NBER WORKING PAPER SERIES

DYNAMIC SEARCH IN A NON-STATIONARY SEARCH ENVIRONMENT: AN APPLICATION TO THE BEIJING HOUSING MARKET

Ying Fan
Ziying Fan
Yiyi Zhou

Working Paper 33528
http://www.nber.org/papers/w33528

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2025

The authors have no relevant or material financial interests that relate to the research described in this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

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

Ying Fan, Ziying Fan, and Yiyi Zhou

NBER Working Paper No. 33528
February 2025
JEL No. D8, L8, R3

ABSTRACT

This paper studies how dynamic changes in the search environment affect consumer search and purchase behavior. We develop a dynamic model that incorporates a non-stationary search environment and propose a feasible estimation procedure to estimate its parameters. We apply our model and estimation procedure to the Beijing housing market, utilizing detailed data on consumers' complete search records. We show that accounting for dynamics is crucial for accurately estimating search costs. Additionally, we find that search environment dynamics have a significant impact on consumer decisions and welfare. Housing supply policies that alter search environment dynamics---by increasing the number of new listings and slowing down price increases---benefit consumers, primarily by incentivizing longer searches, more property visits, and ultimately leading to purchases that yield higher utility.

Ying Fan
Department of Economics
University of Michigan
611 Tappan Street
Ann Arbor, MI 48109
and NBER
yingfan@umich.edu

Yiyi Zhou
Stony Brook University
yiyi.zhou@stonybrook.edu

Ziying Fan
School of Public Economics and Administration
Shanghai University of Finance and Economics
Shanghai 200433, China
fan.ziying@mail.shufe.edu.cn

1 Introduction

In markets ranging from real estate to travel accommodations to automobiles, consumers often conduct a search prior to making a purchase decision. The extent to which a consumer searches thus determines the consumer's purchase choice set. As a result, any search frictions that affect a consumer's search also affect the set of options the consumer has to choose from when making a purchase decision.

Furthermore, in many scenarios, the search environment changes over time as prices fluctuate, 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 shopping during the 2022–2023 U.S. real estate market saw significant increases in mortgage rates. Consumers planning vacations during peak travel times may experience rapid changes in the 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 dynamic changes in the search environment affect consumer search and purchase decisions. Consider a scenario where product availability changes over time due to product entry and exit. While a more extensive search expands a consumer's choice set and may lead her to a more suitable product, a longer search runs the risk that her most preferred option among the previously searched products may no longer be available. Additionally, if the prices of new products increase over time, expanding one's choice set by searching longer to discover new products may offer little benefit, leading to a reduced incentive to search. Overall, dynamic changes in the search environment affect consumer search and purchase behavior.

We develop a dynamic model with a non-stationary search environment. The model has two key features. First, consumers in our model make both search and purchase decisions, and both decisions are dynamic. Specifically, consumers observe a set of product characteristics before searching, search to learn about a consumer-specific match value, and choose a product among the searched products that are available at the time of purchase. The purchase decision is dynamic because consumers can choose to buy immediately or wait. The search decision is also dynamic because more searches in a given time period lead to a larger choice set and a higher value of buying now but a smaller set of unsearched products and a lower value of waiting. Second, the search environment in our model is non-stationary. In particular, existing products may exit the market, new products may enter the market, and the prices of new products entering the market at different times may vary. Our model allows consumers to take into account the changing market environment

in both their search and purchase decisions.

To the best of our knowledge, this paper is the first to incorporate a non-stationary environment into a dynamic search model. Although our model is developed in the context of housing search, it can also be used to study other settings where consumers make dynamic search and purchase decisions in a changing environment. Search models that do not incorporate the dynamics of the search environment can lead to biased estimates of search costs. For example, if both search costs and search environment dynamics contribute to limiting consumer search, ignoring the latter may lead to an overestimate of search costs. Similarly, models that ignore search environment dynamics can also result in incorrect policy evaluations, especially when the policy of interest may affect the dynamics of the environment.

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 purchase decisions for two reasons. First, the Beijing housing market experienced a rapid price increase during the sample period. For example, the average list price of new listings increased by more than 30% during the sample period. Second, the dataset we use for our analysis is novel in that it contains complete and detailed information on consumers' search activities. Specifically, our data come from the largest real estate agency in Beijing. For each consumer in our sample, the dataset provides a complete record of the consumer's search and purchase behavior, including when the consumer starts searching, when she stops searching, how many and which properties she visits in each period, and which property she eventually purchases.

Estimating our dynamic search and purchase model presents two challenges. First, our model allows for product characteristics that are known to economic agents but unobservable to researchers. This model feature is important because it is often difficult for researchers to obtain complete data on all product characteristics observable to consumers and relevant for their decisions (see, for example, Berry, Levinsohn and Pakes (1995) and many subsequent papers on differentiated products). At the same time, this feature introduces an endogeneity problem: the price of a product is likely to be correlated with unobservable product characteristics. Therefore, we need to address this endogeneity issue in estimating our model. Second, the vector of state variables has a very high dimension because it includes the characteristics of both searched and unsearched products as well as the match values of searched products. Moreover, this dimensionality changes endogenously over time, as a consumer's search decision affects the sizes of the searched and the unsearched product sets. Therefore, we also need to address this dimensionality issue.

We address these two challenges with a two-step estimation procedure in which we

back out the mean utility of each property by matching the observed share of visits that a property receives before estimating the dynamic model.[1] The first step requires an extension of the standard invertibility result, as consumers in our setting typically visit *a set of* properties in each period. To this end, we derive the probability that a set of properties is sampled and extend the contraction mapping result in Berry, Levinsohn and Pakes (1995). This two-step procedure allows us to estimate parameters embedded in the mean utility before estimating the dynamic model. As a result, we can include a large number of fixed effects in our utility function specification without significantly increasing the computational burden, and we can further address the price endogeneity issue using an instrumental variable approach. Moreover, the procedure allows us to reduce the number of state variables by replacing a vector of characteristics with the scalar mean utility for each property in estimating the dynamic model.

Our estimation yields intuitive results. We find that consumers prefer larger and newer properties with more living rooms and bedrooms, located on higher floors, and close to a subway station. The estimated standard deviation of match values is equivalent to a value of CN¥96,966, or $14,545 using an exchange rate of 0.15. This value is about 2.5% of the average list price and more than 150% of the annual per capita disposable income in Beijing in 2016.[2] This indicates a significant benefit of searching to learn match values. We also find that consumers incur an average search cost of CN¥3,067, which amounts to an average search cost of CN¥457 per search visit.

In contrast, a static search model estimated using the same dataset yields an unreasonably large search cost, highlighting the importance of accounting for dynamics in our setting. A static search model yields an estimated search cost of about CN¥44,000, or $6,600 per visit, which is almost two orders of magnitude higher than the estimate from our dynamic model. This is because a static search model ignores the dynamics in the search environment and thus must rely on very high search costs to explain the observed number of searches.

Based on the estimated dynamic search model, we conduct counterfactual simulations to quantify how the dynamics of the environment affect consumer behavior and welfare. We find that consumers search longer, visit more properties, and purchase properties that generate higher utility when new listing prices increase more slowly, new listings are more

[1] Many papers use a similar sequential estimation procedure, where certain model features are estimated before others. Examples include Eizenberg (2014), Fan and Yang (2020), and Fan and Yang (2024) for estimating multi-stage static models, and Hendel and Nevo (2006), Chatterjee, Fan and Mohapatra (2024), Bodéré (2023), and Elliott (2024) for estimating dynamic models.

[2] The annual per capita disposable income is CN¥57,275. See "the 2017 Beijing Statistical Yearbook, 9-14 Basic Data on Urban Households" (https://nj.tjj.beijing.gov.cn/nj/main/2017-tjnj/zk/e/indexeh.htm).

frequent, and listing exit rates are lower. Examining the trade-off between the resulting higher search and waiting costs and the higher utility from finding a more desirable property, we find that the average net gain per consumer is CN¥111,012 when consumers face half the new listing price increase speed,[3] CN¥81,756 when the new listing arrival rate is doubled, and CN¥41,013 when the listing exit rate is halved. These increases in utility are significant compared to Beijing's annual per capita disposable income of CN¥57,275 in 2016.

To quantify the relative importance of search environment dynamics versus traditional search frictions such as search costs, we conduct a counterfactual simulation in which we halve the search cost per visit. Unsurprisingly, consumers search longer, visit more properties, and purchase properties with higher utility. The average net gain per consumer is CN¥90,666, which is similar in magnitude to the net gain per consumer from varying the dynamics of the search environment studied above (i.e., reducing the speed of price increase and increasing the arrival rate for new listings).

We also study the effects of housing policies designed to increase housing supply and find that they can significantly benefit consumers through influencing dynamics of the search environment. Housing policies such as taxing vacant properties encourage new listings and slow down price increases for these listings, which, in turn, encourages consumers to search longer and visit more properties, leading to higher search and waiting costs, but better purchase outcomes. We find the net effect to be positive. Specifically, doubling the new listing arrival rate and halving the new listing price increase results in an average net gain of CN¥270,800 per consumer, or about 7% of the average transaction price. A decomposition shows that about 72% of the consumer gain comes from searching longer and visiting more properties, while the remaining 28% is mechanically due to the slower price increase.

This paper contributes to the empirical literature on consumer search. Examples of this literature include Honka (2014) and Murry and Zhou (2020) for simultaneous search models, Hodgson and Lewis (2023) and Moraga-Gonzalez, Sándor and Wildenbeest (2023) for sequential search models, and Santos, Hortaçsu and Wildenbeest (2012) for testing simultaneous versus sequential models. Other examples include Kim, Albuquerque and Bronnenberg (2010), Allen, Clark and Houde (2019), and Brown and Jeon (2024). While the existing papers consider a stable search environment, our paper studies consumer search in a non-stationary search environment where both product availability and price can vary

[3]In the counterfactual scenario where the price increase is halved, consumer welfare increases mechanically due to lower prices. We use actual prices (rather than counterfactual prices) in calculating net gains to remove such a mechanical effect and thereby isolate the effect of inducing more searches.

over time. Our results indicate that search environment dynamics have significant effects on consumers' search and purchase behavior as well as consumer welfare.

This paper also contributes to the empirical literature on search in general. A large branch of this literature focuses on how individuals conduct a job search (see Eckstein and Van den Berg (2007) and French and Taber (2011) for reviews of the literature). Again, the existing literature studies search behavior in a stationary search environment, with the exception of Arcidiacono, Gyetvai, Maurel and Jardim (2022), which estimates a continuous-time non-stationary search model. Moreover, similar to the consumer search literature, the job search literature also assumes no unobservable job characteristics and thus no wage endogeneity. In contrast, we study search in a non-stationary search environment and consider price endogeneity.

The remainder of the paper is organized as follows: Section 2 describes the data. Section 3 develops our dynamic search and purchase model. Section 4 explains our estimation procedure, while 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 search environment dynamics and search costs, while Section 8 quantifies the effects of housing supply policies. A final Section 9 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.[4] The dataset provides information on 225,608 properties listed for sale on Lianjia between August 1, 2015 and July 31, 2016, across all six urban districts of Beijing. These six urban districts are: Chaoyang, Dongcheng, Fengtai, Haidian, Shijingshan, and Xicheng. The dataset also includes the complete property visit and purchase records of 455,774 consumers registered on Lianjia who were actively searching during our sample period.

To construct our sample, we exclude properties with a list price higher than CN¥10 million or less than CN¥1 million, as well as properties with a size of less than 25 square meters as these properties are likely to belong to a separate market than properties in

[4]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/).

Using publicly-available data on Lianjia.com, Habib, Peng, Wang and Wang (2023) and Peng (2023) study how macroeconomic conditions affect the Chinese housing market.

our sample. Accordingly, we drop consumers who visited these excluded properties. We also drop consumers who searched across multiple districts. The vast majority (more than 93%) of consumers searched within one district.[5] Note that the districts in Beijing are quite large. For example, Chaoyang District covers 470.8 square kilometers (181.8 square miles) compared to 59 square kilometers (22.7 square miles) for Manhattan. Finally, we drop the 4% of properties 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 our sample, we observe its list price, address, year of construction, floor level, property size, number of living rooms, and number of bedrooms. We define 221 exclusive segments based on neighborhood and list price range, and assign each property to a segment.[6] In addition to these property characteristics, we also observe the transaction dates and prices for properties sold before the end of the sample period.

A novel feature of our data is that for each consumer in our sample, we have a complete record of all her property visits until she purchases a property or until the end of the sample period. The search record is complete because all consumers in the sample sign a sole agency agreement with Lianjia. Moreover, there are no open houses in China. That is, consumers must work with a realtor agent to visit properties. For each property visit, we observe the date of the visit and the identity of the property. In addition, if a consumer purchases a property during the sample period, we also observe which property she purchases and the transaction date.

2.2 Summary Statistics

2.2.1 Properties

Table 1 presents the summary statistics for the properties in our sample. From Table 1, we can see that the average list price is CN¥4 million ($604,000). In the Chinese housing market, the most salient price is the price per square meter, which averages CN¥49,302 ($18,029) per square meter. Residential properties in Beijing are typically apartments in high-rise buildings. The average property in our sample is 18 years old, 84 square meters in size, with 2 bedrooms and 1 living room. More than 30% of the properties are located on the 10th floor or above. About 80% of the properties are located within 1 km of a

[5]Including the small share of consumers who search in multiple districts increases the computational burden significantly. This is because, in our estimation, we invert the mean utilities district by district. If we were to include consumers who searched in multiple districts, we would have to do the inversion for all districts at once.

[6]Lianjia defines a property's neighborhood based on its location. Neighborhoods differ in terms of transportation, amenities, etc. We consider four price ranges: lower than 3 million, between 3 and 4.5 million, between 4.5 and 6 million, and higher than 6 million CN¥.

subway station, a criterion we use to define our indicator variable of whether a property is close to a subway station.

Of the 202,845 properties in our sample, 85,696 (42%) are sold during our sample period. On average, a property stays on the market for approximately 8 weeks. The average transaction price is CN¥3.7 million ($555,000) with a unit transaction price of CN¥48,367 ($17,501) per square meter.

Table 1: Summary Statistics of Properties

	Mean	SD
List price (million CN¥)	4.024	1.962
List price per m^2 (CN¥)	49,302	18,029
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
Above 10th floor	0.316	0.465
Close to a subway station	0.797	0.402
Indicator of being sold	0.422	0.494
Weeks on market	8.193	7.402
Transaction price (million CN¥)	3.702	1.760
Transaction price per m^2 (CN¥)	48,367	17,501

2.2.2 New Listings, Transactions, and Prices Over Time

Both the number of new listings and the number of transactions are relatively stable over the sample period, as shown in Figure 1(a). In this figure, we omit the eight weeks around the Chinese New Year because, during the Chinese New Year, many economic activities are put on hold as a large number of Chinese return to their hometowns to celebrate and resume economic activities only afterward. For example, 2.9 billion passenger trips were made during the 2016 holiday (Zhou (2016)).[7]

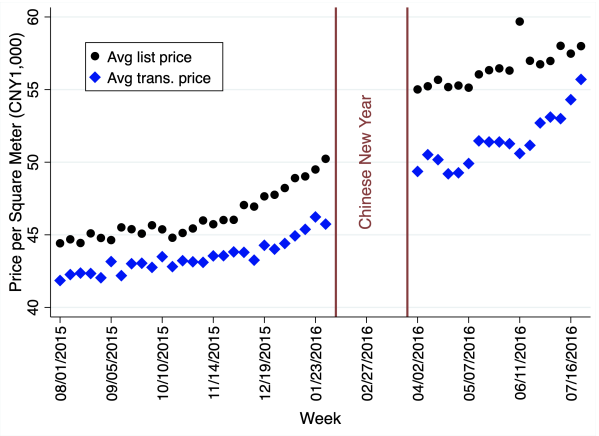
However, both the average list price and the average transaction price increase rapidly during the sample period, as shown in Figure 1(b). For each week in our sample, we calculate the average list price across all new listings in that week and the average transaction price across all transacted properties in that week, and then plot them in Figure 1(b). Since the most salient price in the Chinese housing market is the unit price per square meter, we plot prices in CN¥ per square meter. For this figure, we can see that the

[7]In Supplemental 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, both numbers initially drop to zero and then rise rapidly to about double their pre-holiday levels.

