# Flexible Use of Sequential Search Data: A Partial Ranking Structure

Tinghan Zhang*

*Preliminary*

This version: November 20, 2024

**Abstract**

Sequential search data has become increasingly important in economics and market research. However, the existing structure based on Weitzman's (1979) Optimal Search Rules provides limited support for empirical study, making researchers struggle between using search data with a high computation burden or discarding them completely. This paper reformulates the solution of optimal sequential search with a partial ranking structure, establishing a fully static relationship among product values. This simplifies the model's empirical application while preserving complete search information. With the new structure, I discuss the identification arguments and estimation implementation. I show its flexibility in handling scenarios with partial search information, additional ranking information, and structural changes within the search process (e.g., search and product discovery) with low computational cost.

**Keywords:** Consumer search, choice model, partial ranking

*Tilburg University, Department of Econometrics and Operations Research

# 1 Introduction

Consumer search data has become an indispensable resource in modern economic and marketing research. With increasing market differentiation and the rise of online shopping platforms, the assumption that consumers possess perfect information about all available options at the time of purchase is no longer tenable. Instead, consumers can actively search to reduce uncertainty and make more informed final purchase decisions. Sequential search data provides detailed insights into the search decision-making process, capturing information such as which products were searched, the order of searches, and when the search stopped. This rich dataset is invaluable for understanding consumer preferences and behaviors and serves as a critical tool for developing effective strategies and evaluating market policies.

The Optimal Search Rules proposed by Weitzman (1979) have been a cornerstone for analyzing search data, providing a framework for modeling consumer decision-making during search and purchase. These rules link consumers' choices to their preferences and search costs and offer a foundation for understanding sequential decision-making. Despite the strengths, the empirical application of the Optimal Search Rules remains challenging. As the Optimal Search Rules are based on product values, connecting consumers' sequential decisions helps researchers construct the structure of product values to compute or simulate joint probabilities for maximum likelihood estimation. However, the product values involved in each step of optimal search depend on the outcomes of previous search steps. Since these outcomes are private to consumers and not fully observed by researchers, interdependencies arise between decisions. Consequently, the probabilities derived from these rules are computationally complex. To address these challenges, Armstrong (2017) and Choi et al. (2018) introduce the Eventual Purchase Theorem as an alternative approach. It enables researchers to directly derive consumer demand using a standard discrete choice framework without relying on search data, making it particularly useful when search behavior is not the focus of study or when search data is incomplete. However, this approach comes with limitations: it completely discards information about the search process, making it unsuitable for analyzing search behavior. Researchers are thus forced to choose between models that retain search path information but are computationally complex, and those with simplified computation but forgo search-related insights.

This paper aims to bridge this gap by providing a simplified and flexible approach to the empirical application of sequential search data. I propose a reformulation of the optimal solution to the sequential search process, departing from Weitzman (1979) and introducing the Partial Ranking (PR) structure. This structure captures the static relationships among all product values encountered in the optimal search in a way that is equivalent to, but independent of, the Optimal Search Rules. By adopting the PR structure, search sequences no longer need to be broken down into step-by-step decisions based on the Optimal Search Rules. Instead, the process can be modeled through conditionally independent probabilities centered on the utility of the purchased product, similar to the structure of a multinomial discrete choice model. This reformulation shifts the focus to the final purchase as the anchor of the search process, eliminating the need to condition each search step on the outcomes of earlier ones. As a result, researchers can utilize the observed search sequence to infer consumers' *a priori* ranking of product values, without being encumbered by unobserved private information and computational complexity.

Compared to the Optimal Search Rules (OSR) structure, where product values are incorporated step-by-step based on consumers' sequentially made optimal search decisions, the PR structure offers significant advantages for empirical applications. Its advantage lies in representing the joint probability as a form that can be decomposed into conditionally independent value differences, thus effectively addressing the challenges posed by interdependencies in identification and estimation. In this form, I provide identification arguments for sequential search models that are comparable to those in standard discrete choice models. I find these arguments align with the identification results discovered through simulations in the empirical literature. We then demonstrate how to estimate a sequential search model under the PR structure with a GHK estimation method. Unlike the GHK simulator based on Optimal Search Rules, the core idea behind my simulator is not to decompose and recombine the rules. Instead, it uses the GHK sampling method to replicate the observed partial ranking and calculate its conditional probability given the utility of the purchased product. The implementation of my PR-GHK simulator is much simpler than the existing OSR-GHK methods.

Beyond the baseline model, I demonstrate the PR structure's exceptional flexibility in handling partial search data. First, I show that the discrete choice model based on the Eventual Purchase Theorem is a special case of the sequential search model under the PR structure when ranking information is entirely unobservable. Next, I prove that purchase information can provide inferior ranking insights when the search process is not fully observable and can be integrated with other known ranking information for estimation. Finally, I illustrate the PR-GHK estimator's ability to handle incomplete search information with an example where only the first searched product and the final purchased product are observed in the data. The PR-GHK estimation method effectively and fully utilizes the available search data and performs well in Monte Carlo simulations.

Furthermore, under an independence assumption, the PR structure exhibits remarkable flexibility in incorporating additional information and adapting to observable or tractable structural changes in ranking conditions. For additional information, I demonstrate how the model can incorporate additional signals, such as the commonly used "My Favorite" tag, to supplement preference information beyond the search process. For structural changes, I extend the PR structure to the search-and-product-discovery model (Greminger, 2022, 2024; Zhang et al., 2023), where consumers has additional option(s) to expand their awareness set at each step, increasing the number of available products. Greminger (2022) proves that the optimal solution for this model requires an additional optimal search rule to handle discovery behavior. I further show how to integrate discovery behavior into the PR structure without directly introducing the additional rule but treating it as a choice that induces structural changes in consumer rankings. In this scenario, the GHK method can still be used to reconstruct consumer ranking information between each discovery, effectively incorporating search path and discovery data to estimate the model.

This paper is related to the empirical application methodology of the sequential search model. First, it contributes to the identification arguments of the sequential search model. Current marketing literature mainly discusses which data variations in observed sequences are informative to the corresponding parameters in the model (e.g., Kim et al., 2010). More recent studies (Morozov et al., 2021; Ursu et al., 2024) attempt to formally indicate different parameters correspond to different moments in the data, but these discussions lack completeness due to their detachment from the structure of the search model. Second, it contributes to the estimation of the sequential search model and its derivatives. Empirical research using simulated maximum likelihood estimation has substantial developments based on the Optimal Search Rules. The common simulators include the Crude Frequency Simulator (Chen and Yao, 2017; Ghose et al., 2019), which fits observed sequences by drawing uncertainties extensively to approximate the joint probability of the samples numerically. The simulated likelihood is derived from the average binary outcomes of whether all sequential discrete decisions in the data are satisfied in the complete search process. Another widely-applied method is the Kernel-Smoothed Frequency Simulator, employed by Honka and Chintagunta (2017) (also see Ursu, 2018; Ursu et al., 2020; Yavorsky et al., 2021; Ursu et al., 2023). This approach does not directly fit the joint probability and does not require a formal expression. Instead, it maximizes a constructed kernel-smoothed function of the distances, capturing the extent to which the simulated consumer decisions adhere to the Optimal Search Rules. The third method recently developed is the GHK-style simulator (Jiang et al., 2021; Chung et al., 2024; Greminger, 2024). This approach sequentially samples uncertainties such that subsequent uncertainties satisfy the optimal search rules related to previously sampled variables. Following the Optimal Search Rules, the GHK-style simulator is complicated for implementation (Ursu et al., 2023); while following the PR structure, I eliminate these deficiencies from the Partial Ranking-based GHK-style estimator, significantly enhancing its applicability and extensiveness while retaining its advantages over previous estimation methods.

The rest of the paper is arranged as follows: In Section 2, I define notations and set up the sequential search model, explain the Optimal Search Rules, and illustrate the structure based on the rules. In Section 3, I show the PR structure is an equivalent full description of the optimal solution of the sequential search model to the Optimal Search Rules. In Section 4, I present the joint probability, discuss the identification arguments, and introduce the estimation strategy. In Section 5 I show the structure is compatible with partial search data. In Section 6 I show the structure is compatible with additional information and observable structural changes in the search process with the example of the search-and-product-discovery model proposed in Greminger (2022). I conclude the paper in Section 7.

# 2 Model and Optimal Search Rules (OSR) Structure

This section introduces the baseline optimal sequential search problem and Optimal Search Rules proposed by Weitzman (1979). I first propose the notation restrictions to define the environment and the observation from the data. Then, I introduce the Optimal Search Rules and show the structure set up on these optimal rules. The OSR structure is a straightforward policy-oriented structure. Ursu et al. (2024) contribute a detailed review of the sequential search model under the OSR structure, and I refer the readers to the paper for further understanding.

## 2.1 Notation for Sequence Observation

First, I define notations for the data observed in a sequential search environment. In general, sequential search models are applied to the click-stream data, which records consumers' final purchases and the order of clicks on alternative products before purchasing. Therefore, observations in the click-stream data fully describe purchase decision and all search decisions made for the purchase. I formalize them as the sequence observations as follows.

A typical consumer $i$ has a unit demand for alternatives in a market. We assume that the consumer obtains full knowledge of the market costlessly: she knows the existence of each alternative in the market and the distribution of all market uncertainties. Her information set is partial to product-level information. Denote the set of all products in consumer $i$'s market by $\mathcal{M}_i$. The set $\mathcal{M}_i$ supports consumer $i$'s search and purchase behaviors. I do not consider that the consumer has any foreknowledge on product uncertainties before entering the market, which will be discussed later in Section 5.

The incompleteness of the product information prevents the consumer from knowing the exact utility of any product when she enters the market and, therefore, masks the first best choice. The consumer can reveal product uncertainties to determine the utility, which is realized through the behavior called *inspection*. In a costly search process, inspecting a product is not free. Generally, it is not always optimal to inspect all products in the market regardless of the cost. Only when the utility of a product is determined through inspection will a consumer consider it as an element in her consideration set, from which she purchases one to purchase at the end of a search process. The consideration set expands along with inspections going on in the search process. When the search stops, the consideration set of all inspected products is denoted by $\mathcal{S}_i$, and its size is denoted by $J_i$. The set of uninspected products is defined as $\bar{\mathcal{S}}_i = \mathcal{M}_i \backslash \mathcal{S}_i$.

The way of modeling a consumer's decision-making process in search and purchase largely depends on when the inspection decisions are made. In literature, the products inspected before the purchase can be simultaneously determined when the consumer enters the market, or contingently determined on search outcomes from previous inspections. The sequential search model follows the latter, assuming that consumers make optimal inspection and purchase decisions at every spot of the search process. Therefore, the consumer's observed inspections result from a series of conditional optimal decision-making, and the order of inspections to products in the consideration set matters.

Let the set of all possible inspection orders that lead to the consideration set of $\mathcal{S}_i$ by $\mathcal{K}(\mathcal{S}_i)$. The inspection order of consumer $i$'s search process is noted by $\mathcal{R}_i$. We always have $\mathcal{R}_i \in \mathcal{K}(\mathcal{S}_i)$. Following $\mathcal{R}_i$, we number the products in the final consideration set, i.e., the first inspected product is Product 1, the $j$-th inspected is Product $j$, etc. The last inspected product is numbered $J_i$. Hence, every product with the number $j \leq J_i$ is inspected on $j$-th position in $\mathcal{R}_i$ and is an element of $\mathcal{S}_i$. Hence, inspected products following $\mathcal{R}_i$ are marked by $\{1,2,\cdots,J_i\}_{\mathcal{R}_i}$. Let $H_i$ denote the purchased product, and $h_i$ denote its order number. Notice that $H_i \in \mathcal{S}_i$ and $h_i \leq J_i$ always hold.

Hence, $\{H,\mathcal{S},\mathcal{R},\mathcal{M}\}_i$ is called consumer $i$'s *sequence observation*, which fully captures purchase and inspections before purchase consumer $i$ conducts. For the clearness of representation, I will not stress the inspection order $\mathcal{R}_i$ when using ordered numbers to refer to products in the following of the paper, and I will omit the subscription $i$ in $H_i$, $h_i$, $J_i$, $\mathcal{R}_i$, $\mathcal{S}_i$ and $\mathcal{M}_i$ when they appear individually.

