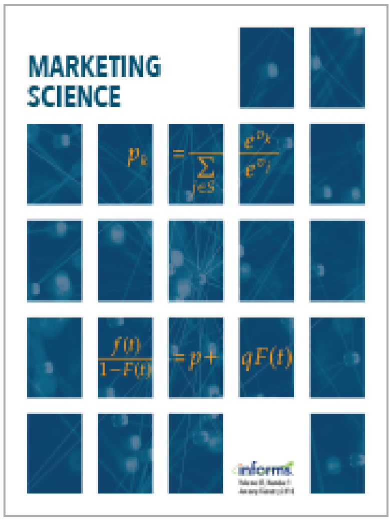


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Online Search and Optimal Product Rankings: An Empirical Framework

Giovanni Compiani, Gregory Lewis, Sida Peng, Peichun Wang

informs
https://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE
Vol. 43, No. 3, May–June 2024, pp. 615–636
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Online Search and Optimal Product Rankings: An Empirical Framework

Giovanni Compiani,[a,*] Gregory Lewis,[b] Sida Peng,[c] Peichun Wang[d]

[a] Booth School of Business, University of Chicago, Chicago, Illinois 60637; [b] Amazon, Boston, Massachusetts 02138; [c] Microsoft, Boston, Massachusetts 02138; [d] Unity, San Francisco, California 94016

*Corresponding author

Contact: giovanni.compiani@chicagobooth.edu, https://orcid.org/0000-0002-4272-5814 (GC); greglewis.work@gmail.com (GL); sida.peng@microsoft.com (SP); peichun21@gmail.com, https://orcid.org/0000-0003-3597-8845 (PW)

Received: February 15, 2022
Revised: June 27, 2023; October 3, 2023
Accepted: October 3, 2023
Published Online in Articles in Advance: December 21, 2023

https://doi.org/10.1287/mksc.2022.0071

Abstract. We study the problem faced by an online retail platform choosing product rankings in order to maximize two distinct goals: consumer surplus and revenues/profits. To this end, we specify a version of the Weitzman sequential search model in which search reveals a consumer's idiosyncratic taste for the product as well as vertical dimensions of its quality, and we derive convenient expressions for consumer surplus and revenues. To optimize consumer surplus, platforms should facilitate product discovery by promoting "diamonds in the rough," that is, products with a large gap between the utility they deliver and what consumers expect based on the presearch information. By contrast, to maximize static revenues, the platform should favor high-margin products, potentially creating a tension between the two objectives. We develop computationally tractable algorithms for estimating consumer preferences and optimizing rankings, and we provide approximate optimality guarantees in the latter case. When we apply our approach to data from Expedia, our suggested consumer surplus–optimizing ranking achieves both higher consumer surplus and higher revenues relative to the Expedia ranking—delivering a Pareto improvement—and also dominates ranking the products in order of utility, which is intuitive but fails to leverage information on what consumers know presearch.

History: Tat Chan served as the senior editor.
Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mksc.2022.0071.

Keywords: consumer search • online platforms • product rankings • recommendation systems

1. Introduction

E-commerce is an important part of the economy and becoming more so. As a result, platforms such as Amazon, Google, and Expedia play an increasingly crucial role in shaping consumer choices. One key tool at their disposal is product rankings, which have the ability to direct consumer attention and, thus, influence purchasing choices. A growing literature explores the impact that rankings have on choices (e.g., Athey and Ellison 2011, Yao and Mela 2011, Ghose et al. 2014, Chen and Yao 2017, Ursu 2018, Choi and Mela 2019, Hodgson and Lewis 2020). Because rankings are often found to be an important driver of consumer decisions, it is natural to investigate what is the best way to rank products on a platform.[1]

Given this, the goal of our paper is to propose and study ranking algorithms that maximize two main objectives: platform profits/revenues—obviously of interest for a platform—and consumer surplus, which is relevant for policy makers as well as for a platform that faces competition and is concerned about consumer churn. One might speculate that having access to top-tier machine learning methods paired with the ability to run A/B tests would mean that this is an already solved problem. Whereas conceptually straightforward, this experimentation strategy runs into two issues. First, the number of possible rankings is very large for typical choice set sizes, and it may be impossible or very costly for platforms to run such a large number of experiments. Second, whereas revenues/profits can be quantified directly from the data, welfare is not immediately observed. This motivates us to pursue a different approach. We first specify a version of the canonical sequential search model of Weitzman (1979) and take that as the ground truth. Then, we use this model to propose algorithms for optimizing consumer surplus and platform revenue. Note that, even with the model primitives in hand, evaluating the target metrics at all possible rankings is intractable given the large number of rankings. Instead, we apply this brute force approach only to the top positions and then proceed greedily for the remaining slots and prove that this strategy offers formal approximation guarantees. Finally, we show how to estimate our model on data from Expedia using the "exploded logit" formula to minimize reliance on simulation methods and explore how our algorithms compare with competing rankings in this real-world setting.

When specifying our model, we aim to strike a balance between two competing objectives: (i) flexibility—because we want the information in the data and not the model assumptions to drive the results of our counterfactual analyses—and (ii) tractability, which is key to being able to estimate the model and then use it for optimization. In our model, consumers search sequentially to reveal the products' utilities.[2] The optimal search rule involves computing, for each good, a search index (sometimes referred to as the reservation value), which encapsulates the good's expected utility presearch as well as the cost of revealing its utility. We propose a parameterization of the utility and search indices that accommodates multiple configurations of which components of utility consumers know presearch versus postsearch. In particular, in the model, consumers reveal their idiosyncratic match values—as is commonly assumed in the literature—but may also uncover vertical dimensions of quality.[3] This creates a potential wedge between the utility of a product for the average consumer and how good that product looks presearch (i.e., its average search index). As is described subsequently, the platform can then increase consumer surplus by steering consumers toward products that deliver high utility relative to what consumers expect ex ante. This informational asymmetry between the platform and the users is at the core of our results.

In order to study optimal rankings, we assume that platforms are able to increase and decrease the search indices of products beyond their baseline levels by either promoting them (giving them prominence on the search page) or burying them in the search results. By deciding what to rank and where to rank it, platforms can thus affect consumer choices. Given our data, we focus on the case of organic product rankings, although the framework could be extended to the case of sponsored search provided that data on sellers' bids is available.

Next, we obtain intuitive expressions for choice probabilities—and thus profits—and consumer surplus. For choice probabilities, we build on a result by Choi et al. (2018) and provide a convenient characterization of purchase behavior in our model, showing that the product ultimately chosen takes a max-min form: it is the product with the highest effective index; the effective index of a product is the minimum of its search and utility indices. For consumer surplus, we show that a key driver is the potential of the products in the choice set, which we define as the difference between the mean utility and search indices after the search indices have been adjusted to account for ranking.

With these expressions in hand, we start by considering the problem of choosing rankings to optimize consumer surplus. Knowing what is best for consumers is important both for platforms with long-run (growth) objectives as well as regulators who want to govern these platforms. To gain intuition, we consider a relaxation of the problem in which the platform can continuously adjust the search indices subject to a "rankings budget" constraint. We find that the platform should act to equate the potentials of all products. Ex ante, a high-potential product is one with a high utility relative to its search index: a diamond in the rough. The product is unlikely to be viewed and chosen by consumers unless the platform promotes it; such promotion helps consumers make better choices. Contrast this with a product that has both high mean utility and search indices, that is, a product that is well-known to be high quality, such as a branded item. This product is often chosen by consumers, and so the platform need not use its limited space promoting it.

An immediate but nontrivial implication of this analysis is that ranking products from highest to lowest utility is not necessarily optimal for consumer surplus because high-utility products need not have high potential.[4] An example is a product from a top brand that has both high salience and utility and needs no promotion to be purchased. In this case, it is preferable to rank a relatively unknown, equally high-quality product at the top. In contrast, existing work that does not explicitly model the consumer search process for possibly vertical components of utility finds that ranking by utility is always optimal (Ghose et al. 2012). Another implication of the framework is that a product that is highly salient but offers low utility (clickbait) should optimally be buried in the rankings because that frees up space to promote more worthy products. Note that these nuances arise because, as mentioned, our model accommodates search over vertical components of utility, so that the platform has more information than consumers do presearch about the average quality of products. The platform can thus use this information to steer consumers toward better options through rankings.

The optimal ranking is different when the platform is concerned with profit/revenue maximization.[5] In a world in which the consumer buys some product regardless of the selection, the platform should optimally offer the consumer no choice at all and only display a single product, the one on which it earns the highest revenue.[6] However, realistically, consumers can shop elsewhere. This competitive force drives the platform toward aligning its rankings with consumer preferences so as to make a sale. We show that the optimal product assortment includes all products that have a (weakly) higher product revenue than the average product revenue; the averaging includes the zero-revenue outside option. Further, the revenue-optimizing ranking is typically different from the consumer surplus-maximizing ranking because the products delivering high revenues need not have high potential.[7]

The continuous relaxation of the problem that we have described so far gives useful intuition, but in practice, the assignment of products to ranks is a discrete optimization problem, and the heuristics from the continuous case cannot be blindly followed. For example, it may not be optimal to place the product with the highest potential at the top of the search rankings because, if it has a low enough search index, it may attract very few purchasers even when top-ranked, which is a waste of that position.[8]

We thus offer a pair of similar algorithms for maximizing consumer surplus and revenue. They combine an exhaustive brute force search of the best ranking for the top K positions with a greedy algorithm for ranking the remaining products. We show that the amount of surplus/profits left on the table by this approach is upper bounded by a function that is increasing in the combined choice probability of all products ranked `below K. Therefore, when this choice probability is small—that is, when products ranked below K are unlikely to be chosen—our algorithms are close to being optimal. Through choosing K, the platform can trade off the tightness of the optimality guarantee against the computational demands of the ranking algorithm. What makes proving these approximation guarantees technically challenging is interference effects among units—ranking one product higher means that it is more likely to be purchased and every other product less so—so that the analysis is not neatly separable across products. Simulation experiments verify that these solutions work well in a range of simulated environments, achieving more than 98% of the available gains in consumer surplus and revenue from rankings (available gains are defined as the difference between the consumer surplus/revenue under the best and the worst rankings).

The second part of the paper takes our approach to the wild. We use two data sets from Expedia: in one data set (the "training" data), rankings were assigned at random, whereas in the other (the "testing" data) they were assigned according to Expedia's algorithm (these data sets were also used in Ursu (2018)). We fit the model only on the training data to avoid concerns related to the endogeneity of product ranks and use the testing data to validate the model out of sample. Identification is aided by the fact that we see both what the consumers clicked on and what they bought, which implies a set of orderings for the search and utility indices. The likelihood of each of these orderings, in turn, can be written using the exploded logit formula, which reduces reliance on simulation methods and thus facilitates estimation.

Finally, we take these estimates and simulate search intensity, consumer surplus, and platform revenue under different ranking algorithms. We compare our ranking algorithm optimized for consumer surplus for $K = 3$ (i.e., we brute force the top three positions, then proceed greedily) to a straight utility ranking, an algorithm optimized for revenue, and the Expedia ranking itself. We find that our algorithm optimized for consumer surplus dominates the Expedia algorithm under our model, raising consumer surplus by \$1.21 per consumer and still raising revenue by 8¢ per customer. Ranking directly by utility is less effective, raising consumer surplus by around 60¢ relative to Expedia and reducing revenue by 15¢. These numbers are at the search impression level; because only around 13% of customers purchase, one should scale the figures up by seven to eight times to obtain numbers per purchasing consumer.

1.1. Related Literature

The paper that is perhaps closest to ours is Derakhshan et al. (2022), who also study optimal rankings in online platforms and propose algorithms based on a microfounded model of consumer behavior. Relative to this recent paper, our main contributions are as follows. First, our search model is different, and in particular, it allows consumers to search in any order, whereas the model in Derakhshan et al. (2022) implies that it is always optimal to search in the same order in which the products are ranked on the page. This is important in our context because more than 80% of the impressions in our data violate this restriction,[9] indicating that other factors beyond rankings affect which product is searched at each step. We capture this by letting the reservation values depend on product attributes immediately visible on the results page and note that this is a key ingredient in our notion of product "potential." Second, we develop an estimation method and show that the primitives needed as inputs to our algorithms can be pinned down from click and purchase data—something that Derakhshan et al. (2022) cite as an avenue for future research. Third, we apply our approach to a real-world setting to quantify the magnitude of the rankings gains as opposed to working with synthetic data.

In addition, our work is related to the broader literature on search. Abaluck et al. (2020) estimate consumer preferences without committing to a specific search model; they show that rankings can be incorporated in their approach but do not discuss optimality of rankings. A related literature in marketing focuses on the determinants of consumers' consideration sets (see, among others, Roberts and Lattin 1991, Mehta et al. 2003, Honka and Chintagunta 2017, Honka et al. 2017). Our search model can be viewed as one way to microfound the consumer's decision of which goods to consider that is amenable to studying optimal rankings.

Our work is also related to the literature on platform design and recommendation systems. Whereas we focus on product rankings, other papers study the customization of email communications to customers (e.g., Ansari and Mela 2003) or the amount and type of information to display on the results page (e.g., Gardete and Hunter 2020, Gu and Wang 2022).

Next, a large literature in operations research studies the assortment optimization problem, that is, deciding which products to stock or which ones to show to consumers as they search (see Kök and Fisher (2007) for a survey, and Jagabathula and Rusmevichientong (2017) and Agrawal et al. (2019) for more recent contributions). Unlike most of this literature, we focus on optimizing not only revenue, but also consumer surplus and on the decision of how to rank products as opposed to just which ones to offer.

A literature in computer science and operations studies how to rank products to maximize directly quantifiable metrics, such as revenue/profits. These papers typically assume stylized models of search, such as the cascade model (Aggarwal et al. 2008, Kempe and Mahdian 2008, Karmaker Santu et al. 2017, Gallego et al. 2020). In contrast, our approach is based on a microfounded model of consumer behavior, enabling us to consider consumer surplus—in addition to revenues/profits—and explore the trade-off between the two goals.