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

(a) New Listings and Transactions

(b) List Prices and Transaction Prices



average list price increases from CN¥44,417 to CN¥57,993 per square meter, representing an annual increase of approximately 30% and an average weekly increase of CN¥261 per square meter during the sample period. For an average-sized property (of 84 square meters), this weekly increase is equivalent to CN¥21,924, or about 40% of the annual per capita disposable income in Beijing in 2016. Figure 1(b) also shows that the average transaction price closely tracks the evolution of the average list price with a stable gap of around CN¥3,000 per square meter.

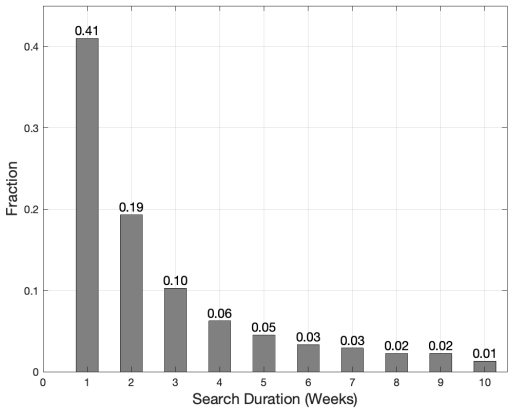
2.2.3 Consumer Search and Purchase Behavior

In our sample, 39,500 consumers start a search and make a purchase during the sample period. Among them, 26,543 consumers either end their search before the Chinese New Year holiday or start their search after the holiday. For these consumers, we observe their complete search record from the beginning to the end of their search.

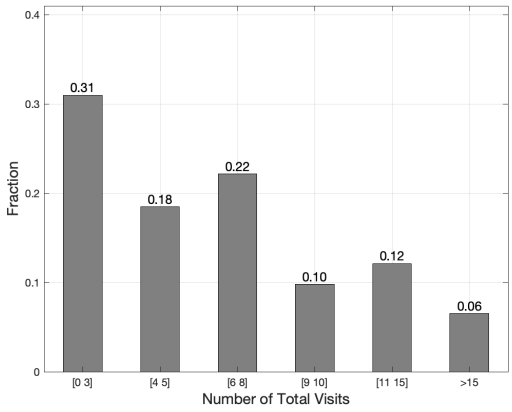
Specifically, we observe these consumers' overall search duration and total number of property visits. Regarding search duration, Figure 2(a) shows that 41% search for one week, 19% search for two weeks, 10% for three weeks, and 30% for more than three weeks. Regarding the number of properties they visit, Figure 2(b) shows that 49% visit five or fewer properties, 32% visit more than five but no more than ten, and 18% visit more than ten properties before making a purchase. The average number of search weeks and property visits are 3.5 weeks and 6.6 visits, respectively. In comparison, U.S. homebuyers, on average, search for 8.2 weeks and visit 10 properties, according to the National Association of Realtor's survey of 2,372 homebuyers from 1987 to 2007 (Genesove and Han (2012)).

Figure 2: Histogram of Search Duration and Total Visits

(a) Search Duration

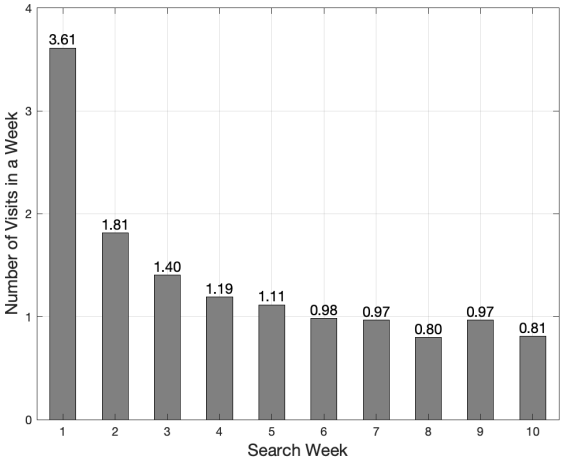


(b) Total Visits



A unique and 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., their search duration and search intensity in each week of the search process). Figure 3 shows the average number of visits by search week. "Search week" refers to the week since a consumer's first property visit. For each search week t, we compute the number of visits a consumer makes in that search week averaged over consumers who search for t or more weeks. From Figure 3, we can see that search intensity decreases over search weeks, from an average of 3.61 visits in the first search week, to 1.81 visits in the second search week, to less than 1 visit after five weeks of searching.

Figure 3: Average Number of Visits by Search Week



In addition to a complete search record for each consumer, we also observe which

property a consumer purchases and when. In our sample, some consumers purchase a property that they searched for in the current period, while others purchase a property that they searched for in a previous period. We call this latter action a recall. The share of recalls in our sample is 17.6%.

Finally, for each consumer, we define her search market as the union of the segments in which she searches. On average, a consumer searches in two segments. The average consumer search market consists of 654 properties in a given week, has 29 new listings per week, and experiences an average weekly exit rate of 12%.

3 Dynamic Search Model

In this section, we develop a model to describe consumers' search and purchase decisions in a non-stationary search environment. Each consumer i arrives exogenously. To simplify the notation, we use t (without the subscript i) to denote the week since a consumer's arrival, which we refer to as the "search week." We use j to denote a property and $m(j)$ to represent the segment to which property j belongs.

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

3.1 Primitives

3.1.1 Utility

The utility that consumer i derives from property j is as follows:

$$u_{ijt} = \boldsymbol{x}_j\boldsymbol{\beta} + \alpha p_{jt} + \xi_j + v_{ij}, \tag{1}$$

where the vector $\boldsymbol{x}_j$ represents observable property characteristics, such as the property size and the number of bedrooms. The term ξ_j captures property characteristics that are known to consumers but are unobservable to researchers. For example, while consumers can infer whether a property receives a lot of natural light from listing photos on Lianjia's website, such information is difficult for researchers to obtain.

The price p_{jt} is the sum of an expected transaction price (p_j^e) and a shock (η_{jt}). The expected transaction price incorporates the list price as well as an expectation of the difference between the transition and list price. Specifically, $p_j^e = p_j^l + g_{m(j)}x_{1j}$, where p_j^l is the

list price, $g_{m(j)}$ is the average price difference per square meter for a property in segment $m(j)$, and x_{1j} is the property size. While p_j^e is time invariant, the price shock η_{jt} is assumed to be i.i.d. across both properties and time and follows a normal distribution with mean 0 and variance equal to its empirical variance.

The idiosyncratic term v_{ij} in (1) captures the match value of property j for consumer i, which consumer i learns after visiting property j. It captures the consumer-specific taste for a product. For example, it includes consumer i's preferences for a particular aspect of the property's floor plan 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_v^2.

We can rewrite the utility as

$$u_{ijt} = \delta_j + \alpha\eta_{jt} + v_{ij} \tag{2}$$

by collecting the property j-specific terms in

$$\delta_j \quad = \quad \alpha p_j^e + \boldsymbol{x}_j\boldsymbol{\beta} + \xi_j, \tag{3}$$

which we label as the mean utility of property j.

3.1.2 Search Costs and Waiting Costs

We assume that the cost of searching n properties at time t is $C_{it}(n) - \vartheta_{itn}$, 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 deterministic component $C_{it}(n)$ is given by

$$C_{it}(n) = \left(\gamma_0 + (\gamma_1 + \gamma_2 m_{it})n + \gamma_3 n^2\right) \mathbb{1}(n > 0), \tag{4}$$

where γ_0 captures the baseline cost of searching (as opposed to no searching). The marginal search cost is $\gamma_1 + \gamma_2 m_{it} + 2\gamma_3 n$, which depends on the cumulative number of searches before time t (denoted by m_{it}). We include the quadratic term n^2 in (4) to capture potential 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 may reflect psychological stress related to not being able to provide a home in a timely manner. For example, in China, it is a culture norm for the groom's family to purchase a home—or at least find one and make the down payment—for the newlyweds prior to marriage (Wei and Zhang (2011)). We assume that w_{it} is i.i.d. and follows a normal distribution with mean w and variance σ_w^2.

3.1.3 Search Set Conditional on the Number of Visits

In each period t, consumer i optimally decides on the number of properties to visit (denoted by n_{it}). We assume that, conditional on n_{it}, the exact set of properties she visits in that period is exogenously drawn. 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 exogeneity assumption does not mean that the set of properties that a consumer visits is completely random. On the contrary, the sampling probability we specify below ensures that a property with a higher mean utility has a higher probability of being sampled. This assumption simply means that the consumer and her agent do not have full control over the set of properties they can visit at a particular time, and that there are exogenous factors influencing whether a consumer can or cannot to visit a given property at a particular time.[8]

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 let $\mathcal{C}^n(\mathcal{A}_{it})$ represent the collection of all subsets of $\mathcal{A}_{it}$ of size n. 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:

$$\Pr(\mathcal{N}|\mathcal{A}_{it}, n) = \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \frac{\sum_{l \in \mathcal{B}} exp(\delta_l)}{\sum_{l \in (\mathcal{A} \setminus \mathcal{N}) \cup \mathcal{B}} exp(\delta_l)} \right]. \tag{5}$$

This sampling probability has several desirable features. First, it is consistent with an extended Logit model. In 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 $n \geq 1$ options. Though seemingly complicated, the expression in (5) is a direct application of the analytic expression for the choice probability in a standard Logit model and the inclusion-exclusion principle.

Second, as we show in Appendix A, the sampling probability in (5) for *a set of properties*

[8] This assumption simplifies both the consumer's decision problem and our estimation. First, if a consumer were to choose the set of properties to visit, she would have $C_{A_{it}}^{n_{it}}$ such sets to choose from, where A_{it} is the number of properties available 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 consumer i has $C_{A_{it}}^{n_{it}} = 6,435$ possible sets to choose from in period t. It seems overly demanding for our model to explain why consumer i chooses a particular set of properties out of the 6,435 possible sets. Second, as explained in Appendix B, this assumption allows us to estimate our model parameters in two steps, which helps address a price endogeneity issue and estimate a large number of fixed-effect parameters.

$\mathcal{N} \subset \mathcal{A}_{it}$ implies the following sampling probability for *a particular property* $j \in \mathcal{A}_{it}$:

$$\begin{aligned}
&\Pr(j|\mathcal{A}_{it}, n) \\
&= \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n = A_{it} \\ \sum_{k=0}^{n-1} \left[(-1)^k \sum_{\mathcal{B} \in \mathcal{C}^{A-n+k}(\mathcal{A} \setminus j)} \left(\frac{C^k_{A-n+k-1} exp(\delta_j)}{\sum_{l \in \mathcal{B}} exp(\delta_l) + exp(\delta_j)} \right) \right] & \text{if } 0 < n < A_{it}, \end{cases}
\end{aligned} \tag{6}$$

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) when $n = 1$, it becomes the choice probability in a standard Logit model, i.e., $\frac{exp(\delta_j)}{\sum_{l \in \mathcal{A}_{it}} exp(\delta_l)}$; and (iii) the sum of this sampling probability over all possible 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

In each period, a consumer makes two decisions: a search decision and a purchase decision. When making these decisions, a consumer considers both the set of properties she has visited and the set of those she has not. We denote the set of properties that consumer i visits before time t and that are still available at time t by $\mathcal{R}_{it}$ (where $\mathcal{R}$ stands for "Recall"). As mentioned, we denote the set of properties she has not visited by time t by $\mathcal{A}_{it}$ (where $\mathcal{A}$ stands for "Available for search").

The timing is as follows:

- At the beginning of the period, consumer i observes δ_j for properties in both $\mathcal{R}_{it}$ and $\mathcal{A}_{it}$ as well as the match value v_{ij} for properties in her recall set $\mathcal{R}_{it}$. Consumer i also observes the search cost shocks ϑ_{itn} for $n = 0, ..., \bar{n}$, where $\bar{n}$ is the maximum number of searches in a period. She decides how many properties to search in time t, which is denoted by n_{it}, 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 (5). Here, $\mathcal{N}$ stands for "Newly searched."

- After visiting the properties in $\mathcal{N}_{it}$, consumer i observes the price shock η_{jt} and the match value v_{ij} for all properties in $\mathcal{R}_{it} \cup \mathcal{N}_{it}$. She now also observes her waiting cost w_{it}. She decides whether to purchase a property in $\mathcal{R}_{it} \cup \mathcal{N}_{it}$ or to continue searching. We denote this decision by y_{it}.

3.1.5 Transition of the Environment

The search environment is non-stationary. Specifically, from time t to time $t+1$, there are three changes in consumer i's search market. Because consumers differ in their search markets, these changes are also consumer-specific.

First, some properties may exit the market at the end of time t. 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 at the beginning of time $t+1$. We assume that the number of new listings 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 new listings in her search market at the beginning of time $t+1$. We further assume that the mean utility of a newly-listed property in consumer i's search market follows a normal distribution $N(\mu_{it}^{new}, (\sigma_i^{new})^2)$.

Third, list prices of new listings in time $t+1$ may be higher than those of new listings in time t. We assume that the trend in the list price of new properties in consumer i's search market is ρ_i. Therefore, in forming an expectation about the next period, consumer i considers the transition of μ_{it}^{new} to be $\mu_{it+1}^{new} = \mu_{it}^{new} + \alpha\rho_i$.

The above three changes determine the transition of consumer i's information set over time. Specifically, at the beginning of time t, consumer i's information set is $(\{\delta_j\}_{j\in\mathcal{A}_{it}}, \{\delta_j, v_{ij}\}_{j\in\mathcal{R}_{it}}, \mu_{it}^{new}, m_{it}, \vartheta_{itn})$. Among these variables, ϑ_{itn} is the i.i.d. search cost shock, while $\Omega_{it} = (\{\delta_j\}_{j\in\mathcal{A}_{it}}, \{\delta_j, v_{ij}\}_{j\in\mathcal{R}_{it}}, \mu_{it}^{new}, m_{it})$ follows a transition determined by the following: the available-to-search set $\mathcal{A}_{it+1} = \mathcal{A}_{it}\backslash\mathcal{N}_{it}\backslash\mathcal{EXIT}_{it} \cup \mathcal{NEW}_{it+1}$, the recall set $\mathcal{R}_{it+1} = \mathcal{R}_{it} \cup \mathcal{N}_{it}\backslash\mathcal{EXIT}_{it}$, the average mean utility of new listings $\mu_{it+1}^{new} = \mu_{it}^{new} + \alpha\rho_i$, and the cumulative number of searches $m_{it+1} = m_{it} + n_{it}$.

3.2 Consumer Decisions

Having described the model primitives, we now describe how consumers make decisions. We describe consumer i's problem in each period backwards: first, we describe her purchase decision after searching, and then we describe her search intensity decision.

