she can search for in the future will have higher prices than the current price level. Thus, the dynamics of the search environment affect the consumer's search and purchase outcomes.

To study such effects, we develop a dynamic search model with a non-stationary search environment. In this model, consumers make search (how much to search in each period) and purchase (when to purchase and which product to purchase) decisions. Specifically, a consumer's utility depends on observable product characteristics (e.g., the location and size of a property) and a match value that is learned only through the search process (e.g., the view and sun exposure of a property). At the beginning of each period, consumers have information about observable characteristics for all products and match values for their searched products. Then, in each period, a consumer decides the extent of her search and learns the match values of her searched products. She then decides whether to make a purchase and, if so, which product to buy, choosing from either her previously searched set or the newly searched set.

Both search and purchase decisions are dynamic in our model. The purchase decision is dynamic because consumers can choose to buy immediately or wait for another period. The search decision is also dynamic because higher search intensity in one period leads to a larger choice set and a higher value of buying now, along with a smaller set of products to visit in future periods and a lower value of waiting. In making both search and purchase decisions, consumers consider the evolving market environment. Although our model is developed in the context of housing search, it can be used to study other settings where consumers make dynamic search and purchase decisions in a changing environment.

To the best of our knowledge, this paper is the first to incorporate a non-stationary environment into a dynamic search model. In a changing environment, consumer search is affected not only by search costs but also by changes in the search environment. A static search model ignores the latter, leading to biased estimates of search costs. For example, in scenarios where the search environment deteriorates over time (e.g., due to an increase in prices or a decrease in product availability), both search costs and the change in the environment contribute to limiting consumer search. By ignoring the dynamics of the environment, a static model would lead to an overestimation of search costs, and thus an overestimation of consumer gains from reducing search costs.

Our empirical setting is the Beijing housing market between August 1, 2015 and July 31, 2016. This setting is ideal for studying how the dynamics of the search environment affect consumers' search and demand for two reasons. First, the period is characterized by rapid price appreciation, with an annual increase of 30%. Second, the dataset we use for our analysis is novel in that it contains information not only on consumers' purchase decisions, but also on their complete search activities. Our data come from the largest real estate agency in Beijing, Lianjia. For each consumer in our sample, the dataset provides a complete record of the consumer's search and purchase behavior. We observe when the consumer starts searching, when she stops searching, how many and which properties she visits in each period, and which property she finally purchases.

Estimating our dynamic search and purchase model is challenging, and a solution used in the literature cannot be easily adapted to meet this challenge. In our model, the state variable, which includes the characteristics of all unsearched products as well as the characteristics and match values of all searched products, has a very large dimension. Such a challenge also arises when estimating a dynamic demand model where the state variable includes the observable characteristics of all products in the current choice set. The dynamic demand literature (e.g., Gowrisankaran and Rysman, 2012) addresses this challenge by reducing the high-dimensional state variable to a one-dimensional inclusive value and assuming that the inclusive value evolves according to a stationary Markov process. However, unlike a dynamic demand model where the state variable describes the choice set and the choice set evolves exogenously, in our model the state variable describes two sets (the searched and unsearched sets) and both sets evolve endogenously. Thus, it is inappropriate to assume that the inclusive value of each set evolves according to a stationary Markov process.

We address this challenge with two ideas. First, before estimating the dynamic model, we back out the mean utility of each property by matching the observed share of visits that the property receives. In estimating the dynamic model, we replace the vector of characteristics with the scalar mean utility for each property. Second, following many papers in the dynamic estimation literature,[1] we further reduce the dimensionality of the state

[1] Many papers approximate a high-dimensional state variable with lower-dimensional statistics. Examples include Collard-Wexler (2013), Sweeting (2013), and Hodgson (2019).

variable by assuming that instead of the mean utilities of all properties, consumers track several highest mean utilities, the average of the mean utilities of the other properties, and the number of properties.

Correspondingly, we carry out the estimation in two steps.[2] In the first step, we back out the mean utility of each property. We do this by matching the model implication for the share of visits that each property receives to the observed share. In this step, we also estimate the coefficients of the property characteristics in the utility function. We refer to these parameters as static parameters. In the second step, we estimate parameters that capture the variance of match values (which determines the magnitude of search benefits), search costs, and waiting costs. We refer to these parameters as dynamic parameters. We estimate these dynamic parameters by solving the dynamic search and purchase model, obtaining the model implications for search duration, number of visits in each period, and purchase decision, and searching for parameters that maximize the likelihood of the observed outcomes.

Such sequential estimation requires that the model implication for the share of visits a property receives, which we use to back out the mean utilities and estimate the static parameters, does not depend on the dynamic parameters. This requirement is satisfied because we assume that while consumers choose the number of properties to search in each period, the set of properties they visit is exogenous and drawn from a distribution that depends on the mean utilities of the properties. A similar exogeneity assumption is made in Hortaçsu and Syverson (2004). However, unlike Hortaçsu and Syverson (2004), where consumers sample a single index fund, consumers in our model sample a set of properties. Therefore, to carry out the first stage of estimation, we extend a standard Logit model from choosing a single option to choosing a set of options. We derive the probability that a set is sampled, which then implies the probability that a particular option is sampled. We also extend the contraction mapping result in Berry, Levinsohn and Pakes (1995) to show that there is a unique vector of mean utilities such that the model-implied visit shares match the observed visit shares even in our extended model.

Our estimation yields intuitive results. According to our estimates, consumers prefer larger and newer properties with more living rooms and bedrooms, located on higher floors and close to subway stations. The estimated standard deviation of match values is equivalent to a value of CNY100,593, which is about 3% of the average list price and more than 150% of the per capita annual disposal income in Beijing in 2016, indicating a significant benefit

[2] Similar sequential estimation procedures are used to estimate multi-stage static models (e.g., Eizenberg, 2014; Fan and Yang, 2020, 2022) and dynamic models (e.g., Chatterjee, Fan and Mohapatra, 2022; Elliott, 2022; Bodéré, 2023).

of searching to learn match values. We also find that consumers incur an average search cost of CNY1,244. In our data, a consumer visits an average of 6.71 properties before making a purchase. Therefore, the average search cost per property is CNY185.

In contrast, a static search model estimated on the same dataset yields an estimate of the standard deviation of match values that is 50% smaller, while yielding a search cost estimate that is 250 times larger. The static model yields a smaller estimate of the standard deviation parameter than the dynamic model because the static model uses less variation in the data to identify this parameter. The static model uses the extent to which observable property characteristics cannot explain purchasing behavior to identify the standard deviation of the unobservable match value. While the static model relies only on such a static feature of the data (which property is purchased), the dynamic model additionally exploits a dynamic feature of the data, i.e., when the purchased property is visited. In particular, a larger recall share (the share of consumers who purchase a property that they visited in an earlier period) implies a larger standard deviation of the match value. This is because if a consumer does not immediately purchase the property she eventually purchases, her decision to continue searching means that she thinks there is a good chance of getting a better draw of the match value. In our data, the recall rate is over 15%. At the same time, the static model yields a much larger estimate of search costs despite the smaller estimate of search benefits (due to a smaller estimate of the standard deviation of the match value). The search cost per visit is about CNY 46,000, or $7,000, which we find unreasonably high. This is because the static model does not account for the dynamics in the search environment. During the sample period, house prices increased rapidly. By ignoring the increased prices as a search friction, the static model overestimates the search cost.

Based on the estimated dynamic search model, we quantify how the dynamics of the environment affect consumers' search and purchase behavior and consumer welfare through counterfactual simulations. We consider both changes in prices and changes in product availability (i.e., entry of new listings and exit of existing listings). We find that halving the price increase, doubling the entry rate of new listings, or halving the exit rate of existing listings leads to longer searches, more total visits before purchase, and purchases of properties that generate more utility. Considering the trade-off between higher search and waiting costs (resulting from more property visits and longer searches) and higher utility from finding a more desirable property (again, resulting from more searches), the net gain is CNY16,696 when consumers face a slower price increase, CNY47,269 when the entry rate is doubled, and CNY18,247 when the exit rate is halved.[3]

[3]In the counterfactual scenario where the price increase is halved, consumer welfare increases mechanically due to lower prices. To remove such a mechanical effect and to isolate the effect of inducing more search, we

To quantify the relative importance of traditional search frictions such as search costs versus search frictions due to dynamics in the search environment, we also conduct a counterfactual simulation in which we reduce the search cost per visit by half. Unsurprisingly, consumers search for more weeks, visit more properties before buying, and buy properties with higher utility. The net gain is CNY16,103. Thus, both traditional search frictions such as search costs and search frictions arising from the dynamics of the search environment significantly affect consumers' search and purchase behavior and welfare. For the same percentage change (halving or doubling), changing the dynamics of the search environment has a greater impact than changing the search cost.

This paper contributes to the empirical consumer search literature. Examples of this literature include Honka (2014) and Murry and Zhou (2020) for simultaneous search models and Moraga-Gonzalez, Sandor and Wildenbeest (2023) and Hodgson and Lewis (2023) for sequential search models, and Santos, Hortaçsu and Wildenbeest (2012) for testing simultaneous versus sequential models.[4] While the existing papers consider stable search environments where both the set of products and their prices are fixed, our paper studies consumer search in changing search environments. Here, both product availability and prices can vary over time, impacting consumers' search and purchase decisions. Our results indicate that these changes have significant effects on consumers' search and purchase behavior as well as consumer welfare.

This paper is also related to the dynamic demand literature. Examples of this literature include Hendel and Nevo (2006), Gowrisankaran and Rysman (2012), Lee (2013), Shcherbakov (2016), and Aguirregabiria (2023). We differ from these papers in two ways. First, we study both dynamic search and dynamic demand. In our model, consumers' search decisions determine the choice set for purchases. Therefore, unlike the papers in the dynamic demand literature, the choice set evolves endogenously. Second, we tackle the issue of a high-dimensional state space differently. Since the transition of the choice set is endogenous, we cannot assume that the inclusive value follows a Markov process and replace a high-dimensional state variable by a one-dimensional inclusive value. Instead, we estimate the mean utilities of properties "offline", i.e., before estimating the dynamic parameters, and consider the state variable to include the mean utilities of the top properties, the average mean utility of the remaining properties, and the number of properties.

The remainder of the paper is organized as follows: Section 2 describes the setting and the data. Section 3 develops the dynamic search and purchase model. Section 4 explains

report the net gain, where the utility of the purchased property is calculated based on the actual prices.

[4]Other examples include Kim, Albuquerque and Bronnenberg (2010), Allen, Clark and Houde (2019) and Brown and Jeon (2020).

the estimation procedure. Section 5 presents the estimation results. Section 6 compares our estimation results with those of a static search model. Section 7 quantifies the effects of the changing environment and search costs through counterfactual simulations. A final Section 8 concludes the paper.

2 Data

2.1 Data Description

Our data come from Lianjia, the largest brokerage company in the second-hand residential housing market in Beijing.[5] The dataset provides information on all second-hand properties listed on Lianjia and all consumers who registered on Lianjia between August 1, 2015 and July 31, 2016.

Our data cover 225,608 properties in the six urban districts of Beijing. Beijing Municipality consists of six urban districts, eight suburban districts and two rural counties. The six urban districts include two core districts (Dongcheng and Xicheng, which occupy the area inside the old walled city) and four surrounding urban districts (Haidian, Chaoyang, Fengtai, and Shijingshan). Our data also cover 455,774 consumers who registered on Lianjia and were actively searching during our sample period.

To construct our sample, we exclude properties with a list price higher than CNY 10 million or less than CNY 1 million, and properties with a size of less than 25 square meters. These properties are likely to belong to a separate market. Accordingly, we drop consumers who visited these properties. We also drop consumers who searched across multiple districts. The districts in Beijing are quite large. For example, Chaoyang District covers 470.8 square kilometers (181.8 square miles). In contrast, Manhattan Island is 59 square kilometers (22.7 square miles). As a result, the vast majority (more than 93%) of consumers searched within one district. At the same time, as will be explained later, including the small share of consumers searched across multiple districts increases the computational burden exponentially. Finally, we drop 4% of the properties in the data that were never visited by any consumer in our sample. In the end, our sample consists of 202,845 properties and 414,166 consumers.

For each property in the sample, we observe its address, construction year, floor number, property size, number of living rooms, and number of bedrooms. We also observe when it was initially listed and the list price. Based on the property's location and list price, we define 221 exclusive segments defined by the neighborhood and price range and assign each

[5]The market share of Lianjian, as measured by the share of total second-hand residential property transactions in Beijing, was 54% in the first half of 2016. In contrast, the market shares of the second, third and fourth real estate companies were 13%, 5% and 4%, respectively. (https://m.sohu.com/n/458280155/)

property to a segment.[6] Finally, if a property was sold before the end of the sample, we observe the transaction date and price.

For each consumer in the sample, we have a complete record of all her property visits. The search record is complete because all consumers in the sample sign a sole agency agreement with Lianjia. For each property visit, we observe the date of the visit and the identity of the property. In addition, if a consumer purchased a property during the sample period, we also observe which property she purchased and when she purchased it.

2.2 Summary Statistics

2.2.1 Properties

Table 1 presents the summary statistics of the properties in the sample. Of the 202,845 properties in our sample, 85,696 (42%) were successfully sold during our sample period. The average list price is CNY4.024 million (USD604,000), and the average transaction price is CNY3.702 million (USD555,000). In the Chinese housing market, the most salient price is the price per square meter, which averages CNY49,302 (\$18,029) and CNY48,367 (\$17,501) per square meter for the list and transaction prices, respectively. On average, a property stays on the market for approximately 8 weeks.

Table 1: Summary Statistics of Properties

	Mean	SD
List price (million CNY)	4.024	1.962
List price per m^2 (CNY)	49,302	18,029
Indicator of being sold	0.422	0.494
Transaction price (million CNY)	3.702	1.760
Transaction price per m^2 (CNY)	48,367	17,501
Weeks on market	8.193	7.402
Property size (m^2)	83.707	35.903
Property age (year)	18.151	8.994
Bedrooms	1.997	0.777
Living rooms	1.142	0.547
Indicator of above 10th floor	0.316	0.465
Close to subway stations	0.797	0.402

The average property is 18 years old, with a size of 84 square meters, 2 bedrooms and 1 living room. About 32% of the properties are located on the 10th floor or above. The

[6] Based on the location of the property, Lianjia divides the market into neighborhoods that differ in terms of transportation, amenities, etc. We consider four price ranges: less than 3 million, between 3 and 4.5 million, between 4.5 and 6 million, and more than 6 million.

majority of them (about 80%) are located within 1 km of a subway station, a criterion we use to define the indicator variable of whether a property is close to a subway station.

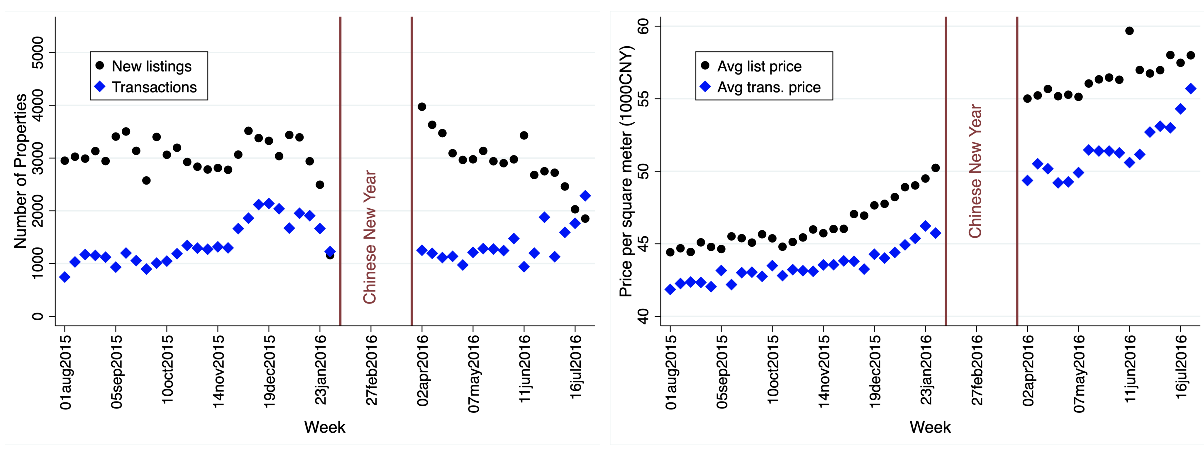
2.2.2 Dynamic Search Environment

Both the number of new listings and the number of transactions are relatively stable over the sample period, as shown in Figure 1(a), which plots these two numbers by week. We omit the eight weeks around the Chinese New Year in early 2016. Many economic activities are put on hold every year during the Chinese New Year, when many Chinese return to their hometowns and resume economic activities only afterward. For example, 2.9 billion passenger trips were made during the 2016 holiday (Zhou (2016)).[7]

However, both list and transaction prices increased rapidly during the sample period, as shown in Figure 1(b). For each week in our sample, we compute the average list price across all new listings in that week and the average transaction price across all transacted properties in that week. Since the most salient price in the Chinese housing market is the unit price per square meter, in calculating these averages we consider prices in CNY per square meter, and plot them in Figure 1(b). The figure shows a clear upward trend in both list and transaction prices. Specifically, the average list price increased from CNY 44,417 to CNY 57,993 per square meter, an annual price increase of about 30% during the sample year. On average, the list price increased by CNY 261 per square meter per week. Similarly, the transaction price increased from CNY41,854 to CNY55,699 per square meter during the sample period, an increase of CNY266 per week and an annual increase of 33%.

Figure 1: New Listings, Transactions, and Prices by Week

(a) New Listings and Transactions (b) List Prices and Transaction Prices



[7] In Supplemental Online Appendix SA, we plot the number of new listings and transactions during the eight weeks around the Chinese New Year in early 2016. It shows that during these eight weeks, the number of new listings and transactions dropped to zero and then rose rapidly to about double the pre-holiday level.