Finally, the paper is related to the notion of *incrementality* in marketing (e.g., Zantedeschi et al. 2017, Ascarza 2018, Hitsch et al. 2018). This literature has found that targeting based on customers' baseline outcomes can perform substantially worse than targeting based on their responsiveness to treatment. Similarly, we show that ranking based on utility levels is suboptimal relative to ranking based on the difference between utility and search indices, which is a measure of how much rankings can incrementally shift consumer choices and welfare.

1.2. Paper Structure

The paper proceeds in three parts. First, we introduce the double index search model in Section 2. Then, in Section 3, we discuss optimal rankings. Finally, we apply the approach to the Expedia data in Section 4 and conclude in Section 5.

2. Model

We consider a setting in which consumers have unit demand for a good in a product category. They log on to an appropriate platform (e.g., Expedia) and enter a search query that describes that category (e.g., "hotel room in New York City on November 5"). They are presented with a finite set of search results. These products are distinguishable by the presearch characteristics presented to the user.

Figure 1 shows an example of the search results on Expedia for the preceding hotel query. The presearch characteristics include the average price per night, photos, the number of reviews, and the number of rooms left. In the case of keyword searches on Bing or Google, for example, the presearch characteristics include whether a link is organic or sponsored as well as the link text itself.

Based on what they see, users may choose to either perform some other search operation (such as refining the query or filtering the results), abandon search, or click on one of the options. If they click, they are taken to a product page, on which they may learn additional product characteristics (e.g., room amenities). They may then stop and purchase, continue search, or abandon search without purchase. At the end of the process, they either have picked the best option from those they considered (clicked on) or have chosen not to buy at all.

We now formalize this setting. A consumer i has a need that may be met by purchasing a single product from a finite set of products $\mathcal{J} \equiv \{1, \ldots, J\}$ plus the outside option (denoted by zero). Consumer i's utility from good j is

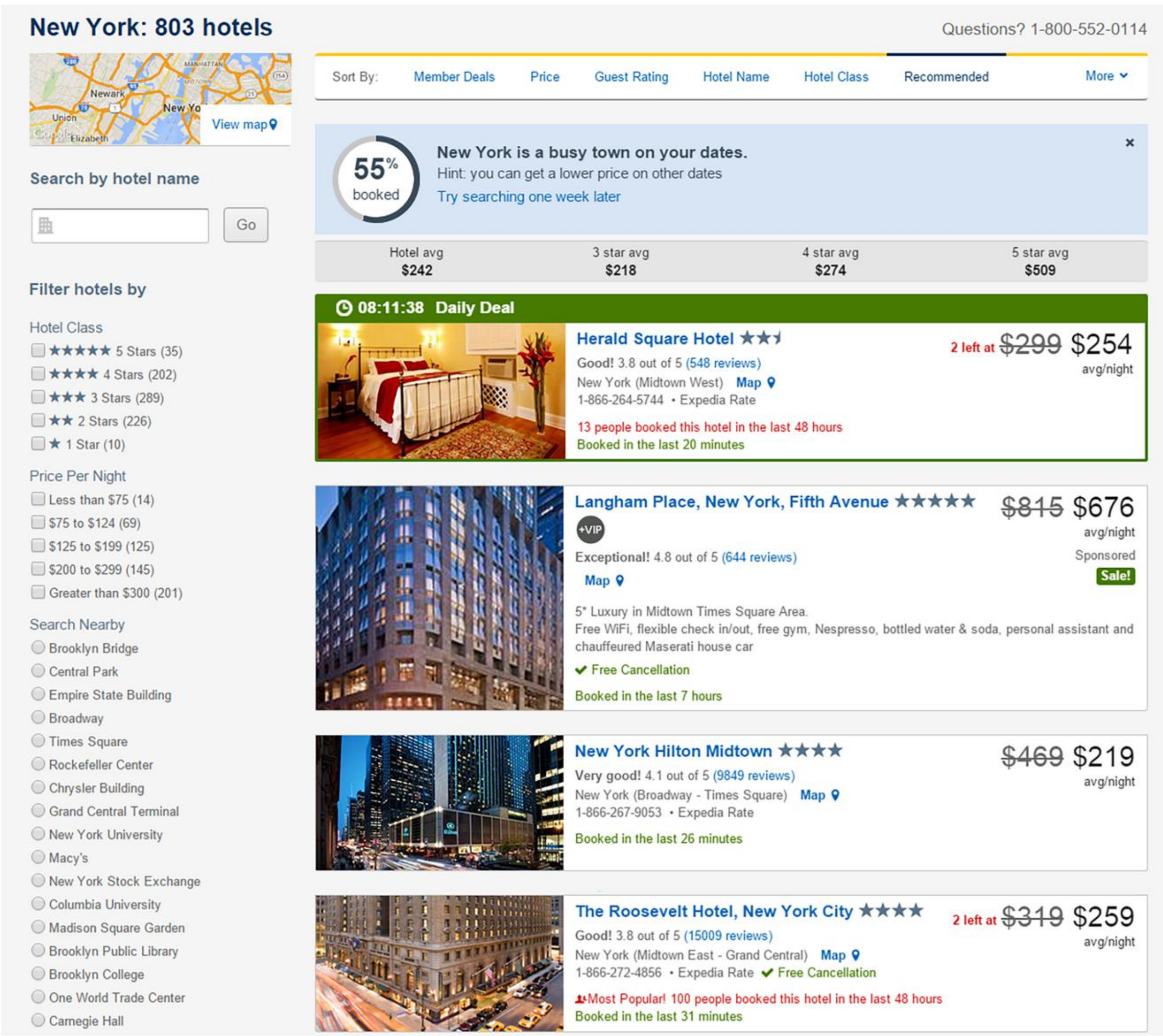
$$u_{ij} = \beta^U x_j + \xi_j^U + \varepsilon_{ij}^{pre} + \varepsilon_{ij}^{post}, \quad (1)$$

where x_j are characteristics of the product that are captured by the data (e.g., a hotel's price and star rating); ξ_j^U represents the vertical component of quality that is unobserved to the researcher but can be estimated via product fixed effects (e.g., the appeal of the photos or the content of reviews); and $\varepsilon_{ij}^{pre}, \varepsilon_{ij}^{post}$ are taste shocks idiosyncratic to consumer i (e.g., i especially likes boutique hotels). Consumer i learns u_{ij} by searching for product j. Depending on the context, the u_{ij} index could represent either the final consumption utility that the consumer derives from the good or (under risk neutrality) the expected utility based on the information on the detailed product description page. We normalize the mean utility of the outside option to zero so that $u_{i0} = \varepsilon_{i0}^{post}$.

We assume that consumer i knows ε_{ij}^{pre} for all products upon landing on the results page and that ε_{ij}^{post} is only revealed when i searches product j by opening the corresponding product page.[10] Thus, we allow consumers to search over their idiosyncratic match value ε_{ij}^{post}, consistent with several papers in the literature. As for x_j and ξ_j^U, we do not take a stand on when they are revealed to consumers. Some of those attributes, such as price, could be immediately visible on the results page, but others—for example, the component of quality captured by ξ_j^U—may only be accessible after clicking on the product page. Our model accommodates both the case in which (x_j, ξ_j^U) are known to the consumers presearch and the case in which they require search as well as any intermediate case in which some of these attributes are known presearch and the others are only revealed after search. Which case is most appropriate for a given empirical setting depends on the structure of the website/app as well as on what is captured by the data available to the researcher. For instance, if the data contains detailed information about product reviews and other features that are only visible on the product page, then those x_j variables would only be visible to consumers postsearch. We formalize this in the following assumptions.

Assumption 1. *The researcher observes x_j, consumer search actions, and consumer choices and can estimate the product fixed effects ξ_j^U.*

Figure 1. (Color online) Search Results Page from Expedia



Assumption 2. *Consumer i knows ε_{ij}^{pre} for all j before engaging in search but needs to search j in order to reveal ε_{ij}^{post}. Prior to search, it may be that the consumer knows all of the (x_j, ξ_j^U) attributes or none of them or only a subset of them for any given good.*

We also maintain the same assumptions on the search process as in Weitzman (1979).

Assumption 3. *Consumers search sequentially with free recall, are forward-looking, and pay cost c_j to open product j's page, which fully reveals u_{ij}. The utilities are independent and identically distributed (iid) across goods conditional on the presearch information.*

The restriction of conditional independence of utilities guarantees that the payoff realization for one product does not cause consumers to update about the payoff distributions of the remaining products.[11]

If we let $\mathcal{I}_{ij}$ be consumer i's information set about product j presearch, the results in Weitzman (1979) yield the optimal search strategy. Specifically, consumer i forms a reservation value s_{ij} for each product j, which is defined as the solution to

$$c_j = \int_{s_{ij}}^{\infty} (u - s_{ij}) f_{u_{ij}|\mathcal{I}_{ij}}(u)du, \tag{2}$$

where $f_{u_{ij}|\mathcal{I}_{ij}}$ is the density of u_{ij} given consumer i's information set. Search optimally proceeds in descending order of reservation values until the highest utility uncovered exceeds the next best value of s_{ij}, at which point i stops searching and chooses the good with the highest utility among those searched. The payoff from the outside option, u_{i0}, is known presearch (i.e., we can set $s_{i0} = \infty$) so that, if u_{i0} is higher than s_{ij} for all $j \geq 1$, then consumer i leaves the results page without searching any products.

The reservation values are to be interpreted as some combination of the visibility, salience, and presearch observable attractiveness of the products. Because the ranking of a product affects how easy it is for consumers to reach its page, the search costs c_j—and, thus, the reservation values s_{ij}—are likely to be a function of rankings. Therefore, we specify the reservation values, or search indices, as

$$s_{ij} = \beta^S x_j + \xi_j^S + f(r_j) + \varepsilon_{ij}^{pre} + \varepsilon_{ij}^S, \tag{3}$$

where r_j is the rank of each product on the page with a higher value indicating a more salient position so that $f(r_j)$ is an increasing function. Note that, in our model, consumer behavior is fully characterized by the set of utility indices (1) and search indices (3).

A few comments are in order on the specification of the search indices. First, s_{ij} depends on ε_{ij}^{pre} in the same way that u_{ij} does, which is consistent with the fact that ε_{ij}^{pre} is known to consumers presearch. In contrast, s_{ij} does not depend on ε_{ij}^{post} because that component of utility is unknown prior to search. Second, notice that x_j enters both utility and search indices but with possibly different coefficients. This allows for a range of possibilities, consistent with Assumption 2. For example, it could be that consumers know x_j prior to search and form expectations based on it regarding the components of utility that they don't yet know. Another possibility is that x_j is not known before search, in which case it would not enter the search index, that is, $\beta^S = 0$. A similar argument applies to the unobserved (to the researcher) terms ξ_j^S and ξ_j^U. Because the relationship between ξ_j^S and ξ_j^U is left unrestricted—they are estimated via two separate sets of hotel fixed effects in the empirical implementation—we allow for the case in which ξ_j^U is known to consumers before search (and is possibly used to form expectations on the unknown parts of utility) as well as the case in which ξ_j^U is unknown prior to search.

In the next sections, we formalize these arguments by explicitly working through three possible microfoundations for the reservation values. In all cases, the intuition is the same: because the reservation values are linear in the presearch components of utility, if we assume that the expectations about the search attributes are also linear, it follows that the search indices are linear functions of the presearch attributes plus a term $f(r_j)$ capturing the effect of search costs. This corresponds exactly to the functional form in (3). We emphasize that the following examples are just three possible ways to microfound our model and that other options are allowed, for instance, hybrid models in which consumers know some of the x_j attributes (e.g., price and brand) prior to search and reveal the remaining x_j attributes (e.g., location of the hotel) as well as ξ_j^U via search. Because, in estimation, we directly target the coefficients in the two indices, we do not have to commit to any of these microfoundations, implying that the estimation is robust to any of these search patterns.

2.1. Example 1: Search over Quality ξ_j^U and Match Value ε_{ij}^{post}

Let utilities take the form in (1). Assume that consumers know x_j and ε_{ij}^{pre} for all j prior to search and pay search cost c_j to learn ξ_j^U and ε_{ij}^{post}. Further, ε_{ij}^{pre} and ε_{ij}^{post} are iid across products and independent of all other variables, and ξ_j^U are independent across products j conditional on the vector of all presearch components $(x_j, \varepsilon_j^{pre})_{j=1}^J$. Then, Equation (2) takes the form

$$c_j = \int_{s_{ij}}^{\infty} (u - s_{ij}) f_{u_{ij}|x_j, \varepsilon_{ij}^{pre}}(u) du. \tag{4}$$

We now show that these search indices take the form we specify in (3). Toward this, assume that the conditional distribution of ξ_j^U belongs to a location family: $\xi_j^U = \gamma x_j + \tilde{\xi}_j^U$ for some parameters γ and a random variable $\tilde{\xi}_j^U$ with $\tilde{\xi}_j^U \perp x_j$. Then, the conditional mean is linear in the product characteristics: $E[\xi_j^U | x_j] = \gamma x_j$. We can interpret γ as capturing the relationship between product characteristics observed on the search page and those that are only observed after clicking through (e.g., price could act as a signal of quality).