3.2.1 Purchase Decision

After visiting properties in the newly-searched set $\mathcal{N}_{it}$ in time t, consumer i observes the match values of the properties in her newly-searched set $\{v_{ij}\}_{j\in\mathcal{N}_{it}}$ in addition to those in her recall set $\{v_{ij}\}_{j\in\mathcal{R}_{it}}$ (from the information set Ω_{it}). She also observes the price shocks

of all properties in her choice set $\{\eta_{jt}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}$, as well as her waiting cost w_{it}. At this point, the choices for consumer i consist of the following options: recalling a property from the previously searched set $\mathcal{R}_{it}$, purchasing a property from the newly-searched set $\mathcal{N}_{it}$, or continuing to search. In other words, her optimization problem at the purchase-decision stage is as follows:

$$\begin{aligned} & \Gamma_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}, \{\eta_{jt}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}, w_{it}) \qquad (7) \\ = \ & \max\{ \underbrace{\max_{j\in\mathcal{R}_{it}} \delta_j + \alpha\eta_{jt} + v_{ij}}_{\text{recall}}, \ \underbrace{\max_{j\in\mathcal{N}_{it}} \delta_j + \alpha\eta_{jt} + v_{ij}}_{\text{buy a newly-searched property}} , \ \underbrace{E_i\left[V_i(\Omega_{it+1})|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}\right] - w_{it}}_{\text{wait}} \}, \end{aligned}$$

where $E_i\left[V_i(\Omega_{it+1})|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}}\right]$ denotes the expected value of continuing to search. This expectation depends on both her information set Ω_{it} and the match values of her newly-searched properties $\{v_{ij}\}_{j\in\mathcal{N}_{it}}$.[9] Her purchase decision (denoted by y_{it}) is the optimizer of the above optimization problem.

3.2.2 Search Decision

A consumer's optimal search intensity in each period, i.e., the number of properties to search (n_{it}), depends on the results from her comparison of the benefits and costs of searching. We specified the search costs in Section 3.1. We now explain the search benefits. Specifically, the benefit of searching n_{it} properties is that consumer i learns about 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 ($\mathcal{R}_{it} \cup \mathcal{N}_{it}$).

Formally, the expected benefit of searching n properties, conditional on the information set Ω_{it}, is:

$$\begin{aligned} & EB_i(n|\Omega_{it}) \qquad (8) \\ = \ & \sum_{\mathcal{N}\in\mathcal{C}^n(\mathcal{A}_{it})} E_{\left(\{v_{ij}\}_{j\in\mathcal{N}}, \{\eta_{jt}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}}, w_{it}\right)}\left[\Gamma_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}}, \{\eta_{jt}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}}, w_{it})\right] \times \Pr(\mathcal{N}|\Omega_{it}, n), \end{aligned}$$

where the first term $E_{...}\left[\Gamma_i(\Omega_{it}, ...)\right]$ 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 a particular set $\mathcal{N}$ is sampled, conditional on both the 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 (5) as $\Pr(\mathcal{N}|\Omega_{it}, n)$ to reflect its dependence on $\{\delta_j\}_{j\in\mathcal{A}_{it}}$, which is a subset of the information in Ω_{it}.

In deciding on her search intensity, consumer i chooses n_{it} to maximize her net gain

[9] The expectation is consumer-specific because the transition of Ω_{it} may differ across consumers.

from searching. That is, the optimal search intensity n_{it} is the solution to the following optimization problem:

$$\max_{0\leq n\leq \bar{n}} \left[EB_i(n|\Omega_{it}) - C(n|\Omega_{it}) + \vartheta_{itn}\right], \tag{9}$$

where, again with a slight abuse of notation, we use $C(n|\Omega_{it})$ to denote the search cost function $C_{it}(n)$ in (4), reflecting its dependence on m_{it} in Ω_{it}.

3.3 Bellman Equation

We define the *ex ante* value function as the expectation of the maximum in (9) over the search cost shocks ϑ_{itn}. Since ϑ_{itn} follows a type-1 extreme value distribution with scale parameter κ, we have:

$$\begin{aligned} V_i(\Omega_{it}) =& E_{(\vartheta_{itn}, n=0,\ldots,\bar{n})}\{\max_{0\leq n\leq \bar{n}} [EB_i(n|\Omega_{it}) - C(n|\Omega_{it}) + \vartheta_{itn}]\} \\ =& \kappa \ln\left(\sum_{n=0}^{\bar{n}} exp\left(\frac{EB_i(n|\Omega_{it}) - C(n|\Omega_{it})}{\kappa}\right)\right) + \kappa\tau, \end{aligned} \tag{10}$$

where τ is the Euler constant. Therefore, by plugging $\Gamma_i(\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}}, \{\eta_{jt}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}}, w_{it})$ from (7) into $EB_i(n|\Omega_{it})$ in (8), and then plugging $EB_i(n|\Omega_{it})$ into (10), we obtain the Bellman equation.

3.4 Discussions

We conclude this section with a discussion of three simplifications we have made in our model. First, we assume that there is no correlated learning. In particular, we assume that visiting one property does not allow consumers to learn about other properties. In contrast, Hodgson and Lewis (2023) allow for correlations among unobservable match values in their model, with a stronger correlation between products more similar in observable characteristics. As a result of such correlations, consumer search in their model exhibits a spatial learning pattern, meaning that a consumer's search tends to converge to the chosen product in the product characteristics space. We assume no correlated learning because we do not observe clear evidence of such a pattern in our setting. For example, in Online Supplemental Appendix SA, we plot the probability that a consumer visits a property within the same residential complex as her final purchase and show that this probability does not increase as she approaches the end of her search. In other words, consumers' searches do not necessarily converge to the property they purchase in terms of location.

Another reason we assume away correlated learning is that it implies dynamic learning, which would make it computationally infeasible to include both search environment dynamics and dynamic learning in our model. Given the focus of the paper (i.e., the effect of search environment dynamics) and the lack of clear evidence on correlated learning, we retain the former and abstract away the latter in our model.

Second, we do not model sellers' decisions. In reality, sellers also make strategic decisions. For example, Merlo, Ortalo-Magné and Rust (2015) study sellers' dynamic decisions regarding initial list prices, list price revisions, and offer acceptance. However, incorporating both sellers' and buyers' dynamic decisions in a non-stationary search environment would be challenging. Instead, we capture the effects of seller decisions in a somewhat reduced-form way. Regarding list prices, we use an instrumental variable approach to address list price endogeneity concerns. Regarding list price revisions, we note that list price revisions in our sample are rare, reducing our concern about not including this seller decision in our model. As for the offer acceptance decision, we capture its effect in the price shock term η_{jt}. For example, if a seller is more willing to accept a low offer, the price shock may be negative. Conversely, if a seller is currently considering a higher offer, the price shock will be positive and large.

Third, we do not model buyer competition. In reality, multiple buyers may bid simultaneously for the same property. However, we do not observe the number of bids or the bid amounts. In our model, the price shock term η_{jt} captures buyer competition to some extent, as intense buyer competition would be reflected by a large price shock.

4 Estimation

The estimation consists of two steps. In the first step, we estimate the parameters of the utility function $(\alpha, \boldsymbol{\beta})$ by matching the observed share of visits that each property receives. We refer to these parameters as the static parameters because they are estimated without solving the dynamic model. We use all properties and consumers in our sample in this step of the estimation.

In the second step, we estimate the remaining parameters which govern a consumer's dynamic search and purchase decisions. These parameters include the search cost parameters $(\boldsymbol{\gamma}, \kappa)$, the waiting cost parameters (w, σ_w), and the standard deviation of the match values (σ_v). We refer to these parameters as the dynamic parameters. In this estimation step, we use a random sample of 1,000 consumers whose entire search record falls within

the sample period.[10]

Appendix B provides a discussion of our two-step estimation procedure, highlighting its advantages (e.g., allowing for price endogeneity and a large set of fixed effects and addressing the issue of high dimensionality of the state space) and clarifying the assumptions needed. Supplemental Appendix SB further provides details on the estimation procedure.

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

We estimate the parameters in the utility function $(\alpha, \boldsymbol{\beta})$ by matching the observed share of visits that each property receives. Specifically, we first invert out the mean utility of each property based on the visit shares and then regress the mean utility of a property on its price and characteristics to obtain estimates for $(\alpha, \boldsymbol{\beta})$. In our sample, all consumers search within one district. Therefore, we partition the sample into six districts and carry out the inversion district by district. 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

$$\tilde{s}_j(\boldsymbol{\delta}_d) = \frac{\sum_{i \in \mathcal{I}_d} \sum_{t=1}^{T_i} \Pr(j|\mathcal{A}_{it}, n_{it})}{\sum_{i \in \mathcal{I}_d} \sum_{t=1}^{T_i} n_{it}}, \tag{11}$$

where $\boldsymbol{\delta}_d = (\delta_j, j \in \mathcal{J}_d)$ represents the mean utilities of properties in district d. In (11), n_{it} is the observed number of properties that consumer i visits at time t. $\Pr(j|\mathcal{A}_{it}, n_{it})$ is the probability that consumer i visits property j at time t, where $\mathcal{A}_{it}$ is the set of properties available for consumer i to search. This probability is 0 for $j \notin \mathcal{A}_{it}$ and is given by (6) 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 of visits is:

$$s_j = \frac{N_j}{\sum_{j \in \mathcal{J}_d} N_j}, \tag{12}$$

[10] We consider only consumers who begin their search after the start of our sample period because we need to know m_{it}, the cumulative number of searches before time t, and $\mathcal{R}_{it}$, the set of properties that consumer i visits before time t and that are still available at time t. However, strictly speaking, we do not need to restrict our sample to those who purchase before the end of the sample period. We do so in our baseline estimation due to the concern that consumers who do not make a purchase during the sample period may not be seriously searching in the housing market. Nevertheless, as a robustness check, we repeat our estimation, including consumers who do not purchase a property before the end of the sample period, in Supplemental Appendix SC, and show that our estimation results are robust.

where N_j is the number of visits that property j receives in the sample. Note that the denominators in (11) and (12) are the same and that both represent the sum of the visits that properties in district d receive.

We invert out the mean utilities $\boldsymbol{\delta}_d$ by matching the model visit shares to their empirical counterparts, i.e.,

$$\tilde{s}_j(\boldsymbol{\delta}_d) = s_j, j \in \mathcal{J}_d. \tag{13}$$

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. This extension allows us to show that the system of equations in (13) 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. As a result, all inverted mean utilities are relative to that of the normalized property.

To estimate $(\alpha, \boldsymbol{\beta})$, we regress the inverted δ_j on the expected price p_i^e and property characteristics $\boldsymbol{x}_j$ according to equation (3).[11] Because the price of a property is likely to be correlated with unobservable property characteristics, the expected price is endogenous. We address this endogeneity issue by using an instrumental variable approach. Specifically, we construct the instrumental variable as the transaction price of properties in the same segment averaged across transactions that occur 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.

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

We estimate the dynamic parameters $(\boldsymbol{\gamma}, \kappa, w, \sigma_w, \sigma_v)$ using maximum likelihood estimation. For each consumer i, we observe her search duration T_i. In addition, for each search week $t = 1, ..., T_i$, we observe her search intensity (i.e., the number of visits n_{it}) and her purchase decision (i.e., $y_{it} = recall$ – to purchase a previously visited property, $y_{it} = j \in \mathcal{N}_{it}$ – to purchase property j which is newly visited by her in the current period,

[11] Since δ_j is relative to the mean utility of the normalized property, the regressors are $p_j^e - p_{j_0^d}^e$ and $\boldsymbol{x}_j - \boldsymbol{x}_{j_0^d}$, where j_0^d is the normalized property in district d.

or $y_{it} = wait$ – to search longer).[12]

Letting $\theta = (\boldsymbol{\gamma}, \kappa, w, \sigma_w, \sigma_v)$ summarize the dynamic parameters in our model, the likelihood of observing the search and purchase path $\{n_{it}, y_{it}\}_{t=1}^{T_i}$ for consumer i is:

$$l_i(\theta) = \int \left(\prod_{t=1}^{T_i} [\Pr_i(n_{it}|\Omega_{it};\theta) \cdot \Pr_i(y_{it}|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}};\theta)] \right) dF_v(\{\{v_{ij}\}_{j\in\mathcal{N}_{it}}\}_{t=1}^{T_i};\theta), \quad (14)$$

where $\Pr_i(n_{it}|\Omega_{it};\theta)$ and $\Pr(y_{it}|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{N}_{it}};\theta)$ are, respectively, the probability that consumer i searches n_{it} properties and the probability that she makes the purchase decision y_{it}, both of which we derive below. In (14), $F_v(\{\{v_{ij}\}_{j\in\mathcal{N}_{it}}\}_{t=1}^{T_i};\theta)$ represents the distribution of the match values of all properties that consumer i has ever visited.[13] We integrate out these match values because they are unobservables in the probabilities. Note that, after recovering δ_j in the first step of the estimation, we observe $\{\delta_j\}_{j\in\mathcal{A}_{it}}$ and $\{\delta_j\}_{j\in\mathcal{R}_{it}}$ in Ω_{it}. The only unobservable variables are $\{v_{ij}\}_{j\in\mathcal{R}_{it}}$ (i.e., the match values of properties in the recall set, which are embedded in Ω_{it} in both probabilities), and $\{v_{ij}\}_{j\in\mathcal{N}_{it}}$ (i.e., the match values of newly-searched properties). Since $\{\{v_{ij}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}\}_{t=1}^{T_i} = \{\{v_{ij}\}_{j\in\mathcal{N}_{it}}\}_{t=1}^{T_i}$, by integrating out the match values of all properties that consumer i has visited, we integrate out all unobservables in the probabilities.

To derive the probability of the search decision, $\Pr_i(n_{it}|\Omega_{it};\theta)$, we note that consumer i chooses n_{it} to maximize her net gain from searching, as described in the optimization problem in (9), after observing Ω_{it} and the search cost shocks ϑ_{itn}. Given that the search cost shock ϑ_{itn} follows a type-1 extreme value distribution with scale parameter κ, the probability that consumer i searches n_{it} properties is:

$$\Pr_i(n_{it}|\Omega_{it};\theta) = \frac{exp\{[EB_i(n_{it}|\Omega_{it};\theta) - C(n_{it}|\Omega_{it};\theta)]/\kappa\}}{\sum_{n=0}^{\bar{n}} exp\{[EB_i(n|\Omega_{it};\theta) - C(n|\Omega_{it};\theta)]/\kappa\}}, \quad (15)$$

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

To derive the probability of the purchase decision, $\Pr(y_{it}|\Omega_{it}, \{v_{ij}\}_{j\in\mathcal{A}_{it}};\theta)$, we note that, after visiting n_{it} properties, consumer i observes the original information set Ω_{it}, the match

[12] 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 a property, rather than which specific property she recalls. This is because the probability of purchasing a particular property in the recall set depends on $\delta_j + v_{ij}$ for all $j \in \mathcal{R}_{it}$ (as well as $\{\delta_j\}_{j\in\mathcal{A}_{it}}$), resulting in high-dimensional state variables. In contrast, the probability of recalling a property depends on a summary statistic, $\max_{j\in\mathcal{R}_{it}}(\delta_j + v_{ij})$ (along with $\{\delta_j\}_{j\in\mathcal{A}_{it}}$). Therefore, for computational reasons, we do not use information on which property a consumer recalls in the estimation.