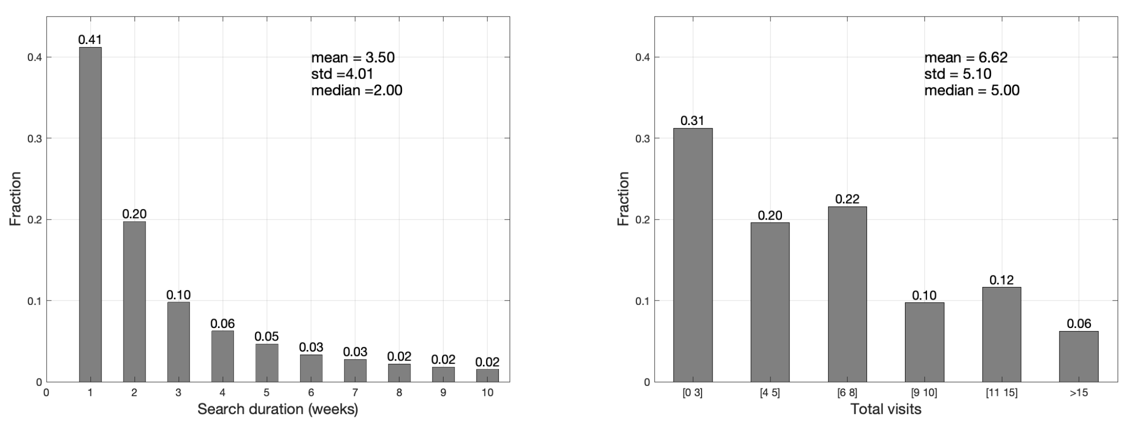
2.2.3 Consumer Search and Purchase

In our sample, 39,500 consumers started a search and made a purchase during the sample period. Among them, 26,543 consumers either ended their search before the Chinese New Year holiday or started their search after the holiday.

We observe these consumers' search duration and total visits. On average, they search for 3.5 weeks and visit a total of 6.62 properties. Figure 2 (a) shows that 41% of them searched for one week, 20% searched for two weeks, 10% searched for three weeks and the remaining 29% searched for more than three weeks. Figure 2 (b) shows that 51% of them visited five or fewer properties, 32% of them visited more than five but no more than ten properties, and the remaining 18% visited more than ten properties before making their purchase.

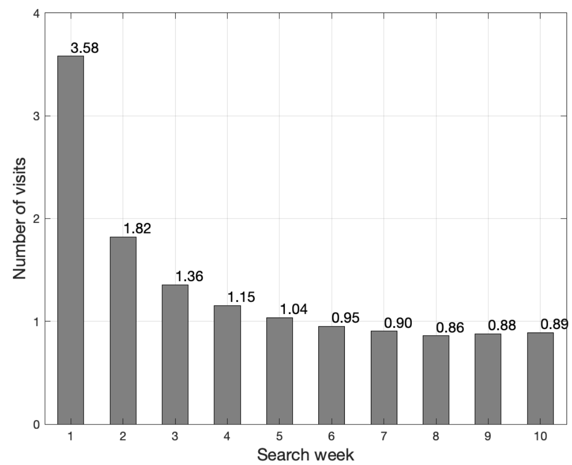
Figure 2: Search Duration and Total Visits

(a) Search Duration (b) Total Visits



An important feature of our data is that we observe not only the total number of visits, but also the dynamic characteristics of consumers' searches (i.e., search duration and search intensity in each week). Figure 3 shows the average number of visits by search week. "Search week" refers to the week since a consumer arrived on the platform. For each search week t, we compute the average number of visits a consumer makes in that search week, where the average is taken over consumers who search for t or more weeks. Figure 3 shows that search intensity decreases over search weeks. The average search intensity starts with 3.58 visits in the first search week, quickly drops to 1.82 visits in the second search week, and continues to drop to less than 1 visit after five weeks of searching.

Figure 3: Number of Visits by Search Week



While Figures 2 and 3 show the distribution of consumers' key endogenous choices, Table 2 reports summary statistics on these endogenous variables as well as exogenous variables describing a consumer's search market. Consumers search for an average of 3.5 weeks and visit a total of 6.6 properties prior to purchase. 17.6% of consumers purchased a previously searched property. For each consumer, we find the segments in which she searched and define her search market as their union. Table 2 shows that on average a consumer searches in two segments with 69% of searches focused on one segment. On average, a consumer's search market consists of 654 properties, sees 29 new listings per week, and experiences an exit rate of 12%. It also faces a weekly increase in the unit list price of CNY228 per square meter. The gap between the list price and the transaction price, which we call the "price discount", decreases by CNY17 per square meter per week.

Table 2: Summary Statistics of Consumers

	Mean	SD
Search duration (weeks)	3.501	4.012
Total visits	6.624	5.097
Recall share	0.176	0.381
Number of segments searched	1.900	1.029
Visit share of Top 1 segment	0.688	0.331
Number of available properties	654	604
Weekly list price change per m^2 (CNY)	228	139
Weekly number of new listings	28.614	22.332
Weekly exit rate	0.122	0.016
Weekly change in price discount per m^2 (CNY)	-16.960	21.946

3 Dynamic Search Model

In this section, we develop a model to describe consumers' search and purchase decisions. Each consumer i arrives exogenously on the platform. To simplify the notation, we use t (without the subscript i) to denote the week since a consumer's arrival on the platform, which we refer to as the "search week."

In each period t, consumer i first decides on the number of properties to visit in the current period and then, after visiting these properties, decides whether to buy one (thus ending the search) and, if so, which one to buy. In what follows, we first describe the primitives of the model and then explain these decisions backwards.

3.1 Primitives

3.1.1 Utility

The utility that consumer i gets from property j depends on a mean utility δ_{jt} and an idiosyncratic value v_{ij} as follows:

$$u_{ijt} = \delta_{jt} + v_{ij}, \tag{1}$$

where the mean utility δ_{jt} is known to the consumer and the idiosyncratic value v_{ij} is learned by search.

The mean utility δ_{jt} is a linear combination of observable characteristics (denoted by $\boldsymbol{x}_j$), the expected transaction price at the time (p^e_{jt}), and an error term (ξ_j) as follows:

$$\delta_{jt} \quad = \quad \boldsymbol{x}_j \boldsymbol{\beta} + \alpha p^e_{jt} + \xi_j, \tag{2}$$

where the vector $\boldsymbol{x}_j$ includes observable property characteristics such as the property size. The error term ξ_j captures unobservable property characteristics that are known to consumers (but unobservable to econometricians). For example, consumers infer whether a property gets a lot of natural light from photos posted on the platform. However, such information is difficult for researchers to capture.

The expected transaction price (p^e_{jt}) is the list price (denoted by p^l_j) minus the expected gap between the list price and the sales price at that time (denoted by p^d_{jt}).[8] The expected

[8] While we observe transaction prices for purchased properties, we do not observe the price that a consumer would have paid for a property that she ultimately did not purchase. Therefore, we use the expected transaction price as the price. A similar approach is used, for example, in Goldberg (1995).

price gap p^d_{jt} of property j at time t depends on its size (in square meters) and the average unit discount (in CNY per square meter) for properties in its segment at that time. Specifically, let x_{1j} be the size of property j, $m(j)$ be the segment of property j, and d_{mt} be the average gap between the list unit price and the transaction unit price in segment m at time t. Then, $p^e_{jt} = p^l_j - p^d_{jt}$, where $p^d_{jt} = x_{1j} d_{m(j)t}$.

Therefore, we can rewrite the mean utility as

$$\delta_{jt} = \delta_j - \alpha p^d_{jt}, \tag{3}$$

where the time-invariant component

$$\delta_j = \boldsymbol{x}_j \boldsymbol{\beta} + \alpha p^l_j + \xi_j \tag{4}$$

depends on the information about property j that is listed on the platform.

The idiosyncratic match value v_{ij} captures what consumer i learns after visiting property j. It includes, for example, consumer i's tastes for the property's sun exposure or a particular amenity in the neighborhood. We assume that v_{ij} is i.i.d. and follows a normal distribution with mean 0 and variance σ^2_v.

3.1.2 Search Costs and Waiting Costs

We assume that the cost of searching n properties is $C(n) - \vartheta_{itn}$, where the search cost shock ϑ_{itn} is i.i.d. and follows a type-1 extreme value distribution with location parameter 0 and scale parameter κ. The observable component is

$$C(n) = \left(\gamma_0 + (\gamma_1 + \gamma_2 m_{it})n + \gamma_3 n^2\right) \mathbb{1}(n > 0). \tag{5}$$

We allow the marginal search cost to depend on the cumulative number of searches before time t (denoted by m_{it}). We also include the quadratic term n^2 to capture the nonlinearity in search costs.

Consumer i also incurs a waiting cost w_{it} in each period in which she does not make a purchase. This waiting cost includes both monetary and non-monetary costs for a consumer. An example of the non-monetary cost could be the psychological stress caused by not having a home for the newlyweds before they get married. In China, it is a cultural norm that the groom's family is expected to buy a home, or at least find a home and contribute to the down payment for the newlyweds before marriage. We assume that w_{it} is i.i.d. and follows a normal distribution with mean w and variance σ^2_w.

3.1.3 Search Set Conditional on the Number of Visits

In each period, consumer i optimally decides the number of properties to visit (denoted by n_{it}). However, conditional on n_{it}, the exact set of properties she visits in period t is exogenous. This assumption is similar to that in Hortaçsu and Syverson (2004) and implies that a consumer's decision is the number, rather than the set, of properties to visit. This assumption simplifies the model because, given the number of searches n_{it}, there are $C_{A_{it}}^{n_{it}}$ possible sets of properties to search if there are A_{it} properties for consumer i to search. The cardinality $C_{A_{it}}^{n_{it}}$ can be very large. For example, if $A_{it} = 15$ and $n_{it} = 8$, then $C_{A_{it}}^{n_{it}} = 6435$. Moreover, this assumption is somewhat justified because there are indeed exogenous reasons why a consumer may or may not be able to visit a property at a given time.

Specifically, let $\mathcal{A}_{it}$ denote the set of properties in consumer i's search market that she has not visited by time t (here, $\mathcal{A}$ stands for "available for search") and $\mathcal{C}^n(\mathcal{A}_{it})$ be the collection of all n-element subsets of $\mathcal{A}_{it}$. For a given number of visits n, we assume that the probability of a particular set $\mathcal{N} \in \mathcal{C}^n(\mathcal{A}_{it})$ being sampled depends on the mean utilities $\{\delta_j : j \in \mathcal{A}_{it}\}$ as follows:

$$\Pr(\mathcal{N}|\mathcal{A}_{it}, n) = 1 + \sum_{k=1}^{n} \left[(-1)^k \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \frac{\sum_{l \in \mathcal{A}_{it} \setminus \mathcal{N}} exp(\delta_l)}{\sum_{l \in \mathcal{A}_{it} \setminus \mathcal{N}} exp(\delta_l) + \sum_{b \in \mathcal{B}} exp(\delta_b)} \right]. \tag{6}$$

This probability has the following desirable features. First, it is consistent with an extended Logit model. In Supplemental Appendix A, we extend a discrete choice model from a setting where a consumer chooses a single option to a setting where a consumer chooses n options. We show that the probability in (6) is consistent with the extended model where consumer i visits the n properties with the highest values of $\delta_j + LogitError_{ij}$ and $LogitError_{ij}$ is i.i.d. and follows a type-1 extreme value distribution. Though seemingly complicated, this expression in (6) is a direct application of the analytic expression of the choice probability in a standard Logit model and the inclusion-exclusion principle.

Second, as we show in Appendix A, this sampling probability for *a set of properties* implies that the following sampling probability for *a particular property* $j \in \mathcal{A}_{it}$:

$$\Pr(j|\mathcal{A}_{it}, n) = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n = A_{it} \\ \sum_{k=0}^{n-1} \left[(-1)^k \sum_{\mathcal{F} \in \mathcal{C}^{A-n+k}(\mathcal{A} \setminus j)} \left(\frac{C_{A-n+k-1}^{k} exp(\delta_j)}{\sum_{l \in \mathcal{F}} exp(\delta_l) + exp(\delta_j)} \right) \right] & \text{if } 0 < n < A_{it} \end{cases} \tag{7}$$

where $A_{it} = \#\mathcal{A}_{it}$. This probability itself has the following three intuitive features: (i) it is increasing in δ_j and decreasing in $\delta_{j'}$ for $j' \neq j$, i.e., property j is more likely to be sampled when its own mean utility increases or other properties' mean utilities decrease; (ii) it becomes the choice probability in a standard Logit model, i.e., $\frac{exp(\delta_j)}{\sum_{l \in \mathcal{A}_{it}} exp(\delta_l)}$, when $n = 1$; and (iii) the sum of the sampling probabilities over j in $\mathcal{A}_{it}$ is n, i.e., $\sum_{j \in \mathcal{A}_{it}} Pr(j|\mathcal{A}_{it}, n) = n$.

3.1.4 Timing

At the beginning of each period, consumer i observes the mean utility δ_{jt} for all properties in the available-to-search set (denoted by$\mathcal{A}_{it}$, where $\mathcal{A}$ stands for "available-to-search") and both the mean utility δ_{jt} and the match value v_{ij} for all properties in the recall set (denoted by $\mathcal{R}_{it}$, where $\mathcal{R}$ stands for "recall"). Recall set $\mathcal{R}_{it}$ consists of properties that consumer i has visited before period t and that are still available in period t. In each period, a consumer makes two decisions: first a search decision and then a purchase decision. The timing is as follows:

- At the beginning of the period, consumer i observes $(\{\delta_{jt}\}_{j \in \mathcal{A}_{it}}, \{\delta_{jt}, v_{ij}\}_{j \in \mathcal{R}_{it}})$ and the search cost shocks $(\vartheta_{itn}, n = 0, ..., \bar{n})$. She decides how many properties to search in time t. We denote the number of searches by n_{it}, which is an integer between 0 and $\bar{n}$.

- A search set $\mathcal{N}_{it} \in \mathcal{C}^{n_{it}}(\mathcal{A}_{it})$ is sampled according to the probability in (6). After visiting the properties in $\mathcal{N}_{it}$, consumer i observes the match values of the properties in the newly searched set, i.e., $\{v_{ij}\}_{j \in \mathcal{N}_{it}}$. Here, $\mathcal{N}$ stands for "Newly searched".

- Consumer i now observes the initial information set $(\{\delta_{jt}\}_{j \in \mathcal{A}_{it}}, \{\delta_{jt}, v_{ij}\}_{j \in \mathcal{R}_{it}})$, match values of the newly searched properties $\{v_{ij}\}_{j \in \mathcal{N}_{it}}$, and the waiting cost w_{it}. She decides whether to buy a property or to continue searching. If she decides to buy in this period, she also decides which property to buy. We denote this decision as y_{it} and explain it in more detail in Section 3.2.

3.1.5 Transition of the Environment

From time t to time $t + 1$, there are three changes in consumer i's search market. First, some properties may exit the market. We assume that the exit rate in consumer i's search market is χ_i and use $\mathcal{EXIT}_{it}$ to denote the set of exited properties in consumer i's search market at time t.

Second, new properties may enter the market. We assume that the number of newly listed properties in consumer i's search market follows a Poisson distribution with an arrival rate λ_i and use $\mathcal{NEW}_{it+1}$ to denote the set of newly listed properties in consumer i's search market at the beginning of time $t + 1$. We assume that the mean utility of a newly listed property in consumer i's search market follows a normal distribution $N(\bar{\delta}_{it}^{new}, (\sigma_i^{new})^2)$. Furthermore, we assume that a consumer, in forming an expectation about the next period, considers the transition of $\bar{\delta}_{it}^{new}$ to be $\bar{\delta}_{it+1}^{new} = \bar{\delta}_{it}^{new} + \alpha\mu_i$,where μ_i captures the trend in the list price of new properties in consumer i's search market.

Third, for properties in $\mathcal{A}_{it} \cup \mathcal{R}_{it}$ that do not exit, a consumer considers the transition of δ_{jt} to be $\delta_{jt+1} = \delta_{jt} + x_{1j}\rho_{m(j)}$, where $\rho_{m(j)}$ captures the trend in the discounts for properties in the segment of property j. Since discounts are measured in terms of price per square meter, we multiply $\rho_{m(j)}$ by the size of the property (x_{1j}) when forming the transition of the mean utility.

Let $\Omega_{it} = (\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}, \{\delta_{jt}, v_{ij}\}_{j\in\mathcal{R}_{it}}, \bar{\delta}_{it}^{new})$ denote the non-transitory variables belonging to the information set that consumer i has at the beginning of the period. The above three changes determine the transition of the information set. Specifically, Ω_{it+1} =($\{\delta_{jt+1}\}_{j\in\mathcal{A}_{it+1}}$, $\{\delta_{jt+1}, v_{ij}\}_{j\in\mathcal{R}_{it+1}}$,$\{\bar{\delta}_{it+1}^{new}\}$), where $\mathcal{A}_{it+1} = \mathcal{A}_{it}\backslash\mathcal{N}_{it}\backslash\mathcal{EXIT}_{it}\cup\mathcal{NEW}_{it+1}$, $\mathcal{R}_{it+1} = \mathcal{R}_{it}\cup\mathcal{N}_{it}\backslash\mathcal{EXIT}_{it}$, $\delta_{jt+1} = \delta_{jt} + x_{1j}\rho_{m(j)}$ for $j \in \mathcal{A}_{it}\backslash\mathcal{N}_{it}\backslash\mathcal{EXIT}_{it}$, δ_{jt+1} is drawn from a normal distribution $N(\bar{\delta}_{it+1}^{new}, (\sigma_i^{new})^2)$ for $j \in \mathcal{NEW}_{it+1}$, and $\bar{\delta}_{it+1}^{new} = \bar{\delta}_{it}^{new} + \alpha\mu_i$.