We can then rewrite the right-hand side of the preceding equation as follows:

$$\begin{aligned}
&\int_{s_{ij}}^{\infty} (u - s_{ij}) f_{u_{ij}|x_j, \varepsilon_{ij}^{pre}}(u) du \\
&= \int_{s_{ij}}^{\infty} (u - s_{ij}) f_{\xi_j^U + \varepsilon_{ij}^U | x_j}(u - \beta^U x_j - \varepsilon_{ij}^{pre}) du \\
&= \int_{s_{ij}}^{\infty} (u - s_{ij}) f_{\tilde{\xi}_j^U + \varepsilon_{ij}^{post}}(u - \beta^U x_j - \gamma x_j - \varepsilon_{ij}^{pre}) du \\
&= \int_{s_{ij} - \beta^U x_j - \gamma x_j - \varepsilon_{ij}^{pre}}^{\infty} (y + \beta^U x_j + \gamma x_j + \varepsilon_{ij}^{pre} - s_{ij}) f_{\tilde{\xi}_j^U + \varepsilon_{ij}^{post}}(y) dy,
\end{aligned} \tag{5}$$

where the first equality follows from the relationship between the conditional densities of u_j and $(\xi_j^U, \varepsilon_{ij}^{post})$, the second from the location family assumption, and the third by the change of variable $y = u - \beta^U x_j - \gamma x_j - \varepsilon_{ij}^{pre}$. Now, let ρ_j be the solution of $c_j = \int_{\rho_j}^{\infty} (y - \rho_j) f_{\tilde{\xi}_j^U + \varepsilon_{ij}^{post}}(y) dy$, and let $\beta^S = \beta^U + \gamma$. Then, if we let the search index take the form $s_{ij} = \beta^U x_j + \gamma x_j + \rho_j + \varepsilon_{ij}^{pre} \equiv \beta^S x_j + \rho_j + \varepsilon_{ij}^{pre}$, we may substitute into (5), and verify that s_{ij} indeed satisfies (4). Note that $\beta^S = \beta^U + \gamma$ is in general different from β^U if $\gamma \neq 0$ because of consumers forming expectations on ξ_j^U based on x_j. For example, if consumers infer that a higher price x_j is associated with better quality ξ_j^U, then we have $\gamma > 0$ and, thus, $\beta^S > \beta^U$. Finally, suppose that search costs are determined by rankings and decrease in rank (i.e., high-rank products have low search costs). Then, it follows that the thresholds ρ_j are a function of rankings, that is, $\rho_j = f(r_j)$ for some unknown increasing function $f(\cdot)$.

Putting this all together, we have $s_{ij} = \beta^S x_j + f(r_j) + \varepsilon_{ij}^{pre}$. This functional form for the search index is subsumed by the specification in (3). Specifically, (3) includes the additional $\xi_j^S + \varepsilon_{ij}^S$ term, which represents anything that affects the salience of product j besides what is captured by the ranking r_j. An example of this is consumer heterogeneity in search costs as well as any mistakes that consumers may make in calculating the optimal reservation values. Alternatively, if different goods have different distributions of $\tilde{\xi}_j^U$, then the term ξ_j^S captures this; for instance, if the variance of $\tilde{\xi}_j^U$ is higher than that of $\tilde{\xi}_k^U$, it is well-known that the reservation value of j is higher (because there is more upside to searching j), which is captured by ξ_j^S being larger than ξ_k^S.[12] In the empirical implementation, the terms ξ_j^S are estimated via fixed effects and thus are left unrestricted, whereas the shocks ε_{ij}^S are helpful to smooth out the likelihood function (see Section 4.2).

2.2. Example 2: Search over Attributes x_j and Match Value ε_{ij}^{post}

Nothing in the derivations from the last section hinges on the fact that x_j is observed by the researcher and ξ_j^U is not. Thus, an analogous argument holds for the case in which the role of the two terms is flipped; that is, consumers know ξ_j^U prior to search and learn x_j upon searching product j. Specifically, let $x_j = \tilde{\gamma}\xi_j^U + \tilde{x}_j$, and let consumers form expectations on x_j based on $E(x_j | \xi_j^U) = \tilde{\gamma}\xi_j^U$. Then, derivations similar to the preceding lead to the functional form $s_{ij} = \xi_j^U(1 + \beta^U\tilde{\gamma}) + \rho_j + \varepsilon_{ij}^{pre}$, where ρ_j solves $c_j = \int_{\rho_j}^{\infty}(y - \rho_j) f_{\beta^U\tilde{x}_j + \varepsilon_{ij}^{post}}(y)dy$. This specification for the search index is subsumed by that in (3) by letting $\xi_j^S \equiv \xi_j^U(1 + \beta^U\tilde{\gamma})$ and $\beta^S = 0$. As before, the functional form in (3) contains the additional error term ε_{ij}^S, which captures heterogeneity in search costs and/or any factors affecting the salience of product j above and beyond what the fully rational model predicts.

2.3. Example 3: Search over Match Value $\varepsilon_{i,j}^{post}$

If consumers know (x_j, ξ_j^U) presearch and only learn $\varepsilon_{i,j}^{post}$ upon clicking on product j, then an analogous derivation to that in Section 2.1 leads to the following expression for the search indices:

$$s_{ij} = \beta^U x_j + \xi_j^U + f(r_j) + \varepsilon_{ij}^{pre} + \varepsilon_{ij}^S,$$

which is a special case of the specification in (3) with $\beta^S = \beta^U$ and $\xi_j^S = \xi_j^U$. Intuitively, if consumers only search over the idiosyncratic shocks $\varepsilon_{i,j}^{post}$—which are assumed to be independent of everything else—then their search indices are simply going to equal the part of utility that they observe presearch (i.e., $\beta^U x_j + \xi_j^U + \varepsilon_{ij}^{pre}$) plus the term $f(r_j)$ capturing the search cost. Note that, in this case, x_j and ξ_j^U enter both the utility and the search indices in the same way because, by independence of $\varepsilon_{i,j}^{post}$, $\mathbb{E}(\varepsilon_{i,j}^{post} | x_j, \xi_j^U) = 0$, and thus, x_j and ξ_j^U do not carry any extra information about $\varepsilon_{i,j}^{post}$ in the search index.[13] As before, we also include the additional search shock ε_{ij}^S, which helps us smooth out the likelihood. This example shows that our model accommodates the case in which consumers only search over their match value $\varepsilon_{i,j}^{post}$, which is commonly considered in the literature.

We reiterate that the models in Sections 2.1–2.3 are just three possible microfoundations for the double index specification. In particular, other Weitzman models in which consumers search a subset of the (x_j, ξ_j^U) variables—on top of ε_{ij}^{post}—are possible.

2.4. Can Rankings Affect Consumers' Expectations?

In the preceding microfoundations, we focus for notational convenience on the case in which rankings only affect search costs but do not directly enter utility. However, our framework can also accommodate the possibility that consumers form expectations on the search attributes based on rankings. For instance, in the first example (Section 2.1), it could be that consumers expect a higher value of quality $\tilde{\xi}_j^U$ when a product is ranked higher. Then, the expectation in (5) is taken over the distribution of $\tilde{\xi}_j^U + \varepsilon_{ij}^{post}$ conditional on r_j. As a result, the term ρ_j showing up in the search index would depend on r_j not just because worse rankings lead to higher search costs, but also via the expectations channel. Because we are modeling ρ_j as a flexible function of rankings—to be estimated via position fixed effects—our framework can accommodate this.

One limitation is that, because we do not model the process whereby consumers form expectations about postsearch attributes based on rankings, we have to assume that this process is unchanged in our counterfactuals. Thus, for example, Expedia users have the same beliefs about the postsearch quality of a hotel ranked in the first position as the ranking algorithm changes. There are two main reasons for this choice. First, changes in the proprietary ranking algorithm used by a platform are usually not announced to the public, and thus, it is reasonable to assume that consumers would not react to those changes in the short-to-medium term. In the long run, we might expect consumers to realize that the ranking algorithm has changed and update their beliefs accordingly—especially if they interact with the platform often—and so we view our analysis as one that best captures patterns in the short-to-medium run. However, even in the long run it is not clear that users would have fully rational expectations given the complexity and lack of transparency of many algorithms. Second, allowing consumers to update their beliefs substantially complicates the task of finding optimal algorithms. In particular, one would need to solve for a fixed point: the optimal ranking is such that consumers' beliefs react to it in a way that makes it optimal for the platform to choose exactly that ranking. Finally, we note that the welfare gains from optimally ranking products are likely to be even larger than our estimates if consumers adjust their beliefs. This is because, if consumers anticipate that rankings are optimized, they need to search less before finding a product that is sufficiently good, implying search cost gains above and beyond our estimates.

3. Optimal Rankings

Online platforms can influence what is bought through their product rankings. In our model, this is captured by the term $f(r_j)$ in the reservation values. With this in mind, we now turn to the problem of optimizing those rankings. We consider two main objective functions: maximizing consumer surplus—an appropriate target for a platform trying to maximize the size of its user base or for a policy-maker—and maximizing revenue.[14] For any given search query, platforms can decide how to order the results that they return. Assuming that each search query corresponds to a fixed set of relevant products $\mathcal{J}$, the problem is then to determine which of those products to rank and how to rank them. The platform may choose not to present all products to the consumer. This is equivalent to allowing the platform to set a product's rank to be zero with $f(0) = -\infty$ so that this product is never considered. However, we rule out "gaps" in the ranking: if there is a product ranked in position L, then positions $1 \dots L-1$ must be filled. The analysis in this section is conditional on a search query and, thus, can seamlessly incorporate additional information that the platform might have on the consumer, delivering personalized rankings.

We assume that the platform knows both the mean baseline search index $\beta^S x_j + \xi_j^S$ and the mean utility $\beta^U x_j + \xi_j^U$ as well as the rankings function $f(\cdot)$. We think it is realistic to assume this given that platforms have access to rich data, including data from experiments, that allow them to obtain good estimates of the quality and salience of each option. Indeed, in the empirical application, we propose one way to estimate these objects based on click and choice data, which are readily available to platforms.

3.1. Choice Probabilities and Consumer Surplus

First, we state a lemma that characterizes the relationship between the search and utility indices and the product eventually purchased. This result is a minor modification of an existing result by Choi et al. (2018) for the Weitzman (1979) model, and we adopt their name for the result (see also Armstrong and Vickers 2015, Kleinberg et al. 2016, Armstrong 2017). Let $\mathbf{s}_i \equiv (s_{i0}, s_{i1}, \dots, s_{iJ})$ and similarly for $\mathbf{u}_i$.

Lemma 1 (Eventual Purchase). *A consumer facing a choice set consisting of $(\mathbf{s}_i, \mathbf{u}_i)$, including an outside option with $s_{i0} = \infty$, purchases a product $j \in \mathcal{J}^*$, where $\mathcal{J}^* = \arg\max_{j\in\mathcal{J}} v_{ij}$ for $v_{ij} = \min\{s_{ij}, u_{ij}\}$.*

Proof. See Online Appendix A.2. □

In words, it is the product with the highest minimum of search and utility indices that gets purchased. We call this quantity $v_{ij} = \min\{s_{ij}, u_{ij}\}$ the effective index of the product. Note that consumers only know u_{ij} for the products they searched, but they still end up choosing the good that maximizes the effective index across all products. One implication of Lemma 1 is that products with high utility but low search indices are rarely purchased. In other words, a product needs to be both good (high u_{ij}) and salient (high s_{ij}) in order to have a high market share.

We now show that Lemma 1 yields convenient expressions for choice probabilities and average consumer surplus under a standard assumption on the distribution of ε_{ij}^{pre}. We begin by fixing the $\varepsilon_{ij}^S, \varepsilon_{ij}^{post}$ shocks; these are integrated out at the end. Let $\delta_{ij}^S = \beta^S x_j + \xi_j^S + \varepsilon_{ij}^S$ be the mean search index (before the effect of rankings), $\delta_{ij}^U = \beta^U x_j + \xi_j^U + \varepsilon_{ij}^{post}$ be the mean utility, and $\delta_{ij}^V(r_j) \equiv \min\{\delta_{ij}^S + f(r_j), \delta_{ij}^U\}$ be the mean effective index. Define $\phi_{ij}(r_j) \equiv \delta_{ij}^U - \delta_{ij}^S - f(r_j)$ to be the potential of product j—that is, the difference between the mean utility and search indices, inclusive of the rankings effect. This potential captures how much better the product is (on average) than it appears to be based on the information displayed on the results page as well as its ranking. We can now derive convenient formulae for the choice probability functions and the average consumer surplus. The latter is defined to be the average utility of a consumer on the platform, gross of any search costs, for which the average is taken over all consumer-specific unobservables $(\varepsilon_i^S, \varepsilon_i^{pre}, \varepsilon_i^{post})$.

Proposition 1 (Aggregate Demand and Consumer Surplus). *Let $\mathbf{r} \equiv (r_1, \dots, r_J)$, and assume that ε_{ij}^{pre} is drawn from a Gumbel distribution independently across goods. Then, the probability of a consumer choosing product j is given by*

$$P(Choose\ j) \equiv q_j(\mathbf{r}) = \int q_{ij}(\mathbf{r}) dF_{\varepsilon_i^S, \varepsilon_i^{post}},$$

where $q_{ij}(\mathbf{r}) \equiv \frac{\exp \delta_{ij}^V(r_j)}{1+\sum_k \exp \delta_{ik}^V(r_k)}$ and $F_{\varepsilon_i^S, \varepsilon_i^{post}}$ is the distribution of $(\varepsilon_i^S, \varepsilon_i^{post})$.

Further, the average consumer surplus is given by

$$CS \equiv C + \int \left[\log\left(1 + \sum_j \exp\, \delta_{ij}^V(r_j)\right) + \sum_{j:\phi_{ij}(r_j)>0} q_{ij}(\mathbf{r})\phi_{ij}(r_j) \right] dF_{\varepsilon_i^S,\varepsilon_i^{post}}, \quad (6)$$

where C is the Euler constant.

Proof. See Online Appendix A.3. □

This proposition says, first, that the choice probabilities are determined by the mean effective indices according to the standard logit form (integrated over $\varepsilon_{ij}^S, \varepsilon_{ij}^{post}$). This is simply a consequence of Lemma 1 and the extreme value distribution assumption on $\varepsilon_{i,j}^{pre}$. The consumer surplus, on the other hand, differs from the standard log-sum form—corresponding to the first term inside the integral in (6)—in that it has an additional term that binds only for products whose utility is higher than the search index or, equivalently, whose potential ϕ_{ij} is positive (diamonds in the rough).[15] The reason is that although choices are based on effective indices δ_{ij}^V, consumer surplus is based on utility, and so whenever the utility exceeds the effective index, this must be counted too (it can't be less than the effective index because of the *min* operator in the definition of effective indices). This characterization of consumer surplus is a key building block in the analysis that follows.