[13] It is $\prod_{j\in\mathcal{N}_{it}} \prod_{t=1}^{T_i} \Phi(v_{ij}/\sigma_v)$, where $\Phi(\cdot)$ is the standard normal distribution function.

values of the newly-searched properties $\{v_{ij}\}_{j\in\mathcal{N}_{it}}$, the price shock η_{jt}, and her waiting cost ω_{it}. She then makes a purchase decision according to the optimization problem in (7). Therefore, the choice probabilities are:

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

$$\begin{aligned}
&\Pr_i(y_{it}=j\in\mathcal{N}_{it}|\Omega_{it},\{v_{ij}\}_{j\in\mathcal{N}_{it}};\theta) \qquad (17)\\
=&E_{\{\eta_{jt}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}}\Bigg\{\mathbb{1}\Bigg(\delta_j+\alpha\eta_{jt}+v_{ij}\geq\max_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}[\delta_j+\alpha\eta_{jt}+v_{ij}]\Bigg)\\
&\times\Phi\Bigg(\Big\{\max_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}[\delta_j+\alpha\eta_{jt}+v_{ij}]-E_i\left[V_i(\Omega_{it+1};\theta)|\Omega_{it},\{v_{ij}\}_{j\in\mathcal{N}_{it}}\right]+w\Big\}/\sigma_w\Bigg)\Bigg\},
\end{aligned}$$

$$\begin{aligned}
&\Pr_i(y_{it}=wait|\Omega_{it},\{v_{ij}\}_{j\in\mathcal{N}_{it}};\theta) \qquad (18)\\
=&E_{\{\eta_{jt}\}_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}}\Bigg\{\Phi\Bigg(\Big\{E_i\left[V_i(\Omega_{it+1};\theta)|\Omega_{it},\{v_{ij}\}_{j\in\mathcal{N}_{it}}\right]-w-\max_{j\in\mathcal{R}_{it}\cup\mathcal{N}_{it}}[\delta_j+\alpha\eta_{jt}+v_{ij}]\Big\}/\sigma_w\Bigg)\Bigg\},
\end{aligned}$$

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

We estimate θ using maximum likelihood 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 solution to the Bellman equation.

The dimension of the vector of the state variables Ω_{it} is high and changes over time both exogenously and endogenously. This is because the state variables $\Omega_{it}=(\{\delta_j\}_{j\in\mathcal{A}_{it}},\{\delta_j,v_{ij}\}_{j\in\mathcal{R}_{it}},\mu_{it}^{new},m_{it})$ include the mean utilities of the properties available to search ($\{\delta_j\}_{j\in\mathcal{A}_{it}}$) as well as both the mean utilities and the match values of the properties available to recall ($\{\delta_j,v_{ij}\}_{j\in\mathcal{R}_{it}}$). Thus, the state space is large. Moreover, the sizes of both the available-to-search set $\mathcal{A}_{it}$ and the recall set $\mathcal{R}_{it}$ change over time due to new listing entries and current listing exits. These sizes are also influenced by consumer i's endogenous decision regarding the number of searches per period. To address this dimensionality issue, we follow much of the literature on dynamic estimation (e.g., Sweeting (2013) and Hodgson (2024)) and approximate the value function of high-dimensional state variables

with a function of the statistics of these variables. We provide more details on dynamic estimation in Supplemental Appendix SB.

Identification The standard deviation of the match value (σ_v) is identified by both the static and dynamic features of consumer purchase patterns. Statically, we observe which property each consumer purchases. This information helps identify σ_v because the gap between the mean utility of a consumer's purchased property and the highest mean utility in the consumer's choice set at the time of purchase informs us about the importance of match values. A larger gap implies a larger standard deviation for match values. Dynamically, we observe when a consumer visits the property she ultimately purchases. This information also helps identify σ_v because if a consumer purchases a property that she visits in an earlier period (i.e., recalling the property), her continued searching reflects her belief that there is a good chance of getting a better draw of the match value in the future. Therefore, a larger recall share also indicates a larger standard deviation of match values.

The search cost parameters $(\boldsymbol{\gamma}, \kappa)$ are identified by the mean and variance of the search intensity. Recall that the deterministic component of the search cost is $(\gamma_0 + (\gamma_1 + \gamma_2 m_{it})n + \gamma_3 n^2)\mathbb{1}(n > 0)$. While the fraction of observations with zero searches identifies γ_0, the variation in the fraction of n searches across n identifies γ_1 and γ_3. Similarly, the way search intensity varies with m_{it} identifies γ_2. Finally, the scale parameter of the search cost shock κ is identified by the observed variance in the search intensity.

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

5 Estimation Results

5.1 Estimates of the Static Parameters

We estimate the parameters of the utility function $(\alpha, \boldsymbol{\beta})$ by regressing the inverted mean utility δ_j on the expected price and property characteristics. The results are presented in Table 2, with the OLS estimates reported in column (I) and the IV regression estimates in column (II). Since properties that are more attractive to consumers in ways unobservable to researchers may have higher prices, prices are potentially endogenous, leading to an upward bias for the price coefficient in the OLS regression. Consistent with this intuition, we find that the IV regression indeed yields a more negative price coefficient.

In the following, we focus on the IV regression results.

Overall, the estimation yields intuitive results. Consumers prefer newer, larger properties with more bedrooms and more living rooms, especially those located on the 10th floor or higher and close to a subway station. For example, on average, an extra bedroom with an average size of 20 square meters is valued at 0.3 (=(0.356+0.047×20)/4.559) million CN¥. Similarly, an additional living room with an average size of 30 square meters is valued at 0.5 (=(1.065+0.047×30)/4.559) million CN¥. Locating on the 10th floor or above is worth CN¥34,900 more than locating below the 10th floor. As a robustness analysis, in Supplemental Appendix SC, we allow the price coefficient to vary by district. We find some heterogeneity in price sensitivity among consumers searching in different districts, though the estimates for parameters common in the baseline and robustness specifications are close.

Table 2: Estimates of Parameters in Mean Utility

	(I) OLS		(II) IV	
	Est	SE	Est	SE
Expected price (million CN¥)	-0.192***	(0.006)	-4.559***	(0.143)
Property age (year)	-0.031***	(0.001)	-0.009***	(0.001)
# Bedrooms	0.206***	(0.007)	0.356***	(0.016)
# Living rooms	0.530***	(0.008)	1.065***	(0.025)
Property size (m^2)	0.005***	(0.000)	0.047***	(0.001)
Above 10th floor	0.089***	(0.007)	0.159***	(0.017)
Close to a subway station	-0.029**	(0.010)	0.209***	(0.023)
Neighborhood FE	yes		yes	

*** $p < 0.01$, ** $p < 0.05$.

5.2 Estimates of the Dynamic Parameters

Table 3 reports the estimation results for the dynamic parameters, including the standard deviation of the match values (σ_v), the waiting cost parameters (w, σ_w), and the search cost parameters ($\boldsymbol{\gamma}, \kappa$).

From Table 3, we see a significant benefit of searching to learn about the match value. The estimated standard deviation of the match value is 0.442. Based on the estimated price coefficient (-4.559), this estimated standard deviation corresponds to a value of CN¥96,966 (=0.442/4.559×10^6), or about 2.5% of the average list price and 169% of the annual per capita disposable income in Beijing in 2016.

From Table 3, we can also see that the waiting cost is, on average, 0.047, which is equivalent to CN¥10,407 (=0.047/4.559×10^6) per week. This relatively high waiting

cost is consistent with the observation that consumers search for 3.5 weeks on average before purchasing a property. That being said, in relative terms, this average waiting cost is only 0.26% of the average list price.

Finally, we see from Table 3 that the baseline search cost ($\hat{\gamma}_0$) is CN¥1,150 (=5.242$\times$0.001/4.559$\times 10^6$) and that the marginal search cost increases with the cumulative number of searches a consumer has made in the previous weeks ($\hat{\gamma}_2 > 0$). The marginal cost also increases with the current number of searches ($\hat{\gamma}_3 > 0$), implying that the search cost is convex in the number of searches. On average, consumers incur a total search cost of CN¥3,067. Given that the average consumer visits 6.7 properties in total, this corresponds to an average search cost of CN¥457 per visit.

Table 3: Estimates of Dynamic Parameters

		Est	SE
SD of match value	(σ_v)	0.442***	(0.027)
Mean waiting cost	(w)	0.047**	(0.019)
SD of waiting cost shock	(σ_w)	0.043*	(0.022)
Search cost: (0.001)			
const.	(γ_0)	5.242***	(0.110)
n	(γ_1)	3.146***	(0.113)
(past searches)$\times n$	(γ_2)	3.372***	(0.156)
n^2	(γ_3)	0.150***	(0.014)
scale parameter	(κ)	6.280***	(0.093)

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

5.3 Model Fit

To assess how well our estimated model fits the data, we simulate each consumer's decision "path", which describes her search duration, weekly number of visits, and the property she purchases. We simulate 50 such paths for each consumer. Details on the simulation are provided in Supplemental Appendix SD.

Table 4 reports the summary statistics for the observed and simulated search and purchase outcomes. The first two columns summarize the observed data, with the summary statistics taken across consumers, while the last two columns summarize the simulated results, with the summary statistics taken across both consumers and simulations. The rows correspond to the summary statistics for 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.[14]

From Table 4, we can see that the estimated model fits the data fairly well. For example, the simulated average search duration and total visits are 3.5 weeks and 6.7 visits, respectively, while their observed counterparts are 3.4 weeks and 6.7 visits. Similarly, the simulated recall share is 14.7%, compared to the observed recall share of 15.5%. The mean and standard deviation of the utility of the purchased property are (2.852, 1.805) in the simulation and (2.844, 1.945) in the data.

Table 4: Model Fit: Summary Statistics – Data vs. Simulation

	Data		Model Simulation	
	Mean	SD	Mean	SD
Search duration (weeks)	3.448	3.900	3.516	3.001
Total visits	6.710	5.116	6.700	4.270
Recall share	0.155		0.147	
Utility of the purchased property	2.844	1.945	2.852	1.805

Our estimated model also successfully captures the dynamic patterns of consumer search in the data. Figure 4 shows how the number of searches varies over search weeks according to both the data and our simulations. According to the data, consumers visit an average of about 4 properties in the first week, with a sharp decline in the second week and a continued decline over time.[15] Our simulation based on the estimated model tracks this dynamic pattern well.

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

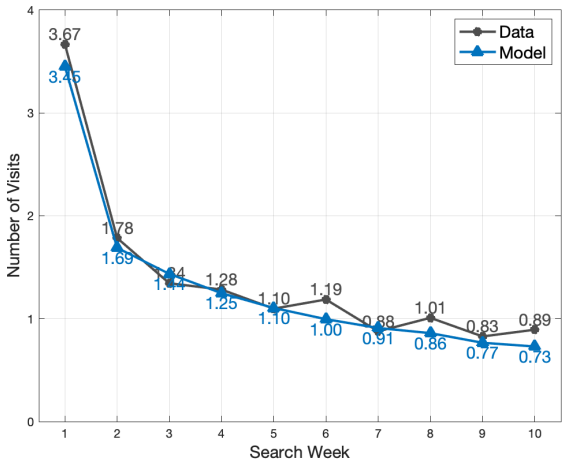
6 Comparison to a Static Search Model

In this section, we estimate a static search model and show that ignoring the dynamics of the search environment leads to unreasonably large search cost estimates. Specifically, we estimate a static simultaneous search model and compare the estimation results to those based on our dynamic model. Note that a model ignoring search environment dynamics typically endogenizes the total number of searches and the property purchased, but

[14]For the summary statistics of the utility of the purchased property, we report the average of $\delta_{j(i)}$ across consumers in the first column, where $j(i)$ indicates the property purchased by consumer i in the data. We report the standard deviation of $\delta_{j(i)}+v_{ij(i)}$, which is the square root of $var(\delta_{j(i)})+\hat{\sigma}_v^2$, in the second column. We report the mean and standard deviation of $\delta_{j^r(i)}+v_{ij^r(i)}$, where $j^r(i)$ indicates the purchased property of consumer i in a simulation indexed by r, in the third and fourth columns.

[15]The average number of visits in the n^{th} week is calculated over consumers who search for n weeks or longer.

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



ignores the number of searches in each period and the timing of purchases. There are two types of models with these features in the literature: the simultaneous search model and the sequential search model. In a simultaneous search model, a consumer searches a set of properties all 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 does not recall (unless she visits all properties). Since more than 15% of the consumers in our data recall, the simultaneous search model provides a more appropriate comparison for our study.

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 per period and the duration of the search. Second, the deterministic component of the search cost in the static model is $C(n) = \gamma_1 n + \gamma_2 n^2$, which omits the constant term $\gamma_0 \mathbb{1}(n > 0)$. This is because all consumers in the sample choose $n > 0$, where n in the static model represents the total number of visits, rather than the number of visits per period. This specification of the search cost also omits the covariate m_{it}, the cumulative number of searches prior to time t, which is undefined in the static model.

The timing of the static model is the same as in the stage game of our dynamic model. First, consumer i observes the mean utilities of properties in consumer i's search market

$\{\delta_j\}_{j\in\mathcal{A}_i}$ and the search cost shocks $(\vartheta_{in}, n = 0, ..., \bar{n})$, and decides on the number of visits n_i. Then, a search set $\mathcal{N}_i \in \mathcal{C}^{n_i}(\mathcal{A}_i)$ is sampled according to the probability given in (5) 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 values for properties in the searched set, i.e., $\{v_{ij}\}_{j\in\mathcal{N}_i}$. Finally, given $\{\delta_j\}_{j\in\mathcal{A}_i}$ and $\{v_{ij}\}_{j\in\mathcal{N}_i}$, consumer i decides which property to purchase.

We estimate the search cost parameters $(\gamma_1, \gamma_2, \kappa)$ and the standard deviation of match values (σ_v) using maximum likelihood estimation. Let θ represent these parameters. The likelihood of observing that consumer i searches n_i properties and purchases property j out of her searched properties in $\mathcal{N}_i$ is given by:

$$l_i(\theta) = \Pr(n_i|\{\delta_j\}_{j\in\mathcal{A}_i};\theta) \cdot \Pr(j|\{\delta_j\}_{j\in\mathcal{N}_i};\theta), \tag{19}$$

where the probability that consumer i purchases property j is

$$\Pr(j|\{\delta_j\}_{j\in\mathcal{N}_i};\theta) = \Pr(\delta_j + v_{ij} \geq \max_{j'\in\mathcal{N}_i} [\delta_{j'} + v_{ij'}]).$$

The probability that consumer i searches n_i properties is similar to that in (15):

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

where the expected benefit of searching n properties is

$$EB(n|\{\delta_j\}_{j\in\mathcal{A}_i};\theta) = \sum_{\mathcal{N}\in C^n(\mathcal{A}_i)} E_{\{v_{ij}\}_{j\in\mathcal{N}}}\left\{\max_{j\in\mathcal{N}} [\delta_j + v_{ij}]\right\} \times \Pr(\mathcal{N}|\mathcal{A}_i, n).$$

6.2 Estimation Results Based on the Static Search Model

Table 5 reports the estimation results for the static search model. We have two findings from comparing these results with those from our dynamic model.

First, the static model yields a smaller estimate of the standard deviation of the match value, implying a smaller benefit from searching for a given set of available properties. The estimated standard deviation of match values from the static model is 0.235 compared to 0.442 in the dynamic model. These estimates correspond to values of CN¥51,546 and CN¥96,966, respectively. Since greater variance in match values leads to a greater benefit from searching, the estimated static model implies a smaller benefit from searching compared to the dynamic model, for a given set of available properties.

