ONLINE APPENDIX
Trade-Offs Between Ranking Objectives: Descriptive Evidence and Structural Estimation

Rafael P. Greminger*

July 27, 2026

A Further Evidence of Position Effects

Table OA.1 presents the parameter estimates for the main regression (1). As in the main text, the coefficient estimates are scaled to represent changes in percentage points.[1] Columns 1 and 4 show the results of a baseline model that does not include the interaction term and thus assumes that position effects are homogeneous. Columns 2 and 5 show the main results. Columns 3 and 6 show the results for the specification that additionally includes hotel fixed effects.[2]

The remainder of this appendix presents results for alternative specifications. Throughout, the results indicate that the main findings are robust to these alternative specifications. First, I replicate the analysis of Figures 2 and 3 for a Probit model and a linear probability model that parametrizes position effects more flexibly. To make estimation feasible, I include observable query characteristics for the Probit model rather than individual fixed effects. The results are shown in Figures OA.1 to OA.4.

Table OA.2 presents coefficient estimates for another linear probability model where the position effect is specific to the first three positions and then follows a linear specification for the remaining hotels. This specification is given by

$$\mathbb{P}(Y_{ij}=1|z_{ij},pos_{ij}) = x_j'\beta + \sum_{h=1}^{3} 1(pos_{ij}\leq h)(\gamma_h + x_j'\theta_h) - 1(pos_{ij}\geq 4)(pos_{ij}-3)(\gamma_4 + x_j'\theta_4) + \tau_i \ . \quad \text{(OA.1)}$$

*UCL School of Management, r.greminger@ucl.ac.uk.

[1] Standard errors are clustered at the consumer query level, capturing that search behavior can induce correlation in the error terms. A large draw for one alternative can mean that consumers are less likely to click on or book another alternative, suggesting a potential negative correlation in the error terms.

[2] Several hotels are displayed only to a single consumer and are therefore excluded from this specification. Only price and whether the hotel is on promotion vary across different search sessions, so the coefficients in β cannot be estimated for other characteristics. The interaction between hotel characteristics and position can still be estimated because it is identified by within-hotel variation in position.

In this specification, γ_h and θ_h capture the position effect on position h. To see this, consider, for example, the effect of decreasing pos_{ij} from $h = 2$ to $h = 1$. (OA.1) and the definition of the position effect in (2) implies that this effect is given by

$$\text{Position effect}_{ij} = (\gamma_1 + \gamma_2) + x_j'(\theta_1 + \theta_2) - \gamma_1 - x_j'\theta_1 = \gamma_2 + x_j'\theta_2\,. \tag{OA.2}$$

Hence, γ_2 and θ_2 capture the expected change in Y_{ij} when moving from the second to the first position. Similarly, γ_4 and θ_4 capture the expected change in Y_{ij} when moving across positions beyond $h = 3$. For example, changing pos_{ij} from $h = 5$ to $h = 4$ implies

$$\text{Position effect}_{ij} = -(4 - 5 - 3)(\gamma_4 + x_j'\theta_4) = \gamma_4 + x_j'\theta_4\,. \tag{OA.3}$$

Finally, I also estimate the following specification, which allows position effects to be fully flexible across all positions:

$$\mathbb{P}(Y_{ij} = 1|z_{ij}, pos_{ij}) = x_j'\beta + \sum_{h=1}^{33} 1(pos_{ij} = h)(\gamma_h + x_j'\theta_h) + \tau_i\;. \tag{OA.4}$$

In this specification, γ_h and θ_h capture the position effect on position h relative to the bottom position $h = 34$.[3]

Figures OA.5 and OA.6 present the estimates for the main parameters θ_h for positions $h \leq 15$. For clicks, the results are consistent with the parsimonious specification from the main text: price mutes position effects, while location score—a desirable attribute that also has a strong effect in the parsimonious specification—amplifies position effects. For the other variables, the effects are not statistically significant, indicating insufficient variation in these attributes in the data.

Given the small booking probabilities and many parameters to estimate, differences in position-specific position effects across different hotels are difficult to identify, as evidenced by the coefficients γ_h not being statistically significant for the less flexible specification in Table OA.2. Nonetheless, the results from the extremely flexible specification continue to imply that lower-priced hotels have larger position effects conditional on other attributes.

[3] I drop sessions with more than 34 hotels to avoid estimating position effects for positions that are rarely observed.

TABLE OA.1 – Coefficient Estimates (LPM, Random Ranking)

	Clicks			Bookings		
	(1)	(2)	(3)	(4)	(5)	(6)
Position	0.1398***	0.0838***	0.0809***	0.0085***	-0.0021	-0.0034
	(0.0019)	(0.0126)	(0.0128)	(0.0005)	(0.0028)	(0.0029)
Price	-0.0184***	-0.0262***	-0.0369***	-0.0016***	-0.0024***	-0.0032***
	(0.0003)	(0.0005)	(0.0006)	(0.0001)	(0.0001)	(0.0001)
On promotion	1.2784***	1.7224***	1.6935***	0.1443***	0.1692***	0.1881***
	(0.0522)	(0.1150)	(0.1291)	(0.0134)	(0.0290)	(0.0338)
Star rating	2.0188***	2.6052***		0.1367***	0.1653***	
	(0.0354)	(0.0711)		(0.0088)	(0.0172)	
Review score	0.2722***	0.3090***		0.0575***	0.1049***	
	(0.0364)	(0.0846)		(0.0083)	(0.0195)	
No reviews	0.4459***	0.2226		0.1786***	0.2467***	
	(0.1682)	(0.3984)		(0.0351)	(0.0838)	
Chain	0.1993***	0.2526**		0.0196*	0.0611**	
	(0.0478)	(0.1004)		(0.0117)	(0.0246)	
Location score	0.6157***	0.8238***		0.0637***	0.0741***	
	(0.0186)	(0.0344)		(0.0040)	(0.0074)	
Position × Price		-0.0005***	-0.0005***		-0.0000***	-0.0000***
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
Position × Star rating		0.0342***	0.0316***		0.0017**	0.0012*
		(0.0030)	(0.0030)		(0.0007)	(0.0007)
Position × Review score		0.0024	0.0009		0.0027***	0.0028***
		(0.0036)	(0.0036)		(0.0008)	(0.0009)
Position × No reviews		-0.0116	-0.0140		0.0039	0.0038
		(0.0170)	(0.0172)		(0.0036)	(0.0036)
Position × Chain		0.0031	-0.0002		0.0024**	0.0017
		(0.0042)	(0.0043)		(0.0010)	(0.0011)
Position × Location score		0.0119***	0.0124***		0.0006**	0.0010***
		(0.0014)	(0.0014)		(0.0003)	(0.0003)
Position × On promotion		0.0262***	0.0274***		0.0015	0.0014
		(0.0050)	(0.0051)		(0.0013)	(0.0013)
FE Hotel	no	no	yes	no	no	yes
N	1,220,909	1,220,909	1,219,242	1,220,909	1,220,909	1,219,242
R2 (adj.)	0.0020	0.0026	0.0154	0.0098	0.0099	0.0054

Notes: Coefficient estimates are scaled to represent changes in terms of percentage points. All specifications include session fixed effects. Standard errors are shown in parentheses and are clustered at the query level. Star ratings are adjusted so that the position effect from the first row is for a hotel with 1 star and other characteristics equal to the minimum observed value. Statistical significance is indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

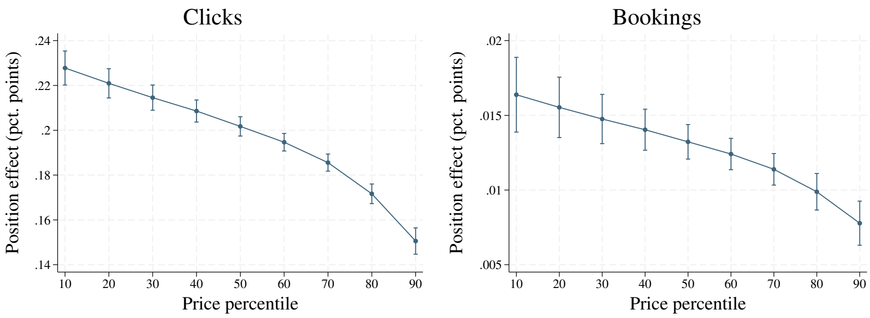

FIGURE OA.1 – Heterogeneous Position Effects: Probit

Notes: This figure replicates Figure 2 for a Probit model.

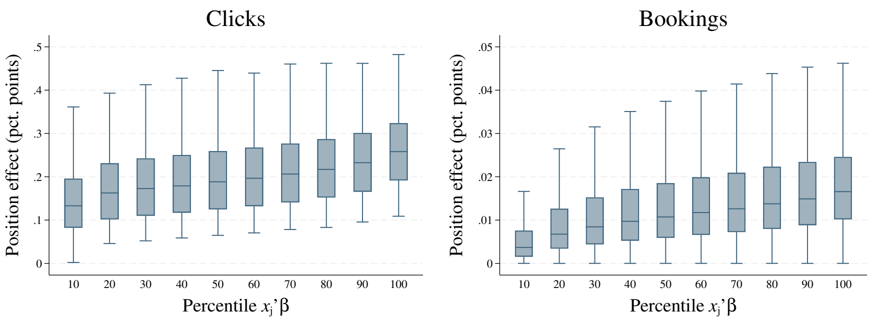

FIGURE OA.2 – Heterogeneous Position Effects: Probit

Notes: This figure replicates Figure 3 with estimates for a Probit model.

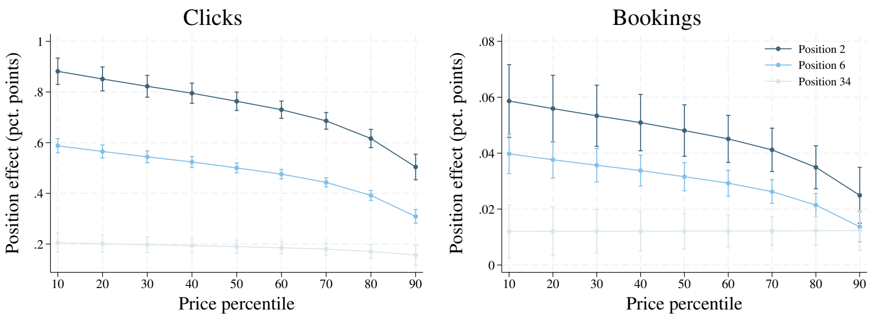

FIGURE OA.3 – Heterogeneous Position Effects: Cubic

Notes: This figure replicates Figure 2 with estimates from a linear probability model that adds pos_{ij}^2, $pos_{ij}^2 x_j$, pos_{ij}^3 and $pos_{ij}^3 x_j$ to the baseline specification.

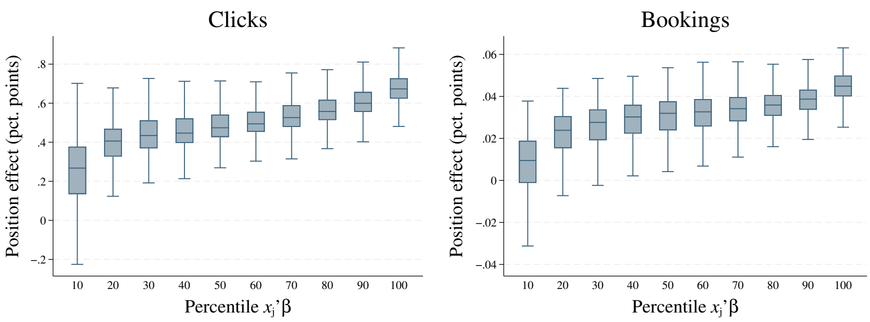

FIGURE OA.4 – Heterogeneous Position Effects: Cubic

Notes: This figure replicates Figure 3 with estimates from a linear probability model that adds pos_{ij}^2, $pos_{ij}^2 x_j$, pos_{ij}^3 and $pos_{ij}^3 x_j$ to the baseline specification. Position effects are computed at $pos_{ij} = 6$.