## 2.2 Values, Optimal Search Rules and OSR Structure

I now describe how sequence observations are taken in an optimal sequential search problem. Consumer $i$'s final evaluation to a product depends on how much it matches consumer $i$ when all product-level uncertainties are resolved through inspection. We define $u_{iH}$ the *purchase value* to refer to the utility consumer $i$ expects her to obtain from purchasing product $H$ had she fully revealed the uncertainty of it. When consumer $i$ enters the market, product-level uncertainty prevents the consumer from knowing the purchase value of any alternative in the market. Consumer $i$ only knows the choice set and the distribution of the purchase value of each product. Inspection with paying the cost $c_{iH}$ fully reveals the uncertainty of product $H$ and determines $u_{iH}$. We do not incorporate consumer learning, so purchase values are invariant throughout the search process once they are determined. In this case, a rational consumer solves a dynamic optimization problem in which she sequentially decides whether to stop the search and make a purchase or inspect another product after every inspection. After inspecting any product $j-1$ in the sequence, a rational consumer with a stopping utility of $\bar{u}_{ij}$ solves the following Bellman equation:

$$W(\bar{u}_{ij},\bar{S}_{i,j}) = \max\left\{\bar{u}_{ij},\ \max_{K\in\bar{S}_{i,j}}\left\{-c_{iK}+F^u_{iK}(\bar{u}_i)W(\bar{u}_{ij},\bar{S}_{ij}\backslash\{K\})+\int_{\bar{u}_{ij}}^{\infty}W(u,\bar{S}_{ij}\backslash\{K\})f^u_{iK}(u)du\right\}\right\} \tag{1}$$

Here, $\bar{S}_{ij}$ is a set of uninspected products at each step of $j$. $F^u_{iK}(\cdot)$ and $f^u_{iK}(\cdot)$ are the cdf and pdf of $u_{iK}$, assumed to be given information to the consumer. At every step $j$, the consumer chooses between taking $\bar{u}_{ij}$ and inspecting another product with the highest expected value. The expected value of inspecting a product $K$ consists of three components: the cost of inspection, the value function when $u_{iK}$ is smaller than $\bar{u}_{ij}$, and the value function when $u_{iK}$ exceeds $\bar{u}_{ij}$.

Weitzman (1979) simplifies the dynamic optimization problem into a quasi-static discrete choice model under an independent assumption: inspecting product $i$ does not obtain any information for those products not inspected. In this case, product purchase values are conditionally independent. In addition, in Equation (1), inspection outcomes from previous steps only affect step $j$'s decision-making through $\bar{u}_{ij}$, which can be interpreted as the maximum purchase value of products inspected up to step $j-1$. Hence, the dynamic optimization collapsed to a stepwise static problem: should the consumer inspect a product with a given outside value?

Since the purchase value of a product is unavailable to consumers, the optimal decision relies on the comparison between expected gains from inspecting the product and the search cost. For a product $K$, suppose the utility that consumer $i$ would obtain had the product not inspected is $\bar{u}$. The consumer has extra gains through inspection only when the purchase value of $K$ exceeds $\bar{u}$. Hence, inspecting product $K$ or not is indifferent when:

$$c_{iK} = \int_{u_{iK}>\bar{u}} (u_{iK} - \bar{u})\, dF^{u}_{iK}(u_{iK}) \tag{2}$$

Similar to Equation (1), $F^{u}_{iK}(\cdot)$ is the cdf of $u_{iK}$. Equation (2) is an implicit function of $\bar{u}$. The right-hand side of the equation is monotonically decreasing with $\bar{u}$, so the solution is unique. Marked by $z_{ij}$, the solution is defined as the *reservation value* of the product $K$ to consumer $i$. When a consumer $i$ has a determined alternative value that is larger than $z_{iK}$, she will not conduct an inspection of product $K$; otherwise, she will inspect product $K$. Intuitively, we can regard $z_{iK}$ as the numerical expression of the value of conducting an additional inspection of product $K$.

With purchase and reservation values defined, Weitzman (1979) proposes four optimal search rules. These rules fully characterize the optimal solution of the baseline sequential search model, and explain the information given in the sequence observation in the data.

1. Optimal Ranking: The reservation value of products inspected earlier is always larger than the reservation value of products inspected later, i.e., consumer searches in decreasing order of reservation values.

   $$t^{1}_{ij} \equiv z_{ij} - \max_{k\in\mathcal{M}_i\setminus\{1,2,\cdots,j\}} z_{ik} \geq 0, \forall j \leq J.$$

   With the Optimal Ranking rule, the observed inspection order $\mathcal{R}_i$ also captures the descending order of reservation values of products inspected. Hence, following the Optimal Ranking rule, we also number products that are off $\mathcal{R}_i$ and not in the consideration set $\mathcal{S}_i$ by $J+1, J+2, \cdots$ according to the descending order of their reservation values. Hence, an uninspected product is numbered with $k > J$, and its reservation value among all products is the $k$-th largest to consumer $i$. Therefore, $\max_{H\in\bar{\mathcal{S}}_i} z_{iH}$ in the above condition can also be expressed by $\max_{k>J} z_{ik}$.

2. Optimal Continuing: consumer continues inspecting another product when the maximum of purchase values among inspected products is smaller than the maximum of reservation values of products not inspected:

$$t_{ij}^2 \equiv \max_{\ell \geq j} z_{i\ell} - \max_{\ell=1}^{j-1} u_{i\ell} \geq 0, \quad \forall j < J.$$

3. Optimal Stopping: consumer stops inspecting another product when the maximum of purchase values among inspected products is larger than the maximum of reservation values of products not inspected:

$$t_i^3 \equiv \max_{j \leq J} u_{ij} - \max_{k>J} z_{ik} \geq 0.$$

   Optimal Continuing and Stopping correspond to consumers' decision to cease the search process. According to the definition of reservation value, continuing is a better choice when the reservation value of product $j+1$ is larger than the maximum purchase value of those inspected. In contrast, stopping is optimal when the maximum purchase value inspected exceeds the reservation value of all uninspected products. Optimal Stopping applies to the last step of the search process, while Optimal Continuing applies to inspections in the middle of search.

4. Optimal Purchasing: Consumer purchases product $H$ (numbered by $h$ given $\mathcal{R}_i$) if and only if it has the largest purchase value among all inspected products.

$$t_i^4 \equiv u_{ih} - \max_{j \leq J, j \neq h} u_{ij} \geq 0.$$

The optimal search rules provide the optimal policy for each decision in the consumer's search process. At each step, the consumer first decides whether to continue searching (Optimal Searching or Continuing). If continuing, they choose which product to inspect next (Optimal Ranking). If stopping, they decide which product to purchase (Optimal Purchasing). These policies fully characterize the optimal solution to the sequential search problem described in Equation (1).

Take Figure 1 as an illustration. When a consumer enters the market, she first inspects the product with the highest reservation value $z_{i1}$ following the Optimal Ranking rule and reveals $u_{i1}$, the purchase value of the inspected product. According to the Optimal Continuing rule, consumers continue inspection when the next largest reservation value in the market and the largest reservation value among uninspected products, $z_{i2}$, is larger than $u_{i1}$. The second inspection reveals $u_{i2}$. She compares $u_{i2}$ with $u_{i1}$ and replaces the maximum purchase value of inspected with $u_{i2}$. In our illustration, $u_{i2}$ is smaller than $z_{i3}$, hence search continues. After product $J$ is in-

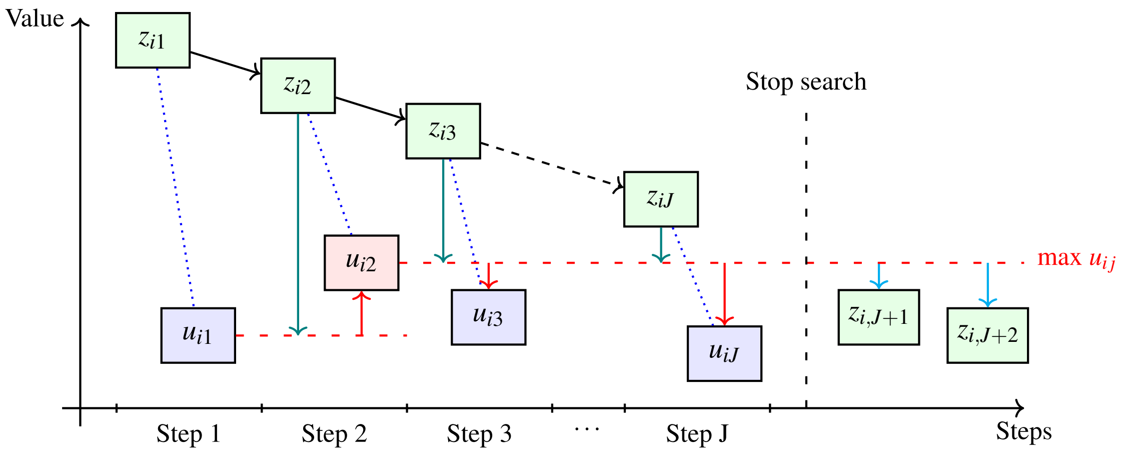


FIGURE 1 – The Optimal Search Rules (OSR) Structure

spected, for the first time, the maximum purchase value among inspected products exceeds the maximum reservation value among uninspected products. The consumer terminates search, and purchases the product with the maximum purchase value among inspected products according to the rule of Optimal Purchasing.

The sequence observations record consumers' complete search and purchase processes. For each observation, we can replicate the search process similar to Figure 1. As the search progresses and stops at a purchase, we can construct a complete step-by-step product value structure following the optimal search rules, noted as the Optimal Search Rules (OSR) structure. In the OSR structure, product values must ensure that each consumer decision aligns with the corresponding optimal policy (shown as solid arrows in Figure 1). Therefore, for any sequence observation, the joint probability is the probability that all optimal policies hold at each step:

$$\Pr(\{h,S,R,\mathcal{M}\}_i|\boldsymbol{X}_i) = Pr(t_{ij}^1 > 0, \forall j \leq J \cap t_{ij}^2 > 0, \forall j \leq J \cap t_i^3 > 0 \cap t_i^4 > 0) \tag{3}$$

In most empirical studies, the joint probability expression based on the OSR structure serves as the empirical foundation of the sequential search model. Although this probability does not have a closed-form solution, researchers can estimate it using simulated maximum likelihood methods.

However, while the OSR structure is intuitive and conceptually clear, it is not entirely suitable for empirical research. The primary reason lies in the dependency of the optimal rules - except for Optimal Ranking - on the purchase value of products. These purchase values are not known *a priori* and can only be determined through inspections within the search process. Consequently, the probabilities of decisions in subsequent search steps are inherently correlated to the outcomes of previous ones, yet it is unclear which specific outcomes drive these dependencies. This makes the joint probability given in Equation 3 difficult to decompose effectively.

As a result, identification in sequential search models must rely on the full joint probability, which significantly increases complexity compared to the common discrete choice model.

It also requires estimation techniques to overcome the computation and implementation challenges when estimating the model, and existing estimation methods often lack the flexibility to accommodate variations in the data or in the model [1].

# 3 Partial Ranking (PR) Structure

Tailoring a more flexible structure of the sequential search model to make it more applicable for empirical research is an important topic. Here, I propose four conditions on the product values in a sequential search process. The following proposition shows that these conditions hold if and only if the optimal rules are fully satisfied, so they equivalently capture the optimal solution of the sequential search model.

**Proposition 1.** *For each sequence observation* $\{H,S,R,\mathcal{M}\}_i$ *in the click-stream data, define* $y_i = \min\{u_{ih}, z_{iJ}\}$*, the minimum between the purchase value of the inspected product and the reservation value of the last inspected product, as the core value of the sequence observation i. Weitzman's Optimal Search Rules hold if and only if the following conditions are fulfilled:*

1. *Distribution Condition:* $u_{ih} \leq z_{iJ}$ *if* $h < J$*;*
2. *Rank Condition:* $z_{i1} \geq z_{i2} \geq \ldots \geq z_{iJ}$*;*
3. *Choice Condition 1:* $u_{ij} \leq y_i$ *for all* $j < J, j \neq h$*;*
4. *Choice Condition 2:* $z_{ik} \leq y_i$ *for all* $k > J$*.*

*Proof.* I first prove in showing the necessity that violating conditions in the proposition always violates Weitzman's Optimal Search Rules.

- When the Distribution Condition is violated, the Optimal Continuing Rule is violated.
- When the Rank Condition is violated, the Optimal Ranking Rule is violated.
- When the first Choice Condition is violated, there would be two cases. When $\exists j, s.t. z_{iJ} < u_{ij}$, the Optimal Continuing Rule is violated. The Optimal Purchasing Rule is violated when $\exists j, s.t. u_{ih} < u_{ij}$.
- When the second Choice Condition is violated, there would be two cases. In the case of $\exists k, s.t. z_{iJ} < z_{ik}$, it violates the Optimal Ranking Rule. If $\exists k, s.t. u_{ih} < z_{ik}$, given the Optimal Purchasing Rule is not violated, i.e., $u_{ih} \geq \max_{j \leq J, j \neq h} u_{ij}$, the Optimal Stopping Rule is violated.

[1] In the introduction, we introduced three mainstream methods for estimation with Equation (3). Among them, the Crude Frequency Simulator is unsuitable for large search models due to its computational burden. The Kernel-Smoothed Frequency Simulator faces challenges in selecting scaling factors, which is critical to its performance but difficult to determine; additionally, model or data variations need to be implemented with additional optimal rules, introducing the need for more scaling factors. Lastly, the OSR-based GHK Simulator is significantly complicated to implement even for the baseline model, and lacks a unified approach to handle model or data variations.