Note that, whereas our model features search costs, the formula for consumer surplus does not account for the reduction in welfare because of the search effort, but only for the effect that search costs have on the quality of the final choice made by consumers.[16] This may be appropriate for online environments, in which the time costs of search are often small. But one may reasonably be concerned that these search costs are important in practice, and so in our application, we track the average number of clicks under different algorithms as an additional performance metric and find that the proposed algorithms essentially do not change the number of searches relative to the status quo. In contexts in which this is not the case—particularly if the proposed consumer surplus–optimizing algorithm yields substantially more searches—one can use one of the microfoundations of the model and the estimated parameters to back out search costs (e.g., from (4)) and then account for the latter in the welfare calculation.

3.2. Optimizing Consumer Surplus

We want to match products to ranks in such a way as to maximize consumer surplus. This matching problem is discrete and, therefore, not amenable to standard techniques. So we begin instead with a relaxation of the problem in which ranks can be assigned continuously subject to a budget constraint that may hold with inequality (because the platform can choose not to list some products at all, leaving some "surplus" ranks). Let $f(\cdot)$ now be defined on the reals with $f' \geq 0$, and let $CS(\mathbf{r}, \delta^S, \delta^U)$ denote the average consumer surplus as a function of the vectors of rankings $\mathbf{r}$ and average indices δ^S, δ^U; that is, we let $\delta_j^S \equiv \beta^S x_j + \xi_j^S$ (so that $\delta_{ij}^S = \delta_j^S + \varepsilon_{i,j}^S$) and similarly let $\delta_j^U \equiv \beta^U x_j + \xi_j^U$. Then, the problem is

$$\max_{r_1 \dots r_J} CS(\mathbf{r}, \delta^S, \delta^U)$$

$$\text{subject to} \sum_{j:j \text{ is listed}} r_j \leq \frac{J(J+1)}{2},$$

where the budget $J(J+1)/2 = 1 + 2 + \cdots + J$ corresponds to the total "ranking power" available.

The optimal solution has the property that the partial derivative of the consumer surplus is equal for all products that are ranked: $\partial CS(\mathbf{r}, \delta^S, \delta^U)/\partial r_j = \lambda \ \forall j \in \mathcal{L}$, where $\mathcal{L}$ is the set of products listed, and when all products are listed, λ is the Lagrange multiplier on the budget constraint. If it is feasible to choose ranks so that $\delta_{ij}^S + f(r_j) \geq \delta_{ij}^U \ \forall j$ and i (i.e., every product has a higher search index than its utility index), it is also optimal for consumer surplus. The reason is that the mean effective indices, $\min\{\delta_{ij}^S, \delta_{ij}^U\}$, are ordered in the same way as mean utilities so that the highest utility products are most often purchased.

However, in most cases, it is not possible to promote all products enough to achieve this, and some products need to be prioritized. To see which products benefit most from higher rankings, we take the derivative of consumer surplus with respect to r_j:

$$\frac{\partial CS(\mathbf{r}, \delta^S, \delta^U)}{\partial r_j} = \int_{\{\varepsilon_{ij}^S - \varepsilon_{ij}^{post} < \delta_j^U - \delta_j^S - f(r_j)\}} q_{ij}(\mathbf{r}) f'(r_j) \times \left(\phi_{ij}(r_j) - \sum_{k:\phi_{ik}(r_k)>0} q_{ik}(\mathbf{r}) \phi_{ik}(r_k) \right) dF_{\varepsilon_i^S,\varepsilon_i^{post}}, \quad (7)$$

where, again, q_{ij} denotes the choice probability of product j for a consumer with shocks $(\varepsilon_i^S, \varepsilon_i^{post})$. We derive this expression in Online Appendix A.4. The intuition for it is in two parts. When $\delta_j^S + f(r_j) + \varepsilon_{i,j}^S \geq \delta_j^U + \varepsilon_{ij}^{post}$ so that the effective index of j is determined by the utility, marginally improving the ranking of the product (i.e., increasing r_j) will not change the choice probabilities and hence has no effect on consumer surplus. This is why the integral in the expression excludes realizations of $(\varepsilon_i^S, \varepsilon_i^{post})$ resulting in the search index being larger than the utility index. On the other hand, when $\delta_j^S + f(r_j) + \varepsilon_{i,j}^S < \delta_j^U + \varepsilon_{ij}^{post}$, improving a product's ranking increases its choice probability on the margin. Whether this is good or bad for consumers depends on the sign of the expression in parentheses in (7), which relates the potential of product j to a weighted sum of the potentials of all products with positive potential.[17] If j offers a potential ϕ_{ij} that is higher than the other products, then consumer surplus increases if product j is promoted (higher r_{ij}). This is because, in this case, product j tends to deliver a positive surprise to consumers when they search it (diamond in the rough); thus, making it more salient on the results page leads more consumers to click on it and eventually choose it. In contrast, when j offers a potential ϕ_{ij} that is lower than the other products, then it is best to reduce the rankings power allocated to it, that is, demote j. Intuitively, in this case, product j tends to deliver a negative surprise—or a not-as-positive surprise—relative to its competitors (clickbait), and thus, it is optimal to steer consumers toward other products.

A key insight from (7) is that, in order to maximize consumer surplus, the platform should promote products with the highest baseline potential, which need not be those with the highest utility levels. As an example, consider two hotels A and B with $\delta^U_A > \delta^U_B$ (e.g., A is a chain hotel with more amenities and a lower price than B) and $\delta^U_A - \delta^S_A < \delta^U_B - \delta^S_B$ (e.g., B is a small boutique hotel with a lower baseline salience δ^S_B). Because the baseline potential of B is higher, it is optimal for consumer surplus to promote B over A even though A has a higher utility. Thus, the optimal ranking need not be the same as the ranking of products based on their utilities. In our empirical application, we indeed find that the utility ranking is suboptimal. Note that this insight hinges crucially on the fact that we allow consumers to uncover vertical components of utility (ξ^U_j and/or elements of x_j) via search because this is what drives the possible wedge between the average utility δ^U_j and the average baseline search index δ^S_j. In fact, if consumers only searched over match value $\varepsilon^{post}_{i,j}$, then the platform would have no private information and thus no scope to use rankings to predictably generate any positive surprises. In this case, it is optimal to rank products by utility.

One aspect that is not captured by the model is the fact that, if a very obvious good choice is not ranked at the top, users may mistrust the platform and leave. For instance, if someone is looking up tourist attractions in Paris and the platform does not rank the Louvre in one of the top positions, this may raise a red flag. This is especially a concern for cases in which products are clearly vertically differentiated (most people agree that the Louvre is a top attraction). On the other hand, in contexts in which products exhibit more horizontal differentiation—such as hotels, for which quality is traded off against price—then this mechanism may not be as prevalent.

The argument so far has treated rankings as continuous. However, in practice, $f(\cdot)$ is bounded, and each rank is discrete and associated with a fixed jump in its impact on a product's search index. It may no longer be possible to equate potentials, for example, some high-utility products may have such low search indices that, even with favorable rankings, they are ignored by consumers, and from the point of view of the platform, this is "wasted" promotion. In view of this, an algorithm that respects the discrete nature of the problem is needed.

3.2.1. The OPT-K Algorithm.

The discrete ranking problem is combinatorial in the number of positions and, therefore, demands a computationally tractable algorithm. We propose the following algorithm, which we label OPT-K: instead of ranking all J products, let us instead consider the simpler problem of assigning products to the first K (or fewer) positions to maximize the consumer surplus from that assignment, which we denote by CS^K, that is,

$$CS^K = C + \int \left[\log\left(1 + \sum_{j: r_j \geq J-K+1} \exp\, \delta^V_{ij}(r_j)\right) + \sum_{j:\phi_{ij}(r_j)>0,\, r_j \geq J-K+1} q_{ij}(\mathbf{r})\phi_{ij}(r_j) \right] dF_{\varepsilon^S_i,\, \varepsilon^{post}_i},$$

where r_j now represents the discrete rank of each product defined as $J - position_j + 1$ (so, again, higher is better). Maximizing CS^K by brute force requires only $\sum_{k=1}^{K} J!/(J-k)!$ evaluations[18] and, thus, scales better than the brute force approach—that is, trying all possible combinations for up to J positions—which requires $\sum_{k=1}^{J} J!/(J-k)!$ evaluations. Further, the computational gap between the two increases the smaller K is relative to J. Of course, only focusing on the top K ranks entails a loss of consumer surplus, and it is natural to consider conditions under which this loss is guaranteed to be reasonably small. The next result addresses this.

Proposition 2 (Approximate Optimality). *Let $K \leq J$, and assume that, for any feasible ranking, (i) $\delta^S_{ij} + f(r_j) < -1$, $\forall r_j < J-K+1$ and (ii) $\sum_{j: r_j < J-K+1} -(\delta^S_{ij} + f(r_j))\exp(\delta^S_{ij} + f(r_j)) < \nu$. Then, the gap between the consumer surplus achieved by optimally allocating all J positions and that achieved by optimally allocating the first K positions can be bounded as follows:*

$$CS(\boldsymbol{r}^*) - CS(\boldsymbol{r}^K) \leq \nu \left(\log\left(1 + \sum_{j=1}^{J} \exp(\delta^U_j)\right) + \frac{2-\nu}{1-\nu} \right),$$

where $\boldsymbol{r}^* \in \arg\max_{\mathbf{r}} \mathrm{CS}$ and $\boldsymbol{r}^K \in \arg\max_{\mathbf{r}} \mathrm{CS}^K$.

Proof. See Online Appendix A.5. □

Proposition 2 says that, if all products placed after position K (i.e., those with low ranks) are unlikely to be bought because their search indices become very low (i.e., ν is small), then optimizing product assignments for the first K positions delivers close to optimal consumer surplus. This is because the bound on the right-hand side of the inequality decreases as ν gets smaller. To interpret assumptions (i) and (ii) in the statement of Proposition 2, notice that, if (i) and (ii) hold simultaneously, it must be that $\sum_{j:r_j<J-K+1}\exp(\delta^S_{ij}+f(r_j)) < \nu$ for any feasible ranking because $-(\delta^S_{ij}+f(r_j)) > 1$. This, in turn, implies $\sum_{j:r_j<J-K+1} q_{ij} < \nu$. Therefore, the assumptions can be viewed as requiring the combined choice probabilities of the products placed after position K to be small enough. Conversely, if products ranked after position K are relatively likely to be purchased (high ν), only focusing on the top K positions may incur a substantial loss of consumer surplus.

What makes the result of Proposition 2 tricky to prove is that the decision of how to rank any one product affects the choice probabilities for all other products. By doing an exhaustive search over the top K positions, we can guarantee approximate optimality whenever the remaining positions bury the products placed there sufficiently so that they are rarely bought. Many online environments have the property that the top-ranked products get the vast majority of clicks, in which case there may be a reasonably small K for which the ν bound is reasonably tight. In Section 3.4, we present simulation evidence to quantify how much welfare is left on the table in practice and, on the flip side, what the computational gains are for various values of K.

3.2.2. Ranking the Remaining $J-K$ Products.

The OPT-K algorithm focuses on the top K positions and does not prescribe how to rank the remaining $J-K$ products. In practice, there may be substantial gains from ranking all the products. We now propose a practical greedy algorithm for the remaining products and call it the OPT-K + Greedy (OPTKG) algorithm.

The greedy algorithm is only needed after the OPT-K algorithm ranks all the top K positions. That is, if the OPT-K algorithm determines only $L < K$ products are needed to maximize consumer surplus, then the OPTKG algorithm also terminates. But, if needed, for the remaining positions, the greedy algorithm begins with the highest remaining rank and myopically assigns the best product to each position holding fixed all the products that have already been ranked. The algorithm terminates when either all positions are assigned or for some rank it is best not to assign any product to that rank. Algorithm 1 formally presents the OPTKG algorithm for consumer surplus optimization, labeled OPTKG-CS.

Algorithm 1 (OPTKG-CS)

Result: Assign a unique rank position $\{K+1, K+2, \ldots, K+k\}$, $\forall k \le J-K$ to k unique products.
Initialization: Let f_i denote the effect of position i's ranking on the search index.
Set N contains all $J-K$ unranked products.
for *position i from* $K+1$ *to* J **do**
 Calculate $CS_{ij} = CS(\{r_j = f^{-1}(f_i), r_{\{1,2,\ldots,J\}\setminus N}\}, \delta^S, \delta^U)$ for each product $j \in N$.
 Calculate $CS_{i0} = CS(\{r_{\{1,2,\ldots,J\}\setminus N}\}, \delta^S, \delta^U)$.
 Assign position i to product j such that $CS_{ij} \in \arg\max_{l\in N\cup\{0\}} CS_{il}$.
 if $j = 0$ **then**
  Break;
 else
  Update $N = N \setminus j$;
 end
end.

3.3. Revenue Maximization

The platform might also consider matching products to ranks so as to maximize the platform's revenues. We begin again with a relaxation of the problem in which ranks can be assigned continuously and then consider the discrete problem.