TABLE OA.2 – Coefficient Estimates (Flexible LPM, Random Ranking)

	Clicks			Bookings		
	(1)	(2)	(3)	(4)	(5)	(6)
1(Position $\leq$ 1)	2.4021***	5.5333***	5.4790***	-0.0072	0.1169	0.1317
	(0.2031)	(1.3558)	(1.3366)	(0.0505)	(0.2986)	(0.2986)
1(Position $\leq$ 2)	1.2725***	1.0553	0.8896	0.0813*	0.0758	0.0016
	(0.1846)	(1.2205)	(1.2071)	(0.0490)	(0.2945)	(0.2965)
1(Position $\leq$ 3)	1.6571***	2.1390**	1.8724**	0.1398***	-0.0189	-0.0060
	(0.1333)	(0.8925)	(0.8834)	(0.0352)	(0.2201)	(0.2229)
1(Position $\geq$ 4) $\times$ (Position - 3)	0.0858***	-0.0012	0.0024	0.0055***	-0.0036	-0.0044
	(0.0021)	(0.0131)	(0.0136)	(0.0005)	(0.0030)	(0.0032)
Price	-0.0184***	-0.0228***	-0.0334***	-0.0016***	-0.0021***	-0.0030***
	(0.0003)	(0.0005)	(0.0006)	(0.0001)	(0.0001)	(0.0001)
On promotion	1.2637***	1.4620***	1.4306***	0.1435***	0.1368***	0.1547***
	(0.0521)	(0.1124)	(0.1270)	(0.0134)	(0.0280)	(0.0328)
Star rating	2.0131***	2.3959***		0.1364***	0.1573***	
	(0.0354)	(0.0688)		(0.0088)	(0.0169)	
Review score	0.2838***	0.5136***		0.0581***	0.0848***	
	(0.0363)	(0.0797)		(0.0083)	(0.0187)	
No reviews	0.4901***	1.0610***		0.1810***	0.2186***	
	(0.1677)	(0.3707)		(0.0351)	(0.0792)	
Chain	0.2090***	0.4151***		0.0201*	0.0252	
	(0.0477)	(0.0968)		(0.0117)	(0.0241)	
Location score	0.6200***	0.6854***		0.0639***	0.0794***	
	(0.0185)	(0.0335)		(0.0040)	(0.0073)	
1(Position $\leq$ 1) =1 $\times$ Price		-0.0083***	-0.0082***		-0.0005	-0.0004
		(0.0018)	(0.0018)		(0.0004)	(0.0004)
1(Position $\leq$ 2) =1 $\times$ Price		0.0012	0.0011		0.0000	-0.0000
		(0.0017)	(0.0017)		(0.0004)	(0.0004)
1(Position $\leq$ 3) =1 $\times$ Price		-0.0042***	-0.0040***		-0.0002	-0.0002
		(0.0013)	(0.0013)		(0.0003)	(0.0003)
1(Position $\geq$ 4) $\times$ (Position - 3) $\times$ Price		-0.0004***	-0.0004***		-0.0000***	-0.0000***
		(0.0000)	(0.0000)		(0.0000)	(0.0000)
1(Position $\leq$ 1) =1 $\times$ Star rating		0.1464			0.0188	
		(0.3004)			(0.0712)	
1(Position $\leq$ 2) =1 $\times$ Star rating		0.1571			0.0446	
		(0.2732)			(0.0684)	
1(Position $\leq$ 3) =1 $\times$ Star rating		0.0949			-0.0337	
		(0.1962)			(0.0489)	
1(Position $\geq$ 4) $\times$ (Position - 3) $\times$ Star rating		0.0287***	0.0267***		0.0015**	0.0013*
		(0.0031)	(0.0032)		(0.0008)	(0.0008)
1(Position $\leq$ 1) =1 $\times$ Review score		-0.5659			-0.0075	
		(0.3882)			(0.0890)	
1(Position $\leq$ 2) =1 $\times$ Review score		-0.0553			-0.0226	
		(0.3481)			(0.0852)	
1(Position $\leq$ 3) =1 $\times$ Review score		-0.2348			0.0619	
		(0.2533)			(0.0617)	
1(Position $\geq$ 4) $\times$ (Position - 3) $\times$ Review score		0.0124***	0.0101***		0.0022**	0.0020**
		(0.0037)	(0.0039)		(0.0009)	(0.0009)
1(Position $\leq$ 1) =1 $\times$ No reviews		-2.5397			-0.0922	
		(1.8275)			(0.3962)	
1(Position $\leq$ 2) =1 $\times$ No reviews		-0.2186			-0.0344	
		(1.6511)			(0.3689)	
1(Position $\leq$ 3) =1 $\times$ No reviews		-0.7859			0.1320	
		(1.1897)			(0.2554)	
1(Position $\geq$ 4) $\times$ (Position - 3) $\times$ No reviews		0.0265	0.0190		0.0032	0.0021
		(0.0174)	(0.0180)		(0.0037)	(0.0039)
1(Position $\leq$ 1) =1 $\times$ Chain		-0.2427			-0.0415	
		(0.4426)			(0.1081)	
1(Position $\leq$ 2) =1 $\times$ Chain		-1.2089***			-0.0408	
		(0.4012)			(0.1041)	
1(Position $\leq$ 3) =1 $\times$ Chain		0.5719**			0.1339*	
		(0.2857)			(0.0749)	
1(Position $\geq$ 4) $\times$ (Position - 3) $\times$ Chain		0.0118***	0.0062		0.0011	0.0002
		(0.0045)	(0.0046)		(0.0011)	(0.0011)
1(Position $\leq$ 1) =1 $\times$ Location score		0.1312			0.0081	
		(0.1428)			(0.0314)	
1(Position $\leq$ 2) =1 $\times$ Location score		0.1355			-0.0154	
		(0.1291)			(0.0311)	
1(Position $\leq$ 3) =1 $\times$ Location score		0.1904**			-0.0102	
		(0.0931)			(0.0225)	
1(Position $\geq$ 4) $\times$ (Position - 3) $\times$ Location score		0.0070***	0.0076***		0.0009***	0.0013***
		(0.0015)	(0.0015)		(0.0003)	(0.0003)
1(Position $\leq$ 1) =1 $\times$ On promotion		-0.0650	-0.0916		-0.1640	-0.1604
		(0.5072)	(0.5053)		(0.1294)	(0.1300)
1(Position $\leq$ 2) =1 $\times$ On promotion		0.5020	0.5675		0.2206*	0.2295*
		(0.4674)	(0.4671)		(0.1246)	(0.1254)
1(Position $\leq$ 3) =1 $\times$ On promotion		0.0370	0.0086		-0.0238	-0.0270
		(0.3366)	(0.3369)		(0.0872)	(0.0870)
1(Position $\geq$ 4) $\times$ (Position - 3) $\times$ On promotion		0.0174***	0.0186***		0.0001	0.0000
		(0.0054)	(0.0055)		(0.0013)	(0.0014)
FE Hotel	no	no	yes	no	no	yes
N	1,220,909	1,220,909	1,219,242	1,220,909	1,220,909	1,219,242
R2 (adj.)	0.0044	0.0052	0.0178	0.0099	0.0100	0.0055

Notes: Coefficient estimates are scaled to represent changes in terms of percentage points. Estimates for query characteristics are omitted. Standard errors are shown in parentheses and are clustered at the query level. Star ratings are adjusted so that the position effect from the first row is for a hotel with 1 star and other characteristics equal to the minimum observed value. Statistical significance is indicated by *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

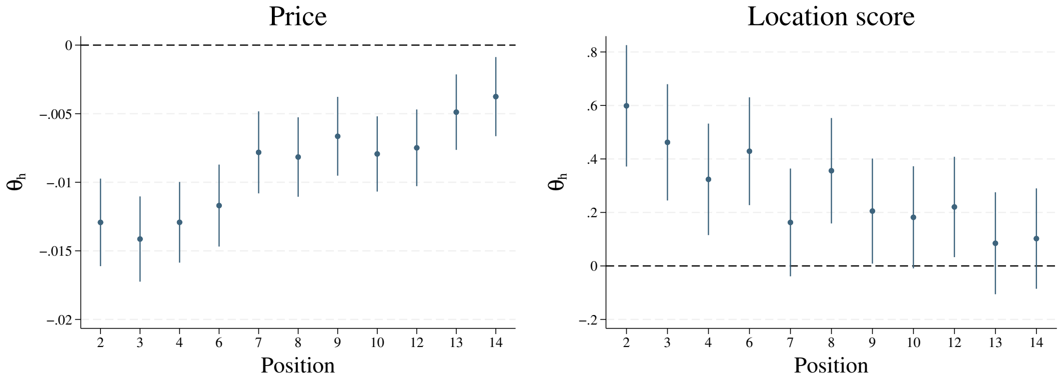

FIGURE OA.5 – Flexible Heterogeneous Position Effects for Clicks

Notes: This figure shows the parameter estimates θ_h in specification (OA.4) for the first 15 positions. A positive estimate implies that the attribute amplifies the position effect.

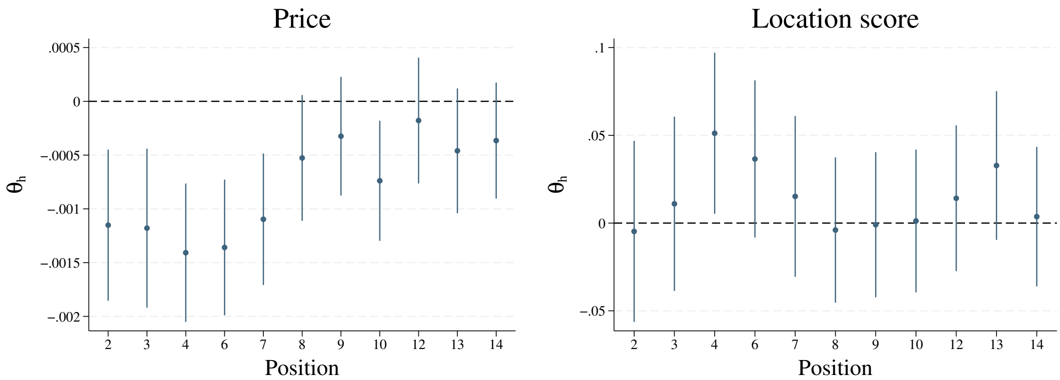

FIGURE OA.6 – Flexible Heterogeneous Position Effects for Bookings

Notes: This figure shows parameter estimates θ_h in specification (OA.4) for the first 15 positions. A positive estimate implies that the attribute amplifies the position effect.

B Comparing Ranking Algorithms

This appendix presents additional results for the comparison of rankings and their trade-offs. First, I provide the procedure used to obtain consumers' updated beliefs under alternative ranking algorithms and show the counterfactual results for the case where consumers' beliefs are fixed. Second, I provide results showing that the proposed Discovery-Value Ranking maximizes consumer welfare and transactions in some cases. Third, I compare different ranking algorithms with a brute-force algorithm that is guaranteed to find the optimal ranking. Because the number of possible rankings grows faster than exponentially in the number of alternatives, this comparison is only feasible for a small number of alternatives. Finally, I compare different ranking algorithms with each other to assess their relative performance for the case considered in the main counterfactual analysis.

B.1 Consumers' Updated Beliefs under Alternative Rankings

In the model, consumers expect to discover alternatives with effective values that have a mean given by equation (6). When they adjust beliefs to the new rankings, μ_1 and ρ are updated to reflect the new ranking of alternatives and the resulting distribution of effective values across positions. To recover these parameters, I use the following procedure.

First, I take draws ν_{ij} and ε_{ij} for each alternative j and consumer i. Then, I use these draws to compute the effective values $\tilde{w}_{ij} = x_j^{l\prime}\hat{\beta} + \nu_{ij} + x_j^{l\prime}\hat{\gamma} + \hat{\tilde{\gamma}}_j + \min\{\hat{\xi}, x_j^{d\prime}\hat{\kappa} - \hat{\mu}_{x\kappa} + \varepsilon_{ij}\}$ using the estimated parameter values.

Next, I estimate the following linear regression on the computed effective values:

$$\tilde{w}_{ij} = \pi_0 + \pi_1 \log(pos_{ij}) + \epsilon_{ij} \ , \tag{OA.5}$$

where pos_{ij} is the position of alternative j in the new ranking for consumer i. Moreover, I compute the mean $\mu_{\tilde{w}}$ as the average of $\tilde{w}_{ij}$ across all alternatives and consumers.

The estimated coefficient $\hat{\pi}_1$ provides the belief parameter ρ^* for the new ranking. Under the assumption that consumers know the distribution of product attributes and shocks, the expected effective value across positions is the same as the unconditional expected effective value. Hence, given the mean $\mu_{\tilde{w}}$ and the functional form in (6), μ_1 and its updated version μ_1^* can be recovered from the following relationship:

$$\mu_{\tilde{w}} = \frac{1}{|J|}\sum_{h=1}^{|J|} \mu(h) = \mu_1 + \frac{1}{|J|}\sum_{h=1}^{|J|} \rho \log(h) \tag{OA.6}$$

Using μ_1 and μ_1^*, I obtain the updated discovery value in the first position as $\tilde{z}^{d*} = \tilde{z}^d + \mu_1^* - \mu_1$. ρ^* and $\tilde{z}^{d*}$ obtained this way fully characterize consumers' updated beliefs under the new ranking, as the functional form in (6) remains unchanged.

Table OA.3 reports the estimates for π_0 and π_1 of the regression (OA.5) for the different rankings. Standard errors are small, so beliefs are estimated precisely. The reported estimates for $\pi_1 = \rho$ suggest that consumers expect alternatives to get worse at a faster rate under the DVR—the heuristic designed to maximize consumer welfare—than under Expedia's own ranking, which is in line with the DVR increasing consumer welfare by ranking high-utility alternatives higher.

TABLE OA.3 – Estimates of Adjusted Beliefs

	ER	DVR	BUR
π_0	5.581	5.903	5.889
	(0.026)	(0.026)	(0.025)
π_1	-0.056	-0.181	-0.175
	(0.010)	(0.009)	(0.009)

Notes: Standard errors in parentheses.

