BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel — 2020-5

Last updated: March 5, 2020

Identifying Consumer-Welfare Changes when Online Search Platforms Change Their List of Search Results

by Ryan Martin

Currency Department
Bank of Canada, Ottawa, Ontario, Canada K1A 0G9

RMartin@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

I wish to thank UCLA professors Rosa Matzkin, John Asker, Denis Chetverikov, Elisabeth Honka, Jinyong Hahn, Andres Santos and Shuyang Sheng for their guidance on this paper. I also wish to thank Jason Allen, Shota Ichihashi, my colleagues in the Bank of Canada's currency research team, participants at the Bank of Canada's microeconometrics and survey methodology conference and all participants of my 2019 job-market talks. An earlier version of this paper was circulated during the 2019 job market under the title "Nonparametric Identification of Average Consumer Welfare Changes when Heterogeneous Consumers Search for Discrete Goods." Any errors are my own.

Abstract

Online shopping is often guided by search platforms. Consumers type keywords into query boxes, and search platforms deliver a list of products. Consumers' attention is limited, and exhaustive searches are often impractical. Thus, the order in which products appear in search results affects the products consumers discover and ultimately purchase. In this setting, I study the identification of consumer-welfare changes in response to exogenous changes in search-result lists. I focus on the case of consumers engaging in costly searches for a single, indivisible (discrete) product among a collection of substitutes. I show that exact consumer-welfare changes—that is, compensating variation and equivalent variation—can be calculated with the use of straightforward integrals of the aggregate demand. I apply my results to shopping data provided by an online travel agency (OTA). I estimate that when the OTA changes search results from random to its proprietary listing structure, welfare improves by an average of $8.84 per user. I estimate an average welfare loss of $20.51 per user when the OTA removes the top five products from all of its search-result lists.

Bank topics: Econometric and statistical methods, Market structure and pricing

JEL codes: *C, C1, C14, D, D11, D12, D6, D8, D83, L, L4, L40*

1 Introduction

Online, consumers often rely on search tools to help them find relevant products to buy. Typically, consumers type keywords into the query box of a search platform, and the platform delivers a list of products; I call this delivered list of products a consumer's "search-result list." The number of substitutes available online is often substantial, and an exhaustive search over all substitutes is impractical. Consumers often start their search from the beginning of the search-result list and stop well before the end. Thus, the order products appear in consumers' search-result lists affects the products consumers discover and ultimately purchase. When search platforms change the order of products that appear in their lists of search results, aggregate demand and consumer welfare are also changed. I study the identification of consumer-welfare changes in response to exogenous changes in search-result lists. I focus on the case of consumers shopping for a single, indivisible product. I show that, in this environment, exact consumer-welfare changes—that is, compensating variation and equivalent variation—can be calculated using straightforward integrals of aggregate demand.

I model consumers as having product knowledge limited to "consideration sets." Each consumer forms her consideration set as a function of her own characteristics and the search-result lists that are returned to her queries. Each consumer's demand is limited to the products that are in her consideration sets: consideration sets censor demand. Aggregate demand is a function of all prices and individual consideration sets. Changes in search-result lists cause discontinuous shifts in demand through their actions on consideration sets. I determine formulas that recover average welfare changes from these shifts in aggregate demand. My results do not require consideration sets to be observed. My model places minimal restrictions on how consumers form their consideration sets. My context of an A/B experiment allows for rich welfare inferences under relatively weak restrictions on consumer primitives.

I consider two applications. In both, I use data from an online travel agency (OTA). For the first application, I measure the welfare consequences of the OTA's changing search-result rules (or algorithms). I take advantage of an experiment that the OTA ran listing hotel bookings for treated consumers in random order and hotel bookings for untreated consumers in a proprietary order. I find that the OTA's proprietary ordering improves average welfare by $8.84 per user over the random ordering. In the second application, I estimate the welfare loss that would result if the OTA were to remove the five most popular products from all search-result lists. I estimate this would lower average welfare by $20.51 per person.

My welfare formulas focus on the utility consumers receive from their final product purchases under different search-result lists. That is, my welfare measures do not explicitly account for the psychic and time costs that arise from changes in search-result lists. My welfare measures account for these costs implicitly through their impact on consideration sets and ultimate product choice. This allows me to leave the consumer's search behavior relatively unrestricted.[1] My demand model makes minimal assumptions about consumer preferences. My most general results require that utility be linear in money but allow the remaining terms to be nonparametric and (potentially) nonseparable in the unobservables. For the special case of a change in search-result lists that causes the (probabilistic) removal or addition of a collection of products from consideration sets, I show that the weaker assumption of a utility that is monotonic in money (and otherwise unrestricted) is sufficient for bounding the resulting welfare changes. With data on the consumer-search process, it should be straightforward to modify my results to include the direct psychic and time costs of searches.

My research is motivated by recent antitrust concerns over the growing concentration of online platforms.[2] Online, most consumers and sellers find each other through the services of these platforms. Without these platforms, consumers and sellers would have a hard time making new connections. Thus, a concentrated platform may have great influence over

[1] To be clear, consumer search is allowed to be costly in my model. Consumers will take their own psychic and time costs into account when deciding how much of the product space to explore. The welfare effects of these costs will then be captured by how they change their final product choice.

[2] For example, the Federal Trade Commission (FTC) is holding ongoing hearings on competition and consumer protection in the 21st century that address this topic: https://www.ftc.gov/policy/hearings-competition-consumer-protection.

who buys what from whom online and at what price. This gives rise to several antitrust concerns over search-platform conduct. For example, a search platform that also sells its own products, such as Amazon, may be tempted to flex its influence over consumers' search results in negotiations with sellers in order to extract high proportions of seller revenue on their site and limit seller behavior off their site.[3] Alternatively, a search platform may be tempted to hide the search results that would lead consumers to other search platforms, in order to protect its own market share. This is especially concerning when the search platform represents its organic[4] search results as unbiased. European antitrust authorities fined Google $2.7 billion for this type of conduct.[5] To the extent that this platform conduct influences the distribution of sellers and the ultimate product choice of consumers, it may also have strong consumer-welfare implications. My model framework allows for a rigorous, minimally restricted study of these consumer-welfare consequences.

My research also highlights the important role A/B tests can play in online search settings. Economists are long familiar with the importance of exogenous variation in price for demand and welfare estimation in classic economic environments. Analogously, exogenous variation in platform listings (in the context of an A/B test or policy intervention) can be essential for accurately measuring welfare changes in search settings. Consider a platform that may choose to list products according to the strategy `list_exploit` or the strategy `list_encourage`. When the platform uses the strategy `list_exploit`, it places the products with the largest markup on the first page. When the platform uses the strategy `list_encourage`, it puts products that generate the most consumer surplus on the first page. Further suppose that consumers may be one of two types: *lazy* or *determined*. *Lazy* consumers only search the first page of results, while *determined* consumers know their surplus-maximizing product and will search all pages, if necessary, starting from the first and continuing until they find it. If the data show consumers only searching the first page of results, then this could be

[3]See Khan (2017) for more discussion on evidence of this behavior.