The static model yields a smaller estimate of σ_v than the dynamic model because it does

Table 5: 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)

*** $p < 0.01$.

not exploit as much variation in the data to identify σ_v. In the static model, σ_v is identified by comparing the mean utility of the purchased property to the highest mean utility among all searched properties. In the dynamic model, the parameter σ_v is identified not only by which property is purchased (a static comparison), but also by when the purchased property is visited (a dynamic feature of the data). In particular, as explained in Section 4, a larger recall share (the share of consumers who purchase a property they previously searched for) implies a larger variance of the match value. In our estimation sample, the recall share is larger than 15%. As a result, our dynamic model yields a larger estimate of the variance of match values than the static model, which ignores this dynamic aspect of the data.

Second, and perhaps more importantly, the static model gives a much higher estimate of search costs, which we believe is unreasonably high. According to the estimated static model, a consumer pays an average search cost of CN¥296,347, almost two orders of magnitude higher than in the dynamic model (CN¥3,067). Given that a consumer in our estimation sample visits an average of 6.7 properties before purchasing, the average search cost per visit is approximately CN¥44,000, or $6,600 according to the estimated static model. We find this estimate unreasonably high.

The static model yields unreasonably high search costs because it does not account for the dynamics in the search environment. It pools properties on the market across time, ignoring the possibility of exits and the fact that some properties enter the market only later. By mistakenly considering all these properties as available for search, it overstates the benefits of searching (despite underestimating the variance of match values). As a result, the static model can explain the observed number of searches only by inflating the estimate of search costs.

7 Effect of Search Environment Dynamics and Search Costs

Having established the importance of considering dynamics in our setting, we now quantify how search environment dynamics (i.e., changes in the search environment over time) affect consumers' search and purchase decisions. The search environment can experience three types of changes over time: new listing price increases, new listing entries, and existing listing exits. We examine the effects of each change through three respective counterfactual simulations (CF1, CF2, CF3). To quantify the relative importance of search environment dynamics versus traditional search frictions such as search costs, we consider another counterfactual scenario (CF4) to assess the effect of search costs and compare these results with those in CF1–CF3.

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

7.1 Effect of Search Environment Dynamics

To examine the effect of new listing price increases, we consider a counterfactual scenario CF1 in which the rate of increase in the list price of new listings is halved. To quantify the effect of new listing entries, we consider a counterfactual scenario CF2 in which the entry rate (i.e., the arrival rate of new listings) is doubled. Finally, to quantify the effect of existing listing exits, we consider a counterfactual scenario CF3 in which the exit rate (i.e., the probability that an existing listing will exit in a week) is halved.

Table 6 reports the average of the main endogenous outcomes in these counterfactual simulations in columns (II)–(IV). For comparison, we also include the outcomes under the actual search environment dynamics in column (I).

We find that, as the price change becomes slower, consumers search longer, visit more properties before purchasing, and purchase properties that generate higher utilities. Comparing columns (I) and (II) of Table 6, we see that the average search duration increases from 3.4 at the observed price increase rate to 4.3 at half the actual price increase rate (row (1)). Similarly, the average total number of visits increases from 6.7 to 7.4 (row (2)). As consumers increase their search duration and number of properties visited, they end up purchasing a property that generates higher utility. Specifically, row (3) shows

that the average utility of the purchased property increases by 1, which is equivalent to an increase in value of CN¥210,792. Part of this increase in utility comes mechanically from a lower price. To remove such a mechanical increase in utility, thus isolating the increase in utility due to searching longer and visiting more properties, we also report in row (3') what the utility of the purchased property would be at the observed price. Under the actual price increase rate in column (I), rows (3) and (3') are identical. At half the price increase rate, the utility in row (3') is, unsurprisingly, smaller than that in row (3). However, even ignoring the mechanical increase in utility, there is an increase in utility equivalent to CN¥118,886 according to row (3'). In other words, searching longer and visiting more properties contributes 56% of the increase in utility when the price change is halved.

While a longer search and a higher number of property visits yields higher utility for a purchased property, doing so also entails higher search and waiting costs. Specifically, the sum of search and waiting costs increases on average by CN¥7,874 (rows (4) and (5)), which is about 7% of the gain from searching longer and visiting more properties (CN¥118,886, row (3')). Therefore, on balance, consumers are better off with a slower price change, even when excluding the mechanical gain from lower prices.

Table 6: Counterfactual Simulation Results

	(I) Actual	(II) CF1 Half Price Increase Rate	(III) CF2 Double Arrival Rate	(IV) CF3 Half Exit Rate	(V) CF4 Half Search Costs
(1) Search duration (week)	3.448	4.305	4.773	3.918	3.793
(2) Total visits	6.710	7.424	7.961	7.120	9.826
(3) Utility	2.844	3.805	3.275	3.049	3.269
(3') at the observed price		3.386			
(4) Search cost (CN¥)	3,067	3,672	4,145	3,413	3,080
(5) Waiting cost (CN¥)	15,460	22,729	27,164	19,067	18,003

Turning to the effect of new listing entries, we find that when consumers anticipate more new listings in the future, they also search longer, visit a greater number of properties before making a purchase, and ultimately purchase properties that generate higher utilities. Specifically, the average search duration increases from 3.4 in column (I) to 4.8 in column (III), and the average number of properties visited increases from 6.7 to 8. As a result, the consumer utility obtained from a purchased property increases by 0.4, which is equivalent to an increase in value of CN¥94,538. Meanwhile, the average search cost increases by CN¥1,078 and the average waiting cost increases by CN¥11,704. Overall, consumers are better off.

Similarly, comparing column (I) and column (IV), we see that decreasing the exit rate for existing listings leads to an increase in search duration as well as higher utility from a purchased property. Overall, the average increase in utility is equivalent to CN¥44,966, which more than offsets the average respective increases in search costs of CN¥346 and waiting costs of CN¥3,607.

In summary, while the comparison with the static search model in the previous section demonstrates the importance of considering dynamics in estimations, the counterfactual simulations in this section highlight the economic significance of search environment dynamics. We find that search environment dynamics have significant effects on consumer behavior and outcomes, mainly because they influence consumers' incentives to search longer and visit more properties.

7.2 Effect of Search Costs

We quantify the effects of search costs by simulating a counterfactual scenario where we halve the search cost function $C_{it}(n)$ in CF4. The results are reported in column (V) of Table 6.

We find that when consumers face a lower search cost per visit, they unsurprisingly extend their search duration by 0.3 weeks and visit 3.1 more properties on average. With more searches, consumers find properties that generate higher utility for them. The average increase in utility is equivalent to CN¥93,222. Interestingly, despite the reduced search cost per visit, consumers end up incurring slightly higher search costs due to more searches. However, the increased search costs (by CN¥13) and the increased waiting costs (by CN¥2,543) are dominated by the increased utility.

Table 7 summarizes and compares the average net gain per consumer to show the relative importance of search environment dynamics versus search costs. To compute the average net gain per consumer, we first compute the changes in average utility, search costs, and waiting costs from the actual scenario (column (I) in Table 6) to a counterfactual scenario (columns (II)–(V) in Table 6) and then calculate (change in utility)$/\hat{\alpha}-$(change in search cost and waiting cost). Table 7 shows that the average net gain per consumer due to a reduction in search costs by half is CN¥90,666. This compares to an average net gain of CN¥111,012 when consumers face half the price increase speed (where utility is calculated based on actual prices), CN¥81,756 when the new listing arrival rate is doubled, and CN¥41,013 when the exit rate is halved.

Overall, these results show that, while search costs affect consumers' search and purchase decisions, search environment dynamics also have significant effects on consumers'

Table 7: Average Net Gain Per Consumer

	CF1: (II)-(I)	CF2: (III)-(I)	CF3: (IV)-(I)	CF4: (V)-(I)
Net Gain (CN¥)	111,012	81,756	41,013	90,666

search and purchase behavior and consumer welfare. At least for the same percentage change (halving or doubling), varying the dynamics of the search environment, especially in terms of the speed of price increase and the rate of arrival of new listings, has a comparable impact as varying search costs.

8 Effect of Housing Supply Policies

Many cities around the world, particularly those experiencing a housing crisis, have implemented policies to encourage new listings in order to increase the supply of housing and address their housing shortages. For example, Melbourne's Vacant Residential Property Tax, Oakland's Vacant Property Tax, Toronto's Vacant Home Tax, and Vancouver's Empty Homes Tax all aim to increase the supply of housing. These policies are likely to both increase the number of new listings and slow down price increases. In other words, they influence the search environment dynamics. In this section, we simulate the effects of such policies by simulating the outcomes under different new listing arrival rates and price trends. Specifically, we scale up the arrival rate of new listings by $\phi_1 \geq 1$ and scale down the weekly increase in the list price of new listings by $\phi_2 \leq 1$.

Figure 5 presents the simulation results in terms of average search duration and number of property visits, showing that as ϕ_1 increases and ϕ_2 decreases, consumers on average search longer and visit more properties before purchasing. Specifically, panel (a) shows that when the price increase trend is reduced from the original rate ($\phi_1 = 1$ and $\phi_2 = 1$) to half the original rate ($\phi_1 = 1$ and $\phi_2 = 0.5$), the average search duration increases from 3.52 to 4.30 (top row of panel (a)). Similarly, when the new listing arrival rate is doubled, i.e., from ($\phi_1 = 1$ and $\phi_2 = 1$) to ($\phi_1 = 2$ and $\phi_2 = 1$), the average search duration increases to 4.77 (rightmost column of panel (a)). A combination of an increase in the new listing arrival rate and a decrease in the price trend ($\phi_1 = 2$ and $\phi_2 = 0.5$) can increase the average search duration to 5.91. Moreover, these two changes reinforce each other: increasing the new listing arrival rate has a stronger effect when the price increase is slower, and conversely, reducing the speed of the price increase has a stronger effect when the new listing arrival rate is higher. For example, as ϕ_1 increases, the average

search duration increases by 1.25 weeks (from 3.52 to 4.77) when $\phi_2 = 1$ and by 1.61 weeks (from 4.30 to 5.91) when $\phi_2 = 0.5$. Regarding the total number of visits, panel (b) shows that as ϕ_1 increases and ϕ_2 decreases, consumers visit more properties before purchasing one. This is largely because they search for a longer period of time (panel (a)), rather than visiting more properties per period. In fact, the average number of visits per week is rather stable, varying between 2.07 and 2.44.

Figure 5: Effects of Housing Supply Policies: Average Search Duration and Total Visits

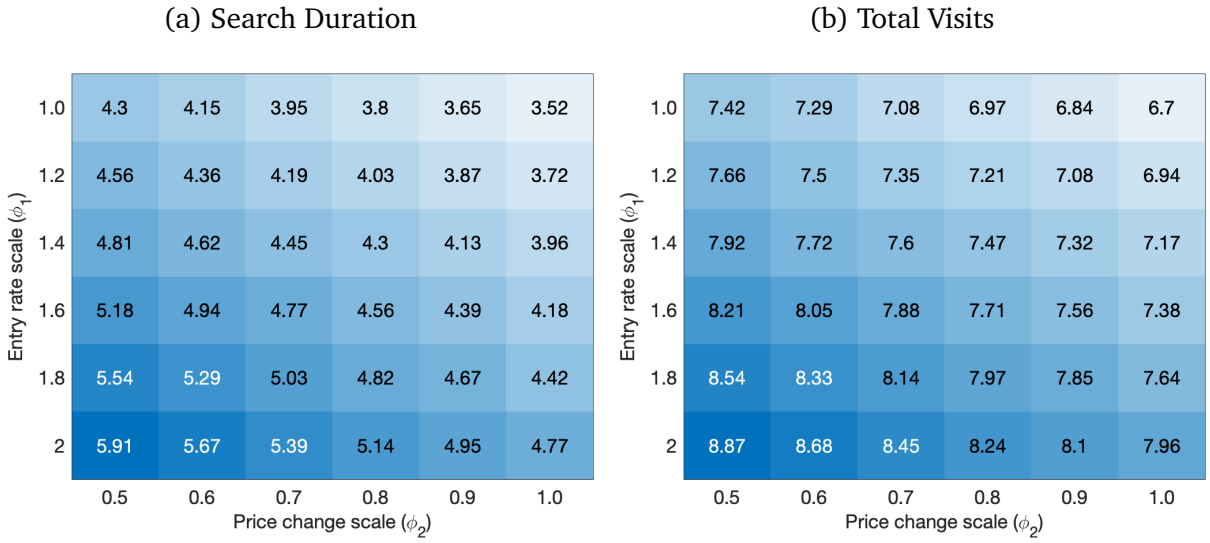


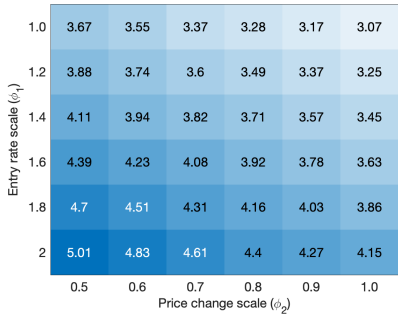
Figure 6 shows that, as ϕ_1 increases and ϕ_2 decreases, the increased search duration and number of visits lead to higher search costs (panel (a)) and waiting costs (panel (b)) but better purchase outcomes (panel (c)). In particular, the average search cost increases from CN¥3,070 to CN¥5,010 and the average waiting cost increases from CN¥15,460 to CN¥38,520 when the new listing arrival rate is doubled and the price trend is halved. At the same time, consumers end up purchasing properties that generate higher utility.

Examining the net effects, panel (a) of Figure 7 shows a positive net effect, with the gains from searching longer and visiting more properties outweighing the increases in search and waiting costs. If the arrival rate is doubled and the price increase is halved, the average net gain per consumer can be as high as CN¥270,800, or about 7% of the average transaction price.

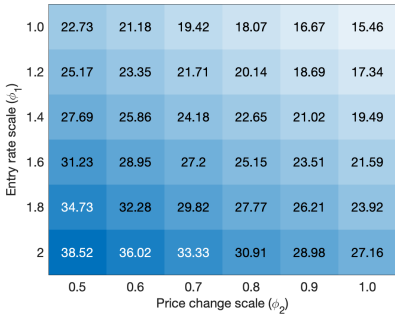
There are potentially three channels through which consumers benefit from such policies. First, consumers are mechanically better off because there are more properties for them to choose from. Second, consumers benefit from lower prices. Third, consumers are better off because their search behavior changes. In our setting, since consumers choose only from properties they have visited, the mere presence of more listings does not nec-

Figure 6: Effects of Housing Supply Policies: Average Search Costs, Waiting Costs, and Utility

(a) Search Costs (1,000 CN¥)



(b) Waiting Costs (1,000 CN¥)



(c) Utility from Purchased Property

essarily increase consumer welfare without a corresponding change in search behavior. However, both the second and third channels are present in our context. To assess their relative importance, we recompute utility using the prices observed in the data,[16] as shown in panel (b) of Figure 7.

Comparing panels (a) and (b) of Figure 7, we see that consumer gains come mainly from the third channel, i.e., from searching longer and visiting more properties. For example, when $(\phi_1 = 2, \phi_2 = 0.5)$, the average net gain per consumer in panel (b) is CN¥194,600, which accounts for 72% of the total net gain in panel (a).

Overall, these simulations suggest that housing supply policies can significantly benefit consumers if they increase the supply of new listings and slow down price growth. This impact is mainly due to their influence on the dynamics of the search environment, which incentivizes consumers to search longer, visit more properties, and ultimately find properties that generate higher utility.