Next, I prove sufficiency in showing that violating Weitzman's Optimal Search Rules also violates conditions in the proposition.

- When the Optimal Ranking Rule is violated. If $\exists j_1 < j_2 < J$ such that $z_{ij_1} < z_{ij_2}$, the Rank Condition is violated; if $\exists j < J$ such that $z_{ij} < \max_{k>J} z_{ik}$, the Choice Condition 2 is violated.

- When the Optimal Stopping Rule is violated. If the Choice Condition 1 holds, we have $u_{ih} = \max_{j \leq J} u_{ij} < \max_{k>J} z_{ik}$. Hence, $\exists k > J, s.t. z_{ik} > u_{ih}$, which violates the Choice Condition 2.

- When the Optimal Continuing Rule is violated, it is to say that $\exists \ell < j \leq J, s.t. z_{ij} < u_{i\ell}$. With the Rank Condition holds, we have $z_{iJ} \leq z_{ij} < u_{i\ell}$. If $\ell \neq h$, the Choice Condition 1 is violated; if $\ell = h$, the Distribution Condition is violated.

- When the Optimal Purchasing Rule is violated, the Choice Condition 1 is violated.

□

Unlike the OSR structure, where product values have complex and hard-to-identify correlations, the product values in these four conditions are centered around the core value, forming a static, centralized structure as in Figure 2. With the equivalence proven in Proposition 1, this structure retains all search information of the OSR structure without any loss.

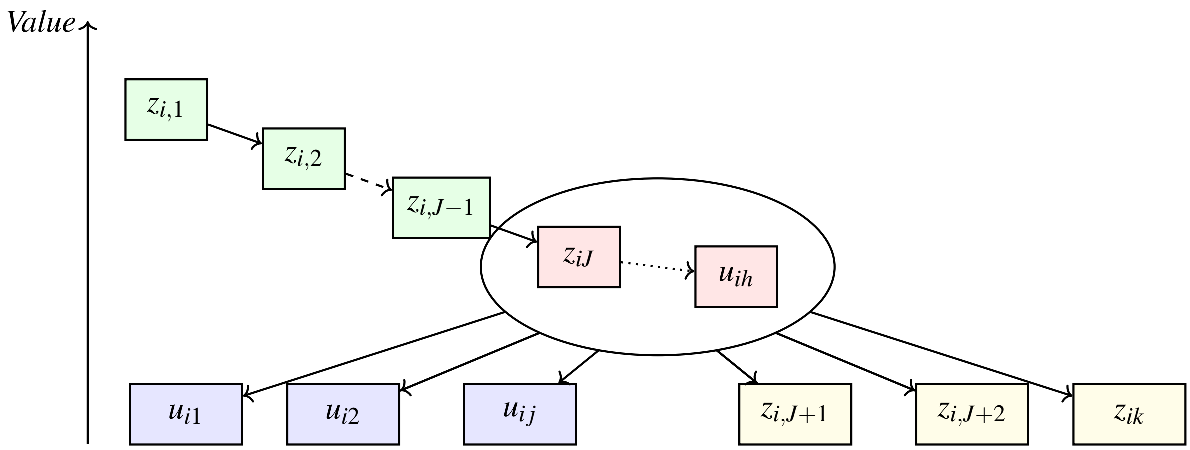


FIGURE 2 – The Partial Ranking Structure

The structure in Figure 2 is similar to a discrete choice structure with a ranking condition. The two Choice Conditions describe the values that rank below the core value, corresponding to the unchosen options in the discrete choice model. The Rank Condition provide the consumer's ranking to the reservation values up to the last inspection. The Distribution Condition specifies the distribution of the core value, which is determined by the feature of the search process.

I term the structure presented in Figure 2 the Partial Ranking structure because it represents optimal search outcomes as a partial ranking of reservation and purchase values for consumers in the market. As proven in Appendix A, this structure is independent of the Optimal Search Rules and instead relies on the optimality results from the multi-armed bandit literature, particularly the branching bandit framework proposed by Keller and Oldale (2003). The authors demonstrate that a Gittins Index policy is optimal in multi-armed bandit problems where actions branch into new actions. In the proposed sequential search problem, we base the framework on two key assumptions. First, inspections reveal uncertainty only for the inspected product (Independence). Second, reservation and purchase values remain constant throughout the search process (Invariance). Independence ensures that selecting the action with the highest reservation or purchase value is optimal at each step. Invariance guarantees that the values of actions (reservation values for inspections and purchase values for purchases) remains unchanged during the search process, referred to as the '*Ranking of Values*.' Under these assumptions, the consumer's sequential decisions align with the ranking of values, analogous to branching decisions in the branching bandit problem. Compared to a discrete choice based on purchase values, the ranking of values captures much richer information from consumer choices, providing greater insight into preferences and search costs [2].

Figure 3 illustrates how the optimal ranking of values forms a partial ranking structure in a three-product case. While the ranking remains stable throughout the search process, only reservation values are available to the consumer at the start, as purchase values are initially unknown. Each inspection reveals the purchase value of a product, which is then incorporated into the ranking without altering the magnitude or order of previously determined components. However, the full ranking cannot be observed from the search outcome, as the purchase decision ends the search and results in a censored observation. This occurs when an exogenous condition is met - specifically, when a revealed purchase value exceeds all remaining reservation values. While purchasing determines when the search ends, it does not affect the underlying ranking; instead, it truncates the ranking by excluding the purchase values of uninspected products.

The cessation of the search transforms the ranking of values into a partial ranking, as the purchase values of unpurchased products and reservation values of uninspected products are censored from the data. However, it is known that these unobserved values are smaller than the core value, which represents the set-identified smallest value among actions taken in the observed sequence. The ranking information of both observed and censored values is fully described by the conditions in Proposition 1, ultimately capturing all available information from the search process.

In the OSR structure, consumers follow optimal search rules step-by-step along the inspection order $\mathcal{R}_i$, constructing the consideration set $\mathcal{S}_i$ and eventually purchasing the product $H_i$. Each decision is made conditional on the previous ones. In contrast, the PR structure takes the search process $\mathcal{R}_i \in \mathcal{K}(\mathcal{S}_i)$ as an observation conditional on the core value $y_i$, given the purchase of product $H_i$. The core value is central to this structure, as it represents the smallest value identifiable from the observed partial ranking. When $h < J$, we know that $u_{ih}$ is smaller than

[2] Ranked data plays a similar role to consumers' "second choice" in Berry et al. (2004).

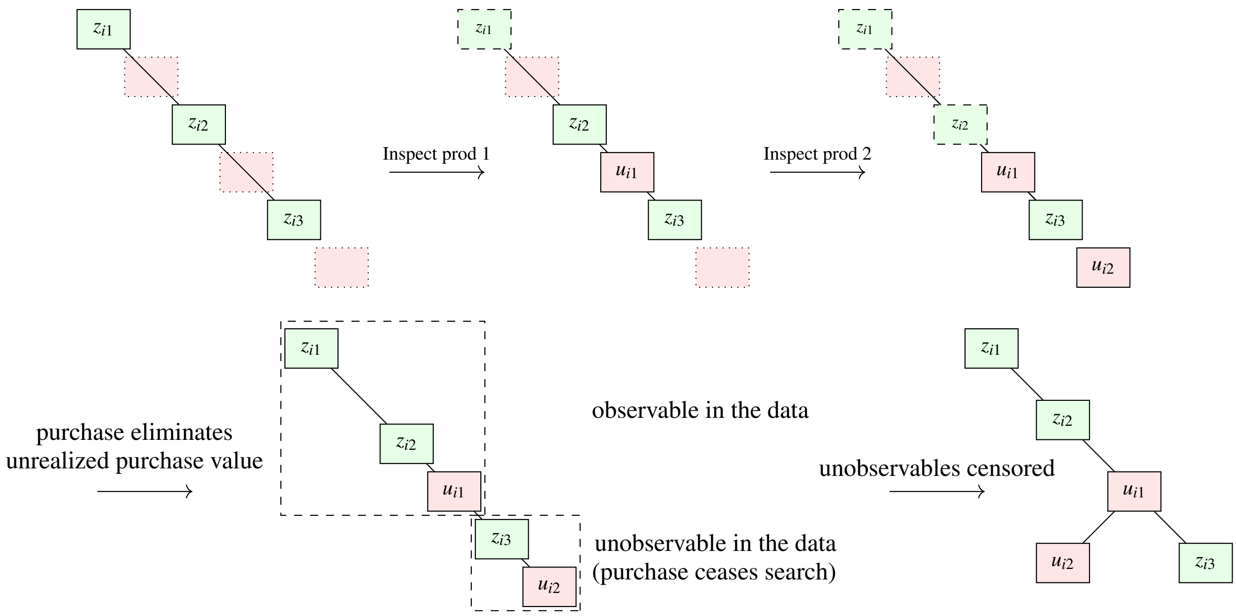


FIGURE 3 – How the Full Ranking of Values Becomes Partial

$z_{iJ}$. However, when $h = J$, this value becomes unidentifiable. In the PR structure, the observed partial ranking of reservation values, the purchase values of unpurchased products, and the reservation values of uninspected products are independent conditional on the core value. This conditional independence makes the core value a sufficient statistic for the conditional distributions of all other components in the structure. As a result, this property simplifies the sequential search process by removing the need to account for uninspected or unpurchased products when considering optimal inspection decisions.

Compared to the OSR structure, the PR structure is particularly well-suited for empirical research on sequential search problems. The PR structure is a static framework where inequality relations connect all product values to at most one component emanated from the core value. These inequality conditions, directly observable from the data, can reproduce the consumer's ranking of values without requiring consideration of uncertainties associated with other products. This characteristic significantly reduces interdependencies within the structure. Moreover, the optimality of the PR structure relies on the Independence and Invariance assumptions, which hold regardless of consumers' decision-making in the search process. This makes the structure sufficiently flexible for empirical analysis. On one hand, even if part of the search information is unknown, the limited observed ranking condition derived from known information remains unchanged and can be used for estimation. On the other hand, any additional information about the unobservable components in the censored part of the ranking can be incorporated into the existing structure. As long as the impact of exogenous or endogenous factors on the partial ranking can be fully traced in data, the PR structure can also accommodate structural changes in the ranking caused by these factors.

# 4 Joint Probability, Identification and Estimation

The PR structure makes it possible to present the joint probability of the sequence observation with a value-difference form similar to the discrete choice data or the ranked data (Hajivassiliou and Ruud, 1994). This provides a different scheme from the existing literature to discuss the identification arguments and estimation implementation of the sequential search model.

## 4.1 Joint Probability

Notice that different from Equation (3), the joint probability following the PR structure can be expressed as:

$$\Pr(\{H,S,R,\mathcal{M}\}_i) = \Pr(z_{iJ} \geq u_{ih} \cap z_{i1} \geq \ldots \geq z_{iJ} \cap \max_{j \leq J} u_{ij} \leq y_i \cap \max_{k>J} z_{ik} \leq y_i) \tag{4}$$

We consider an additive restriction on purchase and reservation values between stochastic randomnesses and the expected value from those observed attributes. The baseline model setup is as follows:

$$\begin{aligned} u_{ij} &= \delta_i^u(X_{ij}^u) + \xi_{ij}^u + \varepsilon_{ij} \\ z_{ij} &= \delta_i^z(X_{ij}^z) + \xi_{ij}^u + \xi_{ij}^z \end{aligned}$$

Let us provide some notes on this setup. First, $\delta_i^u(X_{ij}^u)$ and $\delta_i^z(X_{ij}^z)$ are two deterministic components of product values. Both consumers and researchers can observe $X_{ij}^u$ and $X_{ij}^z$ that directly affect consumers' evaluation of inspecting or purchasing product $j$. The two components are identical in many empirical settings, while they can be further differentiated with additional components in $X_{ij}^z$, such as an advertisement or a featured recommendation, which only affects the inspection but not the purchase.

We incorporate three additional (potentially stochastic) terms into the model, complementing the deterministic components. The first term, $\xi_{ij}^u$, is a shared component influencing both inspection and purchase decisions. This term represents a part of the product's purchase value determined by the consumer before inspection but unidentifiable in the dataset. It is also part of the reservation value, typically interpreted as the consumer's private information or subjective taste for the product. The second term, $\varepsilon_{ij}$, captures product-level uncertainty resolved during inspections. The evaluation of the solved uncertainty is modeled as the unidimensional value added to the deterministic component in this linear specification. Consumers know the distribution of $\varepsilon_{ij}$ of each product $j$ when entering the market. For simplicity and without loss of generality, we assume the distribution of $\varepsilon_{ij}$ is the same across all products and consumers, with a cdf $f^{\varepsilon}(\cdot)$ and a pdf $F^{\varepsilon}(\cdot)$.