B.2 Ranking Effects With Fixed Beliefs

Table OA.4 shows the effects of the two main heuristics for the case where consumers do not update their beliefs after the ranking change. The effects are comparable to those reported in Table 3. Importantly, differences between the rankings remain small, so the main conclusions about the trade-offs between rankings depend little on whether consumers update their beliefs. Also note that consumer welfare effects can be smaller when consumers update to the correct beliefs because the baseline comparison is different: with belief updating, the baseline ranking is the randomized ranking with consumers knowing that the ranking is randomized, whereas with fixed beliefs, the baseline ranking is the randomized ranking with consumers continuing to have the beliefs as estimated.

TABLE OA.4 – Ranking Effects over Random Ranking

	ER	DVR	BUR
Platform			
Total revenues (%)	4.80	12.03	16.32
Number of transactions (%)	5.38	17.33	15.46
Avg. price of booking (%)	-0.55	-4.52	0.75
Consumers			
Consumer welfare ($, average)	0.04	0.14	0.11
Consumer welfare ($, cond. on booking)	2.70	8.15	5.65
Discovery costs ($, cond. on booking)	-0.02	-0.07	-0.06

Notes: Effects of different rankings relative to a randomized ranking when consumers do not adjust their beliefs. ER: Expedia's own ranking. DVR: Discover-Value Ranking (heuristic to maximize consumer welfare). BUR: Bottom-Up Ranking (heuristic to maximize revenues).

B.3 Consumer Welfare Maximization and the DVR

The Discovery-Value Ranking (DVR) balances the utility and the search value of an alternative to maximize consumer welfare (see Section 5.1). Proposition 1 provides two cases in which the DVR is guaranteed to maximize consumer welfare.

Proposition 1. *The Discovery-Value Ranking maximizes consumer welfare when consumers (i) know all alternatives' attributes and search costs at the beginning of search, or (ii) searching an alternative only reveals the shock but no other attributes.*

Proof. To prove this proposition, I use the following logic. Suppose a consumer could choose the order in which to discover products herself, rather than the order being determined by the ranking. In this case, the order that the consumer would optimally choose to discover products in by definition maximizes consumer welfare. Hence, to prove that the DVR maximizes consumer welfare, I show that the consumer's optimal policy would prescribe to discover products in the order of decreasing r_j^d, if that choice were available.

The consumer's choice of the order in which to discover products can be formulated as a search and discovery problem with multiple discovery technologies. Specifically, each product j is a discovery technology that only reveals the respective product on the list, and the consumer sequentially chooses which discovery technology to use next. As Greminger (2022) notes, the optimal policy in such a setting is still characterized by the optimal discovery, search, and stopping rules presented in Section 3.2, with the extension that the optimal discovery rule now prescribes to choose the discovery technology with the largest discovery value z_j^d next.

For the case where the consumer knows the realizations of the product attributes (x_j^l, x_j^d) and search costs (c_s), the definition of the index r_j^d means that it equals the discovery value z_j^d, so that the consumer would optimally choose to discover products in the order of decreasing r_j^d. Hence, the DVR maximizes consumer welfare in this case.

In the case where searching an alternative only reveals its shock and no vertical attributes $(x_j^{l\prime}\gamma = \tilde{\gamma}_j = x_j^{d\prime}\kappa = 0)$, the index r_j^d sorts alternatives in order of first-order stochastic dominance of the effective value $\tilde{W}_j = x_j^{l\prime}\beta + \min\{\xi, \varepsilon_j\}$.[4] Now note that the definition of the discovery value guarantees that when the distribution of some $\tilde{W}_j$ first-order stochastically dominates the distribution of some $\tilde{W}_k$, then $z_j^d \geq z_k^d$. Hence, the consumer would optimally choose to discover j before k and, as a result, that the consumer would optimally choose to discover products in the order of decreasing r_j^d such that the DVR maximizes consumer welfare. Note that this rationale applies independently of whether consumers learn of the change in ranking or not because the optimal policy based on correct beliefs by definition cannot be improved upon, so that the ranking maximizing consumer welfare is the same in both cases.

□

[4]When there are vertical attributes revealed on the detail page, r_j^d does not necessarily sort alternatives in order of first-order stochastic dominance of the effective value $\tilde{W}_j$.

B.4 Additional Ranking Algorithms

Table OA.5 lists the ranking algorithms that I compare with the heuristic algorithms used for the main counterfactual analysis. The first three algorithms are brute-force algorithms that compare all possible rankings and select the best one according to the respective objective. To make these algorithms feasible, they are computed under the assumption that consumers do not update their beliefs about the ranking. However, they are still evaluated for the case when consumers update beliefs.

TABLE OA.5 – Ranking Algorithms

	Name	Objective	Description
WBF	Welfare Brute-Force	Welfare	Compares all possible rankings.
CBF	Transactions Brute-Force	Transactions	Compares all possible rankings.
RBF	Revenues Brute-Force	Revenues	Compares all possible rankings.
DVR	Discovery-Value Ranking	Welfare	Rank by r_j^d defined in (11).
UBR	Utility-Based Ranking	Welfare	Rank by $\mathbb{E}[U_j^l] = x_j^{l\prime}\beta$.
WT3	Welfare Top-3 Ranking	Welfare	OPT3G-CS algorithm of Compiani et al. (2024).
DVR	Discovery-Value Ranking	Transactions	Rank by r_j^d defined in (11).
CT3	Transactions Top-3 Ranking	Transactions	Same as WT3, but for maximizing transactions.
BUR	Bottom-Up Ranking	Revenue	Start ranking from the bottom (see main text).
RT3	Revenues Top-3 Ranking	Revenue	OPT3G-Rev algorithm of Compiani et al. (2024).
PDR	Price-Decreasing Ranking	Revenue	Decreasing order of p_j.

Maximizing Consumer Welfare. The next three algorithms are heuristic algorithms designed to maximize welfare. The DVR balances the utility and search costs of an alternative through the alternative-specific discovery value r_j^d as described in Section 5.1. The WT3 is the OPT3G-CS heuristic algorithm for maximizing consumer welfare proposed by Compiani et al. (2024). This algorithm works by first optimally ranking alternatives in the top three positions using a brute-force algorithm. Then, it fills the remaining positions iteratively by selecting the alternative that increases consumer welfare the most when added in the next position. As Compiani et al. (2024) show, this algorithm performs well in a Weitzman model when the number of positions is small or when consumers mostly search and choose alternatives just from the top of the ranking. Albeit without formal proof, the same logic likely applies in the SD model, so the algorithm provides a potentially good heuristic to maximize welfare in the SD model. Finally, the UBR ranks alternatives by their expected utility ($x_j^{l\prime}\beta + x_j^{d\prime}\kappa$). Proposition 1 shows that when search costs are homogeneous and there are no attributes revealed on the list page, this ranking maximizes consumer welfare. Hence, depending on the estimated model, it provides another potentially good approximation of the consumer-welfare-maximizing ranking.

Maximizing Transactions. The next two algorithms are heuristic algorithms designed to maximize the number of transactions. Maximizing transactions is similar to maximizing consumer welfare because consumers are more likely to search and buy alternatives that provide higher utility. Proposition 2 proves that when consumers do not reveal vertical attributes on the detail page and do not learn of the ranking change, the consumer-welfare-maximizing ranking also maximizes transactions. Hence, provided the DVR approximates the consumer-welfare-maximizing ranking well, it should also provide a good heuristic to maximize transactions.

Proposition 2. *In the SD model without search over vertical attributes ($x_j^{d\prime}\kappa = 0$), the consumer-welfare-maximizing ranking also maximizes transactions when consumers do not update their beliefs about the ranking.*

Proof. The generalized eventual purchase theorem of Greminger (2022) implies that the probability that a consumer purchases an alternative is given by $\mathbb{P}(U_0 \leq \max_{j\in J} W_j)$, where $W_j = \min\{z^d(h_j), \tilde{W}_j\}$. Hence, if the distribution of $\tilde{W}_j$ improves in higher positions in a first-order stochastic dominance sense, then this probability weakly increases, provided consumers do not update their beliefs so that the discovery values remain the same. When search costs are homogeneous and no attributes are revealed on the detail page, then the DVR orders alternatives based on first-order stochastic dominance of their effective values $\tilde{W}_j$, and by Proposition 1, the DVR also maximizes consumer welfare. Hence, the consumer-welfare-maximizing ranking also maximizes transactions in this case. $\square$

The CT3 is another heuristic to maximize transactions. It works in the same way as the WT3, but targets transactions rather than consumer welfare.

Maximizing Platform Revenues. The final three ranking algorithms are heuristic algorithms designed to maximize platform revenues. The BUR starts ranking from the bottom position and is described in detail in Section 5.1. The RT3 is the same as the OPT3G-Rev heuristic algorithm for maximizing revenues proposed by Compiani et al. (2024). It works with the same steps as WT3, but targets revenues rather than consumer welfare. Finally, the PDR ranks alternatives in decreasing order of their prices. This heuristic is motivated by the price effect highlighted in Section 2.5. As Proposition 3 below shows, the PDR maximizes revenues when position effects are homogeneous. Hence, how well the PDR performs relative to other algorithms provides a measure of how important heterogeneous position effects are for revenue maximization.

Proposition 3. *The Price-Decreasing Ranking (PDR) maximizes revenues when position effects are homogeneous.*

Proof. The proposition follows from the fact that position effects are homogeneous in the SD model only when alternatives are homogeneous in the expected list utilities ($u_j^e = x_j^{l\prime}\beta + x_j^{d\prime}\kappa$). Hence, if an alternative has a higher price, other attributes in x_j^l have to adjust so that $x_j^{l\prime}\beta$ remains unchanged. As $x_j^{l\prime}\beta$ and u_j^e are the same for all alternatives, switching two alternatives will only shift demand from the alternative moved down to the one moved up, but does not affect the demand for other alternatives. Hence, whenever a higher-priced alternative can be moved up, moving it up will shift demand to that alternative so that revenues increase. This implies that any ranking that does not order in decreasing order of price cannot be optimal when position effects are homogeneous, which proves the proposition because an optimal ranking always exists with a finite number of alternatives. □

B.5 Ranking Comparison with Reduced Set of Alternatives

To compare the rankings obtained from the different heuristics to the optimal one, I randomly sample 1,000 consumers and reduce the number of alternatives by randomly selecting five for each consumer. For each of the ranking algorithms, I simulate 800,000 search paths and compute the average consumer welfare, revenues, and transactions. I consider two settings: one where consumers update their beliefs following the change in ranking and one where they keep the beliefs as estimated. For the former, I use the procedure described in Online Appendix B.1 to obtain consumers' beliefs under the new ranking.

Table OA.6 reports the results. Relative to the consumer-welfare-maximizing ranking (WBF), the proposed DVR heuristic performs best in both scenarios, outperforming even the WT3, which ranks the first three positions using brute-force. The DVR also performs substantially better than the utility-based ranking (UBR). This result is in line with Compiani et al. (2024), who showed that the UBR is not consumer-welfare-maximizing when consumers search over vertical attributes. Overall consumer welfare differences are small because there are only five alternatives.

Among the heuristic algorithms, the DVR also performs best in terms of transactions, lowering transactions by less than 1% relative to the conversion-maximizing ranking (CBF). Hence, I conclude that the DVR provides a good heuristic to maximize consumer welfare and transactions. Nonetheless, comparing the WBF to the CBF shows that the two objectives are not perfectly aligned, and that the DVR seems to balance the two objectives.

For revenues, the proposed BUR heuristic performs best among the heuristics. Depending on whether consumers adjust beliefs or not, the BUR reduces revenues only by about 0.62% and 0.59% relative to the revenue-maximizing ranking (RBF). The PDR performs by far the worst (-13.09%) because it ignores the underlying demand effects, highlighting the importance of heterogeneous position effects when designing rankings to maximize revenues.

B.6 Ranking Comparison with All Alternatives

Computing the optimal ranking with a brute-force algorithm is not feasible beyond a few alternatives. However, I can still compare the different heuristic algorithms with each other following the same procedure as for the main counterfactual analysis.

Table OA.7 reports the effects of the different rankings relative to Expedia's own ranking. Throughout, the results are in line with the five-alternative comparison. Importantly, the proposed heuristics DVR and BUR continue to outperform all other heuristics for all objectives, confirming they are well-suited to evaluate the trade-offs between ranking objectives.

TABLE OA.6 – Comparison of Rankings (5 alternatives)

	Fixed Beliefs			Updated Beliefs		
	Δ Wel. ($)	Δ Trans. (%)	Δ Rev. (%)	Δ Wel. ($)	Δ Trans. (%)	Δ Rev. (%)
WBF	0.0000	-0.2950	-3.1847	0.0000	-0.2617	-3.2511
CBF	-0.0005	0.0000	-2.4705	-0.0005	0.0000	-2.5661
RBF	-0.0061	-2.4806	0.0000	-0.0062	-2.4253	0.0000
DVR	-0.0016	-0.8100	-3.1152	-0.0014	-0.8036	-3.2200
UBR	-0.0106	-4.9227	-6.7839	-0.0114	-4.3390	-6.2589
WT3	-0.0072	-3.8738	-6.3346	-0.0071	-3.7126	-6.2726
CT3	-0.0077	-3.9172	-6.1223	-0.0076	-3.7624	-6.0659
BUR	-0.0068	-3.2485	-0.5852	-0.0067	-3.2370	-0.6230
RT3	-0.0137	-6.5145	-4.0089	-0.0138	-6.3343	-3.8820
PDR	-0.0407	-20.6330	-13.0901			

Notes: Effects of different rankings for a reduced set of five alternatives. The effects are all relative to the brute-force algorithm for the respective objective. Table OA.5 describes the different ranking algorithms. For the PDR, beliefs would imply that the discovery value increases across positions, so that it is not possible to compute the effects under the updated beliefs.