3.2 The Purchase Decision

After searching, consumer i can either recall from the previously searched set $\mathcal{R}_{it}$, or buy from the newly searched set $\mathcal{N}_{it}$, or continue searching. In other words, her optimization problem at the purchase-decision stage is as follows:

$$
\begin{aligned}
& \mathit{\Gamma}_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}, w_{it}) && (8)\\
= \quad & \max\{ \underbrace{\max_{j\in\mathcal{R}_{it}} \delta_{jt} + v_{ij}}_{\text{recall}}, \underbrace{\max_{j\in\mathcal{N}_{it}} \delta_{jt} + v_{ij}}_{\text{buy a newly searched house}}, \underbrace{E\left[V_i(\Omega_{it+1})|\Omega_{it},\{v_{ij}\}_{j\in\mathcal{N}_{it}}\right] - w_{it}}_{\text{wait}}\},
\end{aligned}
$$

where $E\left[V_i(\Omega_{it+1})|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}\right]$ denotes the expected value of continuing the search at time $t+1$ given the information set at the beginning of period t, Ω_{it}, and the match values of the newly searched properties, $\{v_{ij}\}_{j\in\mathcal{N}_{it}}$. A consumer's purchase decision (denoted by y_{it}) is the optimizer of the above optimization problem.

3.3 The Search Decision

A consumer's optimal search intensity in each period, i.e., the number of properties to search (n_{it}), depends on the comparison between the benefits and costs of searching. We have specified the search costs in Section 3.1. We now explain the search benefits. The benefit of searching n_{it} properties is that consumer i learns the match values of the n_{it} newly searched properties and expands her choice set at the purchase stage from the recall set ($\mathcal{R}_{it}$) to the union of the recall set and the newly searched set (i.e., $\mathcal{R}_{it} \cup \mathcal{N}_{it}$).

Formally, the expected benefit of searching n properties conditional on the information set Ω_{it} is the expectation of $\varGamma_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}}, w_{it})$ where the expectation is taken over the sampled searched set ($\mathcal{N}$), the match values of the properties in that set ($\{v_{ij}\}_{j\in\mathcal{N}}$), and the waiting cost shock (w_{it}). In other words, the expected benefit is

$$EB_i(n|\Omega_{it}) \equiv \sum_{\mathcal{N}\in\mathcal{C}^n(\mathcal{A}_{it})} E_{\left(\{v_{ij}\}_{j\in\mathcal{N}}, w_{it}\right)} \left[\varGamma_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}}, w_{it})\right] \times \Pr(\mathcal{N}|\Omega_{it}, n), \tag{9}$$

where the first term $E_{\left(\{v_{ij}\}_{j\in\mathcal{N}}, w_{it}\right)} \left[\varGamma_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}}, w_{it})\right]$ is the expected value of searching a sampled set $\mathcal{N}$ and the second term $\Pr(\mathcal{N}|\Omega_{it}, n)$ is the probability that the set $\mathcal{N}$ is sampled conditional on information set Ω_{it} and the number of searches n. Here, with a slight abuse of notation, we rewrite $\Pr(\mathcal{N}|\mathcal{A}_{it}, n)$ in (6) as $\Pr(\mathcal{N}|\Omega_{it}, n)$ to reflect its dependence on $\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}$, which is a subset of the information in Ω_{it}.

Consumer i chooses her search intensity n_{it} to maximize her net gain from searching n_{it} properties. That is, the optimal search intensity n_{it} is the solution to the following optimization problem:

$$\underset{0\le n\le \bar{n}}{max} \{EB_i(n|\Omega_{it}) - C(n) + \vartheta_{itn}\}. \tag{10}$$

3.4 Bellman Equation and State Transition

We complete the description of the model by presenting the Bellman equation. We define the *ex ante* value function as the expectation of the maximum in (10) over the search cost shocks:

$$\begin{aligned} V_i(\Omega_{it}) =& E_{(\vartheta_{itn}, n=0,\ldots,\bar{n})} \{\underset{0\le n\le \bar{n}}{max} [EB_i(n|\Omega_{it}) - C(n) + \vartheta_{itn}]\} \\ =& \kappa ln(\sum_{n=0}^{\bar{n}} exp\{[EB_i(n|\Omega_{it}) - C(n)]/\kappa\}) + \kappa\tau, \end{aligned} \tag{11}$$

where τ is the Euler constant. Therefore, plugging $\varGamma_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}, w_{it})$ in (8) into $EB_i(n|\Omega_{it})$ in (9) and then plugging $EB_i(n|\Omega_{it})$ into (11) gives us the Bellman equation.

4 Estimation

The estimation consists of two steps. In the first step, we estimate the parameters in the utility function ($\alpha, \boldsymbol{\beta}$) by matching the observed share of visits that each property receives.

We call these parameters the static parameters. In this step of the estimation, we use all properties and consumers in our sample.

In the second step, we estimate the remaining parameters, which govern the dynamic search and purchase decisions, using maximum likelihood estimation. These parameters include: search cost parameters $(\boldsymbol{\gamma}, \kappa)$, waiting cost parameters (w, σ_w), and the standard deviation of match values (σ_v). We call these parameters the dynamic parameters. In this estimation step, we use a random sample of 1000 consumers who started and ended their searches during the sample period excluding the eight weeks around Chinese New Year. We only consider consumers who started searching after the start of the sample because we need to know how many properties they visited before each period. We do not have to restrict our sample to those who purchased before the end of the sample. We do so in the baseline estimation because without such a restriction, the sample may include consumers who were not seriously searching in the housing market. Nevertheless, as a robustness check, we repeat the estimation including consumers who did not purchase before the end of the sample period.

4.1 Static parameters $(\alpha, \boldsymbol{\beta})$

4.1.1 Estimation of Static Parameters

We estimate the parameters in the utility function $(\alpha, \boldsymbol{\beta})$ by matching the observed share of visits that each property gets. In our sample, all consumers search within one district. Therefore, we partition the sample into six districts and do such a matching district by district. Specifically, let $\mathcal{J}_d$ and $\mathcal{I}_d$ represent the set of properties in district d and the set of consumers searching in district d, respectively.

For a property j in district d, its share of visits according to our model is

$$\tilde{s}_j(\boldsymbol{\delta}_d) = \frac{\sum_{i \in \mathcal{I}_d} \sum_{t=1}^{T_i} \Pr(j|\mathcal{A}_{it}, n_{it})}{\sum_{i \in \mathcal{I}_d} \sum_{t=1}^{T_i} n_{it}}, \tag{12}$$

where $\boldsymbol{\delta}_d = (\delta_j, j \in \mathcal{J}_d)$. In (12), n_{it} is the observed number of properties that consumer i visits in period t. $\Pr(j|\mathcal{A}_{it}, n_{it})$ is the probability that property j is among the n_{it} properties that consumer i visits in period t. It is 0 for $j \notin \mathcal{A}_{it}$ and is given by (7) for $j \in \mathcal{A}_{it}$. The sum is taken over all consumers searching in district d (indexed by $i \in \mathcal{I}_d$) and all periods during a consumer's search (indexed by $t = 1, ..., T_i$), where T_i is the search duration of consumer i.

The empirical counterpart of this share is

$$s_j = \frac{n_j}{\sum_{j \in \mathcal{J}_d} n_j},$$

where n_j is the number of visits that property j receives in the sample.

In Appendix A, we extend the contraction mapping result in Berry, Levinsohn and Pakes (1995) for a single discrete choice model to our setting where a set of options is sampled and show that the system of equations

$$\tilde{s}_j(\boldsymbol{\delta}_d) = s_j, j \in \mathcal{J}_d$$

has a unique solution. Since $\sum_{j \in \mathcal{J}_d} s_j = 1$, we normalize one dimension of $\boldsymbol{\delta}_d$ in each district d to 0. Therefore, all inverted mean utilities are relative to that of the normalized property.

To estimate the parameters in the utility function $(\alpha, \boldsymbol{\beta})$, we regress the inverted δ_j on the list price and property characteristics according to equation (4).[9] One concern is that the list price of a property is likely to be correlated with unobservable property characteristics. We address this potential endogeneity issue by using an instrumental variable approach. Specifically, our instrumental variable is the average transaction price of properties in the same segment where transactions occurred within the three weeks prior to property j's listing. This instrumental variable is relevant because property owners and their agents are likely to choose list prices based on historical transaction prices in the same area and price range. At the same time, it is reasonable to assume that the transaction prices of properties sold in the last three weeks before a property is listed are uncorrelated with the unobservable characteristics of the property.

4.1.2 Discussion of Two-Step Estimation Procedure

In this two-step estimation procedure, we back out the mean utility and estimate the static parameters in the utility function before estimating the dynamic parameters that govern search intensity (i.e., the number of searches in each period), search duration (i.e., the length of the search), and the purchase decision (i.e., which property to purchase). We carry out this first step by matching the model implications of search shares conditional on the observed search intensity and search duration to the observed search shares.

For this two-step procedure to work, we need to be able to write down the model implication of search shares without solving the dynamic search model. We can do this because while a consumer endogenously decides the number and length of her searches, the set of properties visited is exogenous. A similar exogeneity assumption is also made, for example, in Hortaçsu and Syverson (2004). We extend their specification of the sampling probability for a product to the sampling probability for a set of products as in (6). According to our

[9] Since δ_j is relative to the mean utility of the normalized property, the regressors are $p^l_j - p^l_{j^d_0}$ and $\boldsymbol{x}_j - \boldsymbol{x}_{j^d_0}$ where j^d_0 is the normalized property in district d.

sampling probability specification, a property with a higher mean utility has a higher probability of being sampled. The exogenous assumption only means that a consumer and her agent do not have full control over the set of properties they can visit in a given period, certain random factors also play a role in determining the set of sampled properties, and these random factors are unknown to the consumer and thus exogenous. Although not perfect, we believe that this exogeneity assumption is justified.

The advantage of the two-step estimation procedure is twofold. First, following such a two-step procedure, we can allow for unobservable property heterogeneity (i.e., ξ_j in the utility function) and allow list prices to be endogenous. If we were to estimate the utility parameters along with the dynamic parameters using maximum likelihood estimation, we would either have to assume that there is no unobservable property heterogeneity (so that there is no endogeneity in prices) or model how list prices are correlated with the unobservable property heterogeneity (to control for the endogeneity of prices).[10] Second, we back out the mean utilities of all properties in the sample in the first step, which allows us to estimate the transition of the search environment before estimating the dynamic model. For example, as explained in Section 3.1.5, the mean utility of new listings follows a normal distribution $N(\bar{\delta}_{it}^{new}, (\sigma_i^{new})^2)$. We estimate $\bar{\delta}_{it}^{new}$ and σ_i^{new} based on the inverted δ_j for new listings in the sample. Appendix B explains in detail how we estimate the parameters describing the transition of the search environment.

In the end, after the first estimation step, we obtain the transition of the environment in the dynamic model as well as the observable components of the information set Ω_{it} in the dynamic model (i.e., $\delta_{jt} = \delta_j + \alpha p_{jt}^d$).

4.2 Dynamic Parameters $\theta = (\boldsymbol{\gamma}, \kappa, w, \sigma_w, \sigma_v)$

4.2.1 Estimation of Dynamic Parameters

For each consumer i, we observe her search duration T_i. For each search week $t = 1, ..., T_i$, we observe her search intensity (i.e., the number of properties visited n_{it}) and her purchase decision (i.e., whether $y_{it} = recall$ – to buy a previously visited property, $y_{it} = j \in \mathcal{N}_{it}$ – to buy property j which is newly visited by her in the current period, or $y_{it} = wait$ – to search longer).[11]

[10] In this case, the likelihood function would be an expectation of the probability that the observed search path is taken and the observed purchase decision is made. This probability depends on ξ_j's. The expectation is an integral over the distribution of ξ_j's *conditional on the observed list prices.* Therefore, we need a model to describe such a conditional distribution.

[11] We observe the identity of a consumer's purchased property regardless of whether it belongs to the recall set $\mathcal{R}_{it}$ or the newly searched set $\mathcal{N}_{it}$. However, our likelihood function captures whether a consumer recalls, rather than which property she recalls. This is because the probability of purchasing a particular property

The likelihood of observing the search and purchase path $\{n_{it}, y_{it}\}_{t=1}^{T_i}$ for consumer i is

$$l_i(\theta) \quad = \quad \int \left[\Pi_{t=1}^{T_i} Pr_i(n_{it}|\Omega_{it};\theta) \cdot Pr_i(y_{it}|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}};\theta)\right] dF_v(\{\{v_{ij}\}_{j\in\mathcal{N}_{it}}\}_{t=1}^{T_i}) \qquad (13)$$

where $Pr_{it}(n_{it}|\Omega_{it};\theta)$ is the probability of searching n_{it} properties, $Pr(y_{it}|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{A}_{it}};\theta)$ is the purchase choice probability, and $F_v(\{\{v_{ij}\}_{j\in\mathcal{N}_{it}}\}_{t=1}^{T_i}$ is the distribution of match values of all properties that consumer i has visited before purchase. We now derive these two sets of probabilities.

At the beginning of a period, consumer i observes Ω_{it} and chooses n_{it} to maximize her net gain from searching according to the optimization problem in (10). Given that the search cost shock ϑ_{itn} follows the type-1 extreme value distribution with scale parameter κ, the probability that consumer i searches n_{it} is

$$Pr_i(n_{it}|\Omega_{it};\theta) = \frac{exp\{[EB_i(n_{it}|\Omega_{it};\theta) - C(n_{it};\theta)]/\kappa\}}{\sum_{n=0}^{\bar{n}} exp\{[EB_i(n|\Omega_{it};\theta) - C(n;\theta)]/\kappa\}}, \qquad (14)$$

where we add the parameters θ to the expected search benefit $EB_i(n_{it}|\Omega_{it};\theta)$ and search costs $C(n_{it};\theta)$ to make their dependence on the parameters explicit.

After visiting n_{it} properties, consumer i observes the original information set Ω_{it} and the newly acquired information set $\{v_{ij}\}_{j\in\mathcal{N}_{it}}$. Consumer i now makes a purchase decision according to (8). Specifically, the consumer chooses between buying one of the previously searched property ($y_{it} = recall$), buying a newly searched property ($y_{it} = j \in \mathcal{N}_{it}$), or continuing the search ($y_{it} = wait$). The probabilities of these actions are, respectively,

$$\begin{aligned} & Pr_i(y_{it} = recall|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}};\theta) \qquad (15) \\ =& 1\left\{\max_{j\in\mathcal{R}_{it}} \delta_{jt} + v_{ij} \geq \max_{j\in\mathcal{N}_{it}} \delta_{jt} + v_{ij}\right\} \\ & \times \Phi\left(\left\{\max_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}} \{\delta_{jt} + v_{ij}\} - E\left[V_i(\Omega_{it+1};\theta)|\Omega_{it},\{v_{ij}\}_{j\in\mathcal{N}_{it}}\right] + w\right\}/\sigma_w\right) \end{aligned}$$

$$\begin{aligned} & Pr_i(y_{it} = j \in \mathcal{N}_{it}|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}};\theta) \qquad (16) \\ =& 1\left\{\delta_{jt} + v_{ij} \geq \max_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}} \delta_{jt} + v_{ij}\right\} \\ & \times \Phi\left(\left\{\max_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}} \{\delta_{jt} + v_{ij}\} - E\left[V_i(\Omega_{it+1};\theta)|\Omega_{it},\{v_{ij}\}_{j\in\mathcal{N}_{it}}\right] + w\right\}/\sigma_w\right) \end{aligned}$$

in the recall set depends on $\delta_{jt} + v_{ij}$ for all $j \in\mathcal{R}_{it}$ (as well as $\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}$), resulting in a high-dimensional state variable. However, the probability of recall depends on a summary statistic $\max_{j\in\mathcal{R}_{it}} \delta_{jt} + v_{ij}$ (as well as $\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}$).

$$Pr_i(y_{it} = wait|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}; \theta) \tag{17}$$
$$= \Phi\left(\left\{E\left[V_i(\Omega_{it+1};\theta)|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}\right] - w - \max_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}\{\delta_{jt} + v_{ij}\}\right\}/\sigma_w\right),$$

where $\Phi(\cdot)$ is the distribution function of a standard normal distribution.

We estimate θ using maximum likelihood function estimation, where the log-likelihood function is $L(\theta) = \sum_{i=1}^{I} \ln l_i(\theta)$. To compute the likelihood function, we need to compute the value function $V_i(\Omega_{it}; \theta)$ as a fixed point to the Bellman equation. The state variable is the information set $\Omega_{it} = (\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}, \{\delta_{jt}, v_{ij}\}_{j\in\mathcal{R}_{it}})$, which includes the mean utilities of the properties available to search ($\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}$) as well as both the mean utilities and the match values of the properties available to recall ($\{\delta_{jt}, v_{ij}\}_{j\in\mathcal{R}_{it}}$). Thus, the state space is large. Following much of the literature on dynamic estimation (e.g., Collard-Wexler (2013); Sweeting (2013); Hodgson (2019)), we solve this large state space problem by approximating the high dimensional state variable with lower-dimensional statistics. Furthermore, the state variable includes both observable and unobservable variables. We simulate the unobservable variables in estimation. We provide more details on dynamic estimation in Appendix B.

4.2.2 Identification

The dynamic parameters include three sets of parameters: the standard deviation of the match value (σ_v), the search cost parameters ($\boldsymbol{\gamma}, \kappa$), and the waiting cost parameters (w, σ_w).

The standard deviation of the match value is identified by both a static and a dynamic feature of consumer purchase patterns. Statically, we observe which property each consumer purchased. Comparing the mean utility of a consumer's purchased property to the highest mean utility in the consumer's choice set at the time of purchase informs us about the importance of match values. In particular, a larger standard deviation of match values implies that it is more likely to observe a larger gap between the mean utility of the purchased property and the highest mean utility. Dynamically, we observe when when each consumer's purchased property is visited by the consumer. If a consumer visited her purchased property in an earlier period and purchased it in a later period, such a recalling behavior means that the consumer decided to continue the search in the belief that there is a good chance that she gets a better draw of the match value. Therefore, a larger recall share also indicates a larger standard deviation of match values.

The search cost parameters are identified by the mean and variance of the search intensity. Recall that the mean of the search cost is $C(n) = exp(\gamma_0 + (\gamma_1 + \gamma_2 m_{it})n + \gamma_3 n^2)1(n > 0)$. While the fraction of observations with zero searches identifies γ_0, how the fraction with n searches varies as n increases helps us identify γ_1 and γ_3. Finally, how the search intensity varies with m_{it} identifies γ_2. As for the scale parameter of the search cost shock, it is identified by the observed variance in search intensity.