[4]An organic search result is one whose position is not chosen through a payment to the search platform, but rather by the search platform's own algorithm.

[5]http://money.cnn.com/2017/06/27/technology/business/google-eu-antitrust-fine/index.html

explained equally well by either *lazy* consumers and a `list_exploit` strategy or *determined* consumers and a `list_encourage` strategy. An A/B test lets us separate these two cases and draw welfare conclusions in ways data under one search-result-list algorithm cannot.

To the best of my knowledge, this paper is the first to identify compensating variation and equivalent variation from search-result listings as a function of changes in aggregate demand in a general search environment. However, this paper does draw insights and inspiration from several existing strands of literature.

First, this paper is related to that of Bhattacharya (2015). In particular, both his paper and my paper identify the compensating and equivalent variations in a discrete-choice environment where utility is monotonic in money but otherwise unrestricted. While Bhattacharya (2015) examines the welfare consequences of a single price increase, I focus on the welfare consequences of changing the search-result lists. In Bhattacharya (2015), consumer knowledge is perfect and there is no search; the demand and welfare consequences of consideration sets and limited consumer knowledge are not modeled. I show that the additional complications of the search environment require strengthening the utility assumptions to a utility that is linear in money in order to measure generic consumer welfare changes from changes in search-result lists. Further, while my environment is more complex, I find novel and relatively elegant proofs to achieve nonparametric results analogous to his.

Second, my results are related to the literature on the consumer welfare changes that occur from the introduction of new goods. This econometric literature goes back to Hausman (1981, 1996). The above papers and their extensions focus on identifying and estimating the exact consumer-welfare changes that result from the introduction of a single product or product category. All consumers are assumed to have perfect knowledge of all available products. The newness of the product is modeled as the consequence of technological innovation or regulatory decisions and is assumed to be exogenous to the consumer's final purchase decision. In contrast, my paper focuses on the role that search-result lists have in shaping product knowledge and welfare. Individual search behavior is heterogeneous and the consequences of a changing search-result list may be heterogeneous and unpredictable across consumers in my environment.

This paper is also related to Small and Rosen (1981), who developed tools to estimate consumer welfare changes in discrete-choice environments. Small and Rosen (1981) provide tools to estimate the consumer welfare changes that occur in response to a change in price, quality or any variable that varies continuously with indirect utility. In contrast, my paper focuses on changes in consideration sets that, by their very nature, provide discontinuous shifts to indirect utility functions. Thus, the results of this paper represent a significant extension of those of Small and Rosen (1981). Indeed, no simple adaption of the techniques used in their research will lead to correct welfare measures in an environment with changing, heterogeneous consideration sets.

Some recent situation-specific methods have been developed to estimate welfare changes that result from non-idiosyncratic product removal or exit. These include Nevo (2003), Gentzkow (2007), Quan and Williams (2018), and Petrin (2002). My paper provides a generalization of their results, allowing for simultaneous product entry and exit in a flexible utility environment. My results are also shown to be exactly equal to compensating variation or equivalent variation, rather than just being approximations. Finally, my methodology allows for recovery of the welfare changes that were caused by unobservable preference matching rather than just recovering the average welfare components.

In addition, a number of empirical and game-theoretic papers have studied consumer welfare in search markets. The topics studied include welfare changes as a result of search-ranking changes (e.g., Ursu (2017) and Athey and Ellison (2011)), the welfare effects of platform changes (e.g., Lewis and Wang (2013), Dinerstein et al. (2017), Fradkin (2018)), the welfare effects of advertising and search (e.g., Honka, Hortaçsu, and Vitorino (2017) and Seiler and Yao (2017)), and the welfare effects of changing search costs (e.g., Honka (2014), Ershov (2016) and Moraga-Gonzalez, Sándor, and Wildenbeest (2017)). These papers rely on strong modeling assumptions of the search process, parametric utility, and situation-specific measures of consumer welfare. My paper allows for less-restricted search behavior, less-restricted preferences, and welfare measures equivalent to the readily interpretable classical measures of compensating variation and equivalent variation.

Finally, there is a growing body of research that is interested in assessing the value of technology, the internet, and free (digital) goods and services in terms of their contributions to GDP. See, for example, Brynjolfsson, Collis, et al. (2019), Diewert and Feenstra (2017), Diewert, Fox, and Schreyer (2018), Feldstein (2017), Groshen et al. (2017), Syverson (2017), Brynjolfsson and Oh (2012) and Greenstein and McDevitt (2011). These studies use the welfare-analysis tools of Hausman (1996) or Small and Rosen (1981). Their models do not account for searching consumers. Instead, their focus is on measuring the aggregate welfare consequences of products that are available in the digital economy. In contrast, my paper allows for average welfare changes that occur as a result of idiosyncratic changes in shopping behavior.

The rest of this paper is organized as follows. In Section 2, I develop the notation. In Section 3, I present the results under a monotonicity constraint on preferences. In Section 4, I strengthen my assumptions on consumer preferences to quasi-linearity and obtain a general formula to measure the average welfare changes that occur from arbitrary search-result-list changes. I also present a simple example that illustrates the key ideas captured in the formula. In Section 5, I apply my main results to several simple search-result-list changes of practical and empirical interest. In Section 6, I apply my results to data provided by an OTA. Finally, in Section 7, I conclude. All proofs are left to the appendix.

2 Notation

In this section, I develop the notation required for the rest of the paper. I start by discussing the role of search platforms in Section 2.1. In Section 2.2, I discuss product and consumer preference notation. I develop notation and assumptions for consideration sets in Section 2.3.

I define and develop notation for welfare measures in Section 2.4.

2.1 Search-Result Lists

A consumer searching for a product online will type keywords into a search platform's query box. From there, the search platform has a *listing rule* α that determines the order (or layout) in which the relevant products are listed for the consumer to review. The focus of this paper is not on how platforms determine this listing rule; the listing rule is free to depend on observable consumer characteristics and advertising concerns in addition to the consumer's keywords. Rather, this paper focuses on measuring consumer-welfare changes that are a response to changes in the listing rule, say from $\alpha = A$ to $\alpha = B$.

2.2 Preferences

My setup for products and preferences follows the multinomial choice framework with non-separable utility laid out in Bhattacharya (2015). There is an *observable*[6] *set of products* $\mathcal{J} = \{0, 1, \ldots, J\}$.[7] I denote the *observable vector of market prices* $p^m_{\mathcal{J}} = (0, p^m_1, \cdots, p^m_J)$. The price of the outside product is normalized to 0. When discussing product prices that may differ from their market-values, I drop the superscript m.