TABLE OA.7 – Comparison of Rankings (all alternatives)

	Fixed Beliefs			Updated Beliefs		
	Δ Wel. ($)	Δ Trans. (%)	Δ Rev. (%)	Δ Wel. ($)	Δ Trans. (%)	Δ Rev. (%)
DVR	0.1385	17.3350	12.0297	0.1116	12.0933	5.8780
UBR	0.1010	13.2085	8.4352	0.0648	7.5064	3.3675
WT3	0.0804	9.7571	6.7152	0.0510	5.2999	2.7185
CT3	0.0792	9.9339	6.9932	0.0499	5.4531	2.9554
BUR	0.1143	15.4569	16.3182	0.0825	9.7802	10.5441
RT3	0.0627	8.3979	8.2462	0.0354	4.2130	4.1640
PDR	-0.0553	-6.8478	0.7675			

Notes: Effects of different rankings relative to a randomized ranking. Table OA.5 describes the different ranking algorithms.

C Parameter Identification

This appendix formally shows that data without information on how far consumers scrolled is sufficient to identify the model parameters. A main concern is the separate identification of the search and discovery parameters. For example, an increase in either the search or the discovery costs will induce consumers to end discovery sooner and search fewer products. Throughout, I assume that the ranking is randomized, ensuring that the discovery parameters can be separately identified from the preference parameters.

With data on how far consumers scrolled, separate identification follows from moments that condition on the discovered alternatives. For example, larger search costs imply that consumers are less likely to search the alternatives they discovered on the list, whereas larger discovery costs only change the probability where consumers stop scrolling.

Without such data, these identification arguments no longer apply directly because it is unclear how the stopping probabilities and conditional moments could be recovered. However, I now provide two propositions that show how these moments can be recovered in data that does not track how far consumers scrolled. As a result, the same identification arguments that rely on conditional moments still apply even without data on how far consumers scrolled.

The first proposition establishes that the stopping probabilities are identified by the position-specific click probabilities. The second proposition shows that the relevant conditional moments are uniquely determined by the stopping probabilities and the unconditional moments. Hence, these moments are effectively present in the data, even without information on how many alternatives consumers discovered.[5] As a result, the same identification arguments still apply, even in the absence of this information.

C.1 Identification of Stopping Probabilities

Proposition 4. *Suppose there are N alternatives randomly shown in N positions, and let $q_h = \mathbb{P}(\textit{Click in position } h)$ denote the probability of clicking in position h across consumers. The stopping probability in the search and discovery model then is given by*

$$\mathbb{P}(\textit{Stop discovery on position } h) = m(q_1, \ldots, q_N) \ , \tag{OA.7}$$

where the function $m()$ has a known form.

Proof. I first derive two conditions required for the stopping probabilities to be identified by the data. Intuitively, the two conditions guarantee that only the stopping probabilities produce position effects and that there are position effects whenever consumers do not discover all alternatives. After establishing the two conditions, I show that they apply in the model.

[5] These results could also be used to develop a simulated methods of moments estimation approach even without having to observe the discovery process.

To simplify exposition, I introduce the following notation: $q_h = \mathbb{P}(\text{Click in position } h)$; $\Delta q_h = q_h - q_{h-1}$; D_h denotes the event of a consumer ending discovery on position h; R_h denotes the event of a consumer discovering at least position h; R_h^c denotes the complement event of a consumer stopping before reaching position h; and S_{jh} denotes the event of a consumer searching alternative j in position h.

With N alternatives randomly shown across N positions, the expected difference in clicks across two positions is the average across the click-probabilities for each alternative when it is shown in the respective positions. Formally, this is given by

$$\Delta q_h = \frac{1}{N}\left[\sum_j \mathbb{P}(S_{jh}) - \mathbb{P}(S_{jh+1})\right], \tag{OA.8}$$

which is an average across the N possible products that could be shown in the respective positions.

Consumers cannot search products they did not discover such that $\mathbb{P}(S_{jh}|R_h^c) = 0$. Hence, conditioning on reaching position h and differentiating the two cases of stopping on position h or reaching $h+1$, the expected difference in clicks can be written as

$$\begin{aligned} N\Delta q_h = \mathbb{P}(R_h) \Bigg[& \mathbb{P}(D_h|R_h)[\sum_j \mathbb{P}(S_{jh}|D_h)] \\ & + \mathbb{P}(R_{h+1}|R_h)[\sum_j \mathbb{P}(S_{jh}|R_{h+1}) - \sum_j \mathbb{P}(S_{jh+1}|R_{h+1})]\Bigg]. \end{aligned} \tag{OA.9}$$

Now suppose that the following two conditions hold for all positions h:

1. $\Delta q_h = q_h \iff \mathbb{P}(D_h|R_h) = 1$: whenever there are no clicks in the next position ($q_{h+1} = 0$), consumers who reach position h never discover $h+1$.
2. $\Delta q_h = 0 \iff \mathbb{P}(D_h|R_h) = 0$: whenever there is no difference in clicks across the two positions, consumers who reach position h always continue to discover $h+1$.

Applying the first condition and setting $\Delta q_h = q_h$ and $\mathbb{P}(D_h|R_h) = 1$ in (OA.9) yields $\sum_j \mathbb{P}(S_{jh}|D_h) = q_h$. This is because $\mathbb{P}(D_h|R_h) = 1$ implies that $\mathbb{P}(R_{h+1}|R_h) = 1 - \mathbb{P}(D_h|R_h) = 0$. Similarly, applying the second condition and setting $\Delta q_h = 0$ and $\mathbb{P}(D_h|R_h) = 0$ in (OA.9) yields $\sum_j \mathbb{P}(S_{jh}|R_{h+1}) - \sum_j \mathbb{P}(S_{jh+1}|R_{h+1}) = 0$. Substituting these two expressions in (OA.9) then yields

$$N\Delta q_h = \mathbb{P}(D_h|R_h)Nq_h = \mathbb{P}(D_h|R_h)N\frac{1}{N}\left[\sum_j \mathbb{P}(S_{jh}|D_h)\right] \tag{OA.10}$$

$$\Rightarrow \mathbb{P}(D_h|R_h) = \frac{N\Delta q_h}{\sum_j q_j} \tag{OA.11}$$

This relation shows that the probability of stopping on a particular position h conditional on reaching it is nonparametrically identified as long as the position-specific click probabilities can be estimated. The unconditional stopping probabilities then are given by the recursive relation

$$\begin{aligned}\mathbb{P}(D_h) &= \mathbb{P}(D_h|R_h)\mathbb{P}(R_h) + 0 \\ &= \frac{N\Delta q_h}{\sum_j q_j}\left(1 - \sum_{q=1}^{h-1}\mathbb{P}(D_q)\right) ,\end{aligned} \tag{OA.12}$$

where I use that the probability of consumers stopping on position h conditional on not having reached is always zero. This recursive relation characterizes the unconditional stopping probabilities for all positions h starting from $h = 1$. At $h = 1$, $\mathbb{P}(D_0) = 0$ by the assumption that the first discovery is free. Given $\mathbb{P}(D_0)$, (OA.12) yields $\mathbb{P}(D_1)$ from the conditional stopping probability. Given $\mathbb{P}(D_1)$, (OA.12) yields $\mathbb{P}(D_2)$ from the conditional stopping probability, and so on. Hence, once the conditional stopping probabilities are identified, so are the unconditional stopping probabilities.

It remains to show that the two conditions hold in the search and discovery model. The first condition holds as long as consumers cannot search an alternative without discovering it, which is the case in the search and discovery model. This guarantees that $q_{h+1} = 0$ whenever $h + 1$ is not discovered. The random ranking then guarantees that there is a non-zero probability of clicking on an alternative in position $h + 1$ when that position is revealed, which proves that $\Delta q_h = 0$ if and only if $\mathbb{P}(D_h|R_h) = 0$.

The second condition holds because the search and discovery model guarantees that an alternative is searched when shown in some position whenever it will also be searched when shown in the previous position. This follows immediately from the two search implications that provide the conditions under which an alternative is searched, given an effective value for the chosen option. Given that $w_j \leq \tilde{w}_j$, the requirement for an alternative to be searched in earlier positions is weaker than the requirement for it to be searched in later positions. As a result, the second condition also holds in the search and discovery model. □

C.2 Identification of Conditional Moments

Proposition 5. *Moments on clicks and purchases conditional on discovery are uniquely determined by the stopping probabilities and the unconditional moments.*

Proof. Let M be an average quantity of interest determined by either clicks or purchases, and let M_h be the respective position-specific counterparts. For example, M could be the number of clicks consumers make on average and M_h the clicks consumers make on average in position h. The unconditional moments then are $\mathbb{E}[M] = \sum_h \mathbb{E}[M_h]$ and the conditional moments are $\mathbb{E}[M_h|D_h]$, where D_h again denotes the event of ending discovery on position h. Throughout, I assume that sufficient data is available so that $\mathbb{E}[M]$, $\mathbb{E}[M_h]$ and $\mathbb{P}(D_h)$ (through Proposition 4) are precisely estimated and, hence, can be treated as known. The conditional moments

$\mathbb{E}[M_h|D_h]$ then are identified if $\mathbb{E}[M_h|D_h]$ is uniquely determined by these known moments for all positions h.

I now show that this is indeed the case through a recursive relation. The fact that consumers cannot click or choose alternatives from positions they have not discovered implies that $\mathbb{E}[M_h|D_k] = 0$ for all $k < h$. As a result, the law of total probability implies the following:

$$\mathbb{E}[M_h] = \sum_{k \geq h} \mathbb{E}[M_h|D_k]\mathbb{P}(D_k) \ . \tag{OA.13}$$

In the last available position, denoted by $\tilde{h}$, this relation simplifies to $\mathbb{E}[M_{\tilde{h}}] = \mathbb{E}[M_{\tilde{h}}|D_{\tilde{h}}]\mathbb{P}(D_{\tilde{h}})$. Hence, $\mathbb{E}[M_{\tilde{h}}|D_{\tilde{h}}]$ is fully determined by $\mathbb{E}[M_{\tilde{h}}]$ and $\mathbb{P}(D_{\tilde{h}})$. Given $\mathbb{E}[M_{\tilde{h}}|D_{\tilde{h}}]$, $\mathbb{E}[M_{\tilde{h}-1}|D_{\tilde{h}-1}]$ then is also determined through (OA.13), which then allows determining the conditional moment for $\tilde{h} - 2$, and so on. Hence, the recursive relation (OA.13) uniquely determines all conditional moments from the stopping probabilities and the unconditional moments. $\square$

D Estimation Details

This appendix provides details on the estimation approach. First, it shows how to compute the probability of searching an alternative without choosing it. Second, it provides the partitioning of the probability space that guarantees a smooth simulated likelihood function. Finally, it reports results from a Monte Carlo simulation study.

D.1 Probability of Search and Non-Purchase

The probability of searching an alternative k without purchasing it enters the likelihood as $\mathbb{P}(Z_k^s \geq \tilde{w}_j \cap U_k \leq \tilde{w}_j)$, where $\tilde{w}_j$ is the effective value of the eventually chosen option j and is not random. By expressing it as the CDF of a standard normal distribution and the CDF of a bivariate normal distribution, Proposition 6 provides a way to calculate this probability using standard numerical methods, avoiding computationally costly numerical integration.[6]

Proposition 6. *If the two shocks are independent and $\nu_j \sim N(0,1)$ and $\varepsilon_j \sim N(0, \sigma_\varepsilon^2)$, then*

$$\mathbb{P}(Z_j^s \geq q \cap U_j \leq q) = \mathbb{P}\left(Y \leq \frac{q - u_j^e}{\tilde{\sigma}}\right) - \mathbb{P}\left(Y_1 \leq \frac{q - u_j^e}{\tilde{\sigma}} \cap Y_2 \leq q - z_j^e\right) \qquad \text{(OA.14)}$$

for some constant q, where $u_j^e = x_j^{l\prime}\beta + x_j^{d\prime}\kappa$, $z_j^e = x_j^{l\prime}\beta + x_j^{l\prime}\gamma + \tilde{\gamma}_j + \xi$, Y follows a standard normal distribution, $[Y_1, Y_2]$ follows a bivariate normal distribution with correlation $-\frac{1}{\tilde{\sigma}\sigma_\varepsilon}$, and $\tilde{\sigma} = \sqrt{1 + \frac{1}{\sigma_\varepsilon^2}}$.