Let π_j be the revenue that the platform earns from selling product j. We focus on revenue maximization because marginal costs are essentially zero for many platforms, and thus, revenues correspond to profits. However, an analogous argument applies to the case in which marginal costs are nonzero and the platform maximizes profits. The platform's revenue maximization problem can be written as

$$\max_{r_1\ldots r_J} \sum_j q_j(\mathbf{r})\pi_j$$
$$\text{subject to} \sum_{j:j \text{ is listed}} r_j \le \frac{J(J+1)}{2}.$$

We let $\pi(\mathbf{r}, \delta^{\mathbf{S}}, \delta^{\mathbf{U}})$ denote the objective function, that is, the expected (per-customer) revenue. The optimal solution has the property that the partial derivative of expected revenues is equal for all products that are ranked: $\partial\pi(\mathbf{r}, \delta^S, \delta^U)/\partial r_j = \partial\pi(\mathbf{r}, \delta^S, \delta^U)/\partial r_k\ \forall j,k$. However, unlike the consumer surplus, even if it is feasible to choose ranks so that $\delta^S_{ij} + f(r_j) \ge \delta^U_{ij}\ \forall j$ (i.e., every product has higher search than utility index), it might not be optimal for revenues because consumer utilities are not necessarily positively correlated with the revenue from each product. Thus, a short-term revenue-maximizing platform has an incentive to distort ranks (from a consumer surplus perspective) even with unlimited ranking power. To get more intuition, take the derivative of platform revenue with respect to r_j (we derive this in Online

Appendix A.6):

$$\frac{\partial \pi(\mathbf{r}, \delta^S, \delta^U)}{\partial r_j} = \int_{\{\varepsilon_{ij}^S - \varepsilon_{ij}^{post} < \delta_j^U - \delta_j^S - f(r_j)\}} q_{ij}(\mathbf{r}) f'(r_j) \times \left(\pi_j - \sum_{k \le J} q_{ik}(\mathbf{r}) \pi_k \right) dF_{\varepsilon_i^S, \varepsilon_i^{post}}.$$

The expression resembles that of the consumer surplus derivative with the revenue margins π_j taking the place of the potentials ϕ_{ij}. But, because the revenues are fixed even as the rankings are adjusted, the optimal solution is quite different. When there is no outside option, the platform should only display the highest margin good. The intuition is clear: if consumers buy something regardless, steer them in the direction of highest revenues.

With an outside option, the platform has to balance the probability that the consumer buys anything at all with the incentive to push the highest margin products. The sign of the derivative depends on the term $\pi_j - \sum_{k \le J} q_{ik} \pi_k$. This is the revenue from j less the choice-probability weighted revenues from all other products (which includes the zero-margin outside option). So, when most consumers don't purchase anything (i.e., the weighted revenue is close to zero), this term is positive and all products are ranked with high-margin and high–effective index products getting the top spots. But, if consumers who are presented with the full product assortment are likely to purchase something, the weighted revenue exceeds the revenue offered by some of the products in the assortment, and the platform can improve revenues by excluding those low-margin products. Platforms whose customers are unlikely to shop elsewhere (or not buy at all) can therefore afford to choose a product assortment that consists mostly of high-margin products, whereas platforms with less loyal customers cannot.

3.3.1. The OPT-K Algorithm. We propose the analogue OPT-K algorithm for revenue maximization: instead of ranking all J products, we only consider assignments to the first K positions to maximize revenues from those K or fewer products. Define

$$\pi^K = \sum_{j: r_j \ge J-K+1} q_j(\mathbf{r}) \pi_j,$$

and let the ranking that maximizes π^K be $\mathbf{r}^K$. Maximizing π^K again requires only $\sum_{k=1}^{K} J!/(J-k)!$ evaluations.

Proposition 3 (Approximate Optimality). *Let $K < J$, and assume that $\sum_{j: r_j < J-K+1} q_j < \nu$ for any feasible ranking. Then, the gap between the revenue achieved by optimally allocating all J positions and that achieved by optimally allocating the first K positions can be bounded as follows:*

$$\pi(\boldsymbol{r}^*) - \pi(\boldsymbol{r}^K) \le \nu \max_j \pi_j,$$

where $\boldsymbol{r}^* \in \arg\max_{\mathbf{r}} \pi$ and $\boldsymbol{r}^K \in \arg\max_{\mathbf{r}} \pi^K$.

Proof. See Online Appendix A.7. □

Analogously to Proposition 2, Proposition 3 says that, if products placed after position K are unlikely to be bought, then optimizing product assignments for the first K positions can be sufficiently close to the optimal ranking for all products. The intuition for the result is straightforward: the revenue left on the table by not assigning products to positions beyond K is at most equal to the probability that consumers choose products in those positions times the maximum revenue delivered by any one product. As before, it may be useful in practice to rank the remaining $J-K$ products. In the same way as earlier, we define a greedy algorithm OPTKG that maximizes revenues over all possible assignments of the first K products and then greedily assigns each of the remaining products by checking which assignment generates the greatest improvement in revenues over the prior assignment, terminating if nonassignment is ever the best option.

3.4. Simulations

Whereas Propositions 2 and 3 give theoretical guarantees for the performance of the OPT-K algorithm, in practice, the algorithm's performance depends on the number of products J and the choice of K and the distributions of utilities and revenues as well as whether the conditions of the propositions are met. In this section, we illustrate, via simulations, the performance of the OPT-K algorithm under a wide range of conditions. We also discuss in practice how to optimize rankings for the remaining $J-K$ products as well as the runtime of various OPT-K algorithms.

3.4.1. Simulation Environment. We simulate $J = 5$ products to be assigned to ranks $r \in \{5, 4, \dots, 1\}$, in which the effect of ranking on the mean search index exponentially decays; that is, $f(r) = A \cdot exp(r - 6)$ for $A \in \{5, 15, 30\}$. We limit the number of products to five to retain the ability to brute force and find the actual best and worst assignments as benchmark. Product-specific profits are drawn iid from a *Lognormal*(0, 1) distribution, whereas mean search and utility indices of each product are drawn from iid normal distributions, that is, $\delta_j^S \sim i.i.d.N(0, 0.5)$ and $\delta_j^U \sim i.i.d.N(\mu, 0.5)$ for $\mu \in \{-5, 0, 5\}$. For each combination of (A, μ), we simulate 100 times, for a total of 900 simulations. For each simulation draw, we first find the maximum and minimum consumer surplus CS_{max} and CS_{min} by enumerating all $\sum_{j=1}^{J} J!/(J-j)!$ possible rankings and then report the consumer surplus under

Table 1. OPT-K Performance

J = 5; 9×100 simulations	Consumer surplus, %			Platform revenues, %		
Positions $K+1,\ldots J$	None	Random	Greedy	None	Random	Greedy
K = 1	37.4	85.6	98.2	82.3	63.0	98.2
K = 2	62.6	89.0	98.8	93.7	92.5	98.1
K = 3	83.6	92.0	98.3	97.5	96.7	98.5
K = 4	90.5	94.9	99.2	99.7	98.3	99.7
K = 5	100	100	100	100	100	100

Notes. Simulation results of OPT-K algorithm performance for consumer surplus and platform revenues with J = 5 products. Positions $K+1,\ldots,J$ assign (i) no products, (ii) randomly ordered products, or (iii) products according to OPTKG. Performance is first normalized by $(Q-Q_{min})/(Q_{max}-Q_{min})$, where $Q \in \{CS, \pi\}$ and Q_{min} and Q_{max} are obtained via enumeration, and then averaged across nine combinations of parameters $(A,\mu)\times 100$ simulations per each set of parameters.

the OPT-K algorithm as $(CS-CS_{min})/(CS_{max}-CS_{min})$. We follow the same procedure for platform revenues.

3.4.2. Results. Table 1 presents the performance results of the OPT-K algorithm for both consumer surplus and platform revenues. In each simulation, we compare three algorithms that differ in how they assign positions beyond K: (i) the first, labeled "none" in the table, does not assign any products to positions beyond K; (ii) the second ("random") assigns products randomly to positions beyond K; and (iii) the third ("greedy") corresponds to the greedy OPTKG algorithm described earlier.[19] We repeat this for $K=\{1,2,\ldots,5\}$ and note that each of these three algorithms is equivalent to the brute force approach (i.e., trying all possible rankings of the five products) when $K=5$.

The greedy algorithms consistently achieve more than 98% of the consumer surplus and revenue gains. This is true even for $K=1$, that is, when only the top position is brute forced and the remaining four are assigned greedily. In contrast, the none and random algorithms exhibit a much steeper gradient from $K=1$ to $K=5$, highlighting the importance of how the positions beyond K are assigned. Still, all algorithms capture at least 80% of the gains for $K\geq 3$.

We now turn to the computational cost of OPTKG. The number of evaluations required to execute OPT-K is $\sum_{k=1}^{K} J!/(J-k)!$, which scales up quickly with both K and J. On the other hand, the greedy algorithm is cheap: if OPT-K assigns products to all K positions, then the OPTKG algorithm adds at most an additional cost of $(J-K)(J-K+3)/2$ evaluations,[20] depending on how many products end up being assigned. Table 2 compares different OPT-K algorithms in terms of their runtimes. We increase the number of products to $J=15$, which is more realistic in practice: it roughly represents the first full page of search results that are most salient to consumers on a typical platform. With an otherwise similar simulation environment, we run various OPT-K algorithms for five times for each set of parameters. Whereas runtimes increase quickly with K, the OPT-K algorithm remains computationally feasible for small K when $J=15$. The additional cost of the OPTKG algorithm is small and does not meaningfully scale with K. Runtime for consumer surplus and platform revenues are similar. We conclude that, in our simulations, the OPTKG algorithm is computationally feasible and close to optimal for small K. Based on these results, we use the OPTKG algorithm with $K=3$ in our empirical application.

4. Empirical Application

We apply our method to study the customer search and choice data from a hotel booking platform (Expedia).

Table 2. OPT-K Runtime (in Seconds)

J = 15; 9×5 simulations	Consumer surplus		Platform revenues	
Positions $K+1,\ldots J$	None	Greedy	None	Greedy
K = 1	0.06	1.71	0.09	1.86
K = 2	0.37	0.71	0.58	1.32
K = 3	4.28	4.51	5.69	5.95
K = 4	52	52	77	77
K = 5	880	879	891	893

Notes. Simulation results of OPT-K algorithm runtime for consumer surplus and platform revenues in seconds with $J=15$ products. Positions $K+1,\ldots,J$ assign (i) no products or (ii) products according to OPTKG. Runtimes are averaged across nine combinations of parameters $(A,\mu)\times 5$ simulations per each set of parameters.

Table 3. Summary Statistics: Hotel Characteristics

	Observations	Mean	Median	Standard deviation	Minimum	Maximum
Training sample						
Star rating	2,922,728	3.17	3.00	1.06	0	5
Review score	2,922,728	3.69	4.00	1.17	0	5
Chain	2,922,728	0.60	1.00	0.49	0	1
Location score	2,922,728	2.88	2.83	1.55	0	6.98
Price	2,922,728	164.90	128.00	145.20	0	5,000
Promotion	2,922,728	0.20	0	0.40	0	1
Testing sample						
Star rating	6,947,458	3.18	3.00	1.04	0	5
Review score	6,947,458	3.81	4.00	1.01	0	5
Chain	6,947,458	0.65	1.00	0.48	0	1
Location score	6,947,458	2.87	2.77	1.52	0	6.98
Price	6,947,458	149.30	120.00	112.50	0	5,000
Promotion	6,947,458	0.22	0	0.42	0	1

Notes. The table shows summary statistics of hotels. An observation is a hotel impression so that hotels are weighted by their appearance in search results.

Using our demand estimates, we compare our optimal ranking algorithm to both a utility-based ranking and the Expedia ranking, evaluating how they each perform with respect to revenue, consumer surplus, and number of searches.

4.1. Data and Descriptive Evidence

Our data come from Expedia and are made available through Kaggle.com, an online platform on which data miners can use data sets to take part in competitions posted by companies. We refer the reader to Ursu (2018) for a comprehensive discussion of the data; here, we focus on the features that are most directly relevant to our analysis. The Expedia data are composed of two data sets, training and testing data sets. Both data sets contain customer search and choice records from November 2012 to June 2013 among 34 different markets. There are 120,883 search impressions in the training data and 276,644 impressions in the testing data.[21] Both training and testing data contain only impressions for which customers searched at least once, and each impression displays 5 to 38 different hotels. There are 124,561 hotels in the training data and 130,136 hotels in the testing data. We observe their rating, price, country, review score, whether it belongs to a chain or not, and location score. Further, we observe whether the hotel is being clicked on and whether the hotel is booked.

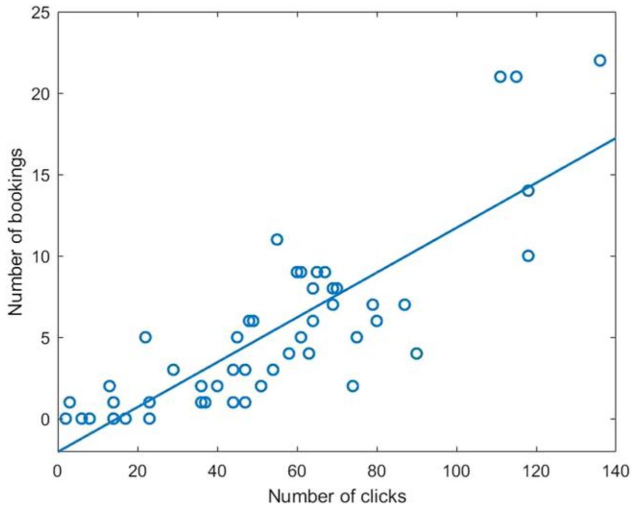
The main difference between the two data sets is that, in the training data, hotels are ranked randomly, whereas in the testing data, hotels are ranked according to the (proprietary) Expedia algorithm. The advantage of the training data is that we can identify the effect of rankings on the customer search index without having to worry about endogeneity of ranks. Although we observe whether a customer searched a hotel or not, we do not observe the order of search. In the estimation, this requires us to integrate out along all possible permutations of search paths. Table 3 shows summary statistics for hotel characteristics in the training and testing samples. One can see that the hotel characteristics have very similar distributions in the two samples except that prices are slightly lower in the testing sample. Table 4 shows summary statistics on consumer behavior. First, consumers face fairly large choice sets,

Table 4. Summary Statistics: Consumer Behavior

	Observations	Mean	Median	Standard deviation	Minimum	Maximum
Training sample						
Number of hotels in choice set	120,883	24.18	28.00	9.19	5	38
Number of searches	120,883	1.10	1.00	0.43	1	5
Indicator for purchase	120,883	0.13	0	0.34	0	1
Testing sample						
Number of hotels in choice set	276,644	25.11	30.00	9.08	5	38
Number of searches	276,644	1.08	1.00	0.38	1	5
Indicator for purchase	276,644	0.94	1.00	0.24	0	1

Note. The table shows summary statistics for consumer behavior at the search impression level.