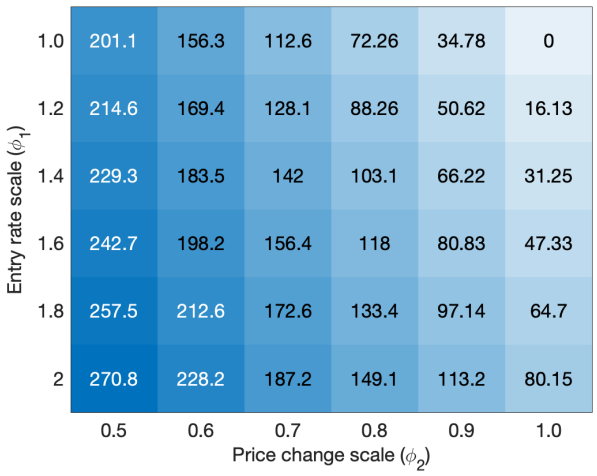
9 Conclusion

In this paper, we study how dynamics in the search environment affect consumers' search and purchase decisions and welfare. We present a dynamic model in which consumers make search and purchase decisions in a non-stationary search environment. We develop a feasible estimation routine to estimate our model and apply it to the Beijing

[16]For a simulated property j that does not appear in the data (for example, in a counterfactual simulation with a higher arrival rate of new listings), we adjust its price by subtracting $(1-\phi_2)\rho_i(\text{EntryWeek}_j - \text{SampleStartWeek})$ from its simulated price, where $(1-\phi_2)\rho_i$ represents the difference in the weekly increase in the list price between the data and the counterfactual scenario.

Figure 7: Effects of Housing Supply Policies: Average Net Gain Per Consumer

(a) Net Gains (1,000 CN¥)



(b) Net Gains at Observed Prices (1,000 CN¥)

housing market between 2015 and 2016. We find that search environment dynamics have a significant effect on consumers' search and purchase decisions and welfare. We also find that a static search model would yield an unreasonably high estimate of search costs. Finally, we find that housing supply policies that increase the number of new listings and slow down price increases benefit consumers, primarily because these changes in the search environment dynamics lead to longer search durations and more property visits.

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

References

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.

Arcidiacono, Peter, Attila Gyetvai, Arnaud Maurel, and Ekaterina S Jardim (2022) "Identification and Estimation of Continuous-Time Job Search Models with Preference Shocks," Working Paper 30655, National Bureau of Economic Research.

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, Yale University.

Brown, Zach Y. and Jihye Jeon (2024) "Endogenous Information and Simplifying Insurance Choice," *Econometrica*, 92 (3), 881–911.

Chatterjee, Chirantan, Ying Fan, and Debi Prasad Mohapatra (2024) "Spillover Effects in Complementary Markets: A Study of the Indian Cell Phone and Wireless Service Markets," working paper.

Crawford, Gregory S and Matthew Shum (2005) "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73 (4), 1137–1173.

Eckstein, Zvi and Gerard J Van den Berg (2007) "Empirical Labor Search: A Survey," *Journal of Econometrics*, 136 (2), 531–564.

Eizenberg, Alon (2014) "Upstream Innovation and Product Variety in the US Home PC Market," *Review of Economic Studies*, 81 (3), 1003–1045.

Elliott, Jonathan T (2024) "Investment, Emissions, and Reliability in Electricity Markets," working paper, Johns Hopkins University.

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.

——— (2024) "Estimating Discrete Games with Many Firms and Many Decisions: An Application to Merger and Product Variety," *Journal of Political Economy*, forthcoming.

French, Eric and Christopher Taber (2011) "Identification of Models of the Labor Market," in *Handbook of Labor Economics*, 4, 537–617: Elsevier.

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

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

Habib, Michel Antoine, Yushi Peng, Yanjie Wang, and Zexi Wang (2023) "The Implementation of Central Bank Policy in China: The Roles of Commercial Bank Ownership and CEO Faction Membership," *Swiss Finance Institute Research Paper* (23-14).

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

Hodgson, Charles (2024) "Information Externalities, Free Riding, and Optimal Exploration in the UK Oil Industry," Working Paper 33067, National Bureau of Economic 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 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.

Keane, Michael P and Kenneth I Wolpin (1994) "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 648–672.

Kim, Jun B, Paulo Albuquerque, and Bart J Bronnenberg (2010) "Online Demand Under Limited Consumer Search," *Marketing Science*, 29 (6), 1001–1023.

Merlo, Antonio, François Ortalo-Magné, and John Rust (2015) "The Home Selling Problem: Theory and Evidence," *International Economic Review*, 56 (2), 457–484.

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

——— (2024) "An Empirical Model of Consideration through Search," working paper.

Murry, Charles and Yiyi Zhou (2020) "Consumer Search and Automobile Dealer Colocation," *Management Science*, 66 (5), 1909–1934.

Peng, Yushi (2023) "Mortgage Credit and Housing Markets," *Available at SSRN 4537500*.

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.

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.

Appendices

A Micro Foundation for Sampling Probabilities

We have three goals in this section. First, we show that the sampling probability $\Pr(\mathcal{N}|\mathcal{A}_{it}, n)$ in (5) 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 sampling probability implies a probability for each option j, i.e., $\Pr(j|\mathcal{A}_{it}, n)$ in (6). Third, we show that there is a unique vector of mean utilities that solve the search share equations in (13). In fact, we show that the mapping used to invert out the mean utilities in Berry, Levinsohn and Pakes (1995) is a contraction mapping even in our extension.

In Sections A.1 and A.2, which correspond to the first two goals, we suppress the subscripts i and t for simplicity.

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

In this section, we show that the sampling probability $\Pr(\mathcal{N}|\mathcal{A}, n)$ in (5) is consistent with an extension of a Logit model from choosing a single option to choosing $n \geq 1$

options.[17] In this probability, $\mathcal{A}$ represents the set of all options and $\mathcal{N} \subseteq \mathcal{A}$ represents a subset of $\mathcal{A}$ of size n. Let $\mathcal{C}^n(\mathcal{A})$ be the collection of all subsets of $\mathcal{A}$ of size n.

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 $\min_{j \in \mathcal{N}}(\delta_j + \epsilon_j) \geq \max_{h \in \mathcal{A} \setminus \mathcal{N}}(\delta_h + \epsilon_h)$. Therefore,

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

In Supplemental Appendix SE, we show that $\Pr(\mathcal{N}|\mathcal{A}, n)$ in (5) can be derived based on the analytic expression for the choice probability in a standard Logit model and the inclusion-exclusion principle.

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

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

The sampling probability $\Pr(\mathcal{N}|\mathcal{A}, n)$ implies the probability that a particular single option is sampled: $\Pr(j|\mathcal{A}, n) = \sum_{\mathcal{N} \in \mathcal{C}^n(\mathcal{A}): j \in \mathcal{N}} \Pr(\mathcal{N}|\mathcal{A}, n)$. When $n = 0$, $\Pr(j|\mathcal{A}, n) = 0$. When $n = A = \#\mathcal{A}$, $\Pr(j|\mathcal{A}, n) = 1$. When $0 < n < A$, we show in Supplemental Appendix SE that:

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

The intuition behind this analytic expression is based on two key points. First, option j is sampled if its rank, determined by the value $\delta_j + \epsilon_j$, is within the top n among all options in $\mathcal{A}$. In other words, $\Pr(j|\mathcal{A}, n) = \sum_{k=0}^{n-1} \Pr(j\text{'s rank is } n-k)$. Second, the probability that j has a particular rank is tied to the probability that j is the best among a subset of options. Correspondingly, in (A.1), the term $\frac{exp(\delta_j)}{\sum_{l \in \mathcal{B}} exp(\delta_l) + exp(\delta_j)}$ represents the probability that j is the best in the subset $\mathcal{B} \cup \{j\}$. In Supplemental Appendix SE, we use a simple example to illustrate these two points and derive the proof for (A.1).

[17] An alternative expression of this probability is given in Moraga-Gonzalez, Sándor and Wildenbeest (2024), which can be derived from a model where a consumer chooses a sampled set $\mathcal{N}$ to maximize "(Inclusive Value of Set $\mathcal{N}$) − (Logit Error)$_{i\mathcal{N}}$", where the Logit error term is i.i.d. and varies at the set level. We choose our model because we think Logit errors corresponding to overlapping sets are unlikely to be independent and those corresponding to sets of different sizes are unlikely to be identically distributed.

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

We now show the three intuitive 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** $k \neq j$**.**

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

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

When $n = 1$, $\Pr(j|\mathcal{A}, n)$ in (A.1) becomes

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

because $\mathcal{C}^{A-1}(\mathcal{A} \setminus j)$ has a singleton element, which is $\mathcal{A} \setminus j$.

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, $\sum_{j \in \mathcal{A}} \Pr(j|\mathcal{A}, A) = A = n$. We now consider the case where $n < A$.

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

Let $\mathcal{D}$ represent $\mathcal{B} \cup \{j\}$. Since $\mathcal{B} \in \mathcal{C}^{A-n+k}(\mathcal{A} \setminus j)$ and $\{j\}$ do not intersect, we can rewrite the above equation as

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

which can be further simplified as follow:

$$\begin{aligned}\sum_{j\in\mathcal{A}}\Pr(j|\mathcal{A},n) &= \sum_{k=0}^{n-1}(-1)^k\left[\sum_{\mathcal{D}\in\mathcal{C}^{A-n+k+1}(\mathcal{A})}C_{A-n+k-1}^{k}\right] \\ &= \sum_{k=0}^{n-1}(-1)^k C_A^{A-n+k+1}C_{A-n+k-1}^{k} \\ &= n.\end{aligned} \tag{A.2}$$

We provide a proof of the last equality, i.e., $\sum_{k=0}^{n-1}(-1)^k C_A^{A-n+k+1}C_{A-n+k-1}^{k}=n$, in Supplemental Appendix SE.

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

In this section, we show that, under certain regularity conditions, the system of equations $(\tilde{s}_j(\boldsymbol{\delta}_d)=s_j, j\in\mathcal{J}_d)$ has a unique solution $\boldsymbol{\delta}_d=(\delta_j, j\in\mathcal{J}_d)$. Recall that the search share function is:

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

The regularity conditions are:

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

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 ($s_j<\frac{\sum_{i\in\mathcal{I}_d}\sum_{t=1}^{T_i}\mathbb{1}(j\in\mathcal{A}_{it})}{\sum_{i\in\mathcal{I}_d}\sum_{t=1}^{T_i}n_{it}}$). This condition is an extension of the condition of $0<s_j<1$ in a single discrete choice model. Conditions 2 and 3 impose constraints on n_{it} in the data. They imply that a consumer always visits properties with a mean utility of ∞ and never visits properties with a mean utility of $-\infty$.

In what follows, we suppress the subscript d for simplicity. Moreover, because we normalize the mean utility of the first property in a district d to 0, we use $\mathcal{J}$ to represent $\mathcal{J}_d\backslash\{1\}$, and use J to denote the cardinality of $\mathcal{J}$. Following Berry, Levinsohn and Pakes (1995), we define a mapping $f: R^J\to R^J$ as

$$f_j(\boldsymbol{\delta})=\delta_j+\ln s_j-\ln\tilde{s}_j(\boldsymbol{\delta}) \tag{A.3}$$

for $j \in \mathcal{J}$. We show that this mapping is a contraction mapping in Supplemental Appendix SE. Therefore, its fixed point is the unique solution to the system of equations $\tilde{s}_j(\boldsymbol{\delta}) = s_j, j \in \mathcal{J}$.

B Discussion of the Two-Step Estimation Procedure

In our two-step estimation procedure, we first back out the mean utilities and estimate the static parameters in the utility function and then estimate the dynamic parameters.

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 so because we assume that, while consumers decide how many properties to search in each period, the specific set of properties they visit is exogenous. A similar exogeneity assumption is made in, for example, Hortaçsu and Syverson (2004). We extend their specification of the sampling probability for a single product to the sampling probability for a set of products. According to our sampling probability specification, properties with higher mean utilities have a higher probability of being sampled. Thus, the exogeneity assumption simply means that a consumer and her agent do not have full control over the set of properties they can visit in a given period. Instead, 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. We believe that this exogeneity assumption is justified in our setting.

The advantage of the two-step estimation procedure is threefold. First, it accommodates unobservable property heterogeneity (i.e., ξ_j in the utility function) and price endogeneity. If we were to estimate the utility parameters along with the dynamic parameters using maximum likelihood estimation, we would either have to assume no unobservable property heterogeneity (and no price endogeneity) or we would have to model how prices correlate with unobservable property heterogeneity in order to control for price endogeneity.[18] In contrast, in our two-step procedure, ξ_j is absorbed in δ_j, which is backed out in the first step of the estimation and treated as an observable in the dynamic estimation.

Second, this procedure allows us to include a large set of neighborhood fixed effects in the utility function without significantly increasing the computational burden. This is because these parameters are estimated in the first stage without solving the dynamic model. While we have data on a number of observable property characteristics, collecting detailed data on neighborhood amenities is difficult. Therefore, it is important to include

[18]This is because consumers' search and purchase decisions depend on the unobservable ξ_j's, so the likelihood function would require an integral over the distribution of ξ_j's *conditional on the observed prices*, necessitating a model to describe this conditional distribution.

neighborhood fixed effects to control for such property heterogeneity at the neighborhood level.

Third, this procedure helps us address the challenge of high dimensionality in the state variables. In our model, the state variables vector includes characteristics of all unsearched properties, as well as the characteristics and match values of all searched properties, resulting in a high-dimensional state space. The dynamic demand literature (e.g., Gowrisankaran and Rysman, 2012) tackles this by reducing the state space 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 variables describe the choice set and the choice set evolves exogenously, in our model with both dynamic search and dynamic demand, the state variables describe two sets (the searched and unsearched sets), both of which evolve endogenously. Thus, it would be inappropriate to assume that the inclusive value of each set in our model evolves according to a stationary Markov process. We address the issue of high dimensionality by estimating the mean utilities of properties "offline" (i.e., before estimating the dynamic parameters) and replacing a vector of property characteristics by a scalar mean utility for each property. We further reduce the state space by considering state variables to include the mean utilities of the top properties, the average mean utility of the remaining properties, and the number of properties.

SA Additional Figures

Figure SA.1 plots the number of new listings, the number of transactions, the average list price, and the average transaction price by week, including the eight weeks around the Chinese New Year. The figure shows that, over the course of these eight weeks, both new listings and transactions initially fall to zero and then quickly rise to around double their pre-holiday levels.