Consumer utility is affected by income y and *observable attributes* Ψ. Both y and Ψ are fixed for each individual. For readability, I suppress the notation for Ψ from utility. All identification results should be interpreted as conditional on Ψ.

Consumers have *unobservable preferences* η. I do not restrict the dimension of these unobservable preferences.[8] However, I restrict utility using the following assumptions:

[6] The term "observable" is always used to mean observable to the researcher, not the consumer.

[7] It is without loss of generality (WLOG) to have the products invariant over the listing rules. For example, if product M becomes available after a change in the listing rule, then we can just disallow M from being in the consideration sets under the initial listing rule.

[8] See Bhattacharya (2015) for a discussion on the importance of leaving the heterogeneity dimension unrestricted in discrete-choice preferences.

Assumption 1.A (Monotonicity). *Utility for product j, $u_j(y - p_j, \eta) \in \mathbb{R}$, is strictly increasing and continuous in its first argument for all products $j \in \mathcal{J}$.*

Assumption 1.B (quasi-linearity). *Utility for product j is linear in its first argument. That is,*

$$u_j(y - p_j, \eta) = y - p_j + \tilde{U}_j(\eta), \tag{1}$$

where $\tilde{U}_j(\eta) \in \mathbb{R}$ for all products $j \in \mathcal{J}$.

Monotonicity is a weak and intuitively appealing assumption. It rules out consumers remaining indifferent between two goods over any price interval and requires consumers to like a good less as its price increases. Quasi-Linearity is stronger but standard in empirical applications.[9]

I maintain monotonicity for the rest of the paper. Under monotonicity, income is assumed observable. Under quasi-linearity, income need not be observable.[10]

2.3 Consideration Sets

I assume that search is costly and thus consumers do not, in general, view all of the products returned to them in platform search lists. Instead, each consumer's product knowledge—and therefore demand—is limited to a sub-collection of $\mathcal{J}$, which is referred to as her consideration set.

Consideration sets are determined through a *consideration function* $\mathcal{C}$. The consideration function takes the following arguments: (1) a listing rule α; (2) observable consumer charac-

[9]All formulas in this paper that hold under quasi-linearity also hold with the addition of a consumer-specific coefficient to money, a_η. That is, for my results, it is WLOG to normalize utility to the form in Equation (1) from a form where $u_j(y - p_j, \eta) = a_\eta(y - p_j) + \tilde{U}_j(\eta)$.

[10]A vector of observable product characteristics $(0, X_1, \ldots, X_J)$ can also be available to the researcher. This information can be treated as being suppressed from utility for readability. If we were to suppress the product characteristics and observable individual characters from utility, we would write $u_j(y - p_j, X_j, \Psi, \eta)$ under monotonicity and $y - p_j + \tilde{U}_j(\eta; X_j, \Psi)$ under quasi-linearity.

teristics Ψ; (3) unobservable consumer preferences η; and (4) an unobservable non-preference characteristic vector ζ. For readability, I suppress Ψ from the consideration-set notation.

Definition 1 (Consideration Sets)**.** *A consumer with characteristics* (y, η, ζ) *has consideration set* $\{0\} \subseteq \mathcal{C}(\eta, \zeta, \alpha) \subseteq \mathcal{J}$ *under listing rule* α*.*

When discussing a fixed consumer (y, η, ζ), I abbreviate her consideration set under listing rules A and B by $\mathcal{C}_A$ and $\mathcal{C}_B$, respectively.

For a consumer (y, η, ζ), the components of ζ capture the unobservable factors that affect her consideration set but that do not enter her utility function. For example, a component of ζ captures a consumer's preference for the act of shopping, itself. Consumers who like to shop will likely have larger consideration sets than consumers who do not like to shop, even when their product preferences are identical. ζ may also capture product and price beliefs that guide the search over different keywords.

As an example, consider a consumer searching for a face moisturizer on *Amazon.com*. Suppose she does not like shopping for very long and chooses to either buy a product from the first page of search results that comes up after her keyword search or not buy anything (captured in ζ). She has a preference for branded products (captured in η) and thinks that "Biore" is likely to be a relatively inexpensive branded product (captured in ζ). Thus, she types "Biore face moisturizer" into the search box and hits the return key. Some of the products in the resulting list are organic and some are sponsored links. Her consideration set is exactly the products listed on this first page of search results, as well as the outside product.

To derive my main results, I rely on the following assumption:

Assumption 2 (Price Independence)**.** *For all products* $j \in \mathcal{J}$*, the consideration function* $\mathcal{C}$ *does not depend on prices* p_j *or income minus prices* $y - p_j$*.*

Note that price independence does not preclude a consumer from considering products according to her *beliefs* about prices. It only requires that the prices she observes do not cause her to change her shopping behavior.[11]

Price independence rules out the possibility of a platform removing a product from its lists due to a change in its price. It also rules out certain consumer behaviors, such as a consumer expanding her consideration set after discovering that all the products in her initial consideration set are unexpectedly expensive. In practice, some price dependence can be tolerated: the key behavior needed is that, as a good's price increases, a consumer's demand for that good falls to 0 for preference reasons before it the price increases causes the good to exit (or another good to enter) her consideration set. That is, as long as the consumer's preferences are more sensitive than the search listing rule, the main results should still hold.

A consumer (y, η, ζ) purchases product j in $\mathcal{C}(\alpha, \zeta, \eta)$ if

$$u_j(y - p_j^m, \eta) > u_k(y - p_k^m, \eta) \text{ for all } k \in \mathcal{C}(\alpha, \zeta, \eta) \setminus \{ m \} .$$

Her *individual demand* is defined by

$$q_j(y, p_j, p_{-j}, \alpha, \eta, \zeta) := \begin{cases} 1 \text{ if } j = \arg\max_{\ell \in \mathcal{C}(\alpha,\eta,\zeta)} u_\ell(y - p_\ell, \eta) \\ 0 \text{ otherwise.} \end{cases} \tag{2}$$

If the price of all products except good j are fixed at their market level, then I denote her *individual demand for product j at market prices* by

$$q_j^m(y, p_j, \eta, \zeta, \alpha) := q_j(y, p_j, p_{-j}^m, \alpha, \eta, \zeta). \tag{3}$$

I will denote the joint distribution of η and ζ by F. Thus, the *average demand for product*

[11] Although not modeled explicitly, it is perfectly fine for a product's non-price characteristics X_j to affect consideration sets. In addition, it is perfectly fine for a consumer's consideration set to depend on her perceptions of her own income level or wealth level, as a part of Ψ. The only complication that might arise here is if a consumer's final product choice changes her perception of her own wealth level. By assumption, this is not allowed.

j is

$$Q_j(y, p_j, p_{-j}, \alpha) = \int q_j(y, p_j, p_{-j}, \alpha, \eta, \zeta) dF \tag{4}$$

for all consumers with income y. Similarly, the *average demand for product j at market prices* is,

$$Q_j^m(y, p_j, \alpha) = \int q_j(y, p_j, p_{-j}^m, \alpha, \eta, \zeta) dF \tag{5}$$

for all consumers with income y.