Figure 2. (Color online) Clicks and Bookings



Notes. The figure shows the relationship between the number of clicks and the number of bookings for the top 50 most displayed hotels in the training data. Each dot corresponds to a hotel, and the best linear fit is plotted.

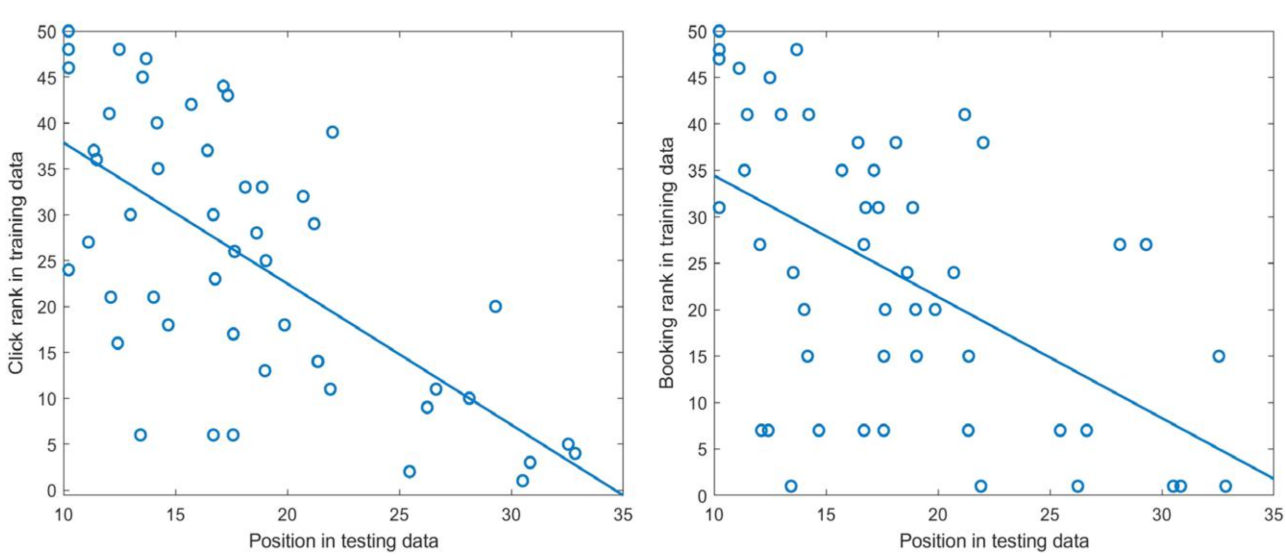
consisting of around 25 hotels on average. In spite of this, consumers only search slightly more than one hotel on average, suggesting that rankings are likely to play an important role in driving final choices. There is, however, heterogeneity in search with some consumers clicking on several hotels. Note that, in order to keep the integration along search paths tractable, we drop impressions for which customers searched more than five times, which corresponds to around 0.5% of the overall sample.

Next, we provide some descriptive evidence to motivate our model. Figure 2 shows the relationship between the number of clicks and the number of bookings for the same hotel in the training data.[22] Whereas hotels that are clicked more often also tend to be booked more, there is also substantial independent variation in the two variables. This suggests that search and choice patterns are driven by two distinct mechanisms, which motivates our double index model. In other words, a model featuring a single index, such as a standard discrete choice model, is not likely to fit the data well.

Second, we look at the ranking algorithm used by Expedia in the testing data. Figure 3 relates a hotel's average position in the testing data with the number of clicks (left panel) and bookings (right panel) in the training data. We expect that hotels that are more often clicked and purchased in the training data (in which ranks are assigned randomly) are the ones that Expedia would choose to rank more favorably in its own algorithm. This is indeed the case.

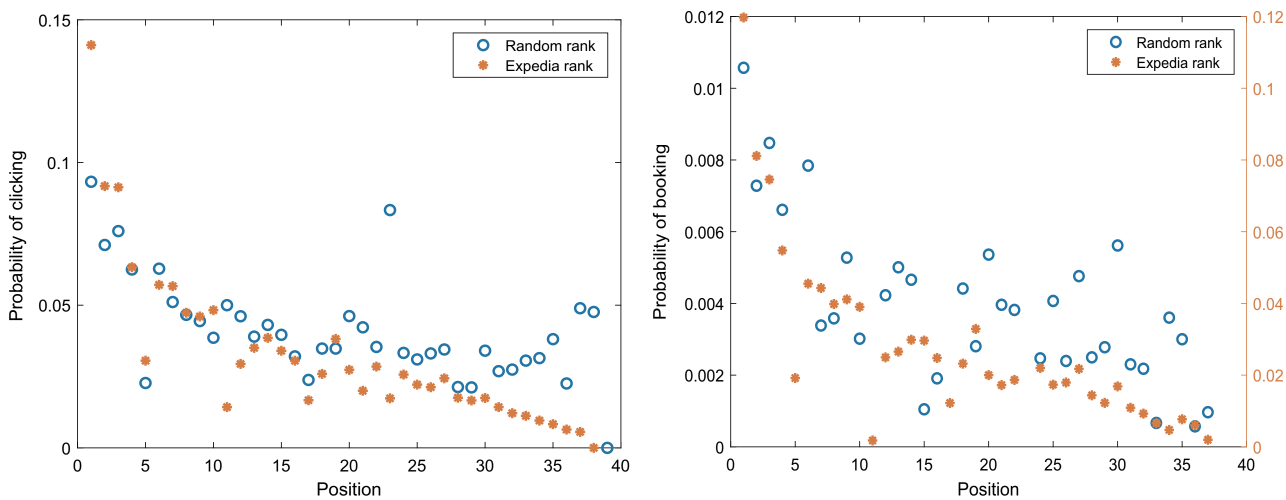
Finally, Figure 4 shows how the probability of clicking and booking a hotel varies with the average hotel position across the two data sets. As expected, better positions are associated with higher click probabilities on average in both data sets (left panel). However, the slope of the relationship is steeper in the testing data. This is consistent with the idea that Expedia is optimizing its rankings so that the hotels ranked in the first few positions are relatively more attractive than random hotels. Notice also that the probability of searching a hotel in position 30 or above (i.e., lower rank) declines to almost zero in the testing data but is still relatively high in the training data. This suggests that, when rankings are not optimized, consumers end up having to search further. Turning to bookings, the right panel of

Figure 3. (Color online) Clicks and Bookings in the Training Data as a Function of Average Expedia Position



Notes. The figure shows the relationship between the average position of a hotel in the testing data and its click rank in the training data (left panel) and its booking rank in the training data (right panel). Each dot corresponds to a hotel, and the best linear fit is plotted.

Figure 4. (Color online) Probability of Clicking and Booking as a Function of Position



Notes. The figure shows the relationship between the position of a hotel and the probability of it being clicked (left panel) and being booked (right panel). Each dot corresponds to a position with unfilled dots referring to the training data and filled dots referring to the testing data. In the right panel, the booking probability for the training and testing data are on separate axes (left and right axes, respectively) as the testing data oversamples searches that conclude in booking.

Figure 4 shows similar patterns: the booking rate declines more rapidly as a function of rank in the testing than the training data, and low-ranked hotels are sometimes booked in the training data but almost never in the testing data.[23]

4.2. Estimation

We now discuss how we estimate our model of search and choice. As discussed in Section 2, our model features two indices per hotel with the following specifications:

$$s_{ijt} = \beta^S x_{jt} + f(r_{jt}) + \xi_j^S + \varepsilon_{ijt}^{pre} + \varepsilon_{ijt}^S$$
$$u_{ijt} = \beta^U x_{jt} + \xi_j^U + \varepsilon_{ijt}^{pre} + \varepsilon_{ijt}^{post}. \tag{8}$$

Note that we now include a t subscript to denote the impression in which a given hotel is shown to consumers. To facilitate estimation, we take the shocks ε_{ijt}^{post} and ε_{ijt}^S (in addition to ε_{ijt}^{pre}) to be distributed iid Gumbel. We also assume that the observed hotel characteristics, including price, are exogenous conditional on the hotel fixed effects ξ_j^U.[24] Note that rankings are exogenous by construction in the training data because they are randomized.

We estimate the model via maximum likelihood. In order to write the likelihood of each consumer's observed click and purchase outcomes, we proceed in three steps. First, we exploit a convenient property of the Gumbel distribution to obtain a closed form for the outcome probability for any given search sequence, conditional on a realization of the vector $\varepsilon_i^{pre} \equiv (\varepsilon_{i1}^{pre}, \ldots, \varepsilon_{iJ}^{pre})$. Then, because the search sequence is not observed in the data, we sum over all possible sequences to obtain the probability of the observed outcome given ε_i^{pre}. Finally, we integrate out ε_i^{pre}.

To illustrate, consider a simple example with two hotels and an outside option, and suppose that the data tells us that consumer i searches both hotels and books hotel 1. This set of outcomes is consistent with two search sequences:

- Search hotel 1, then 2, then book 1 (sequence A).
- Search hotel 2, then 1, then book 1 (sequence B).

Dropping the t subscripts for simplicity, the only possible ordering of the search and utility indices associated with sequence A is

$$s_{i1} \geq s_{i2} \geq u_{i1} \geq u_{i2}, u_{i0}. \tag{9}$$

This is because (i) $s_{i1} \geq s_{i2} \geq u_{i0}$ given that both hotels are searched and 1 is searched before 2; (ii) $s_{i2} \geq u_{i1}$ because 2 is searched after u_{i1} is revealed; and (iii) $u_{i1} \geq u_{i2}, u_{i0}$ because 1 is chosen. Under the maintained assumptions and for fixed $(\varepsilon_{i1}^{pre}, \varepsilon_{i2}^{pre})$, the probability of the ordering in (9) is equal to

$$\frac{\exp(\delta_{i1}^S)}{1 + \exp(\delta_{i1}^S) + \exp(\delta_{i2}^S) + \exp(\delta_{i1}^U) + \exp(\delta_{i2}^U)}$$
$$\times \frac{\exp(\delta_{i2}^S)}{1 + \exp(\delta_{i2}^S) + \exp(\delta_{i1}^U) + \exp(\delta_{i2}^U)}$$
$$\times \frac{\exp(\delta_{i1}^U)}{1 + \exp(\delta_{i1}^U) + \exp(\delta_{i2}^U)}, \tag{10}$$

where, as before, we let $\delta_{ij}^S = \beta^S x_j^S + f(r_j) + \xi_j^S + \varepsilon_{ij}^{pre}$ and $\delta_{ij}^U = \beta^U x_j + \xi_j^U + \varepsilon_{ij}^{pre}$ for $j = 1, 2$. The expression in (10), sometimes referred to as the exploded logit trick (e.g., Chapman and Staelin 1982, Bradlow and Fader 2001), is convenient in that it allows us to write the likelihood in a way that does not rely heavily on computationally costly simulation methods.[25] Specifically, we do not need to numerically integrate over the idiosyncratic shocks $(\varepsilon^U_{i1}, \varepsilon^U_{i2}, \varepsilon^{post}_{i1}, \varepsilon^{post}_{i2})$ although we do have to do so for $(\varepsilon^{pre}_{i1}, \varepsilon^{pre}_{i2})$.

Similarly, for sequence B, there are three possible orderings:

$$s_{i2} \geq s_{i1} \geq u_{i1} \geq u_{i2}, u_{i0}$$
$$u_{i1} \geq s_{i2} \geq s_{i1} \geq u_{i2}, u_{i0}$$
$$s_{i2} \geq u_{i1} \geq s_{i1} \geq u_{i2}, u_{i0}\,.$$

We can write the probability of each of these orderings using the exploded logit formula as in (10); the sum across the three orderings then gives us the outcome probability associated with sequence B. Next, summing the outcome probabilities associated with the two sequences and integrating over $(\varepsilon^{pre}_{i1}, \varepsilon^{pre}_{i2})$, we obtain the likelihood of the observed outcome for consumer i. A similar logic applies to more complicated outcomes involving more than two hotels, and thus, we are able to write the probability of the data in closed form.

Finally, note that the data only covers consumers who clicked on at least one hotel. To account for this sample selection, we divide the likelihood by the probability of clicking on at least one hotel (which again can be written in closed form using the exploded logit formula). This is the (conditional) likelihood that we maximize. More formally, the log-likelihood is given by

$$\ell(data;\theta) = \sum_t \log\left(\frac{\sum_{\text{ord}_t} \int P(\text{ord}_t; \mathbf{x}_t, \theta) dF_{\varepsilon^{pre}}}{P(\text{click at least one hotel}; x_t, \theta)}\right), \tag{11}$$

where $\mathbf{x}_t = (x_{1t}, \ldots, x_{Jt})$, ord_t indexes the different possible orderings of search and utility indices for impression t, $F_{\varepsilon^{pre}}$ denotes the distribution of $(\varepsilon^{pre}_1, \ldots, \varepsilon^{pre}_J)$, and $P(\text{ord}_t; x_t, \theta)$ denotes the probability of ordering ord_t based on the exploded logit formula (e.g., the expression in (10)).

We now provide some intuition for the type of variation in the data that allows us to identify the model. The correlation between hotel characteristics and the probability that hotels are clicked identifies the parameters in the search indices, including the position effects $f(r_j)$. Similarly, the correlation between hotel characteristics and the probability that hotels are chosen, conditional on being clicked on, identifies the parameters in the utility indices. Notice that it is necessary to have a model of search in order to consistently estimate the choice model because the options in the consumers' consideration sets are endogenously determined, and the idiosyncratic errors of options in the consideration set are generally not iid extreme value. In other words, simply estimating a logit model on the observed consideration sets is likely to lead to biased results.