Similarly, the waiting cost parameters are identified by the observed search duration. In particular, the overall level of search duration identifies the mean parameter w and the variance of search duration across consumers identifies the variance of the waiting cost shock σ_w^2.

5 Estimation Results

In this section, we present our estimation results and evaluate the model fit. We begin with the estimates of the parameters in the mean utility function. We then present the estimates of the dynamic parameters, including the standard deviation of match values, the waiting cost parameters, and the search cost parameters. We conclude this session with a discussion of model fit.

5.1 Estimates of Static Parameters

We obtain the estimates of the parameters of the utility function $(\alpha, \boldsymbol{\beta})$ by regressing the inverted mean utility δ_j on the list price and the property characteristics according to equation (4). In Table 3, we report the OLS results in Column (I) and the IV regression results in Column (II). Since properties that are more attractive to consumers in ways that are unobservable to researchers may have higher list prices, list prices are potentially endogenous, leading to an upward bias for the price coefficient in the OLS regression. To address this potential endogeneity issue, we use the average list price of properties in a segment three weeks prior to a property's listing as an instrument for that property's list price. This instrumental variable is highly correlated with the list price. At the same time, it seems reasonable to assume that this instrumental variable, which depends on the list prices of previously listed properties, is uncorrelated with the unobservable characteristics of a property. Comparing the two columns in the table, we find that the IV regression does indeed yield a smaller price coefficient. In the following, we focus on the IV results.

Overall, the estimation yields intuitive results. Consumers like newer and larger properties with more bedrooms, more living rooms, located on the 10th floor or higher, and close to subway stations. For example, suppose the average bedroom size is 20 square meters. Then, on average, an extra bedroom is valued at 0.307 (=(0.344+0.050×20)/4.384) million CNY. Similarly, an additional living room with an average size of 30 square meters is valued at 0.579 (=(1.040+0.050×30)/4.384) million CNY. Locating on the 10th floor or above is worth

CNY35,400 more than locating below the 10th floor. As a robustness analysis, in Appendix SA, we allow the price coefficient to be district specific and find mild heterogeneity in price sensitivity among consumers searching in different districts.

Table 3: Estimates of Parameters in Mean Utility

	(I) OLS		(II) IV	
	Est	SE	Est	SE
List price (million CNY)	-0.193	(0.006)	-4.384	(0.133)
Property age (year)	-0.031	(0.001)	-0.009	(0.001)
# Bedrooms	0.206	(0.007)	0.344	(0.015)
# Living rooms	0.530	(0.008)	1.040	(0.023)
Property size (m^2)	0.005	(0.000)	0.050	(0.002)
Indicator of above 10th floor	0.089	(0.007)	0.155	(0.016)
Close to subway stations	-0.029	(0.010)	0.199	(0.022)
Neighborhood FE	yes		yes	
R square	0.573		0.529	

5.2 Estimates of Dynamic Parameters

Table 4 reports the estimation results for the dynamic parameters, including the standard deviation of match values (σ_v), the waiting cost parameters (w, σ_w), and the search cost parameters ($\boldsymbol{\gamma}, \kappa$).

The estimated standard deviation of the match value is 0.441. Based on the estimated price coefficient ($\hat{\alpha} = -4.384$), this estimated standard deviation corresponds to a value of CNY100,593 ($= 0.441/4.384 \times 10^6$), about 3% of the average list price and 176% of the per-capita disposal income in Beijing in 2016.[12] Therefore, there is a significant benefit of searching to learn about the match value.

The waiting cost is on average 0.051, which is equivalent to CNY11,633 ($= 0.051/4.384 \times 10^6$) per week. This high waiting cost is consistent with the observation that consumers search for 3.5 weeks on average before purchasing a property. The main motivation for buying a property in China is often marriage. The large waiting cost may therefore be explained by the cultural norm in China that the groom's family is expected to purchase a home for the newlyweds before marriage (Wei and Zhang (2011)). However, in relative terms,

[12]The per-capita disposal income in Beijing is CNY57,275 in 2016 according to the 2017 Beijing Statistical Yearbook, 9-14 Basic Data on Urban Housseholds (https://nj.tjj.beijing.gov.cn/nj/main/2017-tjnj/zk/e/indexeh.htm)

this average waiting cost is only 0.28% of the average list price. The standard deviation of the waiting cost is slightly smaller than the mean.

The estimated search cost parameters indicate that there is a baseline search cost that a consumer incurs as long as she searches in a period and that the search cost increases with the number of visits. Specifically, the baseline search cost ($\hat{\gamma}_0$) is CNY703 ($= 3.080 \times 0.001/4.384 \times 10^6$). As for the marginal search cost, it increases with the cumulative number of searches a consumer has done in the previous weeks ($\hat{\gamma}_2 > 0$). It also increases with the current number of searches ($\hat{\gamma}_3 > 0$), implying that the search cost is convex in the number of visits. On average, consumers pay a total search cost of CNY1,244 (i.e., the average of $\sum_{t=1}^{T_i} C(n_{it})/\alpha$ across consumers). Given that an average consumer visits 6.71 properties in total, this implies an average search cost of CNY185 per visit.[13]

Table 4: Estimates of Dynamic Parameters

		Est.	SE
SD of match values	(σ_v)	0.441	(0.029)
Mean waiting cost	(w)	0.051	(0.016)
SD of waiting cost shock	(σ_w)	0.049	(0.027)
Search cost: (0.001)			
constant	(γ_0)	3.080	(0.713)
n	(γ_1)	1.039	(0.442)
(cumulative visits)$\times n$	(γ_2)	4.231	(0.887)
n^2	(γ_3)	0.202	(0.089)
scale parameter	(κ)	3.258	(0.599)

5.3 Model Fit

To assess how well our estimated model fits the data, we simulate for each consumer in the data her decision "path", which describes how long a consumer searches, the number of visits in each week during her search, and which property she purchases. We simulate 50 such paths for each consumer. Appendix C provides details on the simulation. We report summary statistics about consumers' search and purchase decisions according to the data and according to our simulations in Table 5.

Specifically, Table 5 reports summary statistics on search duration, total visits, recall share (i.e., the share of consumers who purchase a previously searched property), and the utility of the purchased property. The first two columns summarize the observed data, where the summary statistics are taken across consumers. We directly observe the search duration

[13] Alternatively, the average of $\frac{\sum_{t=1}^{T_i} C(n_{it})}{\sum_{t=1}^{T_i} n_{it}}/\alpha$ across consumers is CNY174.

and total visits for each consumer as well as the recall share in the data.[14] For "utility of the purchased property", we report summary statistics of $\delta_{jt}+v_{ij}$, where property j is purchased by consumer i in the data, based on the estimated δ_{jt} and the estimated standard deviation of v_{ij}. The last two columns summarize the simulated results, taking summary statistics across consumers and simulations.

Table 5 shows that the estimated model fits the data fairly well. For example, the simulated average search duration and total visits are 3.841 weeks and 7.223 visits while their observed counterparts are 3.448 weeks and 6.710 visits. Similarly, the simulated recall share is 16.1% while the observed recall share is 15.5%. The mean and standard deviation of the utility of the purchased property are (2.704, 1.921) in the data and (2.591, 2.044) according to the estimated model.

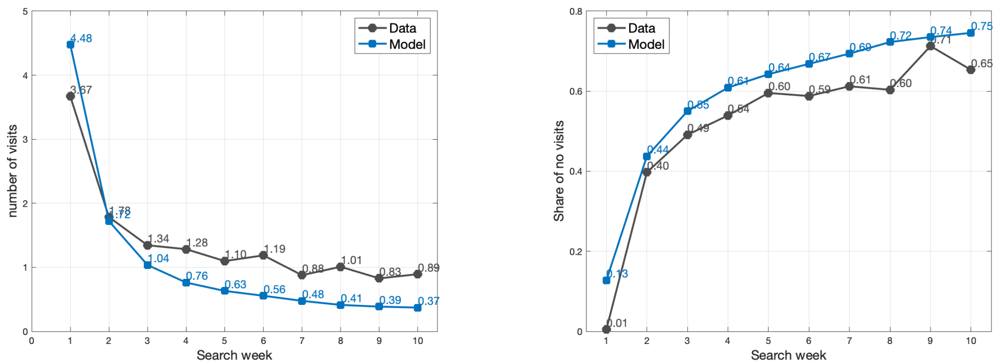
Table 5: Model Fit: Summary Statistics – Data vs. Simulation

	Data		Model Simulation	
	Mean	SD	Mean	SD
Search duration (weeks)	3.448	3.900	3.841	3.143
Total visits	6.710	5.116	7.223	3.574
Recall share	0.155		0.161	
Utility of the purchased property	2.704	1.921	2.730	1.210

Figure 4 shows that our model also successfully captures the dynamic pattern of search intensity. Specifically, Figure 4(a) shows how search intensity changes over time. Take the third search week as an example. This panel shows that the average number of visits in a consumer's third search week is 1.04, where the average is taken over consumers who search for three or more weeks. From this panel, we can see that consumers visit on average about 4 properties in the first week, the number drops sharply in the second week, and it continues to drop as consumers search longer. Our simulation based on the estimated model tracks this dynamic pattern well. Figure 4(b) isolates the extensive margin: how the share of zero visits changes over time. It shows that the longer consumers search, the more likely they are to have zero visits in a week. Again, our simulation follows the data pattern well.

[14] These summary statistics (based on the 1000 randomly sampled consumers) are very close but not identical to those in Table 2 (based on 26,543 consumers in the sample).

Figure 4: Model Fit: Visits by Search Week – Data vs. Simulation



(a) Number of Visits

(b) Share of Zero Visits

Overall, our dynamic search model fits the data well in both the static and the dynamic aspects of the data.

6 Comparison to a Static Search Model

In this section, we estimate a static search model and show that ignoring the dynamics leads to unreasonably large estimates of search costs. A static model endogenizes the total number of searches and the property purchased, but ignores the duration of the search and the number of searches in each period. It also ignores the dynamics in the search environment. There are two types of models with these features in the literature: a simultaneous search model and a sequential search model. In a simultaneous search model, a consumer searches a set of properties at once and then purchases one from the searched set. In a sequential search model, a consumer searches one property at a time and decides whether to continue the search after each visit. As pointed out by Santos, Hortaçsu and Wildenbeest (2012), in a classic sequential search model, a consumer purchases the last property visited and never recalls (unless she visits all properties). Since more than 15% consumers in our data recall, we consider a simultaneous search model in Section 6.1.

6.1 A Static Search Model

The static model is similar to our dynamic search model except for two key differences. First, consumers in the static model choose the total number of visits instead of the number of visits in each period and the duration of the search. Second, the mean search cost in (5) is now $c(n) = \gamma_1 n + \gamma_2 n^2$. We no longer include a constant term $\gamma_0 \mathbb{1}(n > 0)$. This is because all consumers in the sample choose $n > 0$, where n now represents the total number of visits instead of the number of visits in a period. We also do not include the covariate m_{it}, the cumulative number of searches prior to time t, in this static model.

The timing of the static model is the same as in the stage game of the dynamic model. First, consumer i observes the information set Ω_i and the search cost shocks $(\vartheta_{in}, n = 0, ..., \bar{n})$ and decides the number of visits n_i. Recall that in the dynamic model we distinguish between the properties available for search at time t ($\mathcal{A}_{it}$) and those searched before time t ($\mathcal{R}_{it}$). There, the information set Ω_{it} consists of the mean utility δ_{jt} for $j \in \mathcal{A}_{it}$ and both δ_{jt} and the match value v_{ij} for $j \in \mathcal{R}_{it}$. Here, in the static model, the information set Ω_i is simply $\{\delta_j\}_{j \in \mathcal{A}_i}$, i.e., the mean utilities for all properties in her segments in the sample denoted by $\mathcal{A}_i$. Then, a search set $\mathcal{N}_i \in \mathcal{C}^{n_i}(\mathcal{A}_i)$ is sampled according to the probability given in (6) except that the subscript t is dropped for the static model. After visiting the properties in the search set $\mathcal{N}_i$, consumer i observes the match value of the searched set, i.e., $\{v_{ij}\}_{j \in \mathcal{N}_i}$. Finally, given the information $\Omega_i = \{\delta_j\}_{j \in \mathcal{A}_i}$ and $\{v_{ij}\}_{j \in \mathcal{N}_i}$, consumer i decides on which property to buy. This decision is denoted by y_i, and $y_i = j$ means that consumer i chooses property $j \in \mathcal{N}_i$.

We estimate the search cost parameters $(\gamma_1, \gamma_2, \kappa)$ and the standard deviation of match values (σ_v) using MLE. Let θ represent these parameters. The likelihood of observing that consumer i searches n_i properties and buys property j is

$$l_i(\theta) = \Pr(n_i | \{\delta_j\}_{j \in \mathcal{A}_i}; \theta) \cdot \Pr(y_i = j | \{\delta_j\}_{j \in \mathcal{N}_i}; \theta). \tag{18}$$

In (18), the probability that consumer i searches n_i properties is the same as that in (14):

$$\Pr(n_i | \{\delta_j\}_{j \in \mathcal{A}_i}) = \frac{exp\left\{\left[EB(n_i | \{\delta_j\}_{j \in \mathcal{A}_i}) - C(n_i)\right] / \kappa\right\}}{\sum_{n=0}^{\bar{n}} exp\left\{\left[EB(n | \{\delta_j\}_{j \in \mathcal{A}_i}) - C(n)\right] / \kappa\right\}},$$

where the expected benefit of searching n properties is

$$EB(n | \{\delta_j\}_{j \in \mathcal{A}_i}) = \sum_{\mathcal{N} \in C^n(\mathcal{A}_i)} E_{\{v_{ij}\}_{j \in \mathcal{N}}}(\underset{j \in \mathcal{N}}{max}\ \delta_{jt} + v_{ij}) \times \Pr(\mathcal{N} | \mathcal{A}_i, n).$$

Therefore, $\Pr(n_i | \{\delta_j\}_{j \in \mathcal{A}_i}; \theta)$ depends on both the search cost parameters $(\gamma_1, \gamma_2, \kappa)$ and the standard deviation of match values (σ_v).

In (18), the probability that consumer i purchases property j is

$$\Pr(y_i = j|\{\delta_j\}_{j\in\mathcal{N}_i};\theta) = \Pr(\delta_j + v_{ij} \geq \underset{j'\in\mathcal{N}_i}{max}\ \delta_{j'} + v_{ij'}),$$

which only depends on the standard deviation of match values (σ_v).

6.2 Estimation Results Based on the Static Search Model

Table 6 reports the estimation results of the static search model. We have two findings from the comparison of these results with the estimation results of the dynamic model.

First, the static model yields a much smaller estimate of the standard deviation of the match value, implying a much smaller benefit from searching. The estimated standard deviation of match values is 0.235 compared to 0.441 in the dynamic model. These estimates correspond to a value of CNY53,604 and CNY100,593, respectively. Since the variance of the match value determines the benefits of searching (specifically, the larger the variance, the larger the benefits), the benefits of searching according to the estimated static search model are smaller than those according to the estimated dynamic search model.

Table 6: Estimates of the Static Search Model

		Est.	SE
SD of match values	(σ_v)	0.235	(0.014)
Search cost: (0.001)			
n	(γ_1)	301.994	(3.672)
n^2	(γ_2)	-9.489	(0.086)
scale parameter	(κ)	53.192	(0.449)

The static model yields a smaller estimate of σ_v than the dynamic model because the dynamic model exploits more variation in the data to identify σ_v. In the static model, the standard deviation of match values σ_v is identified by comparing the mean utility of the purchased property to the highest mean utility among all searched properties. Specifically, consumer i purchases j if and only if $\delta_j + v_{ij} \geq \underset{j'\in\mathcal{N}_i}{max}\ \delta_{j'} + v_{ij'}$. Therefore, a larger gap between δ_j and $\underset{j'\in\mathcal{N}_i}{max}\ \delta_{j'}$ implies a larger variance for v_{ij}. In a dynamic model, the parameter σ_v is identified not only by such a static comparison (i.e., which property is purchased), but also by a dynamic feature of the data (i.e., when the purchased property is visited). If a consumer purchases a property that was visited in an earlier period, we say that this consumer recalls. As explained in Section 4, a larger recall share implies a larger standard deviation of the match value, because only if there is a large enough chance of getting a large draw of the match value will a consumer decide to continue searching despite having visited a property that turns out to be the best *ex post*. In our estimation sample, the recall share is as high as 15.5%. As a result, our dynamic model yields a larger estimate of the variance of match values.

Second, despite the smaller estimate of search benefits, the static model yields a much larger estimate of search costs, in fact an unreasonably high estimate of search costs. Table 7 reports the summary statistics of the total search cost incurred by a consumer to search the observed number of properties she visited. This total search cost is $c(n_i)/\alpha$ in the static model and $\sum_{t=1}^{T_i} c(n_{it})/\alpha$ in the dynamic model. The average search cost is CNY308,177 in the static model, more than an order of magnitude higher than in the dynamic model (CNY1,244). Given that a consumer in our estimation sample visits an average of 6.71 properties before buying, the average search cost for each property is approximately CNY45,928, or $6,889 per property given the average exchange rate of 0.15 in 2016. We find this estimate unreasonably high.

The static model gives unreasonably high search costs because it does not take into account the dynamics in the search environment. During the sample period, house prices increase rapidly. Therefore, there are two reasons why a consumer stops searching: search costs and increased prices. By ignoring the latter, the static model overestimates search costs.

Table 7: Estimated Search Costs in CNY: Static vs. Dynamic

	Mean	SD
Static search model	308,177	150,543
Dynamic search model	1,244	1,836

7 Impacts of Changing Search Environment and Search Costs

Having established the importance of considering dynamics in our setting, we now quantify how dynamics in the search environment affect consumers' search and purchase decisions. There are three changes to the search environment: price increases, new listings entry, and existing listings exit. We consider the effects of all three changes through counterfactual simulations in Section 7.1. Consumers' search and purchase behavior is also affected by search frictions such as search costs. To quantify the relative importance of the traditional search friction such as search costs versus the search frictions stemming from the dynamics in the search environment, we consider a counterfactual scenario in which consumers' search costs are reduced by half in Section 7.2 and compare the results with those in Section 7.1.