2.4 Welfare Measures

Suppose a platform changes its search listing rule from A to B. The following definitions adapt the classic measures of welfare changes, *compensating variation* and *equivalent variation*, to this situation. For an individual consumer (y, η, ζ), her equivalent variation S^{EV} is the solution in S to

$$\max_{j \in \mathcal{C}_A} u_j(y - S - p_j^m, \eta) = \max_{j \in \mathcal{C}_B} u_j(y - p_j^m, \eta), \tag{6}$$

while her compensating variation S^{CV} is the solution in S to

$$\max_{j \in \mathcal{C}_A} u_j(y - p_j^m, \eta) = \max_{j \in \mathcal{C}_B} u_j(y + S - p_j^m, \eta). \tag{7}$$

S^{EV} is the income loss under the initial listing rule that would harm this consumer as much as the damage done by the new listing rule. S^{CV} is the increase in income under the new listing rule that would return a consumer to the utility level she would have had under the original listing rule. S^{EV} and S^{CV} are both positive if the consumer is harmed by the new listing rule, relative to the older rule, and negative otherwise.

Both S^{EV} and S^{CV} depend on market prices and both listing rules, as well as the individual's unobservable characteristics η and ζ. In the case of monotonicity, both S^{EV} and S^{CV} also depend on individual income. Thus, I denote the functions for equivalent variation and compensating variation by $S^{EV}(y, \eta, \zeta, A, B, p^m_{\mathcal{J}})$ and $S^{CV}(y, \eta, \zeta, A, B, p^m_{\mathcal{J}})$, respectively. Similarly, *average compensating variation* and *average equivalent variation* over all consumers are denoted by μ^{CV} and μ^{EV}. As functions, I write these as $\mu^{CV}(y, A, B, p^m_{\mathcal{J}})$ and $\mu^{EV}(y, A, B, p^m_{\mathcal{J}})$; the average is taken over unobservables η and ζ, while income is fixed.[12] When it is clear from the context, I suppress the arguments for the listing rules and prices from these welfare functions.

Under quasi-linearity, each individual (y, η, ζ) has $S^{CV} = S^{EV}$. In this case, I use S^W to represent both S^{CV} and S^{EV}. S^W is simply the difference between utility under the initial listing rule and utility under the final listing rule.

3 Measuring Welfare Under Monotonicity

This section considers welfare identification under monotonicity. Below, I present the bounds on compensating variation that result from a search-listing change that makes a collection of new goods available for consideration.

Theorem 1 (Probable Addition of Goods)**.** *Suppose the listing rule changes exogenously from A to B such that a previously unconsidered collection of goods $\mathcal{R}$ can enter consideration sets.*[13] *Then, the average compensating variation of this change in listing rule is such that*

$$-\max_{j\in\mathcal{R}} \lim_{p_k\to\infty\forall k\in\mathcal{R}\setminus\{j\}} \int_{p^m_j}^{\infty} Q_j(y, p_j, (p_{\mathcal{R}\setminus\{j\}}, p^m_{-\mathcal{R}}), B)dp_j \geq \mu^{CV} \tag{8}$$

$$\geq -\sum_{j\in\mathcal{R}} \lim_{p_k\to\infty\forall k\in R\setminus\{j\}} \int_{p^m_j}^{\infty} Q_j(y, p_j, (p_{\mathcal{R}\setminus\{j\}}, p^m_{-\mathcal{R}}), B)dp_j.$$

[12]Averages would still be functions of consumer observables Ψ, although the researcher could, of course, average over Ψ (and y) if she desired. Switching the integral order to accommodate earlier averaging is never a problem, either, since demand is nonnegative.

[13]Specifically, I mean for every (y, η, ζ), $\mathcal{C}_A \subseteq \mathcal{C}_B \subseteq \mathcal{C}_A \cup \mathcal{R}$ and $\mathcal{C}_A \cap \mathcal{R} = \emptyset$.

The appendix contains the proof, as well as the analogous bounds for equivalent variation under an exogenous treatment that causes the removal of a collection of goods from search-result lists.

Intuition for the bounds in Theorem 1 can be developed from the following observations. First, a consumer's preference for a good depends on which other goods are in her consideration set. To fix ideas, I focus on a consumer's demand for good 1. I will assume good 1 is initially her favorite among those in her consideration set. However, as the number of goods in her consideration set grows, her willingness to pay for good 1 falls weakly. In particular, if a new good 2 enters her consideration set and she prefers 2 to 1 or her next best choice after 1, her demand line for good 1 will shift downward. Otherwise, no change will occur. The same is true for each additional new good that enters her consideration set.

Second, as the price of a good goes to infinity, it becomes so undesirable that it is "as if" the good were not even included in the consumer's consideration set. Thus, we can simulate the sequential removal of several goods from a consumer's consideration set by examining how her demand changes as the price of these goods goes to infinity. For more intuition in the simpler case of quasi-linearity, see the example in Section 4.

Adding these insights together, we can interpret Equation (8) as follows. The upper bound on welfare is captured by the value of the new good that had been added to the search-result lists that would have shifted demand the most had it been the only good added. Of course, removing the pricing limits decreases the area under the curve, so that

$$-\max_{j \in \mathcal{R}} \int_{p_j^m}^{\infty} Q_j(y, p_j, p_{-j}^m, B) dp_j \tag{9}$$

could also be used as a (looser) upper bound for the average compensating variation.

The lower bound on welfare sums the welfare effects that would result from individually adding each new good to the search-result lists. The sum of individual values exceeds the whole, since each consumer will only purchase, at most, one good from the collection of goods. Unlike the left-hand inequality in Equation (8), the right-hand inequality will no longer hold, in general, if we remove the pricing limits from the right-most terms.

When $\mathcal{R}$ is a singleton, the above bounds collapse to a single point. A little more can be stated with the following definition. Fix A and B and define operator Δ as follows:

$$\Delta Q(y, p_J^m) \equiv Q_j(y, p_J^m, A) - Q_j(y, p_J^m, B), \tag{10}$$

Then we have the following corollary:

Corollary 1 (Probable Addition of a Single Good)**.** *Suppose the listing rule changes exogenously from A to B in a way that makes each consumer more likely to shop for product M.*[14] *Then, the average compensating variation of this change in listing rule is*

$$\mu^{CV} = \int_{p_M^m}^{\infty} \Delta Q_M(y, p_M, p_{-M}^m) dp_M. \tag{11}$$

The proof of Corollary 1 is in the appendix. However, the results are intuitively appealing. They are similar to classic results on the welfare of a new good. In the appendix, I include analogous results that relate the average equivalent variation in the case of a probable removal of a good.