Proof. Under the assumption that the shocks are normally distributed and independent, $\mathbb{P}(Z_j \geq q \cap U_j \leq q)$ can be written as

$$\mathbb{P}(Z_j^s \geq q \cap U_j \leq q) = \int_{q - z_j^e} \Phi\left(\frac{q - u_j^e - \nu}{\sigma_\varepsilon}\right)\phi(\nu)\mathrm{d}\nu \qquad \text{(OA.15)}$$

$$= \int_{-\infty}^{\infty} \Phi\left(\frac{q - u_j^e - \nu}{\sigma_\varepsilon}\right)\phi(\nu)\mathrm{d}\nu - \int_{-\infty}^{q - z_j^e} \Phi\left(\frac{q - u_j^e - \nu}{\sigma_\varepsilon}\right)\phi(\nu)\mathrm{d}\nu\ , \qquad \text{(OA.16)}$$

where u_j^e and z_j^e are defined in the proposition. Using results 10,010.8 and 10,010.1 for normal integrals from Owen (1980), the expression follows.

□

[6]To calculate the bivariate normal CDF, I use the method by Drezner and Wesolowsky (1990) as implemented in the Julia StatsFuns.jl package.

D.2 Partitioning the Probability Space

I calculate the likelihood contribution in (8) as

$$\mathbb{P}(\text{observed choices of } i) = \sum_{R_k \in R} \mathbb{P}(\text{observed choices of } i|R_k)\mathbb{P}(R_k)\ , \tag{OA.17}$$

where $R = \bigcup_k R_k$ is a partition of the probability space of $(\eta, \nu_j, \varepsilon_j)$, the shocks for the chosen option. The partition is defined so that two conditions hold for any region R_k: (i) the number of alternatives the consumer discovered remains fixed, and (ii) ε_j is either above or below the threshold $\xi_j = x_j^{l\prime}\beta + \tilde{\gamma}_j + \xi$.

Recall that the set of alternatives a consumer discovers, $A(\tilde{w}_j)$, is determined by the effective value of the chosen (inside or outside) option, $\tilde{w}_j$. Hence, the number of alternatives the consumer discovered can be fixed by constructing the following intervals:

$$B = (-\infty, z^d(|J|)] \cup (z^d(|J|), z^d(|J|-1)] \cup, \ldots, \cup(z^d(1), \infty]\ . \tag{OA.18}$$

Using these intervals and additionally partitioning on $\varepsilon_j > \xi_j$, I calculate the individual likelihood contributions as

$$\begin{aligned}\mathcal{L}^i(\theta) = \sum_{b\in B} &\mathbb{P}(\tilde{W}_j \in b \cap \varepsilon_j \le \xi_j)\mathbb{E}[p(\tilde{W}_j)|\tilde{W}_j \in b \cap \varepsilon_j \le \xi_j] + \\ &\sum_{b\in B} \mathbb{P}(\tilde{W}_j \in b \cap \varepsilon_j > \xi_j)\mathbb{E}[p(\tilde{W}_j)|\tilde{W}_j \in b \cap \varepsilon_j > \xi_j]\ ,\end{aligned} \tag{OA.19}$$

where $p(\tilde{w}_j)$ denotes the product of the integral in (9). As it has no closed-form solution, I use Monte Carlo integration to compute $\mathbb{E}[p(\tilde{W}_j)|\tilde{W}_j \in b \cap \varepsilon_j \le \xi_j]$, i.e., I take draws $r = 1, \ldots, N_l$ from distributions truncated so that the draws satisfy the respective conditions. This produces a smooth likelihood because all other expressions in the likelihood are smooth functions.

If the outside option is chosen, $\tilde{w}_j = \beta_0 + \eta$ so taking draws for $\tilde{w}_j \in b$ is the same as taking truncated draws $\beta_0 + \eta \in b$. If instead an alternative $j > 0$ is chosen, the expression further depends on the condition $\varepsilon_j \le \xi_j$, and I take draws for $\tilde{w}_j \in b$ as follows:

1. Take a draw ε_j^r from its distribution truncated on $\varepsilon_j \le \xi_j$.
2. Take a draw ν_j^r from its distribution truncated on $\tilde{w}_j(\varepsilon_j^r) = x_j^{l\prime}\beta + \nu_j + x_j^{d\prime}\kappa + \delta_j + \varepsilon_j^r \in b$.
3. Calculate $\tilde{w}_j^r$ using draws ν_j^r and ε_j^r and compute the inner probability $p(\tilde{w}_j^r)$.

Based on draws generated by this procedure, the expression can be calculated as the weighted average

$$\mathbb{P}(\varepsilon_j \le \xi_j) \sum_r \mathbb{P}(\tilde{W}_j(\varepsilon_j^r) \in b) p(\tilde{w}_j^r)\ . \tag{OA.20}$$

I calculate $\mathbb{P}(\tilde{W}_j \in b \cap \varepsilon_j > \xi_j)\mathbb{E}[p(\tilde{W}_j)|\tilde{W}_j \in b \cap \varepsilon_j > \xi_j]$ using the same steps with the truncation on $\varepsilon_j > \xi_j$ in step 1 and the corresponding change of $\tilde{w}_j(\varepsilon_j^r)$ in step 2.

D.3 Monte Carlo Simulation Study

I run a Monte Carlo simulation study to verify whether the estimation procedure recovers the parameters with the present data. To this end, I first simulate data for parameter values close to those obtained from the estimation, and then estimate the model on the simulated data. I repeat this simulation 50 times and average the results across simulations. For the observable product characteristics and the ranking, I use the hotel characteristics from the estimation sample with the randomized ranking. This setup leads to simulated data comparable to the observed data in terms of variation in observable characteristics and consumers' choices. To reduce the computational burden, I randomly sample 5,000 search sessions rather than using the entire estimation sample.

Table OA.8 reports the results. The estimation procedure recovers the parameters well, including the search and discovery costs that are backed out after the estimation.

TABLE OA.8 – Monte Carlo Simulation Results

	True	Estimate	Std. Dev.
β: Price (in \$100)	-0.350	-0.334	(0.051)
β: Star rating	0.150	0.154	(0.046)
β: Review score	0.150	0.138	(0.081)
β: No reviews	0.300	0.418	(0.360)
β: Chain	0.000	0.010	(0.067)
β: On promotion	0.150	0.138	(0.066)
β: Outside option	5.100	5.007	(0.293)
κ: Location score	0.100	0.088	(0.026)
γ: Price (in \$100)	0.150	0.135	(0.048)
γ: Star rating	0.050	0.045	(0.046)
γ: Review score	-0.150	-0.139	(0.081)
γ: No reviews	-0.350	-0.463	(0.370)
γ: Chain	-0.100	-0.107	(0.068)
γ: On promotion	-0.050	-0.036	(0.066)
ξ	2.650	2.587	(0.277)
$\tilde{z}^d$	5.750	5.729	(0.262)
ρ	-0.150	-0.172	(0.031)
$c_d \times 100$	0.006	0.009	(0.007)
c_s	0.001	0.003	(0.005)
Log likelihood		-22413.151	
N consumers		5,000	
No. clicks (average)		1.164	
No. bookings (average)		0.065	

Notes: The table reports averages of parameter estimates and fit measures, and the standard deviation of estimates across 50 simulation runs. Parameter estimates are obtained using 100 simulation draws per region of the partitioned probability space. Search and discovery costs are computed using 10M simulation draws.

E Comparison of Specifications

This appendix compares different specifications of the search and discovery model to determine how they affect the results. To limit the computational burden, I estimate all specifications without fixed effects δ_j and $\tilde{\gamma}_j$.[7]

E.1 Distributions of Idiosyncratic Shocks

I evaluate the sensitivity of the results to using different values for σ_ε (variance of ε) and the upper bound of the distribution of the outside option shock η_i. To this end, I estimate the model under different values for these parameters using the same sample as in the main analysis, simulate how many and where clicks and bookings occur, and predict the revenue and consumer welfare effects of implementing either the Discovery-Value Ranking (DVR) or the Bottom-Up Ranking (BUR) over Expedia's own ranking (ER). By comparing effects to the ER rather than a randomized ranking, I save on computation time, while the results remain qualitatively the same. Other than the different baseline ranking, I use the same sample and procedures as for the main counterfactual analysis.

The results of this comparison are presented in Table OA.9. They reveal that parameter estimates tend to scale with the variance of ε_j. This includes the search and discovery cost estimates, which increase as σ_ε increases. Notably, however, the differences $\tilde{z}^d - \beta_0$ and $\xi - \beta_0$ do not change much, suggesting that the net benefits of the different search actions over taking the outside option change little across specifications. Similarly, search values $z_j^s = x_j^{l\prime}(\beta+\gamma)+\tilde{\gamma}_j+\xi$ change little with σ_ε because both β and γ scale similarly.

Whereas the revenue and consumer welfare effects change with σ_ε, the qualitative results remain the same across all values. Both the DVR and the BUR increase revenues and consumer welfare relative to the ER in all specifications, with the DVR generating larger consumer welfare effects and the BUR generating larger revenue effects. Notably, the differences between the rankings remain quite stable, and the suggested trade-offs between ranking objectives become even smaller compared to the main specification. Hence, the qualitative results are not very sensitive to the assumed variance of the idiosyncratic shocks.

The last columns of Table OA.9 provide results for specifications where the upper bound (b) of the distribution of the outside option shock η differs from the baseline. The results from these specifications show that the search and discovery cost estimates change very little with b. Though the sizes of the effects of the different rankings depend on b, the results also show that both rankings continue to increase both consumer welfare and platform revenues, suggesting that the qualitative results are not very sensitive to this upper bound.

[7] As the fixed effects capture alternative-specific differences, omitting them does not affect how different specifications compare to each other in terms of relative fit and counterfactual effects.

TABLE OA.9 – Comparison of Specifications

	Main	$\sigma_\varepsilon = 5$	$\sigma_\varepsilon = 10$	$b = 0.5$	$b = 5$
Parameter estimates					
β: Price (in \$100)	-0.362	-1.005	-2.018	-0.364	-0.361
β: Star rating	0.168	0.089	0.071	0.170	0.167
β: Review score	0.135	0.671	1.306	0.145	0.130
β: No reviews	0.196	1.201	2.439	0.226	0.178
β: Chain	0.031	0.320	0.823	0.025	0.035
β: On promotion	0.187	0.582	0.709	0.191	0.184
β: Outside option	5.128	13.711	24.978	5.168	5.104
κ: Location score	0.083	0.323	0.749	0.083	0.082
γ: Price (in \$100)	0.164	0.808	1.812	0.163	0.164
γ: Star rating	0.052	0.130	0.148	0.052	0.052
γ: Review score	-0.135	-0.673	-1.288	-0.135	-0.133
γ: No reviews	-0.247	-1.258	-2.407	-0.242	-0.239
γ: Chain	-0.103	-0.390	-0.867	-0.104	-0.103
γ: On promotion	-0.048	-0.443	-0.603	-0.046	-0.047
ξ	5.695	14.272	25.387	5.650	5.701
$\tilde{z}^d$	-0.128	-0.126	-0.091	-0.105	-0.136
ρ	2.578	11.167	22.263	2.580	2.567
$\tilde{z}^d - \beta_0$	0.568	0.560	0.408	0.482	0.597
$\bar{\xi} - \beta_0$	-2.549	-2.544	-2.715	-2.587	-2.537
Discovery costs (\$)	0.023	0.164	0.451	0.025	0.023
Search costs (\$)	0.471	2.302	2.369	0.464	0.488
Log likelihood	-48,893.726	-48,860.509	-49,196.096	-48,951.917	-48,888.534
Fit: bookings					
N, model (cond. on click)	0.057	0.056	0.053	0.057	0.056
N, data	0.055	0.055	0.055	0.055	0.055
Mean position, model	13.758	13.683	13.362	13.947	13.721
Mean position, data	13.468	13.468	13.468	13.468	13.468
Fit: clicks					
N, model	1.170	1.169	1.133	1.178	1.168
N, data	1.134	1.134	1.134	1.134	1.134
Mean position, model	13.622	13.653	13.339	13.841	13.577
Mean position, data	13.305	13.305	13.305	13.305	13.305
Effects of DVR					
Δ total revenues (%)	3.770	4.411	4.953	3.939	3.826
Δ consumer welfare (\$, average)	0.061	0.120	0.154	0.120	0.013
Δ consumer welfare (\$, cond. on click)	7.189	3.288	4.886	7.001	7.582
Δ consumer welfare (\$, cond. on booking)	7.456	4.311	5.511	7.180	8.123
Effects of BUR					
Δ total revenues (%)	8.243	8.860	9.345	8.364	10.006
Δ consumer welfare (\$, average)	0.037	0.094	0.146	0.067	0.010
Δ consumer welfare (\$, cond. on click)	6.467	3.429	6.589	6.359	6.769
Δ consumer welfare (\$, cond. on booking)	5.003	2.709	5.106	5.256	3.599

Notes: The main specification uses $\sigma_\varepsilon = 1$ and $\eta_i \sim Uniform(0, 1)$. The other specifications are the same, except for either σ_ε or $\eta_i \sim Uniform(0, b)$. The welfare effects are all relative to Expedia's own ranking. All specifications are estimated without fixed effects.