In each counterfactual simulation, we simulate 50 decision paths for each consumer and report the mean and standard deviation of outcome variables that capture consumer search behavior (search duration and total visits), purchase outcome (utility of purchased property), and costs (search cost and waiting cost). Details of the simulation are in Appendix C.

7.1 Effects of Changing Search Environment

In this section, we quantify how the three changes in the search environment (price increases, entry of new listings, and exit of existing listings) affect consumers' search and purchase behavior and consumer welfare.

To examine the effects of price increases, we consider a counterfactual scenario with half the actual price increase rate. Specifically, we reduce both the rate at which the list price of new properties increases over time and the rate at which the sales price increases over time. Table 8 reports summary statistics of the main endogenous outcomes in the counterfactual scenario in Column (II). For comparison, we also include the outcomes under the actual observed price change in Column (I).

From Table 8, we can see that as the price change becomes smaller, consumers search longer, visit more properties before buying, and find properties that generate higher utilities. The average search duration increases from 3.448 with the observed price change to 4.970 with half the actual price change (Row (1)). Similarly, the total number of visits increases from 6.710 to 7.892 (Row (2)). As consumers search longer and more, they end up buying a property that generates a higher utility. Specifically, Row (3) shows that the average utility of the purchased property increases from 2.704 to 2.905, corresponds to an increase in value of CNY45,849 ($= (2.905 - 2.704)/4.384 \times 10^6$). Part of the increase in utility comes mechanically from a lower price. To remove such a mechanical change that contributes to the utility increase, or equivalently, to isolate the effect of longer and more searches on consumer utility, we also report in Row (4) what the utility of the purchased property would be at the observed price. Under the actual price change, Rows (3) and (4) are, of course, identical. When the price change is half, the utility at the observed (higher) price in row (4) is, unsurprisingly, smaller than that in Row (3). Nevertheless, even if we ignore the increase in utility due to lower prices, there is an increase in utility equivalent to CNY35,356 ($= (2.859 - 2.704)/4.384 \times 10^6$) according to Row (4). In other words, searching longer and more contributes to 77% of the increase in utility when the price change is reduced. At the same time, both the average search cost and the average waiting cost increase. In total, they increase by CNY18,660 (Rows (5) and (6)), which is only about half of the increase in utility due to longer and more searches (CNY35,356). Therefore, on balance, consumers are better off with a slower price change.

Table 8: Impact of Price Change

		(I) Actual		(II) Half of Actual	
		Mean	SD	Mean	SD
(1)	Search duration (week)	3.448	3.900	4.970	4.226
(2)	Total visits	6.710	5.116	7.892	3.675
(3)	Utility of purchased property	2.704	1.921	2.905	1.290
(4)	at the observed price	2.704	1.921	2.859	1.286
(5)	Search cost (CNY)	1,244	1,836	1,726	2,498
(6)	Waiting cost (CNY)	29,724	13,992	47,902	19,662

In addition to price changes, property availability can also change over time. Therefore, two other changes in the search environment are the entry of new listings and the exit of existing listings. To quantify how they affect consumers' search and purchase decisions, we consider a counterfactual scenario in which the entry rate (i.e., the arrival rate of new listings) is doubled in Column (II) of Table 9 and a scenario in which the exit rate (i.e., the probability that an existing listing will exit) is halved in Column (III) of Table 9. For comparison, we again include the results with the actual entry and exit rates in Column (I). In both counterfactual scenarios, consumers have more incentives to search longer compared to the actual search environment because consumers either expect more new listings in the future or face a lower probability that her previously visited properties will exit the market.

Comparing Column (II) to Column (I) of Table 9, we can see that, when consumers anticipate more new listings in the future, they indeed search longer, visit a greater number of properties before making a purchase, and ultimately purchase properties that generate higher utilities. Specifically, doubling the entry rate increases consumers' average search duration by 1.347 weeks and their average number of visits by 1.083. As a result, their utility from the purchased property increases by 0.271, which is equivalent to an increase in value of CNY61,816. Meanwhile, visiting more properties and searching for a longer time increases the search cost by CNY319 and the waiting cost by CNY14,228. Overall, consumers are better off.

The comparison of Column (III) and Column (II) shows similar patterns: in anticipation of a slower exit of existing listings, consumers search for more weeks, visit more properties in total, and ultimately purchase a more desirable property. Overall, the average increase in utility is equivalent to CNY26,232, which more than offsets the average increase in search costs of CNY159 and waiting costs of CNY7,826.

Table 9: Effects of Entry and Exit

	(I) Actual		**(II) Double Actual Entry Rate**		**(III) Half Actual Exit Rate**	
	Mean	SD	Mean	SD	Mean	SD
Search duration (week)	3.448	3.900	4.795	4.188	4.339	3.626
Total visits	6.710	5.116	7.793	3.760	7.310	3.527
Utility of purchased property	2.704	1.921	2.975	2.221	2.819	1.232
Search cost (CNY)	1,244	1,836	1,563	2,252	1,403	2,043
Waiting cost (CNY)	29,724	13,992	43,952	7,832	37,550	16,443

In the three counterfactual simulations in this section, we vary each of the three changes in the search environment (price changes, entry of new listings, exit of existing listings) while holding the other two changes fixed. We do this to quantify how each aspect of search environment dynamics affects consumer decisions and outcomes. In other words, this is a quantification exercise, not a policy analysis. However, because the price change, the entry of new listings, and the exit of existing listings are unlikely to move independently, we plan to add a counterfactual simulation in which we vary all three simultaneously as a robustness analysis.

7.2 Effects of Search Costs

While Section 7.1 quantifies the effects of changes in the search environment on consumers' search and purchase behavior, we now quantify the effects of the traditional search friction, i.e., search costs. To this end, we simulate the results in a counterfactual scenario where we reduce the search cost function $c(n)$ by half.

Table 10 shows that when consumers face a lower search cost per visit, they unsurprisingly extend their search duration by 0.609 weeks and visit 0.873 more properties on average. In comparison, these changes are 1.522 weeks and 1.182 visits when the price change is halved, 1.347 weeks and 1.083 visits when the entry rate is doubled, and 0.891 weeks and 0.6 visits when the exit rate is halved.

With more searches, consumers find properties that generate higher utilities for them. The average increased utility is equivalent to CNY19,617. Interestingly, despite the reduced search cost per visit, consumers end up incurring higher search costs due to more searches.

However, the increased search costs (CNY85) and waiting costs (CNY3,429) are dominated by the increased utility. The net benefit is CNY16,103. This compares to a net gain of CNY16,696 when consumers face a slower price increase (where utility is calculated based on actual prices), CNY47,269 when the entry rate is doubled, and CNY18,247 when the exit rate is halved.

Overall, these results show that while traditional search frictions such as search costs affect consumers' search and purchase decisions, search frictions arising from the dynamics of the search environment also have significant effects on consumers' search and purchase behavior and consumer welfare. At least for the same percentage change (halving or doubling), varying the dynamics of the search environment has a greater impact than varying search costs.

Table 10: Impacts of Search Cost

	(I) Actual		**(II) Half Estimated Search Cost**	
	Mean	SD	Mean	SD
Search duration (week)	3.448	3.900	4.057	3.374
Total visits	6.710	5.116	7.583	4.036
Utility of the purchased property	2.704	1.921	2.790	1.245
Search cost (CNY)	1,244	1,836	1,325	1,996
Waiting cost (CNY)	29,724	13,992	33,153	15,143

8 Conclusion

In this paper, we study how dynamics in the search environment affect consumers' search and purchase decisions and consumer welfare. We present a dynamic model in which consumers make search and purchase decisions knowing that both product prices and availability change over time. In this model, both the search decision and the purchase decision are dynamic. The search decision is dynamic because the number of searches in this period affects both the value of buying now (the set of choices increases with the number of searches) and the value of waiting (the set of unsearched items decreases with the number of searches). The purchase decision is dynamic because the decision is not only which product to buy, but also when to buy.

We develop a feasible estimation routine to estimate our dynamic search model. Since the choice set for purchase is endogenously determined by the search decision, the standard approach to estimating a dynamic demand model (where the choice set evolves exogenously) does not apply here. Instead, we develop a two-step estimation procedure to address the challenge of a large state space.

We apply our model and estimation approach to the Beijing housing market between 2015 and 2016, and quantify the relative importance of search environment dynamics and traditional search frictions such as search costs. We find that, for some metrics, changes in product prices and availability have a larger impact on consumers' search and purchase decisions and consumer welfare than the traditional search friction due to search costs. We also find that a static search model would yield an unreasonably high estimate of search costs.

While our model is developed for the context of the housing market and the estimation approach is used to study the Beijing housing market, the model and the estimation approach are potentially applicable to any setting where consumers make purchase decisions after searching and the search environment changes over time. Moreover, the estimation procedure involves an extension of a discrete choice model from a single-option choice to a set-of-options choice, which is also generalizable to other settings where consumers choose a set of products instead of a single product.

References

Aguirregabiria, Victor (2023) "Dynamic demand for differentiated products with fixed-effects unobserved heterogeneity," *The Econometrics Journal*, 26 (1), C1–C25.

Allen, Jason, Robert Clark, and Jean-François Houde (2019) "Search Frictions and Market Power in Negotiated-Price Markets," *Journal of Political Economy*, 127 (4), 1550–1598.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995) "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4), 841–890.

Bodéré, Pierre (2023) "Dynamic spatial competition in early education: An equilibrium analysis of the preschool market in Pennsylvania," working paper.

Brown, Zach Y and Jihye Jeon (2020) "Endogenous information and simplifying insurance choice,"Technical report, Mimeo, University of Michigan.

Chatterjee, Chirantan, Ying Fan, and Debi Prasad Mohapatra (2022) "Spillover Effects in Complementary Markets: A Study of the Indian Cellphone and Wireless Service Markets,"Technical report, Working Paper.

Collard-Wexler, Allan (2013) "Demand fluctuations in the ready-mix concrete industry," *Econometrica*, 81 (3), 1003–1037.

Eizenberg, Alon (2014) "Upstream innovation and product variety in the U.S. home PC market," *Review of Economic Studies*, 81 (3), 1003–1045.

Elliott, Jonathan T (2022) "Investment, emissions, and reliability in electricity markets,"Technical report, Working Paper.

Fan, Ying and Chenyu Yang (2020) "Competition, product proliferation, and welfare: A study of the US smartphone market," *American Economic Journal: Microeconomics*, 12 (2), 99–134.

——— (2022) "Estimating discrete games with many firms and many decisions: An application to merger and product variety,"Technical report, National Bureau of Economic Research.

Goldberg, Pinelopi Koujianou (1995) "Product differentiation and oligopoly in international markets: The case of the US automobile industry," *Econometrica: Journal of the Econometric Society*, 891–951.

Gowrisankaran, Gautam and Marc Rysman (2012) "Dynamics of Consumer Demand for New Dynamic of Consumer Demand for New Durable Goods," *Journal of Political Economy*, 120 (6), 1173–1217.

Hendel, Igal and Aviv Nevo (2006) "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 74 (6), 1637–1673.

Hodgson, Charles (2019) *Information externalities, free riding, and optimal exploration in the uk oil industry*: SIEPR, Stanford Institute for Economic Policy Research.

Hodgson, Charles and Gregory Lewis (2023) "You can Lead a Horse to Water: Spatial Learning and Path Dependence in Consumer Search," *Econometrica*, Forthcoming.

Honka, Elisabeth (2014) "Quantifying search and switching costs in the Quantifying Search and Switching Costs in the US Auto Insurance Industry," *RAND Journal of Economics*, 45 (4), 847–884.

Hortaçsu, Ali and Chad Syverson (2004) "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*, 119 (2), 403–456.

Kim, Jun B, Paulo Albuquerque, and Bart J Bronnenberg (2010) "Online demand under limited consumer search," *Marketing science*, 29 (6), 1001–1023.

Lee, Robin S (2013) "Vertical integration and exclusivity in platform and two-sided markets," *American Economic Review*, 103 (7), 2960–3000.

Moraga-Gonzalez, Jose Luis, Zsolt Sandor, and Matthijs R. Wildenbeest (2023) "Consumer Search and Prices in the Automobile Market," *The Review of Economic Studies*, 90 (3), 1394–1440.

Murry, Charles and Yiyi Zhou (2020) "Consumer search and automobile dealer colocation," *Management Science*, 66 (5), 1909–1934.

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

Shcherbakov, Oleksandr (2016) "Measuring consumer switching costs in the television industry," *The RAND Journal of Economics*, 47 (2), 366–393.

Sweeting, Andrew (2013) "Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81 (5), 1763–1803.

Wei, Shang-Jin and Xiaobo Zhang (2011) "The competitive saving motive: Evidence from rising sex ratios and savings rates in China," *Journal of political Economy*, 119 (3), 511–564.

Zhou, Nan (2016) "National Passenger Traffic Exceeds 2.91 Billion During the 2016 Spring Festival," `https://www.gov.cn/xinwen/2016-03/03/content_5048696.htm`.

A Micro Foundation for the Sampling Probabilities

We have three goals in this section. First, we show that the sampling probability $\Pr(\mathcal{N}|\mathcal{A}, n)$ in (6) can be derived from an extension of a discrete choice model from a single-option choice to a set-of-options choice. Second, we show that this sample probability implies a probability for each option j, i.e., $\Pr(j|\mathcal{A}, n)$ in (7). Third, we show that the solution of the search share equations (12) is unique. In fact, we show that the mapping used to find the solution in Berry, Levinsohn and Pakes (1995) is again a contraction mapping even in our extended model.

A.1 Micro Foundation for the Sampling Probability $\Pr(\mathcal{N}|\mathcal{A}, n)$

In this section, we show that the sampling probability $\Pr(\mathcal{N}|\mathcal{A}, n)$ is consistent with an extension of a Logit model. In this probability, $\mathcal{A}$ represents the set of all options, n is the number of options to sample, and $\mathcal{N} \in \mathcal{C}^n(\mathcal{A})$ represents a particular n-element subset of $\mathcal{A}$, where $\mathcal{C}^n(\mathcal{A})$ is the collection of n-element subsets of $\mathcal{A}$.

We assume that the value associated with an option j in $\mathcal{A}$ is $\delta_j + \epsilon_j$, where ϵ_j is i.i.d. and follows a type-1 extreme value distribution. We further assume that the n highest valued options are sampled. In other words, $\mathcal{N} \in \mathcal{C}^n(\mathcal{A})$ is sampled if and only if $\underset{j \in \mathcal{N}}{min}\ \delta_j + \epsilon_j \geq \underset{l \in \mathcal{A} \setminus \mathcal{N}}{max}\ \delta_l + \epsilon_l$. Therefore,

$$\Pr(\mathcal{N}|\mathcal{A}, n) = \Pr(\delta_j + \epsilon_j \geq \max_{l \in \mathcal{A} \setminus \mathcal{N}} (\delta_l + \epsilon_l),\ \forall j \in \mathcal{N}).$$

This model is an extension of a standard Logit model from restricting $n = 1$ to any $n \geq 1$. When $n = 1$, it becomes a standard Logit model. We now derive the analytic expression for $\Pr(\mathcal{N}|\mathcal{A}, n)$ based on the the inclusion-exclusion principle and the analytic expression for the choice probability in a standard Logit model.

We first apply the inclusion-exclusion principle. The probability that $\mathcal{N}$ is not chosen, i.e., $1 - \Pr(\mathcal{N}|\mathcal{A}, n)$, is the probability that there is at least one $j \in \mathcal{N}$ such that $\max_{l \in \mathcal{A} \setminus \mathcal{N}} (\delta_l + \epsilon_l) > \delta_j + \epsilon_j$. This probability equals to the inclusion and exclusion of the following probabilities:

- include the probability that this inequality holds for one option in $\mathcal{N}$
- exclude the probabilities that this inequality holds for two options in $\mathcal{N}$
- include the probability that inequality holds for three options in $\mathcal{N}$
- so on and so forth

In other words,

$$1 - \Pr(\mathcal{N}|\mathcal{A}, n) = \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \Pr(\max_{l \in \mathcal{A} \backslash \mathcal{N}} (\delta_l + \epsilon_l) > \delta_b + \epsilon_b, b \in \mathcal{B}) \right].$$

We then derive the analytic expression for $\Pr(\max_{l \in \mathcal{A} \backslash \mathcal{N}} (\delta_l + \epsilon_l) > \delta_b + \epsilon_b, b \in \mathcal{B})$ using the choice probability in a standard Logit model.

$$\begin{aligned} \Pr(\max_{l \in \mathcal{A} \backslash \mathcal{N}} (\delta_l + \epsilon_l) > \delta_b + \epsilon_b, b \in \mathcal{B}) &= \sum_{l \in \mathcal{A} \backslash \mathcal{N}} \Pr(\delta_l + \epsilon_l > \delta_b + \epsilon_b, b \in \mathcal{B} \cup (\mathcal{A} \backslash \mathcal{N})) \\ &= \sum_{l \in \mathcal{A} \backslash \mathcal{N}} \frac{exp(\delta_l)}{\sum_{l \in \mathcal{A} \backslash \mathcal{N}} exp(\delta_l) + \sum_{b \in \mathcal{B}} exp(\delta_b)}. \end{aligned}$$

In this equation, the analytic expression for $\Pr(\delta_l + \epsilon_l > \delta_b + \epsilon_b, b \in \mathcal{B} \cup (\mathcal{A} \backslash \mathcal{N}))$ is the choice probability in a standard Logit model.