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

(a) New Listings and Transactions

(b) List Prices and Transaction Prices

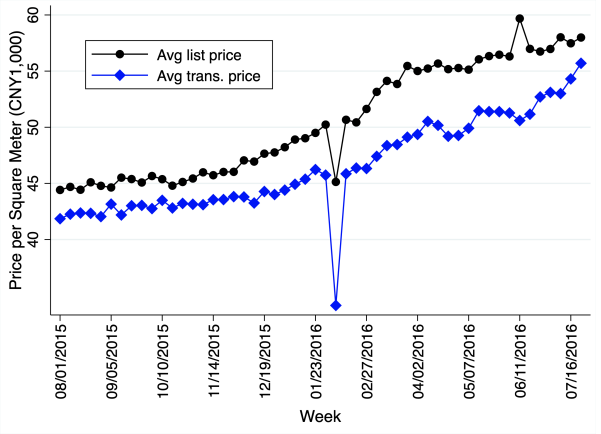
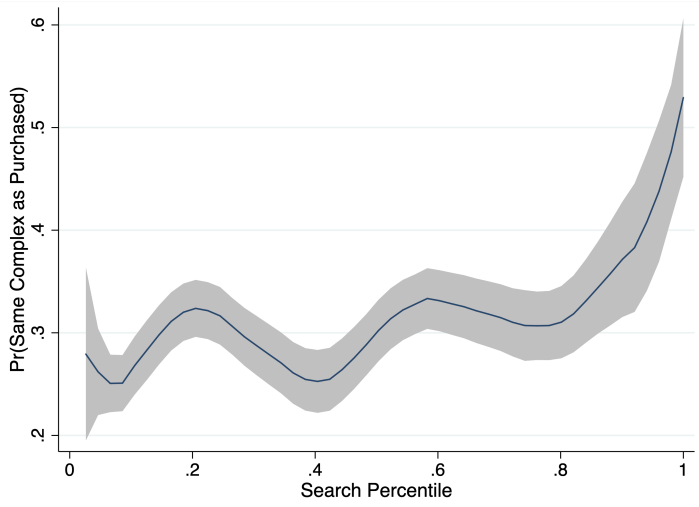


Figure SA.2 plots the probability that a consumer visits a property within the same residential complex as her final purchase against the search percentile. To generate this plot, we first define two variables $\mathbb{1}(\text{Same Complex as Purchased})_{ij}$ and $(\text{Search Percentile})_{ij}$ for each (i, j) combination where consumer i searches property j in the data. The dummy variable $\mathbb{1}(\text{Same Complex as Purchased})_{ij}$ takes the value of 1 if property j is in the same residential complex as the property that consumer i purchases, and 0 otherwise. The variable $(\text{Search Percentile})_{ij}$ is defined as follows. Suppose consumer i searches for T_i weeks, visits $(n_{i1}, n_{i2}, ..., n_{iT_i})$ properties in each week, and visits property j in search week t. Then the search percentile of property j for consumer i is $(\text{Search Percentile})_{ij} = (\sum_{\tau \leq t} n_{i\tau})/(\sum_{\tau \leq T_i} n_{i\tau})$.

We then run a kernel regression of $\mathbb{1}(\text{Same Complex as Purchased})_{ij}$ on $(\text{Search Percentile})_{ij}$ and plot the smoothed values in the solid line and the 95% confidence interval using the shaded band in Figure SA.2. In this regression, we exclude the purchased property for each consumer as well as properties visited by the consumer in a week when there is no available property in the same complex as the purchased property.

This figure supports our assumption of no learning. If a consumer were able to learn about a property by visiting another property in the same area, her search would become more concentrated in the area where her final purchase is located. However, Figure SA.2 shows that the probability of a searched property being in the same complex does not monotonically increase. In fact, this probability fluctuates as the consumer approaches the end of her search. For instance, the probability is roughly the same at search percentiles of 20%, 60%, and 85%. Additionally, this probability remains below 40% even at the 90% search percentile.

Figure SA.2: How $\Pr$(Same Complex as Purchased) Varies During the Search Process



SB Estimation Details

SB.1 Details on Each Consumer's Search Environment

Consumer i's search environment is characterized by the new listing arrival rate λ_i, the listing exit rate χ_i, the trend in the list price of new listings ρ_i, as well as the mean and standard deviation of a new listing's mean utility $(\mu_{it}^{new}, \sigma_i^{new})$.

To define $(\lambda_i, \chi_i, \rho_i, \mu_{it}^{new}, \sigma_i^{new})$, we first define them at the segment level and then take a weighted average across segments within consumer i's search market, where the weight is consumer i's search share in each segment. We define the segment-level new listing arrival rate λ_m as the average number of new listings in segment m per week. Similarly, we define the segment-level exit rate χ_m as the average ratio of the number of exits to the

number of listings, averaged across weeks.[1] To define ρ_m, we first calculate the average list price of new listings in segment m in each week and then define ρ_m as the linear trend in this average price. In other words, ρ_m is the average weekly change in segment m. To define μ_{mt}^{new}, we first calculate the average mean utility of new listings in segment m at the beginning of the sample and then define μ_{mt}^{new} as the sum of this average and $\alpha\rho_m(t - \text{SampleStartWeek})$, where, with a slight abuse of notation, t here represents a calendar week (instead of a search week). Finally, the standard deviation σ_m^{new} is the standard deviation of mean utilities of all new listings in segment m.

SB.2 Value Function Approximation

The vector of state variables in the value function $V_i(\Omega_{it};\theta)$, i.e., $\Omega_{it} = (\{\delta_j\}_{j\in\mathcal{A}_{it}}, \{\delta_j, v_{ij}\}_{j\in\mathcal{R}_{it}}, \mu_{it}^{new}, m_{it})$, is high dimensional. In addition, its dimension changes over time both exogenously and endogenously. To address this dimensionality issue, we approximate the value function $V_i(\Omega_{it};\theta)$ by a function of fewer state variables. Specifically, we first define a set of reduced state variables $\tilde{\Omega}_{it}$, then solve the value function at a set of randomly-drawn grid points for $\tilde{\Omega}_{it}$, and finally interpolate the value function at other points with a polynomial of $\tilde{\Omega}_{it}$ (see, for example, Keane and Wolpin (1994), Crawford and Shum (2005), and Sweeting (2013).)

We define the reduced state variables $\tilde{\Omega}_{it}$ as follows. First, we define $u_{it}^* = \max_{j\in\mathcal{R}_{it}}(\delta_j + v_{ij})$ and replace $\{\delta_j, v_{ij}\}_{j\in\mathcal{R}_{it}}$ by u_{it}^*. 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 $\{\delta_j\}_{j\in\mathcal{A}_{it}}$ by $\{\delta_{it}^{(1)}, ..., \delta_{it}^{(K)}, \bar{\delta}_{it}^{(K+1)}, A_{it}\}$, where $(\delta_{it}^{(1)}, ..., \delta_{it}^{(K)})$ represents the highest K mean utilities among $\{\delta_j\}_{j\in\mathcal{A}_{it}}$; $\bar{\delta}_{it}^{(K+1)}$ denotes the average of δ_j 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 $K = 5$. In the end, the vector of reduced state variables is $\tilde{\Omega}_{it} = \{\delta_{it}^{(1)}, ..., \delta_{it}^{(K)}, \bar{\delta}_{it}^{(K+1)}, A_{it}, u_{it}^*, \mu_{it}^{new}, m_{it}\}$.

SC Robustness

In this section, we report the results from two robustness analyses.

In the first robustness analysis, we allow the price coefficient to differ for consumers who search in different districts. Specifically, in Table SC.1, we report the results from

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

regressing the mean utility of a property on its characteristics, where the price coefficient is district specific. The regression results show some heterogeneity in the price coefficient. The parameters common in this robustness specification (Table SC.1) and the baseline specification (Table 2) are close.

Table SC.1: Estimates of Parameters in Mean Utility: District-Specific Price Coefficient

	(I) OLS		(II) IV	
	Est	SE	Est	SE
Expected price (million CN¥)				
Dongcheng	-0.275***	(0.015)	-5.155***	(0.217)
Xicheng	-0.189***	(0.015)	-3.806***	(0.151)
Chaoyang	-0.220***	(0.009)	-5.195***	(0.136)
Haidian	-0.523***	(0.006)	-3.312***	(0.161)
Fengtai	-0.445***	(0.014)	-4.634***	(0.127)
Shijingshan	-0.461***	(0.026)	-4.983***	(0.153)
Property age (year)	-0.005***	(0.001)	-0.016***	(0.002)
# Bedrooms	0.258***	(0.007)	0.364***	(0.016)
# Living rooms	0.378***	(0.008)	1.094***	(0.022)
Property size (m^2)	0.005***	(0.000)	0.050***	(0.001)
Above 10th floor	0.087***	(0.007)	0.160***	(0.017)
Close to a subway station	0.118***	(0.010)	0.227***	(0.024)
Neighborhood FE	yes		yes	

*** $p < 0.01$.

In the second robustness analysis, we no longer restrict the sample to consumers who purchase a property before the end of the sample in estimation. Specifically, we now include consumers who do not make a purchase during the sample period but visit at least one property during the last four weeks before the end of the sample period.[2] Table SC.2 reports the estimation result using this larger sample. The comparison to Table 3 shows that our main results are robust to including these non-buyers.

[2]We add this restriction to rule out inactive consumers.

Table SC.2: Estimates of Dynamic Parameters: Using a Sample Including Non-Buyers

		Est	SE
SD of match value	(σ_v)	0.438***	(0.009)
Mean waiting cost	(w)	0.057***	(0.004)
SD of waiting cost shock	(σ_w)	0.040***	(0.010)
Search cost: (0.001)			
const.	(γ_0)	5.285***	(0.240)
n	(γ_1)	3.270***	(0.135)
(past searches)$\times n$	(γ_2)	3.384***	(0.099)
n^2	(γ_3)	0.143***	(0.015)
scale parameter	(κ)	6.240***	(0.087)

*** $p < 0.01$.

SD Simulation Details

SD.1 Model Fit Simulation

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

At $t = 1$, the recall set is $\mathcal{R}_{i1} = \emptyset$. Therefore, the vector of state variables is $\Omega_{i1} = (\{\delta_j\}_{j\in\mathcal{A}_{i1}}, \mu_{i1}^{new}, m_{i1})$, where $\mathcal{A}_{i1}$ is the observed available-for-search set and the number of past searches m_{i1} is 0. We simulate the number of searches $n_{i1}^{(r)}$ and the purchase decision $y_{i1}^{(r)}$ in the following three steps: (i) we compute $\Pr_i(n|\Omega_{i1};\hat{\theta})$ according to (15) for $n = 0, ..., \bar{n}$ and simulate $n_{i1}^{(r)}$ based on these probabilities; (ii) we draw $\mathcal{N}_{i1}^{(r)}$ based on $\Pr(\mathcal{N}|\mathcal{A}_{i1}, n_{i1}^{(r)})$ in (5); and (iii) we compute the probability of purchasing a newly-searched property $\Pr_i(y_{i1} = j \in \mathcal{N}_{i1}^{(r)}|\Omega_{i1}, \{v_{ij}^{(r)}\}_{j\in\mathcal{N}_{i1}^{(r)}};\hat{\theta})$ according (17) and the probability of waiting $\Pr_i(y_{i1} = wait|\Omega_{i1}, \{v_{ij}^{(r)}\}_{j\in\mathcal{N}_{i1}^{(r)}};\hat{\theta})$ according to (18) and simulate $y_{i1}^{(r)}$ based on these probabilities. If $y_{i1}^{(r)} \neq wait$, the path ends. Otherwise, we continue to $t = 2$.

At $t = 2$, the recall set is updated to $\mathcal{R}_{i2}^{(r)} = \mathcal{R}_{i1} \cup \mathcal{N}_{i1}^{(r)} \backslash \mathcal{EXIT}_{i1}$, the available-for-search set is updated to $\mathcal{A}_{i2}^{(r)} = \mathcal{A}_{i1} \backslash \mathcal{N}_{i1}^{(r)} \backslash \mathcal{EXIT}_{i1} \cup \mathcal{NEW}_{i2}$, and the number of past searches is updated to $m_{i2}^{(r)} = n_{i1}^{(r)}$. The state variables are $\Omega_{i2}^{(r)} = (\{\delta_j\}_{j\in\mathcal{A}_{i2}^{(r)}}, \{\delta_j, v_{ij}^{(r)}\}_{j\in\mathcal{R}_{i2}^{(r)}}, \mu_{i2}^{new}, m_{i2}^{(r)})$. We simulate $(n_{i2}^{(r)}, y_{i2}^{(r)})$ following the same steps as for $t = 1$ except that, in step (iii), we

additionally consider the probability of recall according to (16). If $y_{i2}^{(r)} \neq wait$, the path ends. Otherwise, we continue to $t = 3$ and repeat the process until $y_{it}^{(r)} \neq wait$.

SD.2 Counterfactual Simulations

The simulation procedure for the counterfactual simulations is similar to that for the model-fit simulation, with three differences.

First, instead of using the estimated value function, we solve for the value function in a counterfactual scenario and compute $\Pr_i(n|\Omega_{it}^{(r)};\hat{\theta})$ and $\Pr_i(y_{it}|\Omega_{it}^{(r)},\{v_{ij}^{(r)}\}_{j\in\mathcal{N}_{it}^{(r)}};\hat{\theta})$ based on the recomputed value function.

Second, in counterfactual scenarios with a different new listing arrival rate from that in the data, we replace the observed $\mathcal{NEW}_{it}$ with the simulated $\mathcal{NEW}_{it}^{(r)}$ and the observed $\mathcal{EXIT}_{it}$ with the simulated $\mathcal{EXIT}_{it}^{(r)}$. We simulate $\mathcal{NEW}_{it}^{(r)}$ as follows. Let λ_i^{CF} ($> \lambda_i$) denote the arrival rate of new listings in consumer i's search market in a counterfactual scenario. We draw a set of additional new listings by first drawing an integer $add_{it}^{(r)}$ from a Poisson distribution with arrival rate $\lambda_i^{CF} - \lambda_i$ and then drawing $\delta_j^{(r)}$ from $N(\mu_{it}^{new}, (\sigma_i^{new})^2)$ for each of the $add_{it}^{(r)}$ additional new listings. We add these additional new listings (denoted by $\mathcal{ADD}_{it}^{(r)}$) to the observed $\mathcal{NEW}_{it}$ to obtain the simulated $\mathcal{NEW}_{it}^{(r)}$. To simulate $\mathcal{EXIT}_{it}^{(r)}$, we draw a set of exited properties from $\cup_{\tau<t}\mathcal{ADD}_{i\tau}^{(r)} \backslash \mathcal{EXIT}_{i\tau}^{(r)}$ based on a Bernoulli distribution with a probability of χ_i. We add these additional exited properties to the observed $\mathcal{EXIT}_{it}$ to define $\mathcal{EXIT}_{it}^{(r)}$.

Third, in counterfactual scenarios with a different exit rate from that in the data, we replace the observed $\mathcal{EXIT}_{it}$ with the simulated $\mathcal{EXIT}_{it}^{(r)}$. We simulate $\mathcal{EXIT}_{it}^{(r)}$ as follows. Let χ_i^{CF} ($< \chi_i$) denote the exit rate in consumer i's search market in a counterfactual scenario. We draw a subset of properties from the observed $\mathcal{EXIT}_{it}$ based on a Bernoulli distribution with a probability of $\chi_i - \chi_i^{CF}$ and define $\mathcal{EXIT}_{it}^{(r)}$ as the difference between $\mathcal{EXIT}_{it}$ and this subset.

SE Proofs

Throughout this section, we suppress the subscripts i and t for simplicity.