Note that the results of Corollary 1 hold regardless of how many consumers' consideration sets hold M under listing rule A. This is in contrast with Theorem 1, where all of $\mathcal{R}$ must be unconsidered under listing rule A.

4 Measuring Welfare Under Quasi-linearity

In this section, I determine how to measure welfare changes as a response to search-listing changes. I start with my most general result: a formula that measures welfare changes from aggregate demand lines under quasi-linearity. I leave the proof for the appendix, but follow

[14]Specifically, I mean for every (y, η, ζ), $\mathcal{C}_B$ is equal to $\mathcal{C}_A$ or $\mathcal{C}_A \cup \{ M \}$.

up with an example that shows the key ideas. I then present some simpler formulas that can be applied to specific cases of welfare changes.

For succinctness, I first denote *total consumer welfare under listing rule* α by Ω_α and define it with the following formula:

$$\Omega_\alpha := \lim_{p_2,\ldots,p_J\to\infty}\int_{p_1^m}^{\infty} Q_1(y,p,p_{-1},\alpha)dp + \lim_{p_3,\ldots,p_J\to\infty}\int_{p_2^m}^{\infty} Q_2(y,p,(p_1^m,p_{-(1,2)}),\alpha)dp$$
$$+\cdots+\int_{p_J^m}^{\infty} Q_J(y,p,p_{-J}^m,\alpha_m)dp. \quad (12)$$

Total consumer welfare under listing rule α—hereafter abbreviated as total welfare—captures the total value to consumers of the products in their (heterogeneous) consideration sets under listing rule α. This total value is relative to the outside good—no purchase. More precisely, the first term in the sum that defines Ω_α calculates the average value of allowing product 1 to enter all consumers' consideration sets; if a consumer does not have product 1 in her consideration set under listing rule α, then this consumer's contribution to the average is 0. This added value is relative to consideration sets that only contain the outside product. The second term in the sum that defines Ω_α adds the average value consumers gain by having product 2 in their consideration sets, relative to consideration sets that contain (at most) product 0 and product 1; consumers without product 2 in their consideration set contribute nothing to this value. Consumers with product 2 but not product 1 in their consideration set will contribute average values to the 2nd term that reflect product 2's value relative to the outside product alone. This process of adding in the average value of one more product is continued from the third term until the Jth term in the sum. By the Jth term, all of the consideration sets will have reached their full size under listing rule α. All the terms together give Ω_α, the total value of sequentially allowing products 1 to J to enter all of the consideration sets.

Lemma 1 below provides some deeper intuition. Fix a consumer (y,η,ζ) and let ω_α

denote her "total welfare." That is,

$$\omega_\alpha(y,\eta,\zeta) := \lim_{p_2,\ldots,p_J\to\infty}\int_{p_1^m}^{\infty} q_1(y,p,p_{-1},\alpha,\eta,\zeta)dp + \lim_{p_3,\ldots,p_J\to\infty}\int_{p_2^m}^{\infty} q_2(y,p,(p_1^m,p_{-(1,2)}),\alpha,\eta,\zeta)dp$$
$$+\cdots+\int_{p_J^m}^{\infty} q_J(y,p,p_{-J}^m,\alpha_m,\eta,\zeta)dp. \tag{13}$$

Then we have the following result:

Lemma 1. *Under quasi-linearity,*

$$\omega_\alpha(y,\eta,\zeta) = \max_{j\in\mathcal{C}_\alpha} u_j(y-p_j^m,\eta) - u_0(y,\eta).$$

This result is proved in the appendix as part of my proof of Theorem 2. Theorem 2 is stated below. Mathematically, Equation (13) becomes a telescoping sum of the utility differences from the utility of the outside good up to the utility of the good the consumer most prefers in her consideration set. The utilities for all goods between the outside good and her most preferred good will difference out. Aggregating Lemma 1 into a statement about average demand lines is straightforward and yields the following conclusion.

Theorem 2. *Under quasi-linearity, the average welfare change μ^W that occurs as a result of a change in the platform listing rule from A to B is*

$$\mu^W = \Omega_A - \Omega_B.$$

In words, by looking at the difference in the total welfare created by a change in the listing rule, we can recover the exact average compensating variation. (The exact average compensating variation is also the exact average equivalent variation, since the terms coincide under quasi-linearity.) Of course, under quasi-linearity, the numbering of products 1 to J can be arbitrarily rearranged and the formula still holds.

Theorem 2 captures several key ideas about welfare analysis in a search environment. The first key is the use of prices. By raising a product's price high enough—beyond consumers' reservation prices—we can effectively turn off the value consumers gain from considering that product. The second key is that, under quasi-linearity, the total value consumers gain from their purchases under a given listing rule can be calculated by a sum across all demand lines. Thus, it is not necessary to know each consumer's idiosyncratic consideration set; it suffices to know the aggregate demand curves.

The final key is the importance of the reference product, product 0. Theorem 2 works by building up the utility around the outside product under the different listing rules. If the utility of the outside product is not comparable across listing rules, then we cannot hope to make meaningful welfare comparisons between the two outcomes. If there is greater homogeneity in the content of the consideration set across the listing rules, then a larger reference collection of products can be used and the welfare formula can be simplified, as illustrated in Section 5.

The proof of Theorem 2 can be found in the appendix. However, the following example captures many of the key ideas at work.

Example: Consider a market with a single consumer (y, η, ζ) who considers a single product, product 1, along with the outside product under listing rule A. That is, this consumer has consideration set $\{0, 1\}$ under listing rule A. Under listing rule B, her consideration set grows to $\{0, 1, 2, 3\}$; she gains *two* additional products in her consideration set. For simplicity, assume all prices are 0 and that utility has the following form:

$$u_0(y, \eta) = y$$
$$u_1(y, \eta) = y + a$$
$$u_2(y, \eta) = y + 10a$$
$$u_3(y, \eta) = y + 10a + \epsilon,$$

where a and ϵ are positive.

Then, we see that the consumer's product choice under listing rule A is 1 and her product choice under listing rule B is 3. Her change in utility is $S^W = u_1(y, \eta) - u_3(y, \eta) = -9a - \epsilon$.