Combining the above two steps yields

$$\begin{aligned} \Pr(\mathcal{N}|\mathcal{A}, n) &= 1 - \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \Pr(\mathop{max}_{l \in \mathcal{A} \backslash \mathcal{N}} (\delta_l + \epsilon_l) > \delta_b + \epsilon_b, b \in \mathcal{B}) \right] \\ &= 1 - \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \frac{\sum_{l \in \mathcal{A} \backslash \mathcal{N}} exp(\delta_l)}{\sum_{l \in \mathcal{A} \backslash \mathcal{N}} exp(\delta_l) + \sum_{b \in \mathcal{B}} exp(\delta_b)} \right], \end{aligned}$$

which can be further simplified as

$$\begin{aligned} \Pr(\mathcal{N}|\mathcal{A}, n) &= 1 - \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \left(1 - \frac{\sum_{b \in \mathcal{B}} exp(\delta_b)}{\sum_{l \in \mathcal{A} \backslash \mathcal{N}} exp(\delta_l) + \sum_{b \in \mathcal{B}} exp(\delta_b)} \right) \right] \\ &= 1 - \sum_{k=1}^{n} (-1)^{k-1} C_n^k + \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \frac{\sum_{b \in \mathcal{B}} exp(\delta_b)}{\sum_{l \in \mathcal{A} \backslash \mathcal{N}} exp(\delta_l) + \sum_{b \in \mathcal{B}} exp(\delta_b)} \right] \\ &= \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \frac{\sum_{b \in \mathcal{B}} exp(\delta_b)}{\sum_{l \in \mathcal{A} \backslash \mathcal{N}} exp(\delta_l) + \sum_{b \in \mathcal{B}} exp(\delta_b)} \right], \qquad \text{(A.1)} \end{aligned}$$

where the last line holds because setting $x = -1$ in the binomial theorem $(1+x)^n = \sum_{k=0}^{n} C_n^k x^k$ yields $0 = 1 - \sum_{k=1}^{n} C_n^k (-1)^{k-1}$.

A.2 Implied $\Pr(j|\mathcal{A},n)$ and Its Properties

A.2.1 The Analytic Expression for $\Pr(j|\mathcal{A},n)$

The probability that option j is sampled conditional on sampling n options is $\Pr(j|\mathcal{A},n) = \sum_{\mathcal{N}\in\mathcal{C}^n(\mathcal{A}):j\in\mathcal{N}} \Pr(\mathcal{N}|\mathcal{A},n)$. When $n=0$, $\Pr(j|\mathcal{A},n)=0$. When $n = A = \#\mathcal{A}$, $\Pr(j|\mathcal{A},n) = 1$. In Supplemental Appendix SB, we show that for $0 < n < A$,

$$\Pr(j|\mathcal{A},n) = \sum_{k=0}^{n-1}\left[(-1)^k \sum_{\mathcal{F}\in\mathcal{C}^{A-n+k}(\mathcal{A}\backslash j)} \left(\frac{C^k_{A-n+k-1} exp(\delta_j)}{\sum_{l\in\mathcal{F}} exp(\delta_l) + exp(\delta_j)}\right)\right]. \tag{A.2}$$

In this section, we provide intuitions for this analytic expression.

Option j is sampled if it's rank, according to value $\delta_j + \epsilon_j$, is no more than n among all options in $\mathcal{A}$. In other words, $\Pr(j|\mathcal{A},n)$ is the sum of $\Pr(j$'s rank is $n-k)$ for $k = 0,...,n-1$. The probability that j has a certain rank is closely related to the probability that j is the best in a subset of options. For instance, j has a rank of 1, then j is the best among all options in $\mathcal{A}$. Similarly, the rank of j is 2 means that j is the best among $A-1$ options.

Specifically, the connection between the probability that j is the best in a subset of options and the probability of j has a certain rank is as follows: $\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k}(\mathcal{A}\backslash j)}\Pr(j$ is the best in $\mathcal{F}\cup j)$ is the sum of $\Pr(j$'s rank is $n-k)$, $\Pr(j$'s rank is $n-k-1)$, ..., and $\Pr(j$'s rank is 1) where each probability is properly weighted. Consider a simple example where $\mathcal{A} = \{1,2,3,4\}$ and $n = 3$. Let $j = 1$. When $k = 1$, $\mathcal{F} \in \mathcal{C}^{A-n+k}(\mathcal{A}\backslash j)$ can be $\{2,3\}$,$\{2,4\}$,$\{3,4\}$. The event "1 is the best among $\{1,2,3\}$" includes the event "1's rank is 1" and partially overlaps with the event "1's rank is 2." The union of the three events "1 is the best among $\{1,2,3\}$", "1 is the best among $\{1,2,4\}$", and "1 is the best among $\{1,3,4\}$" includes the event "1's rank is 2" exactly once but counts the event "1's rank is 1" three times. Therefore, $\sum_{\mathcal{F}=\{2,3\},\{2,4\},\{3,4\}}\Pr(j$ is the best in $\mathcal{F}\cup j) =\Pr(j$'s rank is 2$) + 3\Pr(j$'s rank is 1).

Such a relationship between the probability that j is the best among a subset and the probability that j has a particular rank explains the terms $\frac{exp(\delta_j)}{\sum_{l\in\mathcal{F}} exp(\delta_l)+exp(\delta_j)}$ and $C^k_{A-n+k-1}$ in (A.2), which correspond to probability that option j is the best in $\mathcal{F}\cup j$ and the weight used to adjust double counting.

A.2.2 Three Properties of $\Pr(j|\mathcal{A},n)$

Property 1. $\Pr(j|\mathcal{A},n)$ **is increasing in** δ_j **and decreasing in** δ_k **for** $k \neq j$**.** Option j is sampled if and only if $\delta_j + \epsilon_j$ is among the top n highest values in $\{\delta_{j'} + \epsilon_{j'}\}_{j'\in\mathcal{A}}$. Since an increase in δ_j leads to an increase in the probability that option j is among the top-n options, $\Pr(j|\mathcal{A},n)$ is increasing in δ_j. Similarly, since an increase in δ_k for $k \neq j$ leads to a decrease in the probability that option j is among the top-n options, $\Pr(j|\mathcal{A}, n)$ is increasing in δ_k for $k \neq j$.

Property 2. $\Pr(j|\mathcal{A}, n)$ **becomes the choice probability in a standard Logit model when** $n = 1$**.** When $n = 1$, $\Pr(j|\mathcal{A}, n)$ in (A.2) becomes

$$\Pr(j|\mathcal{A}, 1) = \sum_{\mathcal{F} \in \mathcal{C}^{A-1}(\mathcal{A} \backslash j)} \left(\frac{exp(\delta_j)}{\sum_{l \in \mathcal{F}} exp(\delta_l) + exp(\delta_j)} \right) = \frac{exp(\delta_j)}{\sum_{l \in \mathcal{A}} exp(\delta_l)},$$

because $\mathcal{C}^{A-1}(\mathcal{A} \backslash j)$ has a singleton element, i.e., $\mathcal{A} \setminus j$. In other words, $\Pr(j|\mathcal{A}, 1)$ is indeed the choice probability in a standard Logit model.

Property 3. The sum of $\Pr(j|\mathcal{A}, n)$ **across** j **in** $\mathcal{A}$ **is** n**, i.e.,** $\sum_{j \in \mathcal{A}} Pr(j|\mathcal{A}, n) = n$**.** When $n = A$, $\Pr(j|\mathcal{A}, A) = 1$ and, therefore, $\sum_{j \in \mathcal{A}} \Pr(j|\mathcal{A}, A) = A$. We now consider the case where $n < A$.

$$\begin{aligned} \sum_{j \in \mathcal{A}} \Pr(j|\mathcal{A}, n) &= \sum_{j \in \mathcal{A}} \left\{ \sum_{k=0}^{n-1} \left[(-1)^k \sum_{\mathcal{F} \in \mathcal{C}^{A-n+k}(\mathcal{A} \backslash j)} \left(\frac{C^k_{A-n+k-1} exp(\delta_j)}{\sum_{l \in \mathcal{F}} exp(\delta_l) + exp(\delta_j)} \right) \right] \right\} \\ &= \sum_{k=0}^{n-1} (-1)^k \left[\sum_{j \in \mathcal{A}} \left\{ \sum_{\mathcal{F} \in \mathcal{C}^{A-n+k}(\mathcal{A} \backslash j)} \left(\frac{C^k_{A-n+k-1} exp(\delta_j)}{\sum_{l \in \mathcal{F}} exp(\delta_l) + exp(\delta_j)} \right) \right\} \right]. \end{aligned}$$

Let $\mathcal{D}$ represent the union of $\mathcal{F}$ and $\{j\}$. Since $\mathcal{F} \in \mathcal{C}^{A-n+k}(\mathcal{A} \backslash j)$ and $\{j\}$ have no intersect, we can rewrite the above equation as

$$\sum_{j \in \mathcal{A}} \Pr(j|\mathcal{A}, n) = \sum_{k=0}^{n-1} (-1)^k \left[\sum_{\mathcal{D} \in \mathcal{C}^{A-n+k+1}(\mathcal{A})} \left\{ \sum_{j \in \mathcal{D}} \left(\frac{C^k_{A-n+k-1} exp(\delta_j)}{\sum_{l \in \mathcal{D}} exp(\delta_l)} \right) \right\} \right].$$

Because in the summation over $\mathcal{F} \in \mathcal{C}^{A-n+k}(\mathcal{D})$, the denominators are the same and each term $exp(\delta_l)$ for $l \in \mathcal{D}$ is repeated $C^{A-n+k-1}_{A-n+k}$ times, the above line can be simplified to

$$\begin{aligned} \sum_{j \in \mathcal{A}} \Pr(j|\mathcal{A}, n) &= \sum_{k=0}^{n-1} (-1)^k \left[\sum_{\mathcal{D} \in \mathcal{C}^{A-n+k+1}(\mathcal{A})} C^k_{A-n+k-1} \right] \\ &= \sum_{k=0}^{n-1} (-1)^k C^{A-n+k+1}_{A} C^k_{A-n+k-1} \\ &= n. \end{aligned} \tag{A.3}$$

We provide a proof for the last equality, i.e., $\sum_{k=0}^{n-1} (-1)^k C^{A-n+k+1}_{A} C^{A-n-1}_{A-n+k-1} = n$ in Supplemental Appendix SB.3.

A.3 $\tilde{s}_j(\delta) = s_j, j \in \mathcal{J}_d$ Has a Unique Solution

In this section, we show that, under certain regularity conditions, the system of equations $(\tilde{s}_j(\boldsymbol{\delta}_d) = s_j, j \in \mathcal{J}_d \backslash \{1\})$ has a unique solution $\boldsymbol{\delta}_d = (\delta_j, j \in \mathcal{J}_d)$ such that $\delta_1 = 0$. Recall that we normalize the mean utility of the first property to 0 and that the share function is

$$\tilde{s}_j(\boldsymbol{\delta}_d) = \frac{\sum_{i \in \mathcal{I}_d} \sum_{t=1}^{T_i} \Pr(j|\mathcal{A}_{it}, n_{it})}{\sum_{i \in \mathcal{I}_d} \sum_{t=1}^{T_i} n_{it}}, \tag{A.4}$$

In what follows, we suppress the subscript d for simplicity and use J to denote the cardinality of $\mathcal{J}$.

The regularity conditions are:

1. $0 < s_j < \frac{\sum_{i \in \mathcal{I}} \sum_{t=1}^{T_i} \mathbf{1}(j \in \mathcal{A}_{it})}{\sum_{i \in \mathcal{I}} \sum_{t=1}^{T_i} n_{it}}$.
2. For any it, $\#\{\delta_j : j \in \mathcal{A}_{it}, \delta_j = \infty\} \leq n_{it}$.
3. For any it, $\#\{\delta_j : j \in \mathcal{A}_{it}, \delta_j = -\infty\} \leq A_{it} - n_{it}$.

Condition 1 imposes a constraint on the observed search share s_j. It means that each property is visited at least once ($s_j > 0$) but not to the extent that it is visited whenever it is in a consumer's available-to-search set ($s_j < \frac{\sum_{i \in \mathcal{I}} \sum_{t=1}^{T_i} \mathbf{1}(j \in \mathcal{A}_{it})}{\sum_{i \in \mathcal{I}} \sum_{t=1}^{T_i} n_{it}}$). This condition is an extension to the condition $0 < s_j < 1$ in a single discrete choice model.

Conditions 2 and 3 impose constraints on n_{it} in the data. Condition 2 implies that if a consumer's available-to-search set includes properties with a mean utility of ∞, the consumer visits these properties. Condition 3 implies that if a consumer's available-to-search set includes properties with a mean utility of ∞, the consumer does not visit these properties.

Following Berry, Levinsohn and Pakes (1995), we define a mapping $f : \{0\} \times R^{J-1} \to R^{J-1}$ as

$$f_j(\delta) = \delta_j + \ln s_j - \ln \tilde{s}_j(\delta) \tag{A.5}$$

for $j \in \mathcal{J} \backslash \{1\}$. Recall that we normalize δ_1 to 0. In Supplemental Appendix SB, we show that this mapping is a contraction mapping. Therefore, the fixed point of it is the unique solution to the system of equation $\tilde{s}_j(\delta) = s_j, j \in \mathcal{J} \backslash \{1\}$.

B Estimation Details

B.1 Estimation of the Environment Transition

As explained in Section 3, the transition of the environment in consumer i's search market is described by the new listing arrival rate λ_i, listing exit rate χ_i, the mean and standard deviation of a new listing's mean utility $(\bar{\delta}_{it}^{new}, \sigma_i^{new})$, the trend in new listing's mean utility μ_i, and the trend in discount ρ_m for each segment in consumer i' search market. To estimate these parameters, we start with estimating segment-level parameters.

We estimate the segment-level arrival rate λ_m by first finding the number of new listings in segment m in each week in the sample and then taking an average across weeks in the sample. Similarly, we estimate the segment-level exit rate χ_m by first computing the exit rate in segment m in each week (as the ratio of the number of exits to the number of listings) and then taking an average across weeks.[15] We estimate the segment-level mean $\bar{\delta}_{mw}^{new}$ for new listing's mean utilities as the average of δ_j across new listings in segment m and calendar week w. The estimated standard deviation σ_m^{new} is the standard deviation of all new listings in segment m. We estimate the trend in new listing's mean utility μ_m by first computing $\bar{\delta}_{mw+1}^{new} - \bar{\delta}_{mw}^{new}$ and then taking the average across weeks in the sample. Similarly, we estimate the trend in the discount ρ_m by first computing $d_{mw+1} - d_{mw}$ and then taking the average across weeks in the sample, where d_{mw} is the average gap between the transaction price and the list price for transacted properties in segment m in calendar week w.

To estimate $(\lambda_i, \chi_i, \sigma_i^{new}, \mu_i)$ in consumer i's search market, we take the weighted average of the corresponding segment-level values where the weight is the share share of consumer i's searches in a segment. To estimate $\bar{\delta}_{it}^{new}$, we take the same weighted average across $\bar{\delta}_{mw}^{new}$ to obtain $\bar{\delta}_{iw}^{new}$. We then find the corresponding calendar week w and assign the corresponding $\bar{\delta}_{iw}^{new}$ to $\bar{\delta}_{it}^{new}$.

B.2 Value Function Approximation

B.2.1 Reduced state space

The state variable $\Omega_{it} = (\{\delta_{jt}\}_{j \in \mathcal{A}_{it}}, \{\delta_{jt}, v_{ij}\}_{j \in \mathcal{R}_{it}}, \bar{\delta}_{it}^{new})$ is high dimensional. We approximate the value function $V_i(\Omega_{it}; \theta)$ by a function of a lower dimensional state variable as follows.

First, we define $u_{it}^* = \max_{j \in \mathcal{R}_{it}} \delta_{jt} + v_{ij}$ to be the maximum utility among properties in the recall set and replace $\{\delta_{jt}, v_{ij}\}_{j \in \mathcal{R}_{it}}$ in the state variable by u_{it}^*.

[15] A property is considered to have exited in a particular week if it is either sold in that week or has not been visited ever since two weeks prior to that week.

Second, we assume that instead of tracking δ_j for all properties in $\mathcal{A}_{it}$, a consumer tracks the mean utilities of the top K properties and the average mean utility of the remaining properties. In other words, we replace $\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}$ by $\{\delta_{1t},...,\delta_{Kt},\bar{\delta}_{K+1t},A_{it}\}$, where, with a slight abuse of notations, $\delta_{1t},...,\delta_{Kt}$ represent the highest K mean utilities among $\{\delta_{jt}\}_{j\in\mathcal{A}_{it}}$; $\bar{\delta}_{K+1t}$ denotes the average of δ_{jt} for properties outside the top K in the set $\mathcal{A}_{it}$, and $A_{it} = \#\mathcal{A}_{it}$ is the number of properties available for search. In practice, we set $K = 5$.

Let $\tilde{\Omega}_{it} = \{\delta_{1t},...,\delta_{Kt},\bar{\delta}_{K+1t},A_{it},u^*_{it}\}$ denote the reduced state variable. We approximate the value function with a polynomial of the reduced state variable.

B.2.2 Simulation of $\tilde{\Omega}_{it}$ and Its Transition

The state variable $\tilde{\Omega}_{it}$ includes observable variables $(\delta_{1t},...,\delta_{Kt},\bar{\delta}_{K+1t},A_{it})$ as well as the unobservable variable u^*_{it}. In estimation, we simulate the unobservable u^*_{it} by drawing v_{ij} for $j\in\mathcal{N}_{it}, t = 1,...,T_i$.