Lastly, $\xi_{ij}^z$ is a component that only affects the reservation but not the purchase value. The key aspect of this component is the rent created by the search cost $c_{ij}$. When the search cost is zero, the solution to Equation (2) equals the expectation of the purchase value. However, when the search cost is positive, the difference between the reservation value and the expected utility of the purchase value represents the rent generated by the search cost. In our model, the search cost influences consumers' search behavior solely through the rent it generates. In a linear specification, the rent from the search cost depends only on the distribution of $\varepsilon$ and the magnitude of the search cost. We represent it as $m_\varepsilon(c_{ij})$, where $m_\varepsilon(\cdot)$ is a strictly decreasing function [3]. Additionally, there may be other unobservable random factors unrelated to the search cost that influence consumers' search behavior. These can be interpreted as unobserved platform rankings (Ursu, 2018), search refinement mechanisms (Chen and Yao, 2017), or subjective or behavioral shocks. To summarize, we have:

$$\xi_{ij}^z = m_\varepsilon(c_{ij}) + \zeta_{ij}^z$$

Again, we stress that the following assumptions hold:

Assumption 1: (Independence) Inspecting a product $j$ does not lead to information on $\varepsilon_{ik}$ for any $k \neq j$.

Assumption 2: (Invariance) Product values remain unchanged during the search process.

Assumption 3: Consumer observes the values of $\xi_{ij}^u, \xi_{ij}^z$ at the beginning of search.

Assumption 4: $\varepsilon_{ij}$ distributes independently and identically. Consumers know $F^\varepsilon(\cdot)$ but not the value of each $\varepsilon_{ij}$ until product $j$ is inspected.

To express the joint probability, we categorize the sequence observations in two cases. First, we consider the case where the purchased product $h$ is not the last inspected product $J$. We represent reservation values of inspected products (superscript $k$), reservation values of uninspected products (superscript $n$), and purchase values of inspected products except for the purchased one (superscript $k'$) in ordered vectorized forms as follows:

$$\boldsymbol{z}_i^k := (z_{i,J}, z_{i,J-1}, \cdots, z_{i,1})^\top, \; \boldsymbol{z}_i^n := (z_{i,J+1}, \cdots, z_{i,|\mathcal{M}|})^\top, \; \boldsymbol{u}_i^{k'} := (u_{i,1}, \cdots, u_{i,h-1}, u_{i,h+1}, \cdots u_{i,J})^\top$$
$$\boldsymbol{z}_i^k = \vec{\delta}_i^{z,k}(\boldsymbol{X}_i^{z,k}) + \boldsymbol{\xi}_i^{u,k} + \boldsymbol{\xi}_i^{z,k}, \quad \boldsymbol{z}_i^n = \vec{\delta}_i^{z,n}(\boldsymbol{X}_i^{z,n}) + \boldsymbol{\xi}_i^{u,n} + \boldsymbol{\xi}_i^{z,n}, \quad \boldsymbol{u}_i^{k'} = \vec{\delta}_i^{u,k'}(\boldsymbol{X}_i^{u,k'}) + \boldsymbol{\xi}_i^{u,k'} + \boldsymbol{\varepsilon}_i^{k'}$$

Following Proposition 1, the joint probability of the sequence is, therefore:

$$\Pr(\{H, \mathcal{S}, \mathcal{R}, \mathcal{M}\}_i) = \Pr\left( \underbrace{\hat{D}}_{(J+|\mathcal{M}|-1)\times(J+|\mathcal{M}|)} \begin{pmatrix} u_{ih} \\ \boldsymbol{z}_i^k \\ \boldsymbol{z}_i^n \\ \boldsymbol{u}_i^{k'} \end{pmatrix}_{(J+|\mathcal{M}|)\times 1} \leq \boldsymbol{0} \right), \text{ where } \hat{D} = \begin{pmatrix} \hat{D}_1 & \hat{D}_2 \\ \hat{D}_3 & \hat{D}_4 \end{pmatrix}$$

[3] The additivity of the search cost rent and the monotonicity of $m_\varepsilon(\cdot)$ are proved in Appendix B

The difference matrix $\hat{D}$ consists of four blocks:

$$\hat{D}_1 = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix}_{J\times(J+1)}, \quad \hat{D}_2 = \{0\}_{J\times(|\mathcal{M}|-1)}$$

$$\hat{D}_3 = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & 0 \end{pmatrix}_{(|\mathcal{M}|-1)\times(J+1)}, \quad \hat{D}_4 = I_{(|\mathcal{M}|-1)\times(|\mathcal{M}|-1)}.$$

Hence, $\hat{D}$ is of rank $J+|\mathcal{M}|-1$, and its form is determined by the sequence observation.

Now, we consider the case when the purchased product $h$ is the last inspected. Following the vectorized form in the previous case, the joint probability of the sequence is:

$$\Pr(\{H,\mathcal{S},\mathcal{R},\mathcal{M}\}_i) = \Pr\left( \underbrace{\tilde{D}}_{(J+2|\mathcal{M}|-3)\times(J+|\mathcal{M}|)} \begin{pmatrix} u_{ih} \\ \boldsymbol{z}_i^k \\ \boldsymbol{z}_i^n \\ \boldsymbol{u}_i^{k'} \end{pmatrix}_{(J+|\mathcal{M}|)\times 1} \leq \mathbf{0} \right), \text{where } \tilde{D} = \begin{pmatrix} \tilde{D}_1 & \tilde{D}_2 \\ \tilde{D}_3 & \tilde{D}_4 \\ \tilde{D}_5 & \tilde{D}_6 \end{pmatrix}$$

The difference matrix $\tilde{D}$ consists of six parts, in which:

$$\tilde{D}_1 = \begin{pmatrix} 0 & 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix}_{(J-1)\times(J+1)}, \quad D_2 = \{0\}_{(J-1)\times(|\mathcal{M}|-1)},$$

$$\tilde{D}_3 = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 0 \end{pmatrix}_{(|\mathcal{M}|-1)\times(J+1)}, \quad \tilde{D}_4 = I_{(|\mathcal{M}|-1)\times(|\mathcal{M}|-1)}$$

$$\tilde{D}_5 = \begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & -1 & 0 & \cdots & 0 \end{pmatrix}_{(|\mathcal{M}|-1)\times(J+1)}, \quad \tilde{D}_6 = I_{(|\mathcal{M}|-1)\times(|\mathcal{M}|-1)}$$

Hence, $\tilde{D}$ is also of rank $J+|\mathcal{M}|-1$. In the following of this paper, we represent the difference matrix as $D$, which takes one from $\{\hat{D},\tilde{D}\}$ depending on $\mathcal{R}_i$. The joint probability of an observation the baseline model is thus expressed as:

$$\Pr(\{H,\mathcal{S},\mathcal{R},\mathcal{M}\}_i) = \Pr\left(D\begin{pmatrix} u_{ih} \\ \boldsymbol{z}_i^k \\ \boldsymbol{z}_i^n \\ \boldsymbol{u}_i^{k'} \end{pmatrix} \leq \boldsymbol{0}\right) = \Pr\left(D\begin{pmatrix} \xi_{ih}^u + \varepsilon_{ih} \\ \xi_i^{u,k} + \xi_i^{z,k} \\ \xi_i^{u,n} + \xi_i^{z,n} \\ \xi_i^{u,k'} + \varepsilon_i^{k'} \end{pmatrix} \leq -D\begin{pmatrix} \delta_i^u(X_{ih}^u) \\ \vec{\delta}_i^{z,k}(\boldsymbol{X}_i^{z,k}) \\ \vec{\delta}_i^{z,n}(\boldsymbol{X}_i^{z,n}) \\ \vec{\delta}_i^{u,k'}(\boldsymbol{X}_i^{u,k'}) \end{pmatrix}\right) \tag{5}$$

The probability in Equation (5) is expressed in a value-difference form. To my knowledge, this paper is the first to express the joint probability of the sequential search model in this way. The probability expression in Equation (5) consists of two parts: the deterministic component on the right-hand side, which is identifiable from the data, and the stochastic terms on the left-hand side, which depends on structural assumptions. Notice that while we incorporate multiple sources of stochasticity in this basic model, the inclusion and the structural settings of the other stochastic components other than $\varepsilon_{ij}$ depend on the specific model setup. I will discuss this in detail in the next subsection. This value-difference formulation closely resembles the probability structure in standard discrete choice models. Compared to the probit and ranked data probabilities explored in Sections 5.6.3 and 7.3.2 of Train (2009), the only distinction lies in our use of a different full-rank difference matrix for the partial ranking structure.

## 4.2 Identification

Building on the joint probability above, we explore identification issues in sequential search models. For the current empirical study, we assume a linear deterministic component and concentrate on identifying the structure of the stochastic part. In the literature, identification is often limited to heuristic approaches that link data variations to specific model parameters. Although Morozov et al. (2021) and Ursu et al. (2024) provide more formal discussions, they rely solely on conditional probabilities due to the inherent complexity of joint probabilities under the OSR structure. This limitation results in some ambiguities, as multiple parameters may produce indistinguishable changes in a single conditional probability, while multiple subevents within the search process can contribute to identifying a single parameter. For example, stopping decisions depends on preferences, search costs, and the scale of uncertainties simultaneously, and preferences can also be identified through consumers' ranking decisions.

The joint probability expression in Equation (5) allows us to consider identification in sequential search models analogously to discrete choice models. This approach enables broader identification arguments, clarifies necessary normalizations for empirical applications, and sheds light on additional stochastic components and assumptions needed for model extensions. While detailed identification issues may vary depending on model specifications, we focus on two foundational principles of discrete choice models in this paper: "Only differences in utility matter" and "The scale of utility is arbitrary." The two principles correspond to location and scale normalization in discrete choice models, which remain essential in sequential search models.

Let us consider the first point. In our baseline model specification, the absolute value of reservation and purchase values is irrelevant because of the value-difference form. If we add a constant to the fixed component of all values (i.e., to the right-hand side of the inequality) and cancel it out through the difference matrix, it will affect neither the consumer's behavior nor the joint probability of the observed data.

This irrelevance holds under the assumption that both reservation and purchase values are composed of a deterministic component and a conditionally independent stochastic component. This ensures the validity of the second equality in Equation (5). The stochasticity of the purchase value is guaranteed with $\varepsilon_{ij}$, while the stochasticity of the reservation values needs further assumption. Without the stochasticity, adding a constant to deterministic components is not irrelevant, leading to two critical issues. First, incorporating alternative-specific constants as controls is practically challenging, as both the relative and absolute sizes of these constants affect the estimation of other parameters. Second, the lack of consumer-specific controls causes a uniform shift in all product values, resulting in omitted variable bias in estimation. As in discrete choice models, avoiding these issues is essential when applying sequential search models [4]. In more complicated extensions of the baseline model, any new values introduced to consumers' ranking should also incorporate a stochastic component that remains conditionally independent from other options.

Now, we address the second point, "The scale of utility is arbitrary," with the linearity assumption of the deterministic components. In discrete choice models, scale is typically irrelevant in a linear preference model, and we achieve identification with a scale normalization on the variance of the error component. However, in the case of sequential search models, scale is relevant to the estimation outcomes in search costs and heteroskedasticity, even in a linear deterministic specification. It results from that the distribution of $\varepsilon_{ij}$ affects the reservation value through the search cost rent $m_\varepsilon(c_{ij})$. In general, $m_\varepsilon(c_{ij})$ is not linear. Without additional assumptions on $F^\varepsilon(\cdot)$, a scale change in $m_\varepsilon(c_{ij})$ does not lead to an identical scale change in $c_{ij}$. Hence, even if one can identify $\xi_{ij}^z$, one cannot identify $c_{ij}$ without the identification to $F^\varepsilon(\cdot)$. However, due to various challenges, the identification of $F^\varepsilon(\cdot)$ in empirical studies is often difficult, necessitating reliance on additional distributional assumptions.

We illustrate this difficulty with two examples in empirical works. In Kim et al. (2010), they considered a context where the consumer has a pre-search mean-zero taste shock before revealing the product unknowns via inspection. The taste shock is noted by $\xi_{ij}$ and affects both

[4] Many studies in the search literature note that without additional stochasticity in reservation values, inspection order is fully determined by given parameters, causing identical preferences and search costs to result in identical inspection orders. While this issue does exist, it is not entirely insurmountable. When the number of products is small, introducing more control variables and heterogeneity in preference parameters can increase the uncertainty dimension and mitigate this issue.

the reservation and the purchase values. Specifically, their model is:

$$\begin{aligned} u_{ij} &= X_{ij}\beta_i + \xi_{ij} + \varepsilon_{ij} \\ c_{ij} &= c \\ z_{ij} &= X_{ij}\beta_i + \xi_{ij} + m_\varepsilon(c) \end{aligned} \tag{6}$$

Their specification introduces $\xi_{ij}^u = \xi_{ij}$ as the unobserved stochasticity while assuming zero standard deviation for $\xi_{ij}^z = m_\varepsilon(c)$. Hence, we decompose the joint probability in the stacked vectorized form as following:

$$\Pr(\{H,S,R,\mathcal{M}\}_i) = \Pr\left( D \begin{pmatrix} \xi_{ih} + \varepsilon_{ih} \\ \xi_i^k \\ \xi_i^n \\ \xi_i^{k'} + \varepsilon_i^{k'} \end{pmatrix} \leq -D \begin{pmatrix} X_{ih}\beta_i \\ \boldsymbol{X}_i^k \beta_i + \vec{m}(\boldsymbol{c}) \\ \boldsymbol{X}_i^n \beta_i + \vec{m}(\boldsymbol{c}) \\ \boldsymbol{X}_i^{k'} \beta_i \end{pmatrix} \right).$$

The identification arguments for this specification are detailed in Ursu et al. (2024) and summarized here for comparison. Following Berry and Haile (2014), the random coefficient setup can be identified if the distributions of $\varepsilon_{ij}$ and $\xi_{ij}$ are known. Assuming both follow mean-zero i.i.d. normal distributions, we discuss whether additional assumptions on their standard deviations ($\sigma_\varepsilon$ and $\sigma_\xi$) are required.