Similarly, since $Q_1(y, p, p^m_{-1}, A) = \lim_{p_2, p_3 \to \infty} Q_1(y, p, p_{-1}, B)$ and $Q_2(y, p, p_{-2}, A) = 0 = Q_3(y, p, p_{-3}, A)$ for all price vectors,

$$
\begin{aligned}
\Omega_A - \Omega_B &= -\lim_{p_3 \to \infty} \int_{p_2^m}^{\infty} Q_2(y, p, (0, p_3), B) dp - \int_{p_3^m}^{\infty} Q_3(y, p, p^m_{-3}, B) dp \qquad (14) \\
&= -\int_0^{\infty} 1(9a > p) dp - \int_0^{\infty} 1(\epsilon > p) dp \\
&= -9a - \epsilon,
\end{aligned}
$$

as claimed in Theorem 2.

In the first integral of Equation (14) above, taking the price of product 3 to infinity makes the consumer behave *as if* product 3 were not in her consideration set. This allows us to measure the value of product 2's addition to her initial consideration set of $\{0, 1\}$. The second integral of Equation (14) then takes account of the value of adding product 3 to a consideration set of $\{0, 1, 2\}$. For more intuition, consider Figure 1 and Figure 2.[15]

5 Welfare Results Under Simple Listing Rule Changes

In this section, I look at formulas for measuring the changes in welfare that result from simple changes in the listing rules. I demonstrate that the calculations required in Theorem 2 can be simplified in many situations of practical and counterfactual interest. I consider listing rules that (1) swap a single product in a search-result list with a new product and (2) add or remove a collection of products from a search-result list. While case (2) was already considered in Section 3, the stronger assumption of quasi-linearity allows me to strengthen

[15]Note, if the pricing limit is not included, then we have $\int_0^{\infty} Q_2(y, p, \mathbf{0}, B) dp = 0 < \lim_{p_3 \to \infty} \int_0^{\infty} Q_2(y, p, (0, p_3), B) dp = 9a$. That is, the pricing limits are essential for obtaining the correct welfare conclusions.

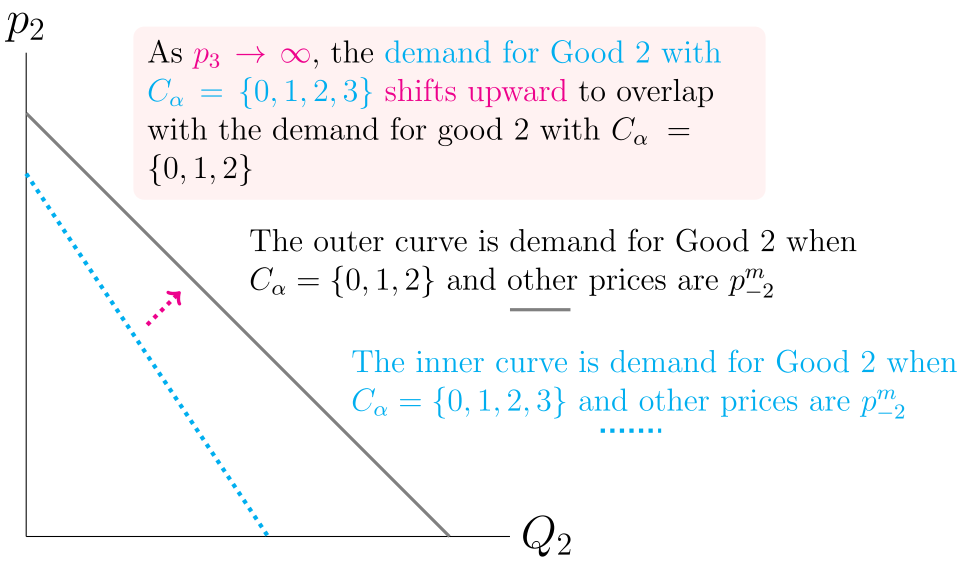


Figure 1: *As the Price of Good 3 Increases, the Demand Curve for Good 2 Shifts Outward Until it is as if no Consumer Considers Good 3.* Both the dotted, blue and solid, black lines show the relation between the average quantity demanded of product 2, Q_2, against the price of product 2, p_2, when the prices of all other products are at market prices. The blue line captures demand under search listing rule B, when the consideration sets are all $\{0, 1, 2, 3\}$, whereas the black line captures the counterfactual demand such that consideration sets are all $\{0, 1, 2\}$. Necessarily, the blue line is (weakly) below the black line at all quantities. If the market price of good 3 went to infinity, then the blue line would shift up to become exactly equal to the black line, as pictured. This figure is the author's own diagram.

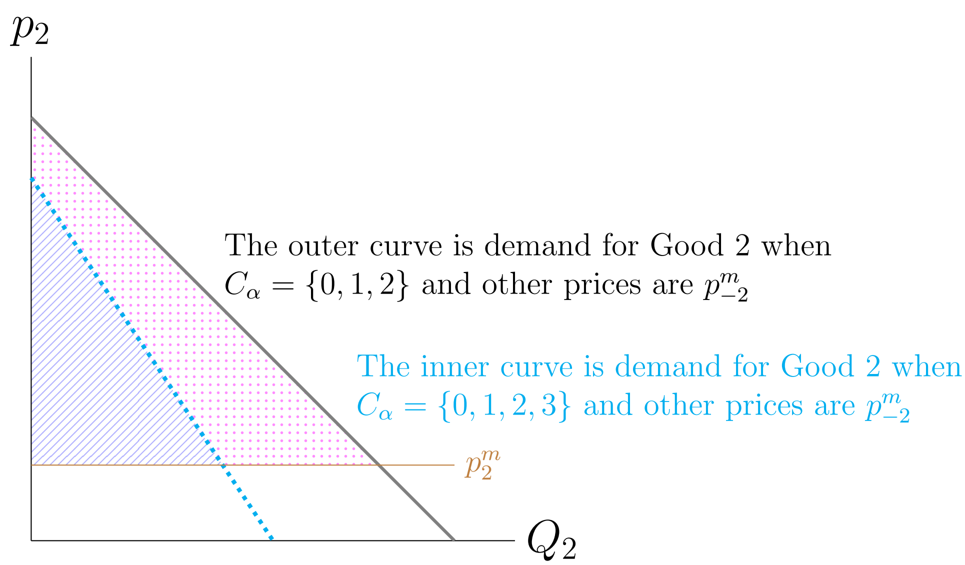


Figure 2: *Calculating the Welfare Benefit of Considering Good 2 and Good 3 Requires A Pricing Limit for Good 3.* The northeast, navy lines indicate the area between the market price for good 2 and the average demand curve for good 2 when all goods are at their market prices. The dotted, blue line is the demand curve for Good 2 when the platform listing rule B is in effect. Under search listing rule B, each consumer's consideration set is $\{0, 1, 2, 3\}$. The solid, black line is the counterfactual average demand for good 2 when all consumers consider $\{0, 1, 2\}$. Under listing rule A, all consumers consider $\{0, 1\}$. The average value that consumers gain in going from listing rule A to B would include the area indicated by the northeast, navy lines and the area indicated by the pink dots. The sum of these areas captures the average welfare gained by consumers who can now consider good 2; that is, the sum captures the gain in going from $\{0, 1\} \rightarrow \{0, 1, 2\}$. To add the welfare gained from going from $\{0, 1, 2\}$ to $\{0, 1, 2, 3\}$, the researcher would add the area under the average demand curve for good 3 when all other goods are constrained to their market prices (not pictured). This figure is the author's own diagram.

the bounds of Theorem 1 into equalities.