SE.1 Proof for $\Pr(\mathcal{N}|\mathcal{A}, n)$ in Equation (5)

We derive $\Pr(\mathcal{N}|\mathcal{A}, n)$ based on the inclusion-exclusion principle and 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 option $j \in \mathcal{N}$ such that $\max_{h \in \mathcal{A} \setminus \mathcal{N}}(\delta_h + \epsilon_h) > \delta_j + \epsilon_j$. This probability reflects the inclusion and exclusion of the following probabilities:

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

In other words,

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

We then derive the analytic expression for $\Pr(\max_{h \in \mathcal{A} \setminus \mathcal{N}}(\delta_h + \epsilon_h) > \delta_l + \epsilon_l, l \in \mathcal{B})$ using the choice probability in a standard Logit model as follows:

$$\begin{aligned} \Pr(\max_{h \in \mathcal{A} \setminus \mathcal{N}}(\delta_h + \epsilon_h) > \delta_l + \epsilon_l, l \in \mathcal{B}) &= \sum_{h \in \mathcal{A} \setminus \mathcal{N}} \Pr(\delta_h + \epsilon_h \geq \delta_l + \epsilon_l, l \in (\mathcal{A} \setminus \mathcal{N}) \cup \mathcal{B}) \\ &= \sum_{h \in \mathcal{A} \setminus \mathcal{N}} \frac{exp(\delta_h)}{\sum_{l \in (\mathcal{A} \setminus \mathcal{N}) \cup \mathcal{B}} exp(\delta_l)}. \end{aligned}$$

Combining the above two steps yields

$$\Pr(\mathcal{N}|\mathcal{A}, n) = 1 - \sum_{k=1}^{n} \left[(-1)^{k-1} \sum_{\mathcal{B} \in \mathcal{C}^k(\mathcal{N})} \frac{\sum_{h \in \mathcal{A} \setminus \mathcal{N}} exp(\delta_h)}{\sum_{l \in (\mathcal{A} \setminus \mathcal{N}) \cup \mathcal{B}} exp(\delta_l)} \right],$$

which can be further simplified as follows:

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

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

SE.2 Intuition and Proof for $\Pr(j|\mathcal{A},n)$ in Equation (6)

Intuition In Appendix A.2.1, we provide the intuition for the analytical expression for $\Pr(j|\mathcal{A},n)$ in (6). The intuition is based on two points. We now use a simple example where $j=1$, $\mathcal{A}=\{1,2,3,4\}$, and $n=2$ to illustrate these two points. In the context of this example, the first point is:

$$\Pr(1|\{1,2,3,4\},2)=\Pr(\text{1's rank is }1)+\Pr(\text{1's rank is }2).$$

To see the second point, note that

$$\Pr(\text{1's rank is }1)=\Pr(\text{1 is the best in }\{1,2,3,4\}).$$

To derive $\Pr(\text{1's rank is }2)$, we first note that

$$\begin{aligned}&\Pr(\text{1 is the best in }\{1,2,3\})\\ =&\Pr(\text{1 is the best in }\{1,2,3,4\})\\ +&\Pr(\text{4 is the best in }\{1,2,3,4\}\text{ and }1\text{ is the best in }\{1,2,3\}),\end{aligned}$$

$$\begin{aligned}&\Pr(\text{1 is the best in }\{1,3,4\})\\ =&\Pr(\text{1 is the best in }\{1,2,3,4\})\\ +&\Pr(\text{2 is the best in }\{1,2,3,4\}\text{ and }1\text{ is the best in }\{1,3,4\}),\end{aligned}$$

$$
\begin{aligned}
&\Pr(\text{1 is the best in } \{1,2,4\}) \\
=&\Pr(1 \text{ is the best in } \{1,2,3,4\}) \\
+&\Pr(3 \text{ is the best in } \{1,2,3,4\} \text{ and } 1 \text{ is the best in } \{1,2,4\}).
\end{aligned}
$$

Therefore, a summation of the above three equations yields:

$$
\begin{aligned}
&\Pr(\text{1 is the best in } \{1,2,3\}) + \Pr(\text{1 is the best in } \{1,3,4\}) + \Pr(\text{1 is the best in } \{1,2,4\}) \\
=&3\Pr(1\text{'s rank is } 1) + \Pr(1\text{'s rank is } 2),
\end{aligned}
$$

which counts $\Pr(1\text{'s rank is } 2)$ once but $\Pr(1\text{'s rank is } 1)$ three times.

In other words,

$$
\Pr(1\text{'s rank is } 2) = \sum_{\mathcal{B}\in\mathcal{C}^2(\{2,3,4\})} \Pr(1 \text{ is the best in } \mathcal{B}\cup\{1\}) - 3\Pr(1\text{'s rank is } 1).
$$

These two points explain why the probability in (6) depends on $\frac{exp(\delta_j)}{\sum_{l\in\mathcal{B}} exp(\delta_l)+exp(\delta_j)}$, which reflects $\Pr(j \text{ is the best among } \mathcal{B}\cup\{j\})$, and constant $C^k_{A-n+k-1}$, which is used to adjust for double-counting.

Proof We derive the analytical expression for $\Pr(j|\mathcal{A},n)$ in (6) when $0<n<A$ based on $\Pr(\mathcal{N}|\mathcal{A},n)$ in (5).

Plugging $\Pr(\mathcal{N}|\mathcal{A},n)$ from (5) 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

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

From (SE.1) to the sum of (SE.2) and (SE.3), we replace $\mathcal{A}\backslash\mathcal{N}$ by $\mathcal{D}$ (so that $\sum_{\mathcal{N}\in\mathcal{C}^n(\mathcal{A}):j\in\mathcal{N}}$ becomes $\sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{A}\backslash j)}$) and decompose $\sum_{\mathcal{B}\in\mathcal{C}^k(\mathcal{N})}$ into two summations: one over $\mathcal{B}$ that includes j (in line (SE.2)), and another over $\mathcal{B}$ that excludes j (in line (SE.3)). In (SE.2), we use $\mathcal{B}'$ to represent $\mathcal{B}\backslash j$. In (SE.3), the summation goes to $n-1$ instead of n because, when $k=n$, $\mathcal{B}\in\mathcal{C}^k(\mathcal{N})$ excluding j does not exist.

From (SE.2) to (SE.4), we use $\mathcal{F}$ to represent $\mathcal{D}\cup\mathcal{B}'$. Because $\mathcal{D}$ and $\mathcal{B}'$ do not intersect, the double summation over $\mathcal{D}$ and $\mathcal{B}'$ in (SE.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 obtain line (SE.5) from line (SE.3).

Line (SE.4) can be further expressed as:

$$\sum_{k=1}^{n}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\setminus j)}\sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{F})}\left(1-\frac{\sum_{l\in\mathcal{D}}exp(\delta_l)}{\sum_{l\in\mathcal{F}}exp(\delta_l)+exp(\delta_j)}\right)\right] \tag{SE.6}$$

$$=\sum_{k=1}^{n}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\setminus j)}\left(C_{A-n+k-1}^{A-n}-\sum_{\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{F})}\frac{\sum_{l\in\mathcal{D}}exp(\delta_l)}{\sum_{l\in\mathcal{F}}exp(\delta_l)+exp(\delta_j)}\right)\right]$$

$$=\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n}$$

$$-\sum_{k=1}^{n}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\setminus j)}\frac{C_{A-n+k-2}^{A-n-1}\sum_{l\in\mathcal{F}}exp(\delta_l)}{\sum_{l\in\mathcal{F}}exp(\delta_l)+exp(\delta_j)}\right] \tag{SE.7}$$

$$=\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n}$$

$$-\sum_{k=1}^{n}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\setminus j)}\left\{C_{A-n+k-2}^{A-n-1}\left(1-\frac{exp(\delta_j)}{\sum_{l\in\mathcal{F}}exp(\delta_l)+exp(\delta_j)}\right)\right\}\right]$$

$$=\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n} \tag{SE.8}$$

$$-\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-2}^{A-n-1} \tag{SE.9}$$

$$+\sum_{k=1}^{n}\left[(-1)^{k-1}C_{A-n+k-2}^{A-n-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k-1}(\mathcal{A}\setminus j)}\left(\frac{exp(\delta_j)}{\sum_{l\in\mathcal{F}}exp(\delta_l)+exp(\delta_j)}\right)\right],$$

where line (SE.7) holds because each term $exp(\delta_l)$ in the numerator in (SE.6) is repeated $C_{\#\mathcal{F}-1}^{\#\mathcal{D}-1}$ times in the summation over $\mathcal{D}\in\mathcal{C}^{A-n}(\mathcal{F})$ and $C_{\#\mathcal{F}-1}^{\#\mathcal{D}-1}=C_{A-n+k-2}^{A-n-1}$.

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

$$\sum_{k=1}^{n-1}\left[(-1)^{k-1}\sum_{\mathcal{F}\in\mathcal{C}^{A-n+k}(\mathcal{A}\setminus j)}C_{A-n+k-1}^{k-1}\right]=\sum_{k=1}^{n-1}(-1)^{k-1}C_{A-1}^{A-n+k}C_{A-n+k-1}^{k-1}. \tag{SE.10}$$

In Supplemental Appendix SE.4, we show that line (SE.8) $=0$, line (SE.9) $=-1$, and line (SE.10) $=1$. Therefore, combining the above simplified expressions for (SE.4) and

(SE.5) yields:

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

SE.3 Proof for the Contraction Mapping in Equation (A.3)

In this section, we first show that the mapping in (A.3) has the following four features: (1) $\frac{\partial f_j}{\partial \delta_k} \geq 0$ for any j, k; (2) $\sum_{k\in\mathcal{J}} \frac{\partial f_j}{\partial \delta_k} < 1$ for any j; (3) There is a value $\underline{\delta}$ such that if $\delta_j < \underline{\delta}$, then $f_j(\boldsymbol{\delta}) > \delta_j$; and (4) there is a value $\bar{\delta}$ such that if $\delta_j > \bar{\delta}$, then $f_j(\boldsymbol{\delta}) < \delta_j$. We then show that these features imply that a truncated version of the mapping is a contraction mapping, which implies that the mapping has a unique fixed point.

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

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 $n = 1$, $\Pr(j|\mathcal{A},1)$ is the choice probability in a standard Logit model. As a result, $\frac{\partial \Pr(j|\mathcal{A},1)}{\partial \delta_j} = \Pr(j|\mathcal{A},1)\,(1-\Pr(j|\mathcal{A},1)) < \Pr(j|\mathcal{A},1)$.

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,

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

where $\Pr(j\text{'s rank is } n)$ 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,

$$\begin{aligned}&\Pr(j\text{'s rank is } n) \qquad \text{(SE.11)}\\ =\;& \sum_{\{(j_1,\cdots,j_{n-1}):j_l\neq j,l=1,\ldots,n-1\}}\left[\Pr(j_1|\mathcal{A},1)\times\cdots\times\Pr(j_{n-1}|\mathcal{A}\backslash\{j_1,\cdots,j_{n-2}\},1)\right.\\ &\left.\times\Pr(j|\mathcal{A}\backslash\{j_1,\cdots,j_{n-1}\},1)\right].\end{aligned}$$

Since the probabilities in (SE.11) are choice probabilities in a standard Logit model, we have $\frac{\partial \Pr(j_l|\mathcal{A}\backslash\{j_1,\cdots,j_{l-1}\},1)}{\partial \delta_j} < 0$ and $\frac{\partial \Pr(j|\mathcal{A}\backslash\{j_1,\cdots,j_{n-1}\},1)}{\partial \delta_j} < \Pr(j|\mathcal{A}\backslash\{j_1,\cdots,j_{n-1}\},1)$. There-

fore, $\frac{\partial \Pr(j\text{'s rank is } n)}{\partial \delta_j} < \Pr(j\text{'s rank is } n)$. 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 $\{\delta_{j'} + \epsilon_{j'}\}_{j' \in \mathcal{A}}$, it decreases as δ_k increases for any $k \neq j$.

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

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

Therefore,

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

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

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

Therefore,

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

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

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

Because increasing the mean utility of every option (including δ_1) by the same amount does not change the sampling probabilities, $\frac{\partial \tilde{s}_j(\boldsymbol{\delta})}{\partial \delta_1} + \sum_{k \in \mathcal{J}} \frac{\partial \tilde{s}_j(\boldsymbol{\delta})}{\partial \delta_k} = 0$, which implies that $\sum_{k \in \mathcal{J}} \frac{\partial \tilde{s}_j(\boldsymbol{\delta})}{\partial \delta_k} = -\frac{\partial \tilde{s}_j(\boldsymbol{\delta})}{\partial \delta_1} > 0$. As a result, $\sum_{k \in \mathcal{J}} \frac{\partial f_j}{\partial \delta_k} = 1 - \frac{1}{\tilde{s}_j(\boldsymbol{\delta})} \sum_{k \in \mathcal{J}} \frac{\partial \tilde{s}_j(\boldsymbol{\delta})}{\partial \delta_k} < 1$.

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

Given Condition 2 in Section A.3, $\delta_j = -\infty$ implies $\Pr(j|\mathcal{A}_{it}, n_{it}) = 0$ and thus $\tilde{s}_j(\boldsymbol{\delta}) = 0$. In other words, $\lim_{\delta_j \to -\infty} f_j(\boldsymbol{\delta}) = \infty$. By continuity of $f_j(\boldsymbol{\delta})$, there exists $\underline{\delta}_j$ such that $f_j(\boldsymbol{\delta}) > \delta_j$ for any $\boldsymbol{\delta}$ 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 $j \in \mathcal{J}$, then $f_j(\boldsymbol{\delta}) < \delta_j$.

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

Features (1)–(4) imply that $f(\boldsymbol{\delta})$ has a unique fixed point.

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

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

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

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

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

where $\varsigma = \max_j \max_{\boldsymbol{\delta}\in[\underline{\delta},\bar{\delta}]^{J-1}} \sum_{k\in\mathcal{J}} \frac{\partial f_j(\boldsymbol{\delta})}{\partial \delta_k}$. 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.

SE.4 Proof of the Combinatorial Identities[3]

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

$$\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-1}^{A-n} = 0 \text{ in line (SE.8)},$$
$$\sum_{k=1}^{n}(-1)^{k-1}C_{A-1}^{A-n+k-1}C_{A-n+k-2}^{A-n-1} = 1 \text{ in line (SE.9)},$$
$$\sum_{k=1}^{n-1}(-1)^{k-1}C_{A-1}^{A-n+k}C_{A-n+k-1}^{k-1} = 1 \text{ in line (SE.10)},$$
$$\sum_{k=0}^{n-1}(-1)^{k}C_{A}^{A-n+k+1}C_{A-n+k-1}^{A-n-1} = n \text{ in line (A.2)}.$$

First, we note that line (SE.10) 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 (SE.9) by $n-1$.

To prove the three identities in (SE.8), (SE.9), and (A.2), we rewrite them as

$$\sum_{h=r}^{m}(-1)^{h-r}C_m^h C_h^r = 0, \tag{SE.12}$$
$$\sum_{h=r+1}^{m}(-1)^{h-r-1}C_m^h C_{h-1}^r = 1, \tag{SE.13}$$
$$\sum_{h=r+2}^{m}(-1)^{h-r}C_m^h C_{h-2}^r = m-r-1, \tag{SE.14}$$

using 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 (SE.12), we take the r^{th}-order derivative of the equation $(1+x)^m = \sum_{h=0}^{m} C_m^h x^h$, which yields

$$C_m^r r!(1+x)^{n-r} = \sum_{h=r}^{m} C_m^h C_h^r r! x^{h-r}.$$

[3]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 (https://math.wvu.edu/~hgould/Vol.2.PDF). The first identity (SE.12) is an application of equation (1.23) in Table II. The proof of the other two identities (SE.13) and (SE.14) is inspired by the proof for this equation.

Setting $x = -1$ leads to the first identity.

To prove the second identity (SE.13), 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 like $a(1+x)^b x^{-c}$ where $b > 0$. Therefore, its value at $x = -1$ is 0. The derivative of the RHS at $x = -1$ is

$$\begin{aligned}\frac{d^r}{dx^r}\left(\frac{1}{x} + \sum_{h=1}^{m} C_m^h x^{h-1}\right)\Bigg|_{x=-1} &= \left[r!(-1)^r x^{-1-r} + \sum_{h=r+1}^{m} C_m^h C_{h-1}^r r! x^{h-1-r}\right]_{x=-1} \\ &= r!\left[-1 + \sum_{h=r+1}^{m} C_m^h C_{h-1}^r (-1)^{h-1-r}\right].\end{aligned}$$

Therefore, $-1 + \sum_{h=r+1}^{m} C_m^h C_{h-1}^r (-1)^{h-1-r} = 0$, which implies the second identity.

To prove the third identity (SE.14), 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

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

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.