We first note that in this specification, $\xi_{ij}$ appears in both purchase and reservation values. Rescaling $\xi_{ij}$ does not alter the relative scale of the linear parameters or $m_\varepsilon(c_{ij})/\sigma_\xi$, making scale normalization on $\xi_{ij}$ necessary. In contrast, $\sigma_\varepsilon$ governs the nonlinear relationship between search costs and the rents in reservation values, while also affecting purchase probabilities. Although $\sigma_\varepsilon$ could theoretically be identified without additional assumptions, empirical studies have consistently reported challenges in its estimation without supplementary information (Yavorsky et al., 2021; Morozov et al., 2021; Greminger, 2024; Ursu et al., 2024). The reason for this difficulty is unclear within the traditional OSR structure but becomes evident in our representation of the joint probability: the presence of $\xi_{ij}$ introduces an unobserved correlation between purchase and reservation values. As shown by Keane (1992) in the context of multinomial probit models, identifying heteroskedasticity in such settings is highly challenging without exclusion restrictions on observable regressors. This means for the identifictaion of $\sigma_\varepsilon$, there must be some observable attributes affecting reservation values but not purchase values. Since the deterministic component of the reservation value shares the expected purchase value for the same product, imposing such restrictions is usually economically unreasonable. Consequently, many empirical studies rely on the assumption that $\varepsilon_{ij}/\sigma_\xi = 1$. An exception is Yavorsky et al. (2021), which introduces additional search cost shifters that affect only reservation values, successfully imposing exclusion restrictions and achieving identification to $\sigma_\varepsilon$.

The other specification is proposed in Chung et al. (2024). They introduce stochasticity to the reservation value through heterogeneous search costs across products. Their model is given by the following:

$$
\begin{aligned}
u_{ij} &= X_{ij}\beta_i + \varepsilon_{ij} \\
c_{ij} &\sim \text{Exp}(c_0) \\
z_{ij} &= X_{ij}\beta_i + \xi_{ij}^{z} = X_{ij}\beta_i + m_{\varepsilon}(c_{ij})
\end{aligned}
$$

Their specification assumes a stochastic $\xi_{ij}^{z}$ with a random search cost. The joint probability can then be written as:

$$
\Pr(\{H,S,R,\mathcal{M}\}_i) = \Pr\left( D \begin{pmatrix} \varepsilon_{ih} \\ \xi_i^{z,k} \\ \xi_i^{z,n} \\ \varepsilon_i^{k'} \end{pmatrix} \leq -D \begin{pmatrix} X_{ih}\beta_i \\ \boldsymbol{X}_i^{k}\beta_i \\ \boldsymbol{X}_i^{n}\beta_i \\ \boldsymbol{X}_i^{k'}\beta_i \end{pmatrix} \right)
$$

In this specification, the preference parameters $\beta_i$ are still not sensitive to the scale change, with their relative scale remaining stable. On the other hand, the independence between $\varepsilon_{ij}$ and $c_{ij}$ eliminates the correlation between reservation and purchase values of the same product, the primary challenge for identification. Hence, the heteroskedasticity can be correctly identified, and the parameter estimation is not affected by weak identification caused by unobservable correlations[5]. However, a new issue arises: the mean and variance of $\xi_{ij}$ are jointly influenced by the variance of $\varepsilon_{ij}$ and the mean and variance of $c_{ij}$. More importantly, the shape of $\xi_{ij}$'s distribution is also affected by the mean and variance of $c_{ij}$, making the identification of $c_{ij}$ exceedingly difficult. These two parameters not only determine the mean and variance of $\zeta$, but also affect the shape of the distribution, which imposes additional difficulty to identification. Therefore, Chung et al. (2024) propose strong assumptions in their working draft and published version, respectively: that $c_{ij}$ follows a log-normal distribution with a known variance, or an exponential distribution. Both assumptions ensure that the distribution of $\zeta_{ij}$ is determined by a single parameter to be estimated, thereby addressing the model's identification issue.

By decomposing Equation (4), we find that identifying sequential search models shares a similar foundation with standard discrete choice models but involves greater complexity due to the stochasticity of reservation values. First, the "only differences in utility matter" condition requires that incorporating reservation values introduces additional stochasticity. Second, the "scale of utility is arbitrary" condition does not extend to search costs or the heteroskedasticity arising from the added stochasticity. Since researchers often rely on additional assumptions about the distribution of $\varepsilon_{ij}$ or $c_{ij}$, which are not neutral for estimating search costs, it is essential to interpret search cost estimation results with caution.

[5]Table 4 in Chung et al. (2024) reports the empirical validation of the comparison.

### 4.3 Estimation

As we introduced earlier, when presented under the PR structure, the sequential search model lies between the standard discrete choice model and the ranked model. Both structures, even if their joint probabilities do not have a closed-form solution, can be estimated with simulated maximum likelihood estimation using a GHK-style simulator. The same approach naturally applies to the sequence search model. Take the specification of Chung et al. (2024) as an example, the implementation procedure is as follows.

1. Draw preference heterogeneity to obtain $\beta_i^d$. Draw $\varepsilon_{ih}$ to determine $u_{ih}^d$ for each draw.
2. If $h \neq J$, draw $c_{iJ}$ conditional on $z_{iJ} > u_{ih}^d$ and compute $p_{i1}^d = \Pr(z_{iJ} \geq u_{ih}^d)$;
   if $h = J$, draw $z_{iJ}$ randomly and assign $p_{i1}^d = 1$. Determine $y_i^d = \min\{u_{ih}^d, z_{iJ}^d\}$.
3. Sequentially draw $c_{i,J-1}, \cdots, c_{i2}$ to determine $z_{i,J-1}^d, \cdots, z_{i2}^d$ conditional on $z_{ij} > z_{i,j+1}^d$. Compute $p_{i2}^d = \prod_{1 \leq j \leq J-1} \Pr(z_{ij} \geq z_{i,j+1}^d)$.
4. Compute $p_{i3}^d = \prod_{J < k \leq |\mathcal{M}_i|} \Pr(z_{ik} < y_i^d)$ and $p_{i4}^d = \prod_{1 \leq j \leq J, j \neq h} \Pr(u_{ij} < y_i^d)$.
5. Compute the likelihood contribution of each draw $L_i^d = p_{i1}^d \cdot p_{i2}^d \cdot p_{i3}^d \cdot p_{i4}^d$. Take average across draws.

Here, we do not compare the performance of the GHK method with other methods used in the empirical literature, such as the Kernel-Smooth Frequency Simulator or the Crude Frequency Simulator, because the performance of our estimator is very close to the GHK method proposed by Chung et al. (2024). We attribute the credit to Ursu et al. (2024) and Chung et al. (2024), which conduct extensive simulation-based validation between estimation methods.

Compared to the other GHK method, our approach was established on an entirely different structure, allowing us to simplify the implementation process, whose complexity was a significant drawback under the OSR structure. The OSR-based GHK simulators construct the likelihood function by enumerating, decomposing, and recombining the inequality conditions from all optimal search rules. However, the implementation complexity depends on whether the decomposition or combination is sufficiently straightforward, which depends on the authors' choice. In Jiang et al. (2021) and Chung et al. (2024), the authors enumerate three or four cases for various sequences and devise unique implementation procedures for each case. In contrast, we find such case-by-case analysis unnecessary under the PR structure.

The core idea of the PR-based GHK method is not to decompose and recombine the optimal search rules but to utilize the sampling techniques of the GHK method to replicate the observed search process in the data and calculate its probability. The implementation is divided into two parts. The first part is based on the Distribution and Ranking Conditions, aiming to replicate the partial ranking relationships revealed in the sequential observations. The second part involves the two Choice Conditions, which calculate the probabilities of product values that do not attract the consumer to inspect or purchase. The two parts are conditionally independent of each other, and the streamlined approach avoids the complex case-specific implementation and enhances the flexibility of the estimation. While the implementation is simpler, the performance of this method is almost identical to that of the OSR-base GHK estimator.

# 5 Extension 1: Partial Search Data

In the full model, the implementation logic of the PR-based GHK method is to replicate the observed search process in the data and calculate its probability. This logic can be extended to scenarios where the search process is not fully observed. Given the core value, the PR structure ensures that the observed and censored parts of the ranking are conditionally independent. Therefore, we only calculate the probability that the reservation values satisfy the observed partial ranking conditions without worrying about additional effects related to the optimal search rules of other product values. In this section, I outline several common cases observed in the data and explain how the PR structure addresses these cases. I also demonstrate how the PR-based GHK method enables simulation-based maximum likelihood estimation that fully utilizes partial information.

We first consider the case where a consumer knows the purchase values of certain products before the search begins. These products do not enter the search process, so their reservation values are missing from observation, and only their purchase values are included in the joint probability. If a known product is chosen, it becomes part of the core value; if not, it is smaller than the core value. We formalize it in the following Corollary:

**Corollary 1.** *When a product is known to the consumer (its purchase value is determined) without inspection observed. If it is not purchased, its purchase value follows Choice Condition 2; if it is purchased, all other products follow conditions in Proposition 1.*

Corollary 1 is straightforward but important. One common example of a known product is the outside option. The search sequence data can incorporate consumers who search in the market but ultimately choose to exit. Without purchasing, these consumers' search data naturally identify the outside option's market share without relying on other assumptions. In practice, we often assume that the purchase value of the outside option is revealed after the first inspection to guarantee at least one inspection, while Corollary 1 enables these consumers to be incorporated into estimation as other purchasers. [6]. Another important example is when the purchase values of all products are known without searching. In this case, the partial ranking collapses, and the core value becomes the purchase value of the purchase product. Conditions in Proposition 1, except Choice Condition 1, become trivial, and the PR structure reduces to a standard discrete

[6] However, with only search and purchase data, we cannot identify consumers who opt for the outside option without any inspection.

choice structure. Therefore, with the supplement of Corollary 1, the discrete choice model can be taken as a particular case of the sequential search model.

The second variation is when researchers only observe the set of inspected products but not the exact search path. In this case, we take the aggregation of the probability of all potential inspection orders that lead to the final purchase, and we obtain the following proposition.

**Proposition 2.** *Define $w_{ij} = \min\{z_{ij}, u_{ij}\}$ the Effective Value of product $j$ to consumer $i$. If $w_{iH} \geq w_{iL}$, $\forall L \in \mathcal{M}_i \backslash \{H\}$, then following Proposition 1, $H$ is always inspected and purchased. On contrary, $w_{ih} \geq w_{ij}, \forall j \neq h$ must hold for any $\{H, \mathcal{S}, \mathcal{R}, \mathcal{M}\}_i$ fulfilling conditions in Proposition 1.*

*Proof.* We separate this proposition into three parts. First, a product with the largest effective value must be inspected in any sequence that fulfills the conditions in Proposition 1; second, given that the highest effective value product is inspected, it must be purchased; third, any purchased product in the Rank Conditional Discrete Framework must have the largest effective value in the market.

We start with proving the first part of the proposition. Suppose there is a consumer $i$, a product $H$ satisfying $\min\{z_{iH}, u_{iH}\} \geq \min\{z_{iL}, u_{iL}\}, \forall L \in \mathcal{M} \backslash \{H\}$ is not inspected in her search sequence. Denote the order of the purchased product by $h'$ and the last inspected product by $J'$. According to the Choice Condition 2, $z_{iH} < y_i = \min\{u_{ih'}, z_{iJ'}\}$. If $h' = J'$, $\min\{u_{ih'}, z_{ih'}\} = \min\{u_{ih'}, z_{iJ'}\} > z_{iH} \geq \min\{u_{iH}, z_{iH}\}$; if $h' < J'$, we know $z_{ih'} \geq z_{iJ'}$ holds according to the Distribution Condition. Hence, $\min\{u_{ih'}, z_{ih'}\} > \min\{u_{ih'}, z_{iJ'}\} > z_{iH} \geq \min\{u_{iH}, z_{iH}\}$. In either case, the effective value of the purchased product is larger than that of product $H$, which contradicts $H$'s effective value condition. Therefore, product $H$ is always inspected.