To simulate the expected value of continuing searching conditional on $\tilde{\Omega}_{it}$ and match values of a newly search set of properties $\mathcal{N}$, i.e., $E\left[V_i(\tilde{\Omega}_{it+1})|\tilde{\Omega}_{it},\{v_{ij}\}_{j\in\mathcal{N}}\right]$, we simulate the transition of the state variable $\tilde{\Omega}_{it}$ according to the transition of the environment. In what follows, we explain how to obtain a simulation draw of $\tilde{\Omega}^r_{it+1}$ based on $\tilde{\Omega}_{it}$ and $\mathcal{N}$, where the superscript r represents a simulation draw.[16]

1. Simulate the available-for-search set in the next period
   (a) We draw a number new^r_{it} from a Poisson distribution with arrival rate λ_i. It represents the number of new listings. We denote the set of new listings by $\mathcal{NEW}^r_{it}$.
   (b) For each house $j\in\mathcal{A}_{it}$, we draw $exit^r_{jt} = 1, 0$ from a Bernoulli distribution with exit rate χ_i and denote the set of exited properties (i.e., j such that $exit^r_j = 1$) by $\mathcal{EXIT}^r_{it}$.
   (c) The available-for-search set in the next period is, therefore, $\mathcal{A}^r_{it+1} = \mathcal{NEW}^r_{it} \cup (\mathcal{A}_{it}\setminus\mathcal{N}\setminus\mathcal{EXIT}^r_{it})$. Correspondingly, $A^r_{it+1} = \#\mathcal{A}^r_{it+1}$.
2. Simulate $\{\delta_{1t+1},...,\delta_{Kt+1},\bar{\delta}_{K+1t+1}\}$
   (a) For $j\in\mathcal{NEW}^r_{it}$, we draw δ^r_{jt+1} from $N(\bar{\delta}^{new}_{it+1},(\sigma^{new}_i)^2)$, where $\bar{\delta}^{new}_{it+1} = \bar{\delta}^{new}_{it} - \alpha\mu_i$.
   (b) For $j\in\mathcal{A}_{it}\setminus\mathcal{N}\setminus\mathcal{EXIT}^r_{it}$, its mean utility transitions according to $\delta_{jt+1} = \delta_{jt} - \alpha x_{1j}\rho_{m(j)}$.

[16] As explained, $\tilde{\Omega}_{it}$ itself is simulated. Therefore, we obtain a set of simulation draws of $\tilde{\Omega}_{it+1}$ for each simulated $\tilde{\Omega}_{it}$. For notation simplicity, we use $\tilde{\Omega}_{it}$ to denote a generic simulated state variable and $\tilde{\Omega}^r_{it+1}$ to denote a specific simulated state variable for next period.

(c) We find the highest K mean utilities for properties in $\mathcal{A}^r_{it+1}$ and compute the average the other $(A^r_{it+1} - K)$ mean utilities.

3. Simulate u^*_{it+1}

   (a) For each $j \in \mathcal{N} \backslash \mathcal{EXIT}^r_{it}$, we simulate the match value v^r_{ij}

   (b) Define $u^{*r}_{it+1} = \max\left\{\{\delta_{jt} + v_{ij}\}_{j\in\mathcal{R}_{it}\backslash\mathcal{EXIT}^r_{it}}, \{\delta_{jt} + v^r_{ij}\}_{j\in\mathcal{N}\backslash\mathcal{EXIT}^r_{it}}\right\}$

C Simulation Details

C.1 Model Fit Simulation

To simulate a path (indexed by r) for each consumer, we first draw match values $v^{(r)}_{ij}$ for all properties in consumer i's search market. We then conduct the forward simulation as follows. A notation with a superscript (r) indicates a simulated outcome and a notation without the superscript represents the observed outcome in the data.

At $t = 1$, the recall set is $\mathcal{R}_{i1} = \emptyset$ and the information set is $\Omega_{i1} = (\{\delta_{j1}\}_{j\in\mathcal{A}_{i1}})$, where $\mathcal{A}_{i1}$ is the observed available-for-search set. We simulate the search intensity $n^{(r)}_{i1}$ and the purchase decision $y^{(r)}_{i1}$ in $t = 1$ in three steps: (1) We compute $Pr_i(n|\Omega_{i1})$ according to (14) for $n = 0, ..., \bar{n}$ and simulate $n^{(r)}_{i1}$ based on these probabilities; (2) we draw $\mathcal{N}^{(r)}_{i1}$ based on $\Pr(\mathcal{N}|\mathcal{A}_{i1}, n^{(r)}_{i1})$ in (6); (3) we compute $Pr_i(y_{i1}|\Omega_{i1}, \{v^{(r)}_{ij}\}_{j\in\mathcal{N}^{(r)}_{i1}})$ according (16) and (17) and simulate $y^{(r)}_{i1}$ based on these probabilities. If $y^{(r)}_{i1} \neq wait$, the path ends. Otherwise, we continue to $t = 2$.

At $t = 2$, the recall set is updated to $\mathcal{R}^{(r)}_{i2} = \mathcal{R}_{i1} \cup \mathcal{N}^{(r)}_{i1}\backslash\mathcal{EXIT}_{i1}$ and the available-for-search set is updated to $\mathcal{A}^{(r)}_{i2} = \mathcal{A}_{i1}\backslash\mathcal{N}^{(r)}_{i1}\backslash\mathcal{EXIT}_{i1} \cup \mathcal{NEW}_{i2}$. The information set is $\Omega^{(r)}_{i2} = (\{\delta_{j2}\}_{j\in\mathcal{A}^{(r)}_{i2}}, \{\delta_{j2}, v^{(r)}_{ij}\}_{j\in\mathcal{R}^{(r)}_{i2}})$. We simulate $(n^{(r)}_{i2}, y^{(r)}_{i2})$ following the same steps except that y_{it} can be "recall" for $t > 1$ and the probability is given by (15). If $y^{(r)}_{i2} \neq wait$, the path ends. Otherwise, we continue to $t = 3$ and repeat the process until $y^{(r)}_{it} \neq wait$.

C.2 Counterfactual Simulations

The simulation procedure for the counterfactual simulations is similar to that for the model-fit simulation, with two differences. First, instead of using the estimated value function, we use backward induction to solve the value function in a counterfactual scenario and compute $Pr_i(n|\Omega^{(r)}_{it})$ and $Pr_i(y_{it}|\Omega_{it}, \{v^{(r)}_{ij}\}_{j\in\mathcal{N}^{(r)}_{it}})$ based on the computed value function. Second, we replace the observed δ_{jt} with the simulated $\delta^{(r)}_{jt}$ in the counterfactual scenario with different price dynamics and we replace the observed $\mathcal{NEW}_{it}$ with the simulated $\mathcal{NEW}_{it}^{(r)}$ in the counterfactual scenario with a different entry rate.

We simulate $\delta_{jt}^{(r)}$ as follows. Let μ_i^{CF} denote the price change of newly listed properties in consumer i's search area and ρ_m^{CF} denote the average discounts of properties in segment m in a counterfactual scenario. For example, both μ_i^{CF} and ρ_m^{CF} are half of their counterparts in the data. For newly listed properties, we draw $\delta_{jt}^{(r)}$ from $N(\bar{\delta}_{it+1}^{new}, (\sigma_i^{new})^2)$, where $\bar{\delta}_{it+1}^{new} = \bar{\delta}_{it}^{new} - \alpha\mu_i^{CF}$. For properties listed in the past that are still on the market for sale, we let their mean utility transit according to $\delta_{jt+1} = \delta_{jt} - \alpha x_{1j}\rho_{m(j)}^{CF}$.

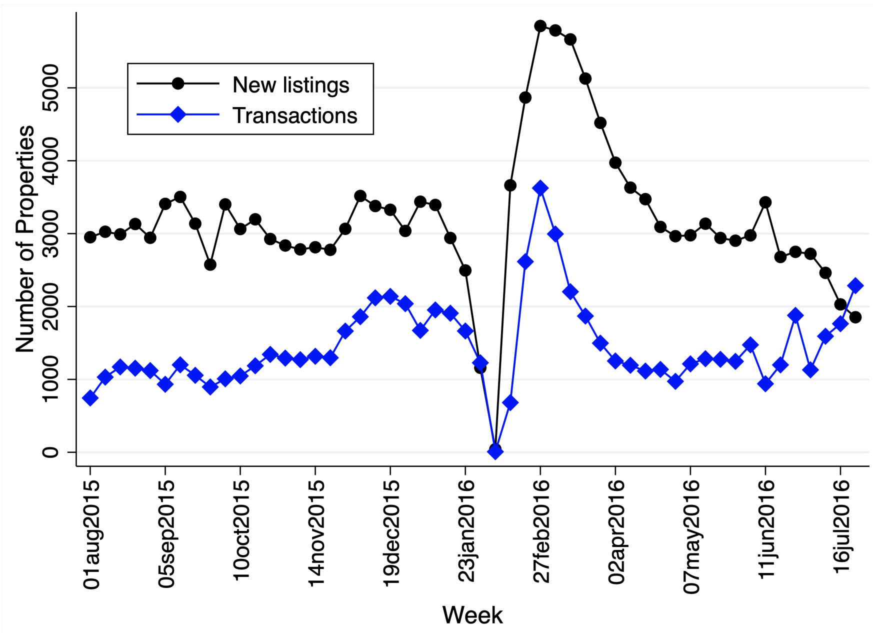
Similarly, we simulate $\mathcal{NEW}_{it}^{(r)}$ as follows. Let λ_i^{CF} denote the arrival rate of new listings in consumer i's search area in a counterfactual scenario. We draw a number new_{it}^r from a Poisson distribution with arrival rate λ_i^{CF} and let new_{it}^r be the number of new listings of consumer i's search area at time t in the counterfactual scenario. For each of the new_{it}^r new listings, we draw δ_{jt+1}^r from $N(\bar{\delta}_{it+1}^{new}, (\sigma_i^{new})^2)$.

SA Additional Tables and Figures

Figure SA.1 shows the numbers of new listings and transactions, and the average list and transaction prices by week, including the eight weeks during the Chinese New Year.

Figure SA.1: New Listings, Transactions, and Prices by Week

(a) New Listings and Transactions



(b) List Prices and Transaction Prices

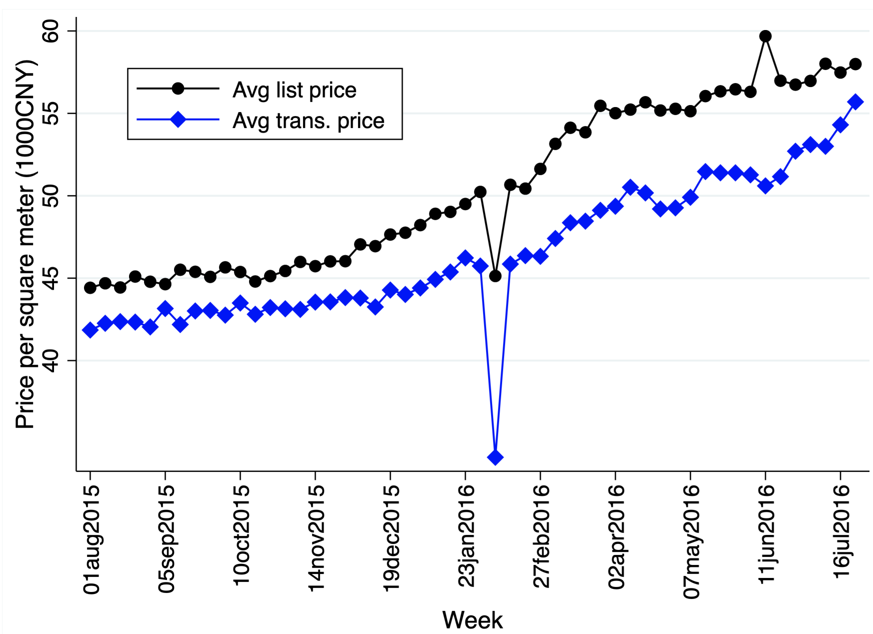


Table SA.1 reports the regression results from regressing the mean utility on property characteristics where we allow the price coefficient to be different for consumers who search in different districts. We find mild heterogeneity in the price coefficients. Parameters common to this table and Table 6 (with homogenous price coefficient) are robust.

Table SA.1: Estimates of Parameters in Mean Utility: District-Specific Price Coefficients

	(I) OLS		(II) IV	
	Est	SE	Est	SE
List price (million CNY)				
Dongcheng	-0.261	(0.021)	-5.497	(0.204)
Xicheng	-0.084	(0.016)	-4.827	(0.164)
Chaoyang	-0.245	(0.009)	-5.301	(0.164)
Haidian	-0.001	(0.012)	-3.661	(0.143)
Fengtai	-0.377	(0.015)	-4.770	(0.155)
Shijingshan	-0.400	(0.027)	-4.066	(0.248)
Property age (year)	-0.004	(0.001)	-0.009	(0.001)
# Bedrooms	0.217	(0.007)	0.367	(0.017)
# Living rooms	0.527	(0.008)	1.087	(0.025)
Property size (sq. meter)	0.005	(0.000)	0.054	(0.002)
Indicator of above 10th floor	0.082	(0.007)	0.171	(0.018)
Close to subway stations	-0.020	(0.010)	0.227	(0.025)
Neighborhood FE	yes		yes	
R square	0.575		0.531	

SB Proofs

SB.1 Proof for $\Pr(j|\mathcal{A}, n)$ in (A.2)

In this section, we show how to derive the analytic expression for $\Pr(j|\mathcal{A}, n)$ in (7), which is also in (A.2).

Plugging (A.1) into $\Pr(j|\mathcal{A}, n) = \sum_{\mathcal{N} \in \mathcal{C}^n(\mathcal{A}): j \in \mathcal{N}} \Pr(\mathcal{N}|\mathcal{A}, n)$ yields

$$
\begin{aligned}
\Pr(j|\mathcal{A},n) &= \sum_{\mathcal{N}\in\mathcal{C}^n(\mathcal{A}):j\in\mathcal{N}} \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B}\in\mathcal{C}^k(\mathcal{N})} \frac{\sum_{b\in\mathcal{B}} exp(\delta_b)}{\sum_{l\in\mathcal{A}\setminus\mathcal{N}} exp(\delta_l) + \sum_{b\in\mathcal{B}} exp(\delta_b)} \right] \\
&= \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{N}\in\mathcal{C}^n(\mathcal{A}):j\in\mathcal{N}} \sum_{\mathcal{B}\in\mathcal{C}^k(\mathcal{N})} \frac{\sum_{b\in\mathcal{B}} exp(\delta_b)}{\sum_{l\in\mathcal{A}\setminus\mathcal{N}} exp(\delta_l) + \sum_{b\in\mathcal{B}} exp(\delta_b)} \right] \qquad \text{(SB.1)} \\
&= \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{A}\setminus j)} \sum_{\mathcal{B}'\in\mathcal{C}^{k-1}(\mathcal{A}\setminus\mathcal{D})} \frac{\sum_{b\in\mathcal{B}'} exp(\delta_b) + exp(\delta_j)}{\sum_{l\in\mathcal{D}} exp(\delta_l) + \sum_{b\in\mathcal{B}'} exp(\delta_b) + exp(\delta_j)} \right] \qquad \text{(SB.2)} \\
&+ \sum_{k=1}^{n-1} \left[(-1)^{k-1} \sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{A}\setminus j)} \sum_{\mathcal{B}\in\mathcal{C}^{k}(\mathcal{A}\setminus\mathcal{D}\setminus j)} \frac{\sum_{b\in\mathcal{B}} exp(\delta_b)}{\sum_{l\in\mathcal{D}} exp(\delta_l) + \sum_{b\in\mathcal{B}} exp(\delta_b)} \right] \qquad \text{(SB.3)} \\
&= \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\setminus j)} \sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{F})} \frac{\sum_{l\in\mathcal{F}\setminus\mathcal{D}} exp(\delta_l) + exp(\delta_j)}{\sum_{l\in\mathcal{F}} exp(\delta_l) + exp(\delta_j)} \right] \qquad \text{(SB.4)} \\
&+ \sum_{k=1}^{n-1} \left[(-1)^{k-1} \sum_{\mathcal{F}\in\mathcal{C}^{A-n+k}(\mathcal{A}\setminus j)} \sum_{\mathcal{B}\in\mathcal{C}^{k}(\mathcal{F})} \frac{\sum_{b\in\mathcal{B}} exp(\delta_b)}{\sum_{l\in\mathcal{F}} exp(\delta_l)} \right] \qquad \text{(SB.5)}
\end{aligned}
$$

From (SB.1) to the sum of (SB.2) and (SB.3), we replace $\mathcal{A}\setminus\mathcal{N}$ by $\mathcal{D}$ and separately consider $\mathcal{B} \in \mathcal{C}^k(\mathcal{N})$ including j in (SB.2) and those excluding j in (SB.3). For $\mathcal{B} \in \mathcal{C}^k(\mathcal{N})$ including j in (SB.2), we use $\mathcal{B}'$ to represent $\mathcal{B}\setminus j$. For $\mathcal{B} \in \mathcal{C}^k(\mathcal{N})$ excluding j in (SB.3), the summation goes to $n-1$ instead of n because when $k = n$, such $\mathcal{B}$ excluding j does not exist. In (SB.4), we use $\mathcal{F}$ to represent the union of $\mathcal{D}$ and $\mathcal{B}'$. Because $\mathcal{D}$ and $\mathcal{B}'$ have no intersection, the double summation over $\mathcal{D}$ and $\mathcal{B}'$ in (SB.2) is equivalent to a double summation over $\mathcal{F}$ and subsets of $\mathcal{F}$ (i.e., $\mathcal{D} \in \mathcal{C}^{A-n}(\mathcal{F})$). We do the same to line (SB.3) to obtain line (SB.5).

Line (SB.4) can be further simplified as

$$
\begin{aligned}
&\sum_{k=1}^{n}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\backslash j)}\sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{F})}\left(1-\frac{\sum_{l\in\mathcal{D}} exp(\delta_l)}{\sum_{l\in\mathcal{F}} exp(\delta_l)+exp(\delta_j)}\right)\right]\\
&=\sum_{k=1}^{n}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\backslash j)}\left(C_{A-n+k-1}^{A-n}-\sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{F})}\frac{\sum_{l\in\mathcal{D}} exp(\delta_l)}{\sum_{l\in\mathcal{F}} exp(\delta_l)+exp(\delta_j)}\right)\right]\\
&=\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n}\\
&-\sum_{k=1}^{n}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\backslash j)}\frac{C_{A-n+k-2}^{A-n-1}\sum_{l\in\mathcal{F}} exp(\delta_l)}{\sum_{l\in\mathcal{F}} exp(\delta_l)+exp(\delta_j)}\right] &\text{(SB.6)}\\
&=\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n}\\
&-\sum_{k=1}^{n}\left[(-1)^{k-1}C_{A-n+k-2}^{A-n-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\backslash j)}\left(1-\frac{exp(\delta_j)}{\sum_{l\in\mathcal{F}} exp(\delta_l)+exp(\delta_j)}\right)\right]\\
&=\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n} &\text{(SB.7)}\\
&-\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-2}^{A-n-1} &\text{(SB.8)}\\
&+\sum_{k=1}^{n}\left[(-1)^{k-1}C_{A-n+k-2}^{A-n-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\backslash j)}\left(\frac{exp(\delta_j)}{\sum_{l\in\mathcal{F}} exp(\delta_l)+exp(\delta_j)}\right)\right]
\end{aligned}
$$

Line (SB.6) holds because in the summation over $\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{F})$, each term $exp(\delta_l)$ for $l\in\mathcal{D}$ is repeated $C_{A-n+k-2}^{A-n-1}$ times.