4.3. Results

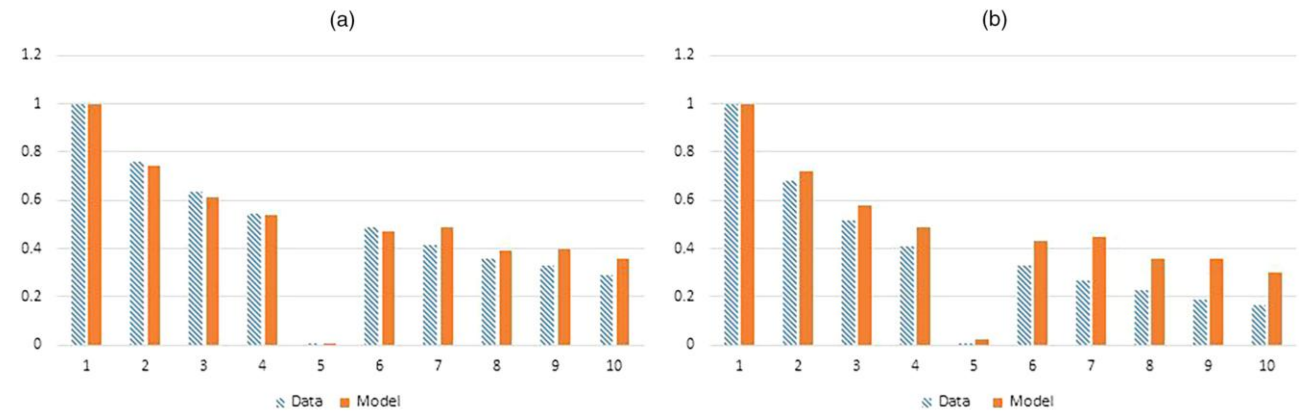
We fit our double index model to the training data set by maximum likelihood. Table 5 presents the results. The first two columns of the table present coefficients and standard errors for the search index, whereas the last two columns present coefficients and standard errors for the utility index. First, notice that, as expected, the rank position coefficients in the search index tend to decrease with position (a striking exception to this pattern is position 11; this is consistent with the fact that Expedia places "opaque offers," that is, offers in which the consumer does not know the name of the hotel before making a purchase, in this position, as well as in position 5—see Ursu (2018) for more on this point). All other coefficient estimates seem reasonable. For example, price coefficients are negative in both the search and the utility indices, which suggests that higher prices not only reduce the probability of booking conditional on searching, but also deters customers from searching in the first place. Similarly, higher review scores positively impact both search and booking.

Table 5. Maximum Likelihood Estimates of the Model

	Search index		Utility index	
	Coefficient	Standard error	Coefficient	Standard error
Position 1	1.34	0.06		
Position 2	0.97	0.06		
Position 3	0.73	0.06		
Position 4	0.58	0.07		
Position 5	0.40	0.38		
Position 6	0.41	0.06		
Position 7	0.49	0.06		
Position 8	0.25	0.07		
Position 9	0.29	0.07		
Position 10	0.20	0.07		
Position 11	−1.76	6.64		
Position 12	0.19	0.08		
Star rating	0.29	0.02	0.12	0.04
Review score	0.09	0.01	0.16	0.04
Chain	0.09	0.04	0.48	0.10
Location score	0.12	0.02	0.02	0.03
Price ($100)	−0.33	0.03	−0.44	0.05
Promotion	0.31	0.04	0.20	0.09

Notes. The table shows the estimates of our model. The first two columns report the coefficients and standard errors for the search index parameters, whereas the last two refer to the utility index parameters. Hotel fixed effects are included in both the search and the utility index but are not reported. Standard errors are bootstrapped.

Figure 5. (Color online) Model Fit



Notes. (a) In sample. (b) Out of sample. This figure shows the probability of clicking for positions 1–10 relative to the probability for position 1. Hashed bars represent the data and solid bars represent the model predictions. Panel (a) refers to the training data, whereas panel (b) refers to the testing data.

4.4. Model Fit

Figure 5 shows the average odds of clicking on a hotel for different positions in the training and testing samples. We can see that the model does a good job at matching the patterns in the data, especially for the first few positions. The model tends to overestimate the odds of search for positions 7–10 in the testing data. Intuitively, in the estimation data—which features random rankings—the hotels in these positions are clicked on fairly often. In contrast, when Expedia uses its proprietary ranking, consumers do not click on those positions as much because it's more likely that they will find a good option at the top of the page and decide to stop searching. As shown in Figure 5, the model captures this difference to a certain extent—the solid bars for positions 7–10 are lower in panel (b) than in panel (a)—but not quite to the same degree as in the data.

4.5. Counterfactual Analysis

Given the estimated model, we perform a range of counterfactual exercises to compare our proposed algorithm to competing algorithms. Specifically, we draw 1,000 customers (i.e., choice sets) at random from the testing data. For each customer, we use the model and the estimated search and utility index parameters to compute the expected consumer welfare, number of searches, and revenue under four rankings.[26] The four rankings are (i) "OPT3G-CS," our approximately optimal algorithm for maximizing consumer surplus with an exhaustive search over the top three positions and a greedy algorithm for rest of the positions; (ii) "OPT3G-Rev," our approximately optimal algorithm for maximizing revenue; (iii) "utility," the algorithm that simply ranks goods in descending order of their utility indices; and finally (iv) the Expedia ranking.

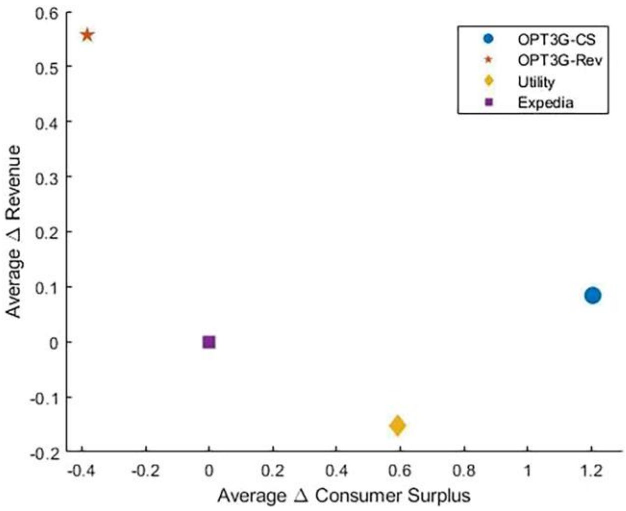
Table 6 shows the average changes in each of these metrics for the OPT3G-CS, OPT3G-Rev, and utility rankings relative to the Expedia algorithm. OPT3G-CS yields an average gain in consumer surplus of more than \$1.20 per customer relative to Expedia. Remember that this is an average over all consumers, including consumers who do not even search, because our model allows for nonsearch as an option. The simulated conversion rate over this population is around 13%, implying a gain in consumer surplus for consumers who purchase of around \$10 (similarly, OPT3G-Rev delivers more than \$4.5 per purchase in additional revenue).

One might wonder whether achieving this higher surplus requires more search on the part of consumers as, in principle, higher search costs could offset the gains from finding a better match. We find that, on average, consumers engage in only 0.0048 additional searches under OPT3G-CS relative to Expedia. Whereas our model does not provide an estimate of search costs, this very small difference suggests that the additional search costs implied by our algorithm are negligible relative to the utility gains.[27] Specifically, the results imply that accounting for search costs would reverse our welfare numbers—that is, make the change in surplus associated with OPT3G-CS (relative to Expedia) negative—only if the per-click search cost were greater than $\$1.205/0.0048 = \$251$, which is unrealistic. We also note that the magnitudes of the consumer welfare numbers should be interpreted with caution because, in the Expedia data set, impressions leading to a transaction were oversampled though it is unclear whether this is

Table 6. Average Changes Relative to the Expedia Ranking

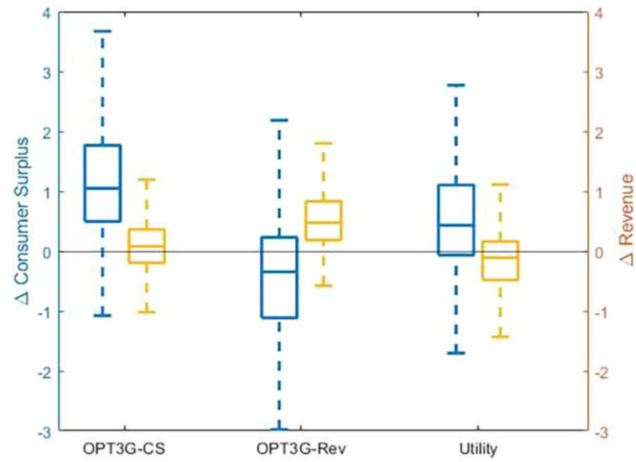
	OPT3G-CS	OPT3G-Rev	Utility
Δ Consumer surplus (\$)	1.2050	−0.3826	0.5937
Δ Search count	0.0048	−0.0009	0.0038
Δ Revenue (\$)	0.0843	0.5575	−0.1520

Figure 6. (Color online) Trade-off Between Consumer Surplus and Revenue Maximization



Note. The figure shows the relationship between the average consumer surplus and the average revenue for the OPT3G-CS, OPT3G-Rev, and utility algorithms relative to the Expedia algorithm.

Figure 7. (Color online) Changes in Consumer Surplus, Number of Searches, and Revenue Relative to the Expedia Ranking



Notes. The figure compares the OPT3G-CS, OPT3G-Rev, and utility algorithms to that used by Expedia. For each of them, the boxplots on the left show the distribution of changes in consumer surplus, and the boxplots on the right show the distribution of changes in revenues. Each box marks the 25th, 50th, and 75th percentiles, and the dotted lines indicate where approximately 95% of each distribution lies.

only true of the testing data (which has a 94% purchase rate versus 13% in the training data) or both. Still, the comparison is informative about the relative performances of the different ranking algorithms.

Next, Figure 6 shows that there is a trade-off between revenue and consumer surplus. Among the four algorithms, OPT3G-CS sacrifices some revenue in order to maximize expected consumer surplus, whereas OPT3G-Rev achieves higher average revenues at the cost of much lower consumer surplus. Notice, however, that OPT3G-CS dominates the Expedia and utility rankings in terms of both consumer surplus and revenue. In this sense, the Expedia and utility rankings are within the Pareto frontier.[28]

Going beyond averages, Figure 7 shows the entire distribution (across impressions) of consumer surplus and revenue for the OPT3G-CS, OPT3G-Rev, and utility algorithms relative to the Expedia ranking. One can see that, whereas there is meaningful heterogeneity across impressions, the results for the averages broadly continue to hold when we look at the full distributions of customers.

In order to shed further light on how the algorithms perform for any given customer, Table 7 reports the fraction of customers for whom each of the four algorithms maximizes expected surplus (relative to the remaining three). As expected, OPT3G-CS maximizes consumer surplus for the vast majority of choice sets. In comparison, the Expedia and utility rankings are optimal only for up to 10% of customers each. OPT3G-Rev comes in last, which is not surprising because it targets a different objective function.

Note that it is to be expected that OPT3G-CS does not always maximize consumer surplus because OPT-K is an approximately optimal algorithm for our choice of $K=3$ (which is smaller than the number of hotels available for ranking in any given impression). As such, it is possible that other algorithms would sometimes dominate it. This being said, the fact that OPT3G-CS still maximizes consumer surplus in more than three quarters of choice sets is reassuring and suggests that we are not too far from the optimal algorithm.

The results so far show that our OPT3G algorithms achieve desirable outcomes relative to two competing

Table 7. Breakdown of Customers by Which Algorithm Maximizes Expected Consumer Surplus

	OPT3G-CS	OPT3G-Rev	Utility	Expedia
Percentage of customers	79.90	2.00	10.70	7.70

Note. For each algorithm, the table reports the fraction of customers in the testing data for whom the algorithm yields a higher consumer surplus than the remaining three.

Table 8. Comparison with Random Rankings

	Random			
	Minimum	Median	Maximum	OPT3G
Δ Consumer surplus	−2.91	−0.87	1.20	1.21
Δ Revenue	−1.32	−0.38	0.61	0.56

Notes. For any given impression, we compare the changes (relative to Expedia) in consumer surplus and revenue achieved by OPT3G-CS and OPT3G-Rev, respectively, with those achieved by 1,000 random rankings. The table reports averages across 1,000 randomly selected impressions.

rankings, the Expedia and utility rankings. However, they do not say anything about the performance of OPT3G in comparison with any of the (many) other possible algorithms. In order to shed some light on this, we perform the following exercise. For a given impression in the testing data, we draw 1,000 rankings at random and compute the corresponding consumer surplus and revenue. We then compare the distribution of surplus and revenue across the random rankings with the performance achieved by our OPTK algorithms. We repeat this exercise for 1,000 randomly drawn impressions and report averages across impressions in Table 8. The results show that OPT3G-CS dominates all of the random rankings, on average, suggesting that it is close to the best possible ranking. On the other hand, OPT3G-Rev—which improves over Expedia by \$0.56, on average—tends to be slightly outperformed by the best random ranking (which improves over Expedia by \$0.61), but its performance is substantially better than the median random ranking, which reduces revenue by \$0.38 relative to Expedia. Thus, OPT3G-Rev appears to be close to the frontier as well.

5. Conclusion

Engagement with the virtual world seems likely to increase over time. How and to what platforms choose to direct consumer attention is a key component of the online world, shaping the choices that consumers make. The model presented here offers one way of formalizing this relationship between platforms and consumers, positing that platforms can affect a search index that determines what consumers choose to view though it cannot directly affect the consumption utility from content. Using this model, we can show that, when a platform wants to optimize consumer surplus, it should aid search discovery by promoting products that are better than the customer would have believed based on the information available presearch (i.e., by surfacing products that would have otherwise been overlooked).