E.2 Functional Form for Beliefs and Number of Alternatives Discovered

To evaluate the impact of the functional form for beliefs, specified in (6), I estimate the model with the following linear specification for $h > 1$:

$$\mu_u(h) = \mu_u(1) + \rho h \ . \tag{OA.21}$$

Figure OA.7 shows the position-specific booking and click probabilities predicted by the model under this linear specification. The results show that the linear specification does not capture that clicks and bookings decrease at a decreasing rate across positions.

To evaluate the impact of the assumptions that consumers discover the first alternative for free and that they discover alternatives one by one, I also estimate two specifications that relax these assumptions. The results in Figures OA.8 and OA.9 show that under these alternative specifications, the model predicts click and booking probabilities that are constant among the positions discovered for free and among positions discovered at the same time, respectively. As the click and booking probabilities decrease across all positions, these results suggest that the main specification better captures the position effects in the data.

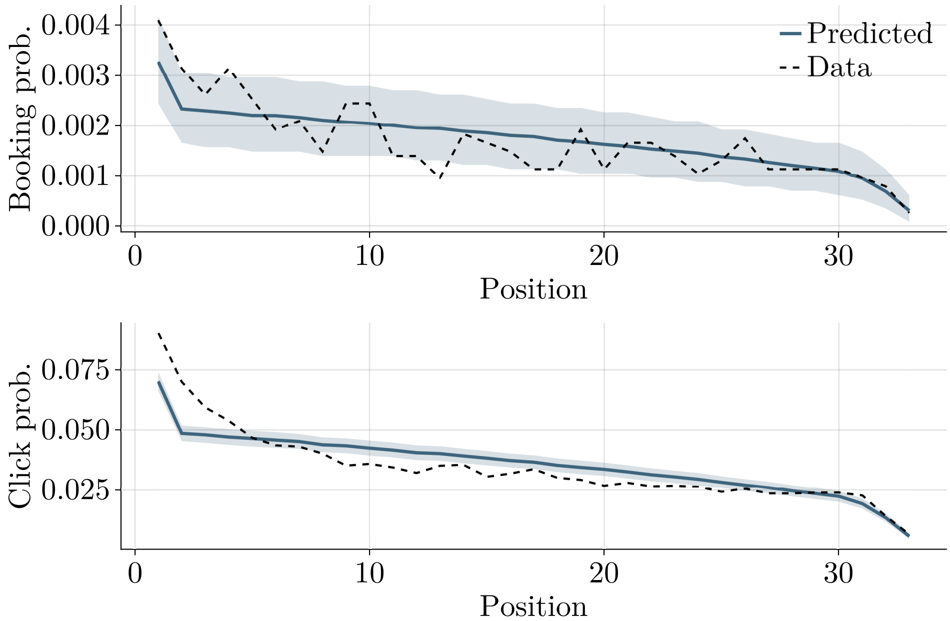

FIGURE OA.7 – Model Fit Under the Linear Belief Specification

Notes: Click and booking probabilities averaged across 10,000 draws per consumer, conditional on consumers searching at least one hotel. The shaded area indicates the 95th percentile of the minimum and maximum number of clicks or bookings across draws and consumers.

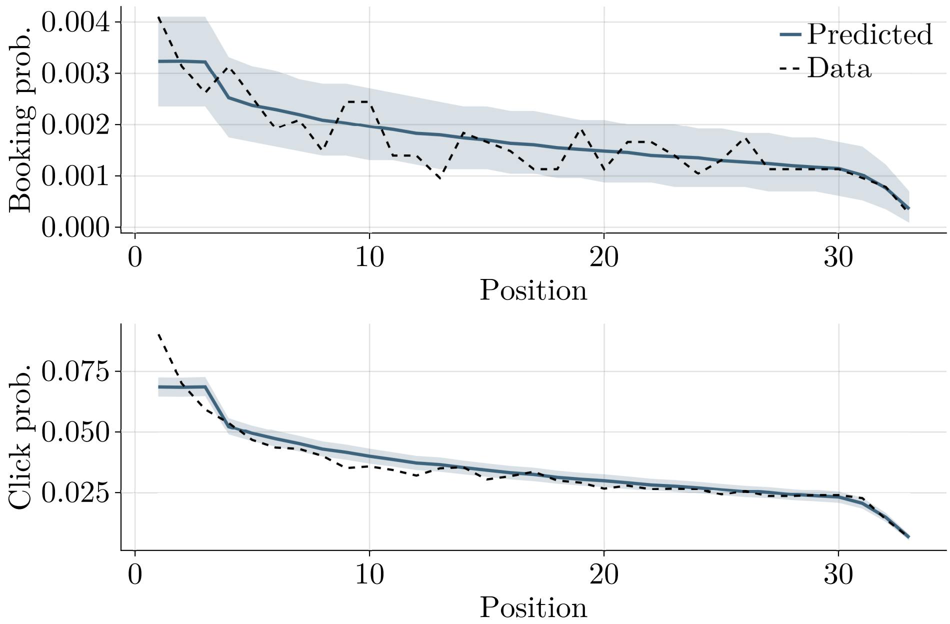

FIGURE OA.8 – Model Fit When Consumers Initially Discover 3 Alternatives

Notes: Click and booking probabilities averaged across 10,000 draws per consumer, conditional on consumers searching at least one hotel. The shaded area indicates the 95% percentile of the minimum and maximum number of clicks or bookings across draws and consumers.

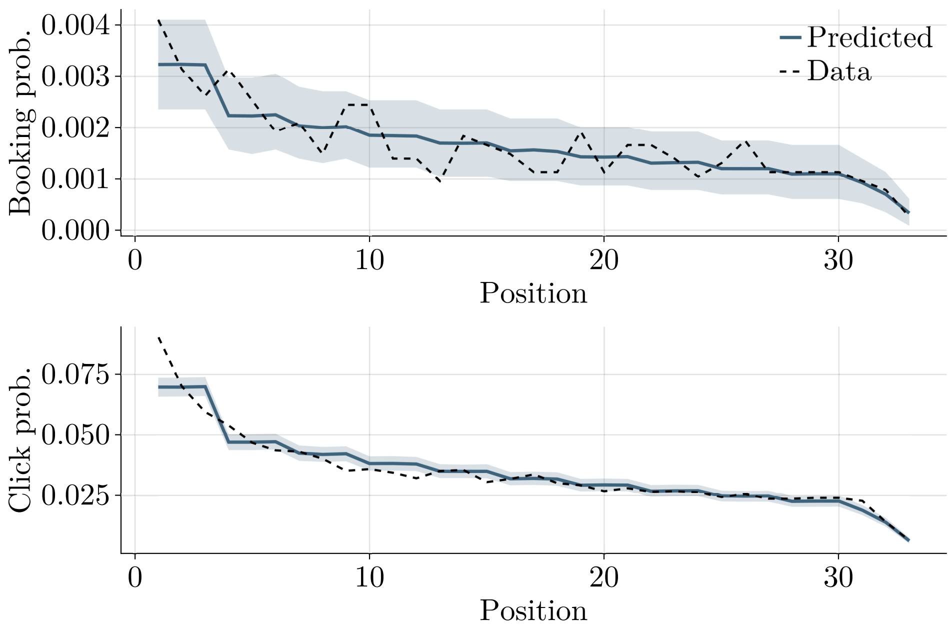

FIGURE OA.9 – Model Fit When Consumers Discover 3 Alternatives at a Time

Notes: Click and booking probabilities averaged across 10,000 draws per consumer, conditional on consumers searching at least one hotel. The shaded area indicates the 95% percentile of the minimum and maximum number of clicks or bookings across draws and consumers.

F Comparison with the Weitzman Model

This appendix provides the specification and estimation approach used for the model comparison in Section 4.2. It also discusses how the position effect mechanism in the Weitzman model limits its flexibility to simultaneously fit position effects in searches and purchases, and provides the full estimation results for this model.

F.1 The Weitzman Model

The *Weitzman* (WM) model is an established approach to quantify the effects of rankings and other changes to the search environment (see Ursu et al., 2024). Unlike in the *search and discovery* (SD) model, consumers in the WM model know all product attributes and preference shocks revealed on the list page (x_j^l, $\tilde{\gamma}_j$, and ν_j) at the beginning of search. Without product discovery, the WM model generates position effects through higher search costs for products in lower positions. Formally, search costs determine $\xi(h)$ defined in (5), which now depends on position h. As search costs increase, $\xi(h)$ decreases, leading to fewer searches and purchases at lower positions. I focus on this position-specific search cost mechanism because it is the standard one in prior work (see Ursu et al., 2024).[8]

The difference in how the SD and WM models generate position effects is not only conceptual but also matters empirically (Greminger, 2022): abstracting from product discovery can introduce biases in the inferred search costs and price elasticities because it assumes that consumers observe the entire list page, an assumption unlikely to hold unless the list displays only a few alternatives. I extend these results by showing how explaining position effects through product discovery affects the model-implied conditional purchase probabilities and, hence, the model's ability to fit position effects in purchases.

F.2 Empirical Specification and Estimation Approach

I use the same utility specification for the WM model as for the SD model. However, I adjust the distribution of the shocks to better fit specifications used in prior work (e.g., Ursu, 2018; Ursu et al., 2024). Specifically, I assume $\eta \sim N(0, \sigma_\varepsilon)$ and $\varepsilon_j \sim N(0, \sigma_\varepsilon)$, and estimate σ_ε. Estimating σ_ε is possible because position on the list acts as a search cost shifter (see Yavorsky et al., 2021). The results are similar (or even less favorable to the WM model) when assuming a standard uniform distribution for η instead. Moreover, using the exact same specification as Compiani et al. (2024) also yields similar results.[9]

[8] More recent work proposes that position-specific beliefs about hidden attributes could also explain position effects (Kaye, 2024; Fong et al., 2024). This alternative mechanism introduces more flexibility into the WM model, potentially allowing it to fit the constant conditional purchase probability. If that is the case, the limitation highlighted here could serve as an identification argument for parameters governing this additional flexibility.

[9] The codes are available as part of the replication package of Compiani et al. (2024).

I introduce position-specific search costs by assuming that the search value depends on position h through the following functional form for $h = 1, \ldots : \xi(h) = \xi + \rho \log(h)$. As in the SD model, I estimate the parameters ξ and ρ, while the assumed shape allows the model to fit the nonlinear decrease in clicks and bookings across positions. Estimating the model this way is equivalent to imposing a functional form on how search costs depend on position and estimating cost parameters as in, e.g., Ursu (2018).

To estimate the WM model, I use the fact that it is a SD model where discovery costs are zero and consumers discover all alternatives. In this case, the likelihood can be calculated by adjusting the likelihood contribution in (9) to

$$\mathcal{L}^i(\theta) = \log \int \prod_{k \in S} \mathbb{P}(Z_k^s \geq \tilde{w}_j \cap U_k \leq \tilde{w}_j) \times \prod_{k \in J \setminus S} \mathbb{P}(Z_k^s \leq \tilde{w}_j) \mathrm{d}H(\eta, \nu_j, \varepsilon_j) \, , \qquad \text{(OA.22)}$$

which can then be computed using the same simulation approach as for the SD model. Matching the specification for the SD model, I also condition the likelihood contribution for the WM model on searching at least once.

This estimation approach is easier to implement and computationally less demanding than the approaches proposed by Jiang et al. (2021) and Chung et al. (2025). It effectively integrates at most over two dimensions (ν_j, ε_j), instead of over as many dimensions as the number of searches each consumer makes. These approaches also cannot be easily used when the search order is not observed. Though integrating over all possible permutations of search sequences is possible when these are not observed (Compiani et al., 2024), this becomes prohibitively costly when consumers make more than just a few searches. In contrast, my approach constructs the likelihood independent of the search sequence, making it straightforward to implement even when consumers make many searches.

F.3 Estimation Results and Model Fit

Table OA.10 reports the full parameter estimates for the WM model. The parameter estimates all have the expected sign, as for the SD model.

Table OA.11 reports additional model fit statistics for both models. The results show that both models fit the attributes of the searched and booked hotels relatively well. However, in line with Figure 4, the table also reveals that the WM model predicts bookings occur on average two positions lower than in the data, whereas the SD model predicts the average position of bookings close to the one in the data. The WM model has a lower AIC and BIC, indicating a better overall model fit. However, these fit measures are based on the conditional likelihood, which does not capture how the models fit when consumers who visit the platform do not search any products. Hence, these fit measures should be interpreted with caution. Moreover, position effects in purchases and their heterogeneity are more relevant for the trade-offs between ranking objectives, which is why I focus these measures.