Similarly, line (SB.5) can be further simplified as

$$
\begin{aligned}
&\sum_{k=1}^{n-1}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k}(\mathcal{A}\backslash j)}C_{A-n+k-1}^{k-1}\right]\\
=&\sum_{k=1}^{n-1}\left[(-1)^{k-1}C_{A-1}^{A-n+k}C_{A-n+k-1}^{k-1}\right]. &\text{(SB.9)}
\end{aligned}
$$

Combining the simplified expressions for (SB.4) and (SB.5) yields

$$
\Pr(j|\mathcal{A},n)=\sum_{k=0}^{n-1}\left[(-1)^{k}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k}(\mathcal{A}\backslash j)}\left(\frac{C_{A-n+k-1}^{k}exp(\delta_j)}{\sum_{l\in\mathcal{F}} exp(\delta_l)+exp(\delta_j)}\right)\right] \tag{SB.10}
$$

This is because line (SB.7) $=0$, line (SB.8) $=-1$, and line (SB.9) $=1$. We provide a proof for line (SB.7) $=0$ and line (SB.8) $=-1$ in Supplemental Appendix SB.3. Line (SB.9) can be rewritten as $\sum_{k=1}^{n-1}\left[(-1)^{k-1}C_{A-1}^{A-n+k}C_{A-n+k-1}^{A-n}\right]$ and, therefore, can be proved by replacing n in line (SB.8) by $n-1$.

SB.2 Proof for the Contract Mapping in (A.5)

In this section, we first show that the mapping in (A.5) has the following four features: (1) $\frac{\partial f_j}{\partial \delta_k} \geq 0$ for any $j, k \neq 1$, (2) $\sum_{k\in\mathcal{J}\backslash\{1\}} \frac{\partial f_j}{\partial \delta_k} < 1$ for any $j \neq 1$, (3) $\min_{j\in\mathcal{J}\backslash\{1\}} \inf_\delta f(\delta) > -\infty$, and (4) there is a value, $\bar{\delta}$, with the property that if for any $j \neq 1$, $\delta_j > \bar{\delta}$, then $f_j(\delta) < \delta_j$. We then show that these features imply that a truncated version of the mapping is a contraction mapping, which establishes the result that the mapping has a unique fixed point and thus the system of equations $\tilde{s}_j(\delta) = s_j,\ j \in \mathcal{J}\backslash\{1\}$ has a unique solution.

(1) $\frac{\partial f_j}{\partial \delta_k} \geq 0$ for any $j, k \neq 1$

We prove this inequality in three steps.

Step 1. We show $\frac{\partial \Pr(j|\mathcal{A},n)}{\partial \delta_j} < \Pr(j|\mathcal{A},n)$ using induction. When $n=1$, $\Pr(j|\mathcal{A},n)$ becomes the choice probability in a standard multinomial Logit model. As a result, $\frac{\partial \Pr(j|\mathcal{A},n)}{\partial \delta_j} < \Pr(j|\mathcal{A},n)$ hold.

To show that $\frac{\partial \Pr(j|\mathcal{A},n-1)}{\partial \delta_j} < \Pr(j|\mathcal{A},n-1)$ implies $\frac{\partial \Pr(j|\mathcal{A},n)}{\partial \delta_j} < \Pr(j|\mathcal{A},n)$ for any $n \geq 2$, we first note that option j is sampled if and only if its rank in terms $\delta_j + \epsilon_j$ is no more than n. Therefore,

$$\Pr(j|\mathcal{A},n) = \Pr(j|\mathcal{A},n-1) + \Pr(j\text{'s rank is } n).$$

In the above equation, $\Pr(j\text{'s rank is } n)$ is the probability that some options $(j_1,\cdots,j_{n-1})$ are the top $n-1$ options while j is the n^{th} best option. In other words,

$$\begin{aligned}\Pr(j|\mathcal{A},n) = {} & \Pr(j|\mathcal{A},n-1)+ \\ & \sum_{\{(j_1,\cdots,j_{n-1}):j_l\neq j\}} \Pr(j_1|\mathcal{A},1)\Pr(j_2|\mathcal{A}\backslash\{j_1\},1)\cdots\Pr(j_{n-1}|\mathcal{A}\backslash\{j_1,\cdots,j_{n-2}\},1)\Pr(j|\mathcal{A}\backslash\{j_1,\cdots,j_{n-1}\},1)\end{aligned} \tag{SB.11}$$

Since the probabilities in the second line are choice probabilities in a standard multinomial Logit model, $\frac{\partial \Pr(j_l|\mathcal{A}\backslash\{j_1,\cdots,j_{l-1}\},1)}{\partial \delta_j} < 0$ and $\frac{\partial \Pr(j|\mathcal{A}\backslash\{j_1,\cdots,j_{n-1}\},1)}{\partial \delta_j} < \Pr(j|\mathcal{A}\backslash\{j_1,\cdots,j_{n-1}\},1)$. Therefore,

$$\begin{aligned}\frac{\partial \Pr(j|\mathcal{A},n)}{\partial \delta_j} < {} & \frac{\partial \Pr(j|\mathcal{A},n-1)}{\partial \delta_j}+ \\ & \sum_{\{(j_1,\cdots,j_{n-1}):j_l\neq j\}} \Pr(j_1|\mathcal{A},1)\Pr(j_2|\mathcal{A}\backslash\{j_1\},1)\cdots\Pr(j_{n-1}|\mathcal{A}\backslash\{j_1,\cdots,j_{n-2}\},1)\Pr(j|\mathcal{A}\backslash\{j_1,\cdots,j_{n-1}\},1)\end{aligned}$$

As a result, $\frac{\partial \Pr(j|\mathcal{A},n-1)}{\partial \delta_j} < \Pr(j|\mathcal{A},n-1)$ implies $\frac{\partial \Pr(j|\mathcal{A},n)}{\partial \delta_j} < \Pr(j|\mathcal{A},n)$.

Step 2. We show $\frac{\partial \Pr(j|\mathcal{A},n)}{\partial \delta_k} < 0$ for any $k \neq j$. Since $\Pr(j|\mathcal{A},n)$ is the probability that $\delta_j + \epsilon_j$ is among the top n highest values in $\{\delta_{j'} + \epsilon_{j'}\}_{j' \in \mathcal{A}}$, for any $k \neq j$, it decreases as δ_k increases.

Step 3. We show $\frac{\partial f_j}{\partial \delta_k} \geq 0$ using the results from steps 1 and 2. For $k = j$, the result in step 1 implies

$$\begin{aligned}
\frac{\partial f_j}{\partial \delta_j} &= 1 - \frac{1}{\tilde{s}_j(\delta)} \frac{\sum_{i \in \mathcal{I}} \sum_{t=1}^{T_i} \frac{\partial \Pr(j|\mathcal{A}_{it}, n_{it})}{\partial \delta_j}}{\sum_{i \in \mathcal{I}} \sum_{t=1}^{T_i} n_{it}} \\
&> 1 - \frac{1}{\tilde{s}_j(\delta)} \frac{\sum_{i \in \mathcal{I}} \sum_{t=1}^{T_i} \Pr(j|\mathcal{A}_{it}, n_{it})}{\sum_{i \in \mathcal{I}_d} \sum_{t=1}^{T_i} n_{it}} \\
&= 1 - \frac{1}{\tilde{s}_j(\delta)} \tilde{s}_j(\delta) = 0.
\end{aligned}$$

For $k \neq j$, the result in step 2 implies

$$\frac{\partial \tilde{s}_j(\delta)}{\partial \delta_k} = \frac{\sum_{i=1}^{I} \sum_{t=1}^{T_i} n_{it} \frac{\partial \Pr(j|\mathcal{A}_{it}, n_{it})}{\partial \delta_k}}{\sum_{i=1}^{I} \sum_{t=1}^{T_i} n_{it}} < 0.$$

Therefore,

$$\frac{\partial f_j}{\partial \delta_k} = -\frac{1}{\tilde{s}_j(\delta)} \frac{\partial \tilde{s}_j(\delta)}{\partial \delta_k} > 0.$$

These three steps complete the proof that $\frac{\partial f_j}{\partial \delta_k} \geq 0$ for any $k \in \mathcal{J}$.

(2) $\sum_{k \in \mathcal{J} \backslash \{1\}} \frac{\partial f_j}{\partial \delta_k} < 1$ **for any** $j \in \mathcal{J} \backslash \{1\}$

Because increasing the mean utility of every option does not change the sampling probabilities, $\tilde{s}_j(\delta) = \tilde{s}_j(\delta + a)$. Total differentiation of this equation with respect to a implies that $\frac{\partial \tilde{s}_j(\delta)}{\partial \delta_1} + \sum_{k \in \mathcal{J} \backslash \{1\}} \frac{\partial \tilde{s}(\delta)}{\partial \delta_k} = 0$. Since $\frac{\partial \tilde{s}_j(\delta)}{\partial \delta_1} < 0$ for $j \neq 1$, we have $\sum_{k \in \mathcal{J} \backslash \{1\}} \frac{\partial \tilde{s}(\delta)}{\partial \delta_k} > 0$. As a result, $\sum_{k \in \mathcal{J} \backslash \{1\}} \frac{\partial f_j}{\partial \delta_k} = 1 - \frac{1}{\tilde{s}_j(\delta)} \sum_{k \in \mathcal{J} \backslash \{1\}} \frac{\partial \tilde{s}_j(\delta)}{\partial \delta_k} < 1$.

(3) There is a value $\underline{\delta}$ such that if $\delta_j < \underline{\delta}$ for any $j \in \mathcal{J} \backslash \{1\}$, then $f_j(\delta) > \delta_j$.

Given Condition 2, $\delta_j = -\infty$ implies $\Pr(j|\mathcal{A}_{it}, n_{it}) = 0$ and thus $\tilde{s}_j(\delta) = 0$. In other words, $\lim_{\delta_j \to -\infty} f_j(\delta) = \infty$. By continuity of $f(\delta)$, there exists $\underline{\delta}_j$ such that $f_j(\delta) > \delta_j$ for any δ where $\delta_j < \underline{\delta}_j$. Let $\underline{\delta} = \min_j \underline{\delta}_j$.

(4) There is a value $\bar{\delta}$ such that if $\delta_j > \bar{\delta}$ for any $j \in \mathcal{J}\backslash\{1\}$, then $f_j(\delta) < \delta_j$.

Given Condition 3, $\delta_j = \infty$ implies $\Pr(j|\mathcal{A}_{it}, n_{it}) = 1$ and thus $\tilde{s}_j(\delta) = \frac{\sum_{i=1}^{I}\sum_{t=1}^{T_i} \mathbf{1}(j \in \mathcal{A}_{it})}{\sum_{i=1}^{I}\sum_{t=1}^{T_i} n_{it}} > s_j$. In other words, $\lim_{\delta_j \to \infty} f_j(\delta) < \delta_j$. By continuity of $f(\delta)$, there exists $\bar{\delta}_j$ such that $f_j(\delta) < \delta_j$ for any δ where $\delta_j > \bar{\delta}_j$. Let $\bar{\delta} = \max_j \bar{\delta}_j$.

Features (1)–(4) imply that $f(\delta)$ is a contraction mapping

First, we show $\underline{\delta} < \bar{\delta}$ by contradiction. Suppose $\underline{\delta} = \bar{\delta}$. Then features (3) and (4) contradict each other. Suppose $\underline{\delta} > \bar{\delta}$. Then, feature (3) implies that $f_j(\bar{\delta}) > \delta_j$, which is contradiction to feature (4).

Then, we define another mapping $\hat{f}(\delta) : 0 \times [\underline{\delta}, \bar{\delta}]^{J-1} \to [\underline{\delta}, \bar{\delta}]^{J-1}$ as

$$\hat{f}_j(\delta) = \max\{\underline{\delta},\ \min\{f_j(\delta), \bar{\delta}\}\}.$$

For any $\delta, \delta' \in 0 \times [\underline{\delta}, \bar{\delta}]^{J-1}$, we define $\vartheta = ||\delta - \delta'||$. We have

$$\hat{f}_j(\delta') - \hat{f}_j(\delta) \leq \hat{f}_j(\delta + \vartheta) - \hat{f}_j(\delta) \leq f_j(\delta + \vartheta) - f_j(\delta) = \int_0^{\vartheta} \sum_{k \in \mathcal{J}\backslash\{1\}} \frac{\partial f_j(\delta + z)}{\partial \delta_k} dz \leq \varsigma\vartheta,$$

where $\varsigma = \max_j \max_{z \in [0, \bar{\delta} - \underline{\delta}]} \sum_{k \in \mathcal{J}\backslash\{1\}} \frac{\partial f_j(\delta + z)}{\partial \delta_k}$. The first inequality holds because of feature (1). The second inequality holds because $\hat{f}$ is a truncated version of f.

By feature (2), $\varsigma < 1$. Therefore, $\hat{f}$ is a contraction mapping and has a unique fixed point. Since features (3) and (4) imply that the fixed point of f is in $\hat{f}$'s domain, f also has a unique fixed point.

SB.3 Proof of the Combinatorial Identities

In this section, we prove that for $A > n > 1$,[17]

$$\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n} = 0 \text{ in line (SB.7)},$$
$$\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-2}^{A-n-1} = 1 \text{ in line (SB.8)},$$
$$\sum_{k=0}^{n-1}(-1)^{k}C_{A}^{A-n+k+1}C_{A-n+k-1}^{A-n-1} = n \text{ in line (A.3)}.$$

First, we note that these three identities can be rewritten as

$$\sum_{k=r}^{m}(-1)^{h-r}C_m^hC_h^r = 0,$$
$$\sum_{k=r+1}^{m}(-1)^{h-r-1}C_m^hC_{h-1}^r = 1.$$
$$\sum_{k=r+2}^{m}(-1)^{h-r}C_m^hC_{h-2}^r = m-r-1$$

where the change of variables $(m = A-1, r = A-n, h = A-n+k-1)$ in the first identity, $(m = A-1, r = A-n-1, h = A-n+k-1)$ in the second identity, and $(m = A, r = A-n-1, h = A-n+k+1)$ in the third identity.

To prove the first identity, we take the r^{th}-order derivative of the equation $(1+x)^m = \sum_{h=0}^{m} C_m^h x^h$. The derivative of the LHS is

$$\frac{d^r}{dx^r}(1+x)^m = C_m^r r!(1+x)^{n-r}.$$

The derivative of the RHS is

$$\frac{d^r}{dx^r}\sum_{h=0}^{m} C_m^h x^h = \sum_{h=r}^{m} C_m^h C_h^r r! x^{h-r}.$$

Therefore, $C_m^r r!(1+x)^{n-r} = \sum_{h=r}^{m} C_m^h C_h^r r! x^{h-r}$. Setting $x = -1$ yields the first identity.

To prove the second identity, we take the r^{th}-order derivative of the equation $\frac{(1+x)^m}{x} = \frac{1}{x} + \sum_{h=1}^{m} C_m^h x^{h-1}$. The derivative of the LHS is the sum of terms $a(1+x)^b x^{-c}$ where $b > 0$.

[17] We thank Pierre-Louis Blayac and Sergey Fomin at the University of Michigan for recommending "Tables of Combinatorial Identities Based on Seven Unpublished Manuscript Notebooks of Henry Gould" edited by Jocelyn Quaintance. The first identity in this section is an application of equation (1.23) in Table II (https://math.wvu.edu/~hgould/Vol.2.PDF). The proof of the other two identities is inspired by this equation.

Therefore, its value at $x = -1$ is 0. The derivative of the RHS is

$$\frac{d^r}{dx^r}\left(\frac{1}{x} + \sum_{h=1}^{m} C_m^h x^{h-1}\right) = r!(-1)^r x^{-1-r} + \sum_{h=r+1}^{m} C_m^h C_{h-1}^r r! x^{h-1-r}.$$

Its value at $x = -1$ is $-r! + \sum_{h=r+1}^{m} C_m^h C_{h-1}^r r!(-1)^{h-1-r}$. Therefore,

$$-r! + \sum_{h=r+1}^{m} C_m^h C_{h-1}^r r!(-1)^{h-1-r} = 0,$$

which implies the second identity.

To prove the second identity, we take the r^{th}-order derivative of the equation $\frac{(1+x)^m}{x^2} = \frac{1}{x^2} + \frac{m}{x} + \sum_{h=2}^{m} C_m^h x^{h-2}$. The derivative of the LHS evaluated at $x = -1$ is again 0. The derivative of the RHS is

$$\begin{aligned}&\frac{d^r}{dx^r}\left(\frac{1}{x^2} + \frac{m}{x} + \sum_{h=2}^{m} C_m^h x^{h-2}\right)\\ =&(r+1)!(-1)^r x^{-2-r} + mr!(-1)^r x^{-1-r} + \sum_{h=r+2}^{m} C_m^h C_{h-2}^r r! x^{h-2-r}\end{aligned}$$

Its value at $x = -1$ is $r![(r+1) - m + \sum_{h=r+2}^{m} C_m^h C_{h-2}^r r!(-1)^{h-r}]$. Therefore,

$$(r+1) - m + \sum_{h=r+2}^{m} C_m^h C_{h-2}^r r!(-1)^{h-r} = 0,$$

which implies the third identity.