Then, we prove the second part of the proposition. Given that product $H$ is inspected in any sequence $i$, we denote $H$'s position in the sequence by $h$.

- Suppose a product $j' \leq J, j' \neq h$ with $u_{ij'} > \min\{u_{ih}, z_{iJ}\}$ exist. If $h = J$, according to the largest effective value assumption, we have $u'_{ij} > \min\{u_{ih}, z_{ih}\} > \min\{u_{ij'}, z_{ij'}\}$, hence $z_{ij'} < \min\{u_{ih}, z_{ih}\} < z_{ih}$. This violates the Rank Condition. If $h < J$, $u_{ij'} > \min\{u_{ih}, z_{iJ}\}$ is equivalent to $u_{ij'} > u_{ih}$. Because of the Distribution and the Rank Conditions, we know that $z_{ij'} \geq z_{iJ} > u_{ih}$. Combine the two inequalities, we have $\min\{u_{ij'}, z_{ij'}\} > u_{ih} \geq \min\{u_{ih}, z_{ih}\}$. This contradicts the largest effective value assumption on product $H$.

- Suppose a product $k' > J > h$ with $z_{ik'} > \min\{u_{ih}, z_{iJ}\}$ exist. Following Choice Condition 2, we have $\min\{u_{ih}, z_{iJ}\} < z_{ik'} \leq y_i$. The condition contradicts itself when $h$ is purchased, so we assume that a product $h' < J, h' \neq h$ is purchased. In this case, with the Rank Condition and Choice Condition 1, we have $\min\{u_{ih'}, z_{ih'}\} \geq \min\{u_{ih'}, z_{iJ}\} \geq u_{ih} \geq \min\{u_{ih}, z_{ih}\}$, which violates the largest effective value assumption.

Hence, whether product $H$ is the last inspected product or not. $\min\{u_{ih}, z_{iJ}\}$ fulfills the Choice Conditions of the Core Value. Because purchase in the Rank Conditional Discrete Choice framework is unique, $H$ is the purchased product.

Last, we prove the third part of the proposition. When product $h$ is purchased, any inspected product $j \neq h$ fulfills $\min\{u_{ij}, z_{ij}\} \leq u_{ij} < y_i = \min\{u_{ih}, z_{iJ}\}$. Because $h \leq J$, following the Rank Condition, $\min\{u_{ih}, z_{iJ}\} \leq \min\{u_{ih}, z_{ih}\}$. Therefore, $\min\{u_{ij}, z_{ij}\} < \min\{u_{ih}, z_{ih}\}$. Similarly, for any uninspected product $k > J$, we have $\min\{u_{ij}, z_{ij}\} \leq z_{ij} < \min\{u_{ih}, z_{iJ}\} < \min\{u_{ih}, z_{ih}\}$. $\square$

Proposition 2 corresponds to the Eventual Purchase Theorem proposed in Armstrong (2017) and Choi et al. (2018), showing that in the optimal search outcome, a product is purchased if and only if its effective value dominates that of other products. Hence, it is possible to establish a standard discrete choice structure that summarizes consumers' purchasing outcomes without knowing the exact search process. We stress that the PR structure also captures the discrete choice structure. Intuitively, we separate the types of sequence observations conditional on whether the reservation or purchase value dominates the effective value of the purchased product. When $\min\{z_{iH}, u_{iH}\} = u_{iH}$, any product $L \in \mathcal{M} \backslash \{H\}$ satisfies either $z_{iL} \geq u_{iH} \geq u_{iH}$ or $u_{iH} \geq z_{i\ell}$. Let $S_i$ be the set of products following the first condition, and $\bar{S}_i$ be the set of products fulfilling the second condition. This case corresponds to sequences where the consumer purchases a product inspected before the last step, or purchases the last product whose purchase value is smaller than its Reservation Value. When $\min\{z_H, u_H\} = z_H$, any product $\ell \neq h$ satisfies either $u_{i\ell} \geq z_{iH} \geq z_{i\ell}$ or $z_{iH} \geq u_{i\ell}$. Let $S_i$ be the set of products following the second condition and $\bar{S}_i$ be the set of products satisfying the first. This case corresponds to sequences where the consumer purchases the last product whose purchase value exceeds its reservation value.

The Eventual Purchase Theorem is often used to directly derive consumer demand in sequential search contexts. However, its application typically discards the use of search data. The likelihood of the discrete choice model based on the Eventual Purchase Theorem calculates the probability of a single purchase decision among products' effective values, which differs fundamentally from the joint probability in the OSR structure (Equation (3)). This disconnect poses structural challenges when incorporating incomplete search data into the former. Proposition 2 provides a solution to bridge this gap. It demonstrates an intrinsic connection between the sequential search model and the discrete choice model based on the Eventual Purchase Theorem. In the extreme case where search data is entirely unavailable, the effective value of the purchased product serves as the only observable information of the consumer's ranking, indicating that all other products' effective values are lower than that of the purchased product.

On the other hand, the effective value of the purchased product can also act as a weaker ranking criterion when the last inspected product is not observed. Consider the scenario where we only know which products the consumer inspected but have no information about the inspection order. Notice that all sequences satisfying conditions in Proposition 1 that lead to the consideration set $S$ and the purchase of product $H$ if and only if the following conditions are satisfied:

$$\begin{cases} \min\{u_{iH}, z_{iH_0}\} > u_{iL}, \forall L \in S\backslash\{H\} \\ \min\{u_{iH}, z_{iH_0}\} > z_{iL'}, \forall L' \in \bar{S} \\ z_{iH_0} > u_{iH} \text{ if } H_0 \neq H \end{cases} \tag{7}$$

Here, product $H_0$ has the smallest reservation value among all products in $S$. Without loss of generality, we assume that the product is unique. The following proposition holds:

**Proposition 3.** *Conditions (7) are satisfied if and only if the following conditions are satisfied:*

$$\begin{cases} w_{iH} > u_{iL}, \forall L \in S\backslash\{H\} \\ w_{iH} < z_{iL}, \forall L \in S\backslash\{H\} \\ w_{iH} > z_{iL'}, \forall L' \in \bar{S} \end{cases} \tag{8}$$

*Proof.* First, notice that $\min\{u_{iH}, z_{iH}\} = w_{iH}$ in Conditions (8) satisfy $w_{iH} \geq w_{iL}, \forall L \in \mathcal{M}\backslash\{H\}$. With Proposition 2, product $H$ is also inspected and purchased under Conditions (8).

Consider the sufficiency. If $H_0 = H$, the first and third inequalities in Conditions (8) are immediately satisfied. For any product $L$ with $L \in S\backslash\{H\}$, $L$ is inspected. Following Proposition 1, $L$ is inspected before $H_0$ and $z_{iL} > z_{H_0} = z_{iH} \geq \min\{u_{iH}, z_{iH}\}$. If $H_0 \neq H$, we have $z_{iH} > z_{iH_0}$. So $\min\{u_{iH}, z_{iH}\} \geq \min\{u_{iH}, z_{iH_0}\} > u_{iL}, \forall L \in S\backslash\{H\}$ and $\min\{u_{iH}, z_{iH}\} \geq \min\{u_{iH}, z_{iH_0}\} > z_{iL'}, \forall L' \in \bar{S}$. For any product $L$ with $L \in S\backslash\{H\}$, $L$ is inspected. Following Proposition 1, $L$ is inspected before $H_0$ and $z_{iL} > z_{H_0} > u_{iH} \geq \min\{u_{iH}, z_{iH}\}$.

Consider the necessity. If $H_0 = H$, the first and second inequalities in Conditions (7) are immediately satisfied. If $H_0 \neq H$, according to the definition of $H_0$, we have $z_{iH_0} < z_{iH}$; $H_0$ is inspected, so $H_0 \in S\backslash\{H\}$. Therefore, $\min\{u_{iH}, z_{iH}\} < z_{iH_0} < z_{iH}$, implying $u_{iH} < z_{iH_0} < z_{iH}$, which is the third inequality in Condition (7). In addition, $\min\{u_{iH}, z_{iH}\} = \min\{u_{iH}, z_{iH_0}\} = u_{iH_0}$. The first and second inequalities in Conditions (7) are immediately fulfilled. □

Proposition 3 demonstrates that when the last inspected product is unknown, the effective value of the purchased product can provide incomplete search information. For a product with no specific ranking information, the relationship between its reservation value and core value determines when it is searched, while its relationship with the effective value of the purchased product determines whether it is searched at all. The latter can still be incorporated into estimation with other incomplete search information when the former is unavailable, though its information quality is less precise.

Propositions 1, 2, and 3 provide sufficient flexibility for us to make full use of the incomplete search information observed in the data. We conclude this extension with a specific scenario: we can only observe the first product consumers inspect and the product they eventually. In this case, the search data is partially missing, and applying the optimal search rules seems impossible without a full simulation of the search process. To utilize the remaining part of the search path, we can reconstruct the conditions in Proposition 1 and implement the estimator as follows:

1. Draw heterogeneities to determine $\delta_i(\cdot)$. Draw $\varepsilon_{ih}$ to determine $u_{ih}^d$ for each draw.
2. If $h \neq 1$, draw $z_{ih}$ conditional on $z_{ih} > u_{ih}^d$ and draw $z_{i1}$ conditional on $z_{i1} > z_{ih}^d$ and compute $p_{i1}^d = \Pr(z_{ih} \geq u_{ih}^d) \cdot \Pr(z_{i1} \geq z_{ih}^d)$;
   if $h = 1$, draw $z_{i1}$ randomly and assign $p_{i1}^d = 1$. Determine $w_i^d = \min\{u_{ih}^d, z_{ih}^d\}$.
3. If $h \neq 1$, compute $p_{i2}^d = \Pr(u_{i1} \leq u_{ih}^d)$; if $h = 1$, assign $p_{i2}^d = 1$.
4. Draw $c_{ij}$ for all $j \neq 1$ and $j \neq h$ conditional on $z_{ij} < z_{i1}^d$, compute $p_{i3}^d = \prod_{j\neq 1,h} \Pr(z_{ij} \geq z_{i1}^d)$.
5. Compute $p_{i4}^d = \prod_{k\in\{j:z_{ij}^d > w_i^d\}} \Pr(u_{ik} \leq w_i^d)$.
6. Compute the likelihood contribution of the draw $L_i^d = p_{i1}^d \cdot p_{i2}^d \cdot p_{i3}^d \cdot p_{i4}^d$. Take average across draws.

We observe that, even without knowing the last inspected product, combining the effective value of the purchased product with the reservation value of the first inspected product helps reconstruct the observable ranking information available in the data. We validate the effectiveness of the method with the following specification:

$$
\begin{aligned}
u_{ij} &= \sum_{s=1}^{3} \gamma_t x_j^s + \beta_i p_{ij} + \varepsilon_{ij}, \quad \text{where } \beta_{ij} \sim N(\bar{\beta}, \sigma_\beta^2); \\
c_{ij} &= \exp(c_{ij}^0), \quad \text{where } c_{ij}^0 \sim N(\bar{c_0}, \sigma_c^2); \\
z_{ij} &= \sum_{s=1}^{3} \gamma_t x_j^s + \beta_{ij} p_j + m_\varepsilon(c_{ij}).
\end{aligned} \tag{9}
$$

Here, $x_j^t$ are dummy product attributes, while $p_j$ indicate product prices. Following Section 4.2, we impose distributional assumptions of $\sigma_\varepsilon = 1$ and $\sigma_c = 0.25$. We obtain the following Monte Carlo simulation results as follow:

TABLE 1 – Monte Carlo Simulation Results with Purchase and First Inspection

| | True value | Estimates | |
|---|---|---|---|
| $\gamma_1$ | 1 | 0.984 | (0.094) |
| $\gamma_2$ | 0.5 | 0.504 | (0.027) |
| $\gamma_3$ | -0.2 | -0.177 | (0.071) |
| $\bar{\beta}$ | -0.6 | -0.584 | (0.096) |
| $\sigma_\beta$ | 0.2 | 0.215 | (0.206) |
| $\bar{c_0}$ | -1.5 | -1.407 | (0.092) |
| N: | | 10000 | |
| D: | | 1000 | |

*Notes:* Data are simulated for 10,000 consumers, and the reported results are obtained after averaging across 100 estimations with different seeds and with 1,000 error draws each. The standard deviation of the mean estimate across these simulations is reported in parentheses.

This process effectively utilizes information from the initial inspection with only one additional sampling (to $z_{i1}$). It offers a practical approach to studying consumer search behavior when search data is incomplete, enabling researchers to leverage available search information while avoiding excessively computation-intensive simulations.

It is important to note that our method requires the availability of final purchase information in the data. If such information is missing, it may be necessary to infer the final purchased product through sampling. For instance, if only three products in the consumer's search sequence are known, researchers would need to compute four joint probabilities: three for the consumer purchasing each of the three products and one for not purchasing any of them. These probabilities would then be summed up for estimation.