Table OA.10 – Parameter Estimates: Weitzman Model

	Estimate	Std. error
Preference parameters in $u_j^l = x_j^{l\prime}\beta + \nu_j^l$		
β: Price (in \$100)	-0.312***	(0.024)
β: Star rating	0.164***	(0.025)
β: Review score	0.132***	(0.035)
β: No reviews	0.255	(0.187)
β: Chain	0.019	(0.034)
β: On promotion	0.121***	(0.033)
β: Outside option	6.416***	(0.150)
Preference parameters in $u_j^d = x_j^{d\prime}\kappa + \delta_j + \varepsilon_j$		
κ: Location score	0.053***	(0.011)
Search value parameters in $z_j^s = u_j^l + x_j^{l\prime}\gamma + \tilde{\gamma}_j + \xi + \rho\log(h)$		
γ: Price (in \$100)	0.108***	(0.023)
γ: Star rating	0.054	(0.025)
γ: Review score	-0.105***	(0.035)
γ: No reviews	-0.194	(0.185)
γ: Chain	-0.057	(0.033)
γ: On promotion	-0.033	(0.033)
ξ	2.201***	(0.126)
ρ	-0.136***	(0.003)
Variance and search cost parameters		
σ_ε	0.722***	(0.014)
c_s	0.032	
Log likelihood	-47,858.384	
N consumers	11,467	

Notes: Parameter estimates obtained using 3,000 simulation draws per region of the partitioned probability space. Product fixed effects δ_j and $\tilde{\gamma}_j$ are included for a limited set of products in the model but not shown in the table. Asymptotic standard errors are shown in parentheses. Search costs are computed using 10M simulation draws. Statistical significance is indicated by * p < 0.1, ** p < 0.05, *** p < 0.01.

TABLE OA.11 – Model Fit Measures

	Bookings			Clicks		
	Data	SD Model	WM Model	Data	SD Model	WM Model
N (per consumer)	0.055	0.057	0.065	1.134	1.172	1.156
Mean position	13.468	13.766	15.810	13.305	13.619	13.406
Price	145.166	148.186	149.495	162.208	163.062	162.832
Star rating	3.495	3.491	3.477	3.508	3.502	3.509
Review score	3.965	3.960	3.966	3.917	3.904	3.921
No reviews	0.009	0.009	0.010	0.018	0.017	0.017
Location score	3.788	3.600	3.521	3.611	3.402	3.378
Chain	0.667	0.654	0.671	0.643	0.625	0.647
On promotion	0.413	0.431	0.397	0.359	0.376	0.358
AIC		97638.713	95750.767			
BIC		97763.616	95875.670			

Notes: This table compares moments in the estimation sample with those predicted by the SD and the WM model and reports the AIC and BIC fit statistics. The reported moments are the average for the number of clicks and bookings per consumer; the average position at which consumers clicked or booked a hotel; and the averages for the different attributes of hotels that consumers searched or chose. The simulated search paths were generated with 10,000 draws for each consumer, conditional on searching at least one alternative.

F.4 Conditional Purchase Probabilities

As highlighted in Figure 4 in Section 4.2, the WM model does not capture well the positions from which consumers book hotels. Figure OA.10 reveals the reason for this result. The panels show the percentage of clicks that result in a booking for the estimation sample and data simulated from the respective models. The first panel shows that this percentage is stable across positions in the data, as Ursu (2018) highlighted. The middle panel shows that the SD model predicts a similarly stable conditional purchase probability across positions. In contrast, the right panel shows that the estimated WM model implies a conditional purchase probability that is three to four times larger for lower positions compared to the top ones. As a result, the WM model overpredicts bookings in lower positions despite matching clicks at those positions, creating the contrast highlighted in Figure 4.

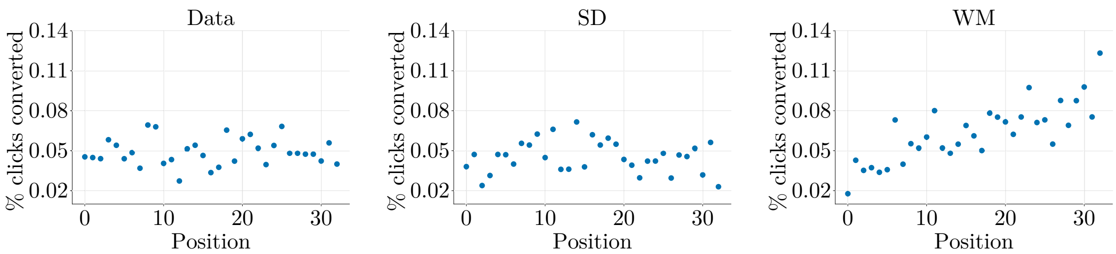

FIGURE OA.10 – Position-Specific Conversion Rate for Clicks

Notes: This figure shows the percent of clicks with a booking across the different positions in either the estimation sample or data simulated from the respective models.

To determine where the difference in the model-implied conditional purchase probabilities comes from, I derive and compare how an alternative's conditional purchase probability depends on its position in the two models.

To get a convenient expression for the conditional purchase probability in the WM model, I use that it is a special case of the SD model (see Greminger, 2022). Hence, the implications in Proposition 2 also apply. Let $\bar{w}_{-j} = \max_{k \in J \setminus j} \tilde{w}_k$ denote the maximum effective value across all alternatives except for some alternative j in position h. The search and discovery implications require $z^s_j(h) \geq \bar{w}_{-j}(h)$ for j to be searched, and the choice implication requires that $u_j \geq \bar{w}_{-j}(h)$ for j to be chosen. The conditional purchase probability for an alternative j in position h therefore is given by

$$\mathbb{P}_{\text{WM}}(\text{choose } j|\text{search } j) = \mathbb{P}(U_j \geq \bar{W}_{-j}|Z^s_j(h) \geq \bar{W}_{-j}) \tag{OA.23}$$

$$= \mathbb{P}(U^l_j + U^d_j \geq \bar{W}_{-j}|U^l_j + x^{l\prime}_j\gamma + \tilde{\gamma}_j + \xi(h) \geq W_{-j}) \,, \tag{OA.24}$$

where $u^l_j = x^{l\prime}_j\beta + \nu_j$ and the probability is taken only over the idiosyncratic shocks ν_j and ε_j, not over alternative-specific attributes that are not random across consumers.

Assuming continuous distributions for the idiosyncratic shocks and taking the derivative of (OA.24) with respect to $\xi(h)$ reveals that an alternative's conditional purchase probability (weakly) decreases in its search value and thus increases in its search cost.[10] This is because with a higher search cost, the alternative is only searched if it offers a sufficiently large list utility u_j^l, which also increases the probability that it is chosen after being searched.

As search costs increase with position h, this effect suggests that the conditional purchase probability increases in h. However, as h increases, the distribution of the maximum effective value $\bar{w}_{-j}$ among alternatives other than j also improves in a first-order-stochastic-dominance sense. Depending on the distributional assumptions, this effect can be either positive or negative. Nonetheless, this second effect will rarely offset the first effect because the change in the distribution of $\bar{W}_{-j}(h)$ tends to be small relative to the direct change in $\xi(h)$, unless there are very few alternatives or position effects in searches are small. As a result, the WM model typically predicts a conditional purchase probability that increases across positions, as the one in Figure OA.10.

The same does not hold for the SD model. In this model, an alternative's position determines its clicks and purchases by determining the likelihood of it being discovered. This probability depends only on the search values and utilities of the alternatives discovered before it rather than on the alternative's own search value and utility. Formally, the conditional purchase probability in the SD model is given by

$$\mathbb{P}_{SD}(\text{choose } j|\text{search } j) = \mathbb{P}(U_j^l + U_j^d \geq \bar{V}_{-j}|U_j^l + x_j^{l\prime}\gamma + \tilde{\gamma}_j + \xi \geq \bar{V}_{-j}) \;, \tag{OA.25}$$

where $\bar{v}_{-j} = \max_{k\in J\setminus j}\min\{\tilde{w}_k, z^d(h_k)\}$. The expression follows from the generalized eventual purchase theorem of Greminger (2022).

Unlike in (OA.24), position h does not enter the search value directly through ξ in this expression. Instead, an increase in h only improves the distribution of $\bar{V}_{-j}$ in a first-order-stochastic-dominance sense, which tends to have a small effect on the conditional purchase probability and can go in either direction depending on the distributional assumptions and parameters such as ξ. Hence, the SD model can generate a conditional purchase probability that is constant across positions and thus consistent with the Expedia data (see Figure OA.10).

[10] Note that (OA.24) has the form $P(X+Y\geq c|X+\xi\geq c) = \frac{\int_{c-\xi}^{\infty}(1-F_Y(c-t))f_X(t)\mathrm{d}t}{\int_{c-\xi}^{\infty}f_X(t)\mathrm{d}t}$, where X and Y are independent random variables and F_Y is the CDF of Y and f_X the PDF of x. The derivative of the numerator and denominator with respect to ξ are given by $1-F_Y(\xi)f_X(c-\xi)$ and $f_X(c-\xi)$, respectively. The derivative of the entire expression with respect to ξ then is given by $\frac{f_X(c-\xi)}{(\int_{c-\xi}^{\infty}(1-F_Y(c-t))f_X(t)\mathrm{d}t)^2}\int_{c-\xi}^{\infty}(F_Y(c-t)-F_Y(\xi))f_X(t)\mathrm{d}t \leq 0$.

G Consumers' Beliefs and Discovery Values

This appendix provides details on how consumers' beliefs about the ranking affect the discovery values, how the discovery values are identified by the data, and how they then can be used to recover discovery costs and the belief parameter. Finally, the appendix shows that any functional form that implies that $\mu(h)$ weakly increases in h contradicts the data.

G.1 Position-Specific Discovery Values

Greminger (2022) shows that the position-specific discovery value $z^d(h)$ is implicitly defined by

$$c_d = \int_{z^d(h)}^{\infty} \left[1 - \tilde{G}_h(t)\right] \mathrm{d}t \, , \tag{OA.26}$$

where $\tilde{G}_h$ is the CDF of the effective value $W_j(h) = X_j^l(h)'\beta + \nu_j + \min\{X_j^l(h)'\gamma + \tilde{\gamma}_j(h) + \xi, X_j^d(h)'\kappa + \delta_j(h), \varepsilon_j\}$. The distribution of this term given the position h is determined by consumers' beliefs over the products' attributes and the shocks, reflected in the respective random variables depending on h.

In Online Appendix EC.2, Greminger (2022) further shows that this is equivalent to $z^d(h) = \mu(h) + \Xi(h)$, where $\Xi(h)$ is implicitly defined by

$$c_d = \int_{\Xi(h)}^{\infty} \left[1 - G_h(t)\right] \mathrm{d}t \, , \tag{OA.27}$$

and G_h now is the CDF of the demeaned effective value $\tilde{W}_j(h) = W_j(h) - \mu(h)$.

The variance and the shape of the distribution of these demeaned effective values does not change with positions. This is because consumers know the distributions of the two shocks and the overall distribution of attributes. Moreover, consumers only expect the mean of $W_j^l(h)$ to depend on the position on the list, whereas the shape of the distribution does not depend on the position. Hence, the distribution of the demeaned effective value does not change across the positions such that $\Xi(h)$ is constant. As a result, the discovery value $z^d(h)$ is given by $z^d(h) = \mu(h) + \Xi$, which is the expression given in the main text.

G.2 Stopping Probabilities Uniquely Determine $z^d(h)$

Given the search parameters, the probabilities of stopping discovery across the different positions determine the discovery value through the stopping rule. To see this, consider the case with only two alternatives, A and B. In this case, the respective probability is given by

$$\mathbb{P}(\text{Stop discovery on position } 1) = \frac{1}{2} \sum_{j \in \{A,B\}} \mathbb{P}(\max\{\tilde{W}_j, U_0\} \geq z^d(2)) \, , \tag{OA.28}$$

which is an expectation over the random ranking. Apart from $z^d(2)$, the search parameters (e.g., β) in (OA.28) are already determined. Hence, the stopping probability identified from the data fully determines $z^d(2)$ through (OA.28). With more than two alternatives, the logic remains the same and the respective probability of stopping discovery on a position determines the associated discovery value.

G.3 Discovery Values Uniquely Determine c_d and ρ

The overall level of the discovery values $z^d(h)$ uniquely determines c_d when consumers know the overall attribute distribution. Intuitively, the discovery values capture the net benefits of discovering more alternatives. These net benefits are determined by consumers' beliefs about the alternatives they will discover and the discovery costs. Hence, the discovery values could be large because consumers expect to discover good alternatives or because the cost of discovering alternatives is small. Whereas it is generally not possible to distinguish the two from the data, the assumption that consumers know the overall distribution they are sampling from fixes consumers' beliefs to the attribute distribution in the data.[11] Given these beliefs, the discovery value fully determines the discovery costs c_d. The procedure described in Online Appendix G.4 shows how the discovery costs can be recovered from the discovery values given the assumptions on consumers' beliefs.

Differences in $z^d(h)$ across positions h then uniquely determine the belief parameter ρ. For example, the functional form (6) that I impose implies

$$\Delta z^d(h) = z^d(h) - z^d(h+1) = \mu(h) - \mu(h+1) = \rho\frac{\log(h)}{\log(h+1)}\ , \qquad \text{(OA.29)}$$

which pins down ρ given the difference in the discovery values across the positions. Note, however, that this functional form is not necessary for the identification of ρ because $z^d(h)$ is identified for each position $h > 1$ by the data. Hence, in principle, more flexible functional forms for the beliefs about the ranking could be used.

Combined, this shows that c_d and ρ are uniquely determined by the discovery values given the other parameters. Hence, whenever the discovery values are identified, c_d and ρ are also identified. Moreover, given some estimate for $\tilde{z}^d$ and the other parameters, c_d can be recovered from $\tilde{z}^d$ without having to estimate it directly.