5.1 Single Product Swap

In this section, I present measures for the welfare change that results from swapping two goods in the search-result list. That is, if good 1 is being swapped for good M, good M is not included in search-result lists under rule A while good 1 is not included in search-result lists under rule B.

Theorem 3 (Probable Product Swap)**.** *Under quasi-linearity, when the listing rule changes from A to B such that good 1 is swapped for good M, then the average welfare change is*

$$\mu^W = \int_{p_1^m}^{\infty} Q_1(y, p, p_{-1}^m, A)dp - \int_{p_M^m}^{\infty} Q_M(y, p_M, p_{-M}^m, B)dp_M.$$

In Theorem 3, good 1 and good M are not contemporaneously present in any individual's consideration set. The welfare comparison can still be made with the $\mathcal{C}_A \setminus \{1, M\} = \mathcal{C}_B \setminus \{1, M\}$ being the reference collection for each consumer. This is why quasi-linearity is essential for Theorem 3 or any time a change in the search-result list simultaneously increases some products' probability of being shopped while decreasing other products' probability of being shopped. The proof is provided in Section A.5.

5.2 Adding a Collection of Goods to Search-Result Lists

I conclude this section with a formula for measuring welfare changes that result when a collection of goods is added to search-result lists. These results are also applicable when the cost of searching products (weakly) decreases for each consumer. The results for the removal of a collection of goods from the search-result lists are analogous and included in the appendix.

First, for a collection of products $\mathcal{R} = \{r_1, \ldots, r_R\}$, define the *total consumer value of* *products in* $\mathcal{R}$ by

$$\Gamma_\alpha(\mathcal{R}) := \lim_{p_{r_2},\ldots,p_{r_R}\to\infty}\int_{p^m_{r_1}}^{\infty}\Delta Q_{r_1}(y,p,p_{-r_1},\alpha)dp + \lim_{p_{r_3},\ldots,p_{r_R}\to\infty}\int_{p^m_{r_2}}^{\infty}Q_{r_2}(y,p,(p^m_{r_1},p_{-(r_1,r_2)}),\alpha)dp + \cdots + \int_{p^m_{r_R}}^{\infty}Q_{r_R}(y,p,p^m_{-r_R},\alpha)dp.$$

Note that $\Gamma_\alpha(\mathcal{J}) = \Omega_\alpha$.

Then, the following formula can be used to calculate the exact welfare changes that result from changes in the search-result list that weakly increase the probability that each good in $\mathcal{R}$ enters the consideration sets.[16]

Corollary 2. *Let* $\mathcal{R} = \{r_1,\ldots,r_R\}$. *Suppose that a change in the listing rule from A to B increases the probability that each of the products in* $\mathcal{R}$ *is considered. Then, under quasi-linearity,*

$$\mu^W = \Gamma_A(\mathcal{R}) - \Gamma_B(\mathcal{R}).$$

The proof is provided in the appendix.

6 Data Application

6.1 Data Overview

In this section, I estimate welfare changes from listing rule changes for a data set that details the click and purchase behavior of a collection of consumers booking hotels using an online travel agent (OTA). The data is from the 2013 data challenge for the IEEE's International Conference on Data Mining (ICDM).[17] The competition was open to the public through the online data science community, Kaggle. Data for the contest was provided by Expedia.

[16] This result does not allow for the displacement of other goods.

[17] IEEE is the Institute of Electrical and Electronics Engineers. The ICDM is considered the world's premier research conference in data mining. Data challenges are typically held annually.

The data is centered on a collection of *search impressions* that OTA users interacted with, primarily on *Expedia.com*. To understand the term search impression, first consider a consumer searching on *Expedia.com* for vacation accommodations in 2013. This consumer would initially face a page as pictured in Figure 4. Here, the consumer would enter her vacation destination, the days she planned to spend at her vacation destination, the number of rooms she wanted to book and the number of adults and children that she will be traveling with. All this information, pictured in the blue boxes in Figure 4, is collected by Expedia and used to produce a sequence of listings of available hotel rooms. The user is promptly directed to this listing sequence upon entering her information and clicking the button that says "Search for Hotels."

An example of a single hotel listing is given in Figure 5a. The blue boxes identify the information that Expedia collects and that is included in the data set for each hotel listing. Each search is likely to produce several listings. The number of individual listings will vary, depending upon the destination city, the availability of hotel rooms on the given date, and the presence of advertising. In the data, the number of listings on the first page of results varies between 1 and 34. A *search impression* is then defined to be the first page of search results for a given user query. For a large fraction of searches, the number of relevant listings is greater than the length of the search impression. This fact is suggested by the high frequency of long search distributions; see Figure 3.

When a consumer clicks on a listing, a new page opens and provides more details about the property's available rooms. In particular, when a user clicks on a hotel listing, she receives information on the available hotel such as the size of the beds, the parking fees, some pictures of the room interiors, the availability of free breakfast, the room amenities and any hidden fees. An example of this final page is shown in Figure 5b. From here, the consumer may choose to book with the hotel.

The data provides information on the entire first page of the search results each consumer faces. It also tells us the listings each consumer clicked on as well as the listing the consumer eventually booked (if any). Thus, if we assume the products that enter a consumer's consideration set are exactly the products in a consumer's search impression, then we can readily estimate demand. Moreover, Expedia provides us with data from one of their experiments: this data set includes search impressions where the hotel-listing order was determined by Expedia's proprietary ranking algorithm as well as search impressions that resulted in listings being ordered randomly over the pages of results. This provides us with an excellent opportunity to study the welfare consequences of moving from random rankings to Expedia's proprietary ranking system.

While the data set provides an excellent opportunity to study the relationship between prices and consideration sets, a few important caveats should be pointed out. First, search impressions only list the first page of results for each user query. Thus, a consumer who searches beyond the first page of listing results (should there be additional listings) will not have her full consideration set observed in this study.[18] Her behavior on the second page of results would be treated as a separate (unassociated) search impression, if included at all. The same would be true for a consumer who searched over multiple start and end dates. Thus, to the extent that these consumers viewed multiple search-result pages or considered alternative booking dates, the results of this study will underestimate the size of the individual OTA user's consideration sets. Ursu (2017) provides some evidence from a companion data set[19] where more than 40% of *Expedia.com* users only look at the first page of results. Thus, for a large fraction of *Expedia.com* users, it would be reasonable to assume that each consumer has a consideration set that is exactly equal to her search impression. Second, for competitive reasons, Expedia would not verify how representative the sample was. However, Ursu (2017) was able to verify that the Expedia data set was representative of the largest shopping groups on Expedia. Thus, the results should be interpreted as averages for this collection of typical consumers rather than averages over all consumers.