# 6 Extension 2: Additional Information and Structural Change

So far, our focus has been primarily on the baseline model. Beyond the baseline, the framework of sequential search models can incorporate additional components, such as extra consumer ranking information, multiple choices, and structural changes in ranking conditions. For traditional OSR structures, accommodating such additional information can be challenging, as it may require imposing new optimal search rules to describe consumer behavior under the given decision environment, thereby complicating the joint probability and the estimation process. However, this limitation does not apply to the PR structure. Under the Independence assumption, these additional pieces of information can be flexibly utilized in the estimation process.

Let us first consider the case of additional consumer ranking. Imagine a scenario where, during the search process, a consumer adds a product $\ell$ to their "My Favorites" list for the first time. Since no other products in "My Favorites" are available for comparison, and the action incurs no cost, we can assume that, up to this point, product $\ell$ has the highest purchase value among all inspected products. As an additional ranking condition, we can construct the estimator as follows:

1. Draw preference heterogeneities to determine $\beta_i$. Draw $\varepsilon_{ih}$ to determine $u_{ih}^d$ for each draw.
2. If $h \neq J$, draw $c_{iJ}$ conditional on $z_{iJ} > u_{ih}^d$ and compute $p_{i1}^d = \Pr(z_{iJ} \geq u_{ih}^d)$;
   if $h = J$, draw $z_{iJ}$ randomly and assign $p_{i1}^d = 1$. Determine $y_i^d = \min\{u_{ih}^d, z_{iJ}^d\}$.
3. Sequentially draw $c_{i,J-1}, c_{i,J-2}, \cdots, c_{i1}$ to determine $z_{iJ}^d, \cdots, z_{i1}^d$ conditional on $z_{ij} > z_{i,j+1}^d$. Compute $p_{i2}^d = \prod_{j \leq J-1} \Pr(z_{ij} \geq z_{i,j+1}^d)$.
4. If $h = \ell$, follow the rest of the baseline estimator implementation procedure; if $h > \ell$, proceed with the following.
5. Draw $\varepsilon_{i\ell}$ conditional on $u_{i\ell} < y_i^d$.
6. Compute $p_{i3}^d = \prod_{k>J} \Pr(z_{ik} < y_i^d)$, $p_{i4}^d = \prod_{j<\ell} \Pr(u_{ij} < u_{i\ell}^d)$, $p_{i5}^d = \prod_{\ell<j<J, j\neq h} \Pr(u_{ij} < y_i^d)$.
7. Compute the likelihood contribution of the draw $L_i^d = p_{i1}^d \cdot p_{i2}^d \cdot p_{i3}^d \cdot p_{i4}^d \cdot p_{i5}^d$. Take average across draws.

Following the previous specification in (10), the Monte Carlo simulation results as follows:

TABLE 2 – Monte Carlo Simulation Results with First My Favorite

| | True value | Estimates | |
|---|---|---|---|
| $\gamma_1$ | 1 | 0.966 | (0.137) |
| $\gamma_2$ | 0.5 | 0.507 | (0.041) |
| $\gamma_3$ | -0.2 | -0.179 | (0.096) |
| $\bar{\beta}$ | -0.6 | -0.557 | (0.137) |
| $\sigma_\beta$ | 0.2 | 0.226 | (0.302) |
| $\bar{c_0}$ | -1.5 | -1.356 | (0.114) |
| N: | | 10000 | |
| D: | | 1000 | |

*Notes:* Data are simulated for 10,000 consumers, and the reported results are obtained after averaging across 50 estimations with different seeds and with 1,000 error draws each. The standard deviation of the mean estimate across these simulations is reported in parentheses.

It is important to note that while new information is introduced, our method does not require additional optimal search rules to process this information, which can be a notable challenge for the other estimators based on the OSR structure.

Finally, we note that the estimation method based on the PR structure can be applied in scenarios where pure structural changes occur during the search process. Such structural changes are characterized by modifications to the consumer's ranking conditions without altering their search strategy before and after the change.

A representative example is the consumer search-and-product-discovery model proposed by Greminger (2022). In this model, when consumers are dissatisfied with the remaining products in their market awareness, they can choose to incur a cost to expand their awareness set ($\mathcal{M}_i$) through a specific route, such as visit the sales page of rolling down the current list. The expansion of the awareness set is defined as the behavior of *discovery*. This action introduces reservation values for the newly discovered products at the point of discovery, integrating them into the consumer's ranking of values. Consequently, this update creates a structural change in the original ranking conditions with an updated discovery option.

For this extended model, Greminger (2022) provides the corresponding optimal search rules. First, he establishes that other products' purchase and reservation values remain unaffected by the possibility of future discovery actions. Beyond these two values, he demonstrates that the discovery action itself can also be characterized by a discovery reservation value, which satisfies the following condition:

$$c_{ir}^d = \int_{q_{ir}^d}^{\infty} [1 - G_{ir}(w)]dw$$

Here $q_{ir}^d$ is the discovery reservation value, $c_{ir}^d$ is the discovery cost of a discovery route $r$ for consumer $i$, and $G_{ir}(w)$ is the cdf of consumers' expectation of the largest effective value obtained in one discovery. Greminger (2022) proves that consumer $i$'s discovery reservation value on route $r$ for the $t$-th time takes the form of:

$$q_{irt} = \Theta_i(\mathrm{E}(X_{ijr}), \mathrm{Var}(X_{ijr}), c_{ijr}^{ins}, c_{ir}^{dis}, n_r) + \tau_{irt}, \quad \text{where } \Pr(\tau_{irt} < x) = F^{\tau}(x)$$

Here, $\Theta_i(\cdot)$ is a deterministic function of the empirical mean of product characteristics on route $r$, the empirical variance of product characteristics on route $r$, the inspection search cost, the discovery cost, and the number of discovered products within one discovery. Notice that we allow an external stochasticity of $\tau_{irt}$ with known distribution for the discovery behavior. The reason is the same as the stochastic reservation value specified in Section 4.2: to maintain the identification of the modified model.

Greminger (2022)'s extension transforms the sequential search model into one with multiple structural changes throughout the search sequence. Despite this complexity, it can still be addressed using the estimation method under the PR structure. As long as the timing of each discovery action is observable in the data, the consumer's search process can be segmented with discovery, and the ranking conditions for each segment are observed. In this approach, we calculate a core value for each segment of the search process, starting from the final segment and working backward to reconstruct the ranking conditions. It is important to note that the ranking conditions in each segment depend only on the products discovered up to that point. Similarly, the purchase values of products not purchased in each segment are tied to the core value of the current and following segments. This segment-based decomposition ensures that the estimation remains manageable even when multiple structural changes occur during the search process.

We show the effectiveness of our estimation with a Monte-Carlo simulation with the following specification:

$$
\begin{aligned}
& u_{ij} = \sum_{s=1}^{3} \gamma_s x_j^s + \beta p_{ij} + \xi_{ij}^u + \varepsilon_{ijr}, \quad \text{where } \xi_{ij}^u \sim \mathcal{N}(0,1); \\
& c_{ij}^{ins} = c^{ins} = \exp(c^0); \\
& z_{ij} = \sum_{s=1}^{3} \gamma_s x_j^s + \beta p_{ij} + \xi_{ij}^u + m_\varepsilon(c^{ins}); \\
& \log(c_{irt}^{dis}) \sim N(c^1, \sigma_c^2 = 0.25^2); \\
& q_{irt} = \Theta_i(\mathrm{E}_r(X_{ir}, P_{ir}, n_d), \mathrm{Var}_r(X_{ir}, P_{ir}, n_d), c^{ins}, c_{irt}^{dis}).
\end{aligned}
\tag{10}
$$

As the discovery cost $c_{irt}^{dis}$ is stochastic, we do not need $\tau_{irt}$ for identification.

We assume 2000 consumers search in a market with 1000 products; among them, 600 products are in Route 1, and 400 in Route 2, with Route 2 having lower prices but larger variances. Each consumer initially meets with a market of only 1 product and an outside option while can discover up to a maximum of 15 products randomly assigned from the two routes. Each time of discovery reveals $n_d = 2$ products unless there remains only 1 product not discovered on this route. Appendix C summarizes the detailed implementation procedure. The Monte Carlo Estimation results are shown in the following:

TABLE 3 – Monte Carlo Simulation Results with First My Favorite

| | True value | Estimates | |
|---|---|---|---|
| $\gamma_1$ | 0.3 | 0.292 | (0.034) |
| $\gamma_2$ | 0.2 | 0.180 | (0.058) |
| $\gamma_3$ | 0.1 | 0.096 | (0.037) |
| $\beta$ | -0.6 | -0.572 | (0.017) |
| $c^0$ | -2 | -1.953 | (0.047) |
| $c^1$ | -2.5 | -2.474 | (0.052) |
| | | | |
| N: | | 2000 | |
| D: | | 1000 | |

*Notes:* Data are simulated for 2,000 consumers, and the reported results are obtained after averaging across 100 estimations with different seeds and with 1,000 error draws each. The standard deviation of the mean estimate across these simulations is reported in parentheses.

The empirical applications of the search and product discovery model can also be found in Greminger (2024) and Zhang et al. (2023). Greminger (2024) uses a different data structure, as his dataset does not contain consumer search sequences, while Zhang et al. (2023) employs the Kernel Smooth Frequency Simulator. Since Greminger (2022) provides alternative Optimal Search Rules for structural changes related to product discovery, it is feasible to construct the Kernel Smooth Frequency Simulator for estimation. However, its estimation performance is inferior to the PR-based GHK estimation method, particularly in estimating search costs and discovery costs.

Finally, we emphasize that our estimation method applies not only to cases where additional optimal search rules can be derived but also to situations with observable structural changes where no extra optimal search rules are available. Klein et al. (2024) examine the process of preference discovery in a sequential search context, assuming that consumers receive an unpredictable signal after making a choice (e.g., entering the checkout page), leading to a shift in their preferences. Such changes alter all product values as well as consumers' ranking in the market, requiring them to revise their search decisions based on updated preferences. For such a complex composite sequential search process, the computational and implementation shortcomings of existing estimation methods are amplified, rendering them ineffective for estimation. In contrast, the Partial Ranking-based GHK estimation method provides a uniquely effective and computationally efficient solution.

# 7 Conclusion

The primary goal of this paper is to reduce the empirical constraints researchers face when studying consumers' search and purchase behavior. While significant efforts have been made to enhance the rigor and feasibility of the widely applied sequential search model, its lack of simplicity and flexibility often places researchers in a dilemma between using search data with complex estimation challenges and ignoring it altogether.

This paper reconstructs Weitzman (1979)'s description of the optimal solution for search models, introducing a partial ranking structure to represent the solution. Its most practical contribution is removing the reliance on Optimal Search Rules for handling consumer search data in practice, instead translating consumers' sequential actions directly into their rankings to product values.

For empirical researchers, this means that the interdependencies within the search model, given the final purchase, are no longer an obstacle. When analyzing subsequent decisions, researchers no longer need to account for whether later inspections or purchases were made upon uncertainties observed in earlier inspections, transforming the search model into a simple static framework. This transformation not only formalizes the joint probability for discussing identification and simplifying estimation, but also offers flexibility in handling various types of data, such as incomplete search data, supplemental data, and structural changes in the search process.

The partial ranking structure lays a new foundation for applying sequential search models in empirical studies, particularly in analyzing strategies or policies that influence the search process. The PR-based GHK estimation method proposed under this structure is not only simple to implement but also significantly enhances the flexibility and adaptability of utilizing observed search data. Leveraging search data for research holds great potential in the digital age, and the work in this paper lowers the barrier to broader applications of search data.

# References


Armstrong, M. (2017). Ordered consumer search. *Journal of the European Economic Association*, 15(5):989–1024.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of political Economy*, 112(1):68–105.

Berry, S. T. and Haile, P. A. (2014). Identification in differentiated products markets using market level data. *Econometrica*, 82(5):1749–1797.

Chen, Y. and Yao, S. (2017). Sequential search with refinement: Model and application with click-stream data. *Management Science*, 63(12):4345–4365.

Choi, M., Dai, A. Y., and Kim, K. (2018). Consumer search and price competition. *Econometrica*, 86(4):1257–1281.

Chung, J. H., Chintagunta, P., and Misra, S. (2024). Simulated maximum likelihood estimation of the sequential search model. *Quantitative Marketing and Economics*, pages 1–60.

Ghose, A., Ipeirotis, P. G., and Li, B. (2019). Modeling consumer footprints on search engines: An interplay with social media. *Management Science*, 65(3):1363–1385.

Greminger, R. P. (2022). Optimal search and discovery. *Management Science*, 68(5):3904–3924.

Greminger, R. P. (2024). Heterogeneous position effects and the power of rankings. *arXiv preprint arXiv:2210.16408v4*.

Hajivassiliou, V. A. and Ruud, P. A. (1994). Classical estimation methods for ldv models using simulation. *Handbook of econometrics*, 4:2383–2441.