G.4 Recovering Search and Discovery Costs

c_d and ρ are not estimated directly. Instead, they are recovered from the estimated discovery value for the first position $\tilde{z}^d = z^d(1)$ and the other parameters. Similarly, search costs c_s are

[11] Fixing beliefs by assuming rational expectations is the common approach in empirical search models. For example, virtually all studies that estimate a Weitzman model assume that consumers know the distribution of the shock that will be revealed when searching an alternative.

not estimated directly and instead recovered from the estimated search values. The following shows the procedure to recover these costs after estimating the other parameters of the model.

Recovering Discovery Costs. To recover c_d given estimates for $\tilde{z}^d = z^d(1)$ and the other parameters, I apply the following procedure:

1. Obtain an estimate $\hat{\mu}_{\tilde{w}}$ for $\mu_{\tilde{w}} = \mathbb{E}[\tilde{W}_j]$. To obtain this estimate, I use the following Monte Carlo integration procedure:
   (a) Obtain estimate $\hat{\mu}_{x\kappa}$ by taking the average of $x_j^{d\prime}\hat{\kappa}$ across all products in the data.
   (b) Take N_{c_d} draws (x_q^l, x_q^d, $\hat{\tilde{\gamma}}_q$, ν_q, ε_q) from the respective empirical and assumed distributions.
   (c) For each draw, compute $\tilde{w}_q = x_q^l\hat{\beta} + \nu_q + x_q^l\hat{\gamma} + \hat{\tilde{\gamma}}_q + \min\{\hat{\xi}, x_q^d\hat{\kappa} - \hat{\mu}_{x\kappa} + \varepsilon_q\}$.
   (d) Compute $\hat{\mu}_{\tilde{w}}$ as the average across $\tilde{w}_q$.
2. Substitute (6) in $\frac{1}{|J|}\sum_{h=1}^{|J|}\mu(h) = \hat{\mu}_{\tilde{w}}$ to obtain an estimate $\hat{\mu}_1$ given the estimate $\hat{\rho}$. In my application, I set $|J| = 34$, which is the maximum number of positions in the data.
3. Obtain $\hat{\Xi}$ from $\hat{\mu}_1$ and $\hat{\rho}$ through $\hat{\tilde{z}}^d = \hat{\mu}_1 + \Xi$ implied by the definition of the discovery value and the functional form in (6).
4. Compute the estimate $\hat{c}_d$ from $\hat{\Xi}$ by applying the definition of Ξ given in (7), which is equivalent to the following:

$$c_d = \mathbb{E}\left[\max\{0, \tilde{W}_j - \hat{\mu}_{\tilde{w}} - \hat{\Xi}\}\right] . \tag{OA.30}$$

   To compute the right-hand side, I use the following standard Monte Carlo integration procedure:
   (a) Take N_{c_d} draws (x_q^l, x_q^d, $\hat{\tilde{\gamma}}_q$, ν_q, ε_q) from the empirical distribution of attributes and the assumed distributions of the two taste shocks.
   (b) For each draw, compute $\tilde{w}_q = x_q^l\hat{\beta} + \nu_q + x_q^l\hat{\gamma} + \hat{\tilde{\gamma}}_q + \min\{\hat{\xi}, x_q^d\hat{\kappa} - \hat{\mu}_{x\kappa} + \varepsilon_q\}$.
   (c) Compute $\hat{c}_d$ as the average of $\max\{0, \tilde{w}_q - \hat{\mu}_{\tilde{w}} - \hat{\Xi}\}$ across draws.

Note that by directly sampling from the attribute distribution in the last step, I avoid having to impose parametric assumptions on the distribution of product attributes. Moreover, by computing c_d from $\tilde{z}^d$ instead of estimating it directly, this procedure avoids having to do the costly computation of Ξ given c_d during estimation. Finally, note that consumers are assumed to believe that $U_j^d|x_j^l = x_j^{l\prime}\gamma + \tilde{\gamma}_j + \tilde{\varepsilon}_j$, which leads to the adjustment of the mean of the empirical distribution of $x_j^{d\prime}\kappa$ by $\hat{\mu}_{x\kappa}$ in the first and last step.

Recovering Search Costs. Search costs c_s are recovered from the search values z_j^s implied by the estimates. The definition of the search values is given in equation (5), which integrates over the distribution of the shock $\tilde{\varepsilon}_j$ that has zero mean conditional on the list attributes in x_j^l. To compute the search costs, note that (5) is equivalent to

$$c_s = \mathbb{E}[\max\{0, \tilde{\varepsilon}_j - \xi\}] . \tag{OA.31}$$

The assumptions on consumers' beliefs imply that $\tilde{\varepsilon}_j = X_j^{d\prime}\kappa + \Delta_j + \varepsilon_j - \mu_{x\kappa}$. $\mu_{x\kappa}$ is the mean of $X_j^{d\prime}\kappa + \Delta_j + \varepsilon_j$, which is the sum of the random variables capturing consumers beliefs over x_j^d, δ_j, and ϵ_j. It is subtracted so that $\tilde{\varepsilon}_j$ has zero mean conditional on x_j^l.

To obtain the search cost estimate $\hat{c}_s$, I use the following Monte Carlo integration procedure to compute the right-hand side:

1. Obtain estimate $\hat{\mu}_{x\kappa}$ by taking the average of $x_j^{d\prime}\hat{\kappa}$ across all products in the data.
2. Take N_{c_s} draws (x_q^d, ε_q) from the respective empirical and assumed distributions.
3. For each draw, compute $u_q^d = x_q^{d\prime}\hat{\kappa} + \varepsilon_q - \hat{\mu}_{x\kappa}$.
4. Compute the search cost estimate $\hat{c}_j^s$ as the average across $\max\{0, u_q^d - x_j^{l\prime}\hat{\gamma} - \xi\}$.

G.5 Weakly Increasing $\mu(h)$

Any functional form that implies that $\mu(h)$ weakly increases in h contradicts the data. Hence, even if the model were estimated without the restriction that $\mu(h)$ weakly decreases in h, it would necessarily produce estimates that imply that the assumption holds. To show this, I first establish the following proposition:

Proposition 7. *If $\mu(h)$ weakly increases in h and the ranking is randomized, then there are no position effects on positions $h > 1$ for consumers who take the outside option.*

Proof. The optimal policy is based on the three reservation values independent of how $\mu(h)$ depends on h, and $z^d(h)$ will continue to in- or decrease in h depending on how $\mu(h)$ in- or decreases in h (see Greminger, 2022). The result then follows from the stopping rule. This rule implies that consumers who take the outside option continue discovering as long as $U_0 > z^d(h)$. As a result, if $\mu(h)$ and, hence, $z^d(h)$ weakly increase in h, consumers who take the outside option always discover either only the first—which is discovered for free—or all available alternatives. This already implies the result because position effects can only arise in the SD model through consumers stopping discovery before reaching the last position, unless the ranking is not randomized. □

The proposition makes a prediction for the case when $\mu(h)$ is weakly increasing in h. This prediction contradicts the data because there are position effects for consumers who take the outside option, including for positions $h > 1$. Specifically, re-running the analysis from Section 2.2 while excluding the first position continues to produce significant position effects. Hence, a weakly increasing $\mu(h)$ contradicts the data. Moreover, note that when consumers know that they are facing a randomized ranking, $\mu(h)$ will be constant. Hence, the proposition also implies that consumers do not adjust their beliefs immediately to the randomized ranking; otherwise, the position effects would disappear for consumers who take the outside option.

H Software Implementation

I provide the estimation approach as Julia and Python packages. The packages are available on GitHub and can be installed using the usual package managers. The Julia package is the main implementation and includes all the functionalities, while the Python package provides a convenient interface to this package for Python users. Both packages are open-source and freely available. The links to the repositories are as follows:

- Link to GitHub repository for Julia package: *https://github.com/rgreminger/StructuralSearchModels.jl*.
- Link to GitHub repository for Python wrapper: *https://github.com/rgreminger/structuralsearchmodels*.

In its current form, the Julia package implements the estimation procedure developed in this paper, as well as methods to simulate data from the models. It is designed to be user-friendly, and comes with documentation and examples. Moreover, it is optimized for performance and efficiently estimates the two models using parallelization and automatic differentiation.

The Julia package builds on various open-source software and packages. In particular, the main dependencies are the following:

- *Julia* by Bezanson et al. (2017): high-level programming language used for the implementation.
- *Distributions.jl* by Besançon et al. (2021): implements distribution types that make it easy to work with different distributions.
- *Optimization.jl* by Dixit and Rackauckas (2023): flexible framework that allows to define optimization problems and solve them using different solvers and options.
- *Optim.jl* by Mogensen and Riseth (2018): provides numerical optimization routine used by default.
- *ForwardDiff.jl* by Revels et al. (2016): provides automatic differentiation to compute gradients and Hessians for gradient-based optimization.

The fit plots in this paper are not provided as part of the package to keep the dependencies limited. However, the respective codes are available as part of the replication package. These codes rely on the *Makie.jl* plotting package by Danisch and Krumbiegel (2021).

I Data

The original dataset from Kaggle.com contains 9,917,530 observations on a hotel-session level. Following Ursu (2018), I exclude sessions with at least one observation satisfying any of the following criteria:

1. The implied tax paid per night either exceeds 30% of the listed hotel price (in $), or is less than $1.
2. The listed hotel price is below $10 or above $1,000.
3. There are less than 50 consumers looking for hotels at the same destination throughout the sample period.
4. The consumer observed a hotel in positions 5, 11, 17, or 23, i.e., the consumer did not have opaque offers (Ursu (2018) provides a detailed description of this feature in the data).

The final dataset contains 4,503,128 observations, which is 85 fewer observations than Ursu (2018) reports. The difference stems from three sessions (IDs 79921, 94604, 373518). For these sessions, the criteria above yields a different result due to differences in the numerical precision. Specifically, I evaluate the criteria with double precision calculations, whereas Ursu (2018) uses the Stata default, single precision.

Completing information on the dataset, Table OA.12 provides a detailed description of each variable.

TABLE OA.12 – Variable Descriptions

Hotel-level	
Price (in $)	Gross price in USD
Star rating	Number of hotel stars
Review score	User review score, mean over sample period
No reviews	Dummy whether hotel has zero reviews (not missing)
Chain	Dummy whether hotel is part of a chain
Location score	Expedia's score for desirability of hotel's location
On promotion	Dummy whether hotel on promotion
Session-level	
Number of items	How many hotels in list for consumer, capped at first page
Number of clicks	Number of clicks by consumer
Made booking	Dummy whether consumer made a booking
Trip length (in days)	Length of stay consumer entered
Booking window (in days)	Number of days in future that trip starts
Number of adults	Number of adults on trip
Number of children	Number of children on trip
Number of rooms	Number of rooms in hotel

References

Besançon, M., T. Papamarkou, D. Anthoff, A. Arslan, S. Byrne, D. Lin, and J. Pearson (2021). Distributions.jl: Definition and modeling of probability distributions in the JuliaStats ecosystem. *Journal of Statistical Software 98*(16), 1–30.

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah (2017). Julia: A fresh approach to numerical computing. *SIAM Review 59*(1), 65–98.

Chung, J. H., P. K. Chintagunta, and S. Misra (2025). Simulated maximum likelihood estimation of the sequential search model. *Quantitative Marketing and Economics 23*, 105–164.

Compiani, G., G. Lewis, S. Peng, and P. Wang (2024). Online search and optimal product rankings: An empirical framework. *Marketing Science 43*(3), 615–636.

Danisch, S. and J. Krumbiegel (2021). Makie.jl: Flexible high-performance data visualization for Julia. *Journal of Open Source Software 6*(65), 3349.

Dixit, V. K. and C. Rackauckas (2023). Optimization.jl: A unified optimization package. https://doi.org/10.5281/zenodo.7738525.

Drezner, Z. and G. O. Wesolowsky (1990). On the computation of the bivariate normal integral. *Journal of Statistical Computation and Simulation 35*(1-2), 101–107.

Fong, J., O. Natan, and R. Pantle (2024). Consumer Inferences from Product Rankings: The Role of Beliefs in Search Behavior. Working Paper, https://www.ssrn.com/abstract=4896993.

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

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

Kaye, A. (2024). The personalization paradox: Welfare effects of personalized recommendations in two-sided digital markets. Working Paper, https://apkaye.github.io/aaronpkaye.website/Kaye_Aaron_JMP.pdf.

Mogensen, P. K. and A. N. Riseth (2018). Optim: A mathematical optimization package for Julia. *Journal of Open Source Software 3*(24), 615.

Owen, D. B. (1980). A table of normal integrals. *Communications in Statistics - Simulation and Computation 9*(4), 389–419.

Revels, J., M. Lubin, and T. Papamarkou (2016). Forward-mode automatic differentiation in Julia. Working Paper, https://arxiv.org/abs/1607.07892.

Ursu, R., S. Seiler, and E. Honka (2024). The sequential search model: A framework for empirical research. *Quantitative Marketing and Economics 23*, 165–213.

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

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