Whereas we base our model on the canonical sequential search protocol, several other search strategies share the same structure featuring two indices per good. Generalizing beyond Weitzman search would be an interesting avenue for future research. This may require more complicated expressions for the search and utility indices—possibly leading to nonparametric specifications—and it may also lead to a strategy for testing which search protocol consumers follow.

A second direction would be to incorporate sponsored ads into the model. The current framework focuses on organic rankings in a platform without sponsored ads. However, firms selling high-potential products may be willing to pay for better ranks, generating an additional source of revenue that the platform should take into consideration when optimizing its organic rankings.

Finally, one might wonder whether the tension between maximizing revenues and optimizing consumer surplus is a consequence of the fact that the model is static. In a dynamic model in which Expedia takes into account the effect of its rankings not just on present, but also on future outcomes, it's possible that the two objectives might be more aligned. For instance, if consumer surplus is a good predictor of the likelihood of customer retention, then maximizing consumer surplus may also be optimal from a revenue perspective in the long run. Because the data does not track consumers over time, we are unable to explore this question here, but it would be an interesting avenue for future research.

Acknowledgments

The authors are grateful for discussions with Bart Bronnenberg, Tat Chan, Jean-Pierre Dubé, Pedro Gardete, Rafael Greminger, Carl Mela, Sridhar Moorthy, Ilya Morozov, Stephan Seiler, Adam Smith, Raluca Ursu, Giorgos Zervas, and Fan Zhang as well as participants at the 2020 Marketing Science Conference, Netflix, UCL, Tilburg, 2022 Workshop on Consumer Search and Switching Costs, Virtual Quantitative Marketing Seminar, Wharton, Stanford, the University of Toronto and Wash University Junior Faculty Development Forum. Adam Sy provided excellent research assistance. This project was started when G. Compiani was an intern at Microsoft Research New England. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial or nonfinancial interest in the subject matter or materials discussed in this manuscript. The authors have no funding to report.

Endnotes

[1] De los Santos and Koulayev (2017) propose a utility-based ranking that maximizes click-through rates.

[2] In our framework, the act of searching a product conflates two distinct actions: scrolling to the product in the results page and clicking on it. Greminger (2022a) develops a model that distinguishes these two steps, and Greminger (2022b) estimates it on the same data as the present paper.

[3] Papers that model consumers as searching over their match value include Kim et al. (2010) and Ursu (2018). A paper that, like ours, allows for vertical dimensions of quality to be revealed with search is Jiang et al. (2021).

[4] Using a different model, Derakhshan et al. (2022) find a similar result.

[5] Because the marginal cost of making a sale are essentially zero for many platforms (sellers are responsible for the physical delivery of the good), revenues often coincide with profits, and we use the two interchangeably.

[6] The revenue the platform earns from selling each product depends on its individual agreements with each seller but is typically a fixed percentage of the sale price.

[7] This is consistent with the insights in Ursu and Dzyabura (2020) although the mechanism through which this plays out is different. In Ursu and Dzyabura (2020), the retailer/platform may have an incentive to promote products with lower utility because this can lead consumers to search more products and, thus, be more likely to purchase one of them. In contrast, in our case, even for fixed search intensity, the platform may prefer to promote products that have lower utility as long as they yield higher revenues. Indeed, in our application, we find that there is a trade-off between consumer surplus and revenue even though the two algorithms lead to essentially the same number of searches on average.

[8] Placing products with low search indices and high utilities (i.e., high potential) in top positions may also have the unfortunate effect of diminishing consumer trust in the search algorithm if the consumers never sample them. Good recommendations are those that are followed.

[9] Whereas we do not observe the exact order of search in our data, we see on which products a given consumer searched. In more than 80% of impressions, the consumer searched a product A but not some of the products ranked higher than A, which implies that they did not search in the order in which the products were ranked.

[10] The distinction between ε_{ij}^{pre} and ε_{ij}^{post} corresponds to that between presearch and postsearch shocks in Ursu et al. (2022).

[11] For a relaxation of this assumption, see Bronnenberg et al. (2016) and Hodgson and Lewis (2020).

[12] Note that we maintain the restriction that ξ_j^S and $f(r_j)$ are additively separable; that is, we rule out interactions between rankings and additional factors, such as the variance of $\tilde{\xi}_j^U$, in driving the reservation value.

[13] However, if $\varepsilon_{i,j}$ has different variances across goods, this induces differences in search indices above and beyond $f(r_j)$. These differences are captured by ξ_j^S, which then are different from ξ_j^U. For instance, if goods j and k have the same ξ^U but the variance of $\varepsilon_{i,j}$ is larger than that of $\varepsilon_{i,k}$ (implying a bigger upside to searching j), then we have $\xi_j^S \geq \xi_k^S$, all else equal. This is allowed in the model because we estimate ξ^S and ξ^U via two separate sets of fixed effects.

[14] We believe that the algorithms we develop extend to the case in which the platform maximizes a convex combination of consumer surplus and revenue, but we have not formally analyzed this.

[15] This expression for consumer surplus arises in logit models whenever there is a distinction between anticipated and realized utility (Allcott 2013, Train 2015).

[16] Choi et al. (2018) derive a formula for consumer surplus that accounts for search costs. Note, however, that their definition of expected surplus differs from ours. Choi et al. (2018) consider expected surplus from the perspective of a consumer who has not searched yet; that is, the expectation is taken over the realizations of the components of utility that are revealed upon search (which could include x_j and ξ_j^U as well as ε_{ij}^{post} in our model). In contrast, we take the perspective of the researcher/platform and focus on the average surplus across consumers, with which the average is taken over realizations of the consumer-specific shocks $(\varepsilon_{i,j}^S, \varepsilon_{i,j}^{pre}, \varepsilon_{ij}^{post})$. These are the only terms in the model that the researcher cannot pin down after estimating the model, and it is, thus, intuitive to take an average over them when defining the platform problem.

[17] Note that the sign of (7) is the same as the sign of the term in parentheses because q_{ij} and $f'(r_j)$ are both positive.

[18] All possible assignments of the J products to the first k positions take $J!/(J-k)!$ evaluations, and because the algorithm may assign up to position K, we must sum over all assignments that have $k=1$ to $k=K$.

[19] To calculate consumer surplus or platform revenues under random assignment, we enumerate all possible assignments for the remaining $J-K$ positions and average the consumer surplus or platform revenues associated with each assignment.

[20] Assignment of all J products requires $(J-K+1)+(J-K)+\cdots+2 = (J-K)(J-K+3)/2$ evaluations.

[21] An impression is a ranked list of hotels, together with their attributes, that the platform returns in response to a user query.

[22] For the descriptive evidence, we focus on the top 50 most often displayed hotels in the training data.

[23] We plot the probability of booking on different axes here because the testing data oversamples searches that terminate in booking so that plotting them on the same axes may be misleading.

[24] Ursu (2018) also maintains exogeneity of hotel characteristics.

[25] Jiang et al. (2021) propose a Geweke–Hajivassiliou–Keane–type estimator that also simplifies the computation of the likelihood for the Weitzman model. One difference is that our estimation approach is somewhat agnostic as to which components of utility consumers discover upon search as discussed in Section 2.

[26] We approximate the expectation by Monte Carlo simulation: for each consumer, we draw 10,000 vectors of the idiosyncratic shocks to the search and utility indices for all products, determine which products they search and purchase and their consumer surplus, and then output the average across all draws.

[27] In contexts in which this is not the case—for example, in which the proposed consumer surplus–optimizing algorithm yields substantially more searches—one could use one of the microfoundations of the model and the estimated parameters to back out search costs (e.g., from (4)) and then account for the latter in the welfare calculation.

[28] Consistent with this, Greminger (2022a) uses the same data and finds that it is possible to simultaneously increase both revenues and consumer surplus relative to the status quo.

References

Abaluck J, Compiani G, Zhang F (2020) A method to estimate discrete choice models that is robust to consumer search. Working paper, National Bureau of Economic Research, Cambridge, MA.

Aggarwal G, Feldman J, Muthukrishnan S, Pál M (2008) Sponsored search auctions with Markovian users. Papadimitriou C, Zhang S, eds. *Internat. Workshop Internet Network Econom.* (Springer, Berlin, Heidelberg), 621–628.

Agrawal S, Avadhanula V, Goyal V, Zeevi A (2019) MNL-bandit: A dynamic learning approach to assortment selection. *Oper. Res.* 67(5):1453–1485.

Allcott H (2013) The welfare effects of misperceived product costs: Data and calibrations from the automobile market. *Amer. Econom. J. Econom. Policy* 5(3):30–66.

Ansari A, Mela CF (2003) E-customization. *J. Marketing Res.* 40(2):131–145.

Armstrong M (2017) Ordered consumer search. *J. Eur. Econom. Assoc.* 15(5):989–1024.

Armstrong M, Vickers J (2015) Which demand systems can be generated by discrete choice? *J. Econom. Theory* 158(A):293–307.

Ascarza E (2018) Retention futility: Targeting high-risk customers might be ineffective. *J. Marketing Res.* 55(1):80–98.

Athey S, Ellison G (2011) Position auctions with consumer search. *Quart. J. Econom.* 126(3):1213–1270.

Bradlow ET, Fader PS (2001) A Bayesian lifetime model for the “hot 100” billboard songs. *J. Amer. Statist. Assoc.* 96(454):368–381.

Bronnenberg BJ, Kim JB, Mela CF (2016) Zooming in on choice: How do consumers search for cameras online? *Marketing Sci.* 35(5):693–712.

Chapman RG, Staelin R (1982) Exploiting rank ordered choice set data within the stochastic utility model. *J. Marketing Res.* 19(3):288–301.

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

Choi H, Mela CF (2019) Monetizing online marketplaces. *Marketing Sci.* 38(6):948–972.

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

De los Santos B, Koulayev S (2017) Optimizing click-through in online rankings with endogenous search refinement. *Marketing Sci.* 36(4):542–564.

Derakhshan M, Golrezaei N, Manshadi V, Mirrokni V (2022) Product ranking on online platforms. *Management Sci.* 68(6):4024–4041.

Gallego G, Li A, Truong V-A, Wang X (2020) Approximation algorithms for product framing and pricing. *Oper. Res.* 68(1):134–160.

Gardete P, Hunter M (2020) Guiding consumers through lemons and peaches: An analysis of the effects of search design activities. Stanford University Graduate School of Business Research Paper No. 3669, Stanford, CA.

Ghose A, Ipeirotis PG, Li B (2012) Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Sci.* 31(3):493–520.

Ghose A, Ipeirotis PG, Li B (2014) Examining the impact of ranking on consumer behavior and search engine revenue. *Management Sci.* 60(7):1632–1654.

Greminger RP (2022a) Heterogeneous position effects and the power of rankings. Preprint, submitted October 28, https://arxiv.org/abs/2210.16408.

Greminger RP (2022b) Optimal search and discovery. *Management Sci.* 68(5):3904–3924.

Gu C, Wang Y (2022) Consumer online search with partially revealed information. *Management Sci.* 68(6):4215–4235.

Hitsch GJ, Misra S, Zhang W (2018) Heterogeneous treatment effects and optimal targeting policy evaluation. Preprint, submitted February 6, http://dx.doi.org/10.2139/ssrn.3111957.

Hodgson C, Lewis G (2020) You can lead a horse to water: Spatial learning and path dependence in consumer search. Cowles Foundation Discussion Paper 2246, New Haven, CT.

Honka E, Chintagunta P (2017) Simultaneous or sequential? Search strategies in the US auto insurance industry. *Marketing Sci.* 36(1):21–42.

Honka E, Hortaçsu A, Vitorino MA (2017) Advertising, consumer awareness, and choice: Evidence from the US banking industry. *RAND J. Econom.* 48(3):611–646.

Jagabathula S, Rusmevichientong P (2017) A nonparametric joint assortment and price choice model. *Management Sci.* 63(9):3128–3145.

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

Kempe D, Mahdian M (2008) A cascade model for externalities in sponsored search. Papadimitriou C, Zhang S, eds. *Internat. Workshop Internet Network Econom.* (Springer, Berlin, Heidelberg), 585–596.

Kim JB, Albuquerque P, Bronnenberg BJ (2010) Online demand under limited consumer search. *Marketing Sci.* 29(6):1001–1023.

Kleinberg R, Waggoner B, Glen Weyl E (2016) Descending price optimally coordinates search. Preprint, submitted March 24, https://arxiv.org/abs/1603.07682.

Kök AG, Fisher ML (2007) Demand estimation and assortment optimization under substitution: Methodology and application. *Oper. Res.* 55(6):1001–1021.

Mehta N, Rajiv S, Srinivasan K (2003) Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Sci.* 22(1):58–84.

Roberts JH, Lattin JM (1991) Development and testing of a model of consideration set composition. *J. Marketing Res.* 28(4):429–440.

Santu SKK, Sondhi P, Zhai C (2017) On application of learning to rank for e-commerce search. *Proc. 40th Internat. ACM SIGIR Conf. Res. Development Inform. Retrieval* (ACM, New York), 475–484.

Train K (2015) Welfare calculations in discrete choice models when anticipated and experienced attributes differ: A guide with examples. *J. Choice Model.* 16:15–22.

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

Ursu RM, Dzyabura D (2020) Retailers' product location problem with consumer search. *Quant. Marketing Econom.* 18:125–154.

Ursu R, Seiler S, Honka E (2022) The sequential search model: A framework for empirical research. Preprint, submitted October 15, https://dx.doi.org/10.2139/ssrn.4236738.

Weitzman ML (1979) Optimal search for the best alternative. *Econometrica* 47(3):641–654.

Yao S, Mela CF (2011) A dynamic model of sponsored search advertising. *Marketing Sci.* 30(3):447–468.

Zantedeschi D, Feit EM, Bradlow ET (2017) Measuring multichannel advertising response. *Management Sci.* 63(8):2706–2728.