Honka, E. and Chintagunta, P. (2017). Simultaneous or sequential? search strategies in the us auto insurance industry. *Marketing Science*, 36(1):21–42.

Jiang, Z., Chan, T., Che, H., and Wang, Y. (2021). Consumer search and purchase: An empirical investigation of retargeting based on consumer online behaviors. *Marketing Science*, 40(2):219–240.

Keane, M. P. (1992). A note on identification in the multinomial probit model. *Journal of Business & Economic Statistics*, 10(2):193–200.

Keller, G. and Oldale, A. (2003). Branching bandits: a sequential search process with correlated pay-offs. *Journal of Economic Theory*, 113(2):302–315.

Kim, J. B., Albuquerque, P., and Bronnenberg, B. J. (2010). Online demand under limited consumer search. *Marketing Science*, 29(6):1001–1023.

Klein, T., Walsh, C., and Zhang, T. (2024). Do i really want to buy this? preference discovery and consumer search. *Working Paper*.

Morozov, I., Seiler, S., Dong, X., and Hou, L. (2021). Estimation of preference heterogeneity in markets with costly search. *Marketing Science*, 40(5):871–899.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Ursu, R., Seiler, S., and Honka, E. (2024). The sequential search model: A framework for empirical research. *SSRN Electronic Journal*.

Ursu, R. M. (2018). The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions. *Marketing Science*, 37(4):530–552.

Ursu, R. M., Wang, Q., and Chintagunta, P. K. (2020). Search duration. *Marketing Science*, 39(5):849–871.

Ursu, R. M., Zhang, Q., and Honka, E. (2023). Search gaps and consumer fatigue. *Marketing Science*, 42(1):110–136.

Weitzman, M. L. (1979). Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654.

Yavorsky, D., Honka, E., and Chen, K. (2021). Consumer search in the us auto industry: The role of dealership visits. *Quantitative Marketing and Economics*, 19:1–52.

Zhang, X. L., Ursu, R., Honka, E., and Yao, Y. O. (2023). Product discovery and consumer search routes: Evidence from a mobile app. *SSRN Electronic Journal*.

# A Optimal Search Strategy in the Sequential Search Model

This appendix proves that searching and purchasing along the descending order of the reservation values of all products and the purchase values of inspected products at the end of search is the optimal strategy. We start from the "branching bandit" problem in Keller and Oldale (2003), which regards the general sequential search model as a process to find the best outcome of many multi-stage projects. In their framework, inspection action in an alternative decision-making project can be understood as a branching action to a sub-project with information about an alternative revealed, while purchase action in the project is a branching action that leads to a termination and an outcome. Theorem 1 in Keller and Oldale (2003) demonstrates that a myopic strategy that purchases the action with the largest Gittins Index in any project of the problem is optimal, as long as the branching bandit problem satisfies the independence assumption - choosing any branching action does not provide additional information about actions not emanating from it. The optimality applies to our baseline model since it adheres to the independence assumption: inspecting a product with uncertainty does not reveal any information about the purchase value of other products.

It remains to be shown that search and purchase following the ranking of the reservation and purchase values coincide with the Gittins Index Policy. Greminger (2022) highlighted that the definitions of reservation and purchase values proposed in Weitzman (1979) are equivalent to the Gittins index values for inspection and purchase actions. Specific to our baseline model, the Gittins Indices of all actions in the problem remain unaffected by time or other exogenous factors. Additionally, when taking any action at any project, the Gittins index of any other alternative action in branched-off sub-projects remains unchanged, while the action taken can never be chosen in any sub-project. As a result, if the consumer follows the Gittins Index Policy and purchases an action in one project, the Gittins Index of the purchased action must be greater than those of all alternative actions in any project of the problem. Hence, the rank condition applies to the whole problem due to the invariance of the Gittins indices. Combining the rank conditions from all steps, we obtain the complete rank condition for the Gittins Indices observed, i.e., reservation values of all products and the purchase values of inspected products. Hence, searching following the rank condition is the optimal strategy in our baseline model.

# B The Search Cost Rent in a Linear Specification

Denote $\delta_{ij}^u(X_{ij}^1) + \xi_{ij}^u$ in Equation (**??**) by $v_{ij}$, which represent the value of the observed part of product $j$ before inspection. Taking it into Equation (2) leads to:

$$
\begin{aligned}
c_{ij} &= \int_{u_{ij} > \bar{u} - v_{ij}} (\varepsilon_{ij} - (\bar{u} - v_{ij}))\, dF^{\varepsilon}(\varepsilon_{ij}) \\
&= \left(1 - F^{\varepsilon}\left(\frac{\bar{u} - v_{ij}}{\sigma_{\varepsilon}}\right)\right) \int_{\varepsilon_{ij} > (\bar{u} - v_{ijr})}^{\infty} (\varepsilon_{ij} - (\bar{u} - v_{ij})) \frac{f(\varepsilon_{ij})}{1 - F\left(\frac{\bar{u} - v_{ij}}{\sigma_{\varepsilon}}\right)} d\varepsilon_{ij} \\
&= \left(1 - F^{\varepsilon}\left(\frac{\bar{u} - v_{ij}}{\sigma_{\varepsilon}}\right)\right) \cdot \mathrm{E}(\varepsilon_{ij} - (\bar{u} - v_{ij}) \mid \varepsilon_{ij} > (\bar{u} - v_{ij})) \\
&= \left(1 - F^{\varepsilon}\left(\frac{\bar{u} - v_{ij}}{\sigma_{\varepsilon}}\right)\right) \cdot \left[\sigma_{\varepsilon} \cdot \frac{f^{\varepsilon}\left(\frac{\bar{v} - v_{ij}}{\sigma_{\varepsilon}}\right)}{1 - F^{\varepsilon}\left(\frac{\bar{v} - v_{ij}}{\sigma_{\varepsilon}}\right)} - \sigma_{\varepsilon} \cdot \frac{\bar{v} - v_{ij}}{\sigma_{\varepsilon}}\right] \\
&= \sigma_{\varepsilon}\left[f^{\varepsilon}\left(\frac{\bar{v} - v_{ij}}{\sigma_{\varepsilon}}\right) - \frac{\bar{v} - v_{ij}}{\sigma_{\varepsilon}}\left(1 - F^{\varepsilon}\left(\frac{\bar{v} - v_{ij}}{\sigma_{\varepsilon}}\right)\right)\right]
\end{aligned}
$$

We see the above equation is only about $\frac{\bar{u} - v_{ij}}{\sigma_{\varepsilon}}$. In addition, notice that

$$
\frac{\partial \sigma_{\varepsilon}\left[f^{\varepsilon}(x) - x\left(1 - F^{\varepsilon}(x)\right)\right]}{\partial x} = -\sigma_{\varepsilon}\left(1 - F^{\varepsilon}(x)\right)
$$

which is always negative with a finite $x$. Notice that the left-hand side has a positive derivative, it implies a bijection between $\bar{v}$ and $c_{ij}$. Therefore, we have a unique solution of $\bar{v}$, denoted by $z_{ij}$. Define $m_{\varepsilon}(x) = \sigma_{\varepsilon}\left[f^{\varepsilon}(x) - x\left(1 - F^{\varepsilon}(x)\right)\right]^{-1}$, we can represent the expression of reservation value in Equation (**??**) by $z_{ij} = v_{ij} + m_{\varepsilon}(c_{ij})$, with $m_{\varepsilon}(c_{ij})$ a strictly decreasing function.

# C GHK Estimator of the Search and Product Discovery Model

We segment the search sequence as follows: We segment the search sequence as follows: for the $J$ inspected products, denote the last inspected product before the $a$-th discovery as $J(a-1)$. Specifically, $J(0)$ represents the last inspected product, $J(1)$ corresponds to the last inspected product before the second-to-last discovery, and so on. We start from Segmentation 0.

1. Draw $\xi_{ih}^u$ to determine $z_{ih}^d$ for each draw.
2. If $h < J(0)$ or $J(0) = J(1)$, draw $\varepsilon_{ih}$ conditional on $u_{ih} < z_{ih}^d$ to determine $u_{ih}^d$ for each draw, calculate $p_{i,0,0} = \Pr(u_{ih} < z_{ih}^d)$;
   else, draw $\varepsilon_{ih}$ randomly to determine $u_{ih}^d$ for each draw, assign $p_{i,0,0} = 1$;
3. Check whether $J(0) > J(1)$. If not, jump to Step 6.
4. If $h \neq J(0)$, draw $\xi_{i,J(0)}^u$ conditional on $z_{i,J(0)} > u_{ih}^d$ and compute $p_{i,1,0}^d = \Pr(z_{i,J(0)} \geq u_{ih}^d)$; if $h = J(0)$, draw $z_{i,J(0)}$ randomly and assign $p_{i,1,0}^d = 1$. Determine $y_i^d = \min\{u_{ih}^d, z_{iJ}^d\}$.

5. If $h \leq J(1)$, sequentially draw $\xi^u_{i,J(0)-1}, \xi^u_{i,J(0)-2}, \cdots, c_{i,J(1)+1}$ to determine $z^d_{i,J(0)}, \cdots, z^d_{i,J(1)+1}$ conditional on $z^d_{i,h} > z_{i,j} > z^d_{i,j+1}$. Compute $p^d_{i,2,0} = \prod_{J(1)+1\leq j\leq J(0)-1} \Pr(z_{i,h} > z_{i,j} \geq z^d_{i,j+1})$;
   If $J(1) < h < J(1)$, sequentially draw $\xi^u_{i,J(0)-1}, \xi^u_{i,J(0)-2}, \cdots, \xi^u_{i,h-1}$ to determine $z^d_{i,J(0)}, \cdots, z^d_{i,h-1}$ conditional on $z^d_{i,h} > z_{i,j} > z^d_{i,j+1}$, and sequentially draw $\xi^u_{i,h+1}, \cdots, \xi^u_{i,J(1)-1}$ to determine $z^d_{i,h+1}, \cdots, z^d_{i,J(1)-1}$ conditional on $z_{i,j} > z^d_{i,j+1}$. Compute $p^d_{i,2,0} = \prod_{J(1)+1\leq j\leq h-1} \Pr(z_{i,h} > z_{i,j} \geq z^d_{i,j+1}) \prod_{h+1\leq j\leq J} \Pr(z_{i,j} \geq z^d_{i,j+1})$.

6. Compute $y_{i0} = \min\{u_{ih}, z_{i,J(0)}\}$ if $J(0) > J(1)$ and $h = J(0)$, otherwise $y_{i0} = u_{ih}$.

7. Draw $q_{ira}$ for all $r$ such that $q_{ira} < y_{i0}$. These are the discovery values that are not realized. Calculate $p_{i,5} = \prod_r \Pr(q_{ira} < y_{i0})$.

So far, we have accomplished the ranking condition reconstruction for the last segmentation. We want to point out that the Step 5 is complicated because we introduce the heterogeneity of reservation value through $\xi^u_{ij}$ but not search cost $c^{dis}_{ij}$, which leads to correlation between search reservation value and purchase value. This complicates the implementation but largely increase the estimation process efficiency.

8. First, draw the discovery value $q^d_1$ realized in this segmentation with several conditions satisfied:

   - It needs to be larger than the drawn reservation values of products already discovered before or in this segmentation but inspected in later segments;
   - It needs to be larger than all discovery values of other unselected routes;
   - It needs to be smaller than $z^d_{i,h}$ if product $h$ is inspected in this segmentation.
   - It needs to be larger than $u^d_{ih}$ if product $h$ is inspected in or before this segmentation.

   Compute the probability of all these conditions being satisfied with $p_{i,0,1}$.

9. Check whether $J(1) > J(2)$. If not, jump to Step 12.

10. Draw $\xi^u_{i,J(1)}$ conditional on $z_{i,J(1)} > q^d_1$ and compute $p^d_{i,1,1} = \Pr(z_{i,J(1)} \geq u^d_{ih})$;

11. Draw reservation values of products inspected in this segmentation following Step 5, obtain $p^d_{i,2,1}$.

12. Compute $y_{i1} = \min\{q^d_1, z^h_{ij}\}$ if $h$ is discovered but not inspected, otherwise $y_{i1} = q^d_1, z^h_{ij}$.

13. Replace the corresponding discovery value on the route discovered in this segmentation in $q_{ira}$ with $q_{1a}$.

14. Repeat Step 8 to 12 until exhausting the observed sequence upto the last Segmentation $T$.

So far, we finished the recovery of the ranking conditions observed in the search sequence.

15. Compute core value $Y_{it}$ for each $t$ with the minimum $y_{it}$ among all current and later segments.
16. Compute $p_{i,3}$ for the probability of the purchase value of all products inspected but not purchased smaller than the corresponding core value.
17. Compute $p_{i,4}$ for the probability of the search reservation value of all products discovered but not inspected smaller than the corresponding core value.
18. Take products of all $p_{i,0,r}, p_{i,1,r}, p_{i,2,r}, p_{i,3}, p_{i,4}$ and $p_{i,5}$ as the simulated likelihood contribution of the draw. Take average across draws.