[18]Of course, Expedia would have been able to tell this in their own data. This information had simply been left out of the competition data.

[19]The companion data set is from the Wharton Customer Analytics Initiative and contains several statistics on *Expedia.com* users.

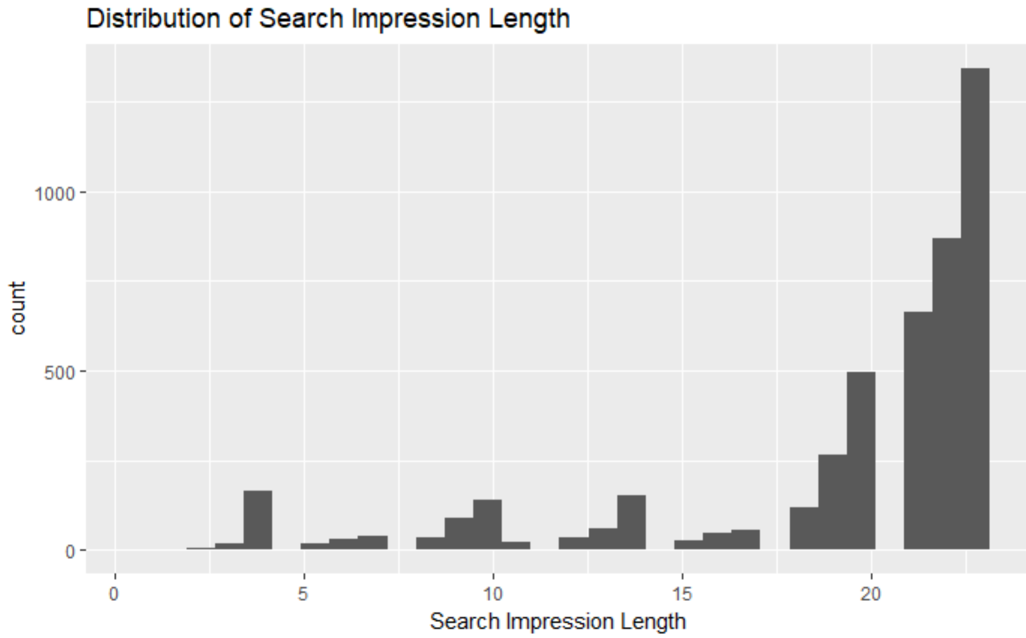


Figure 3: *The Distribution of Search Impression Length After Data Cleaning.* This figure shows the distribution of the length of search impressions over all users and listings in the data provided by Expedia after the author's data cleaning. The data contains a sampling of search impressions of OTA users from 2012 and 2013. The uncleaned data is available to the public on *Kaggle.com*. The author generated this figure using R.

6.2 Demand Estimation Strategy

I fit a model of product choice where, for product j in consumer i's consideration set at time t,

$$u_{ijt} = \alpha(y - p_{ijt}) + \beta' X_j + \eta_{ijt}$$

$$u_{i0t} = \alpha y + \eta_{i0t}$$

Here, η_{ijt} is a standard Type I extreme value distributed random variable that is independent over j and t, given i's consideration set. The vector X_j contains the product characteristics of good j. Income y is not observed but is not needed since it differences out of product decisions. I assumed that each search impression was a unique user and that consideration sets were exactly the products listed in the search impression.[20] For simplicity of analysis, I assumed that the prices and product characteristics are independent of η_{ijt}. In order to ensure my demand parameters could be estimated, I dropped all of the properties that were chosen less than 50 times and all search impressions where these infrequently booked properties were chosen. As there were a small number of observations where the recorded prices were much higher than the actual prices consumers observed (as discussed in Ursu (2017)), I also drop all bookings with a listed nightly price above $1500USD and all search ids that choose a booking with a recorded price over this amount; there are only four of the latter and a handful of the former. In the end, 90,533 rows containing 4,698 search impressions remained. Consideration sets, not including the outside product, have an average length of 19.27 listings and a median length of 21 listings; see Figure 3. The outside product was chosen 38.6% of the time. The average price of a room is $129.29 per night.

I choose components of X_j with the aid of previous studies that looked into the covariates of hotel booking choice in this data set. Following the results of Liu et al. (2013), I included property star ratings, property branding information, a property-location score[21] and an indicator variable for promotions. I ran the regression in R (R Core Team 2017), using the `mlogit` package (Croissant 2019). The results are shown in Table 1. I find that all of the included coefficients are highly significant. As expected for demand, the coefficient on price is negative and highly significant.

6.3 Welfare Change from Random Rankings to Purchase Rankings

As discussed in the data overview, the data contains information over two different listing rules. The first listing rule is a "random" listing rule. Under the random listing rule, the list order is filled in using an *almost* random ordering of products matching consumer's first

[20] These assumptions are supported in a previous study that used this data set (namely, Ursu (2017)).

[21] This is calculated by Expedia, using information about the user's query and the property's locations.

Table 1: *Demand Parameter Estimates and Standard Errors for Online Searches and Bookings of Hotel Rooms in Expedia data set (2012-2013).* This table includes demand parameter estimates for the regression outlined in Section 6.2. Standard errors are listed in parenthesis, below coefficient estimates.

	Dependent variable:
	Hotel Booked
property star rating	0.513^{***}
	(0.048)
property brand boolean	0.418^{***}
	(0.053)
property location score 1	-0.922^{***}
	(0.043)
price in usd	-0.010^{***}
	(0.001)
promotion flag	0.266^{***}
	(0.047)
Observations	4,694
Note:	*p<0.1; **p<0.05; ***p<0.01

page selections. It is almost random because some of these listing positions are reserved for sponsored search impressions.[22] The second listing rule ranks products using a proprietary algorithm known to Expedia. The rule ranks products (at least in part) by their relevance (or probability of purchase). Indeed, the goal of the data challenge was to produce an algorithm that could learn to rank the products in order of their purchase-and-click likelihood.[23]

In order to estimate the change in welfare between the two listing rules, I used demand estimates from the previous section. Given the consideration groups observed over the two listing rules, I could the predict average demand for each group and over all prices. Using

[22]For this paper's study, it is not essential the ranking be perfectly random. The key is that this (imperfectly) random listing rule has welfare consequences that are measurably different from the alternative listing rule.

[23]These algorithms are called LeToR algorithms and the data science community has a literature around them.

Equation (12) and Theorem 2, I found the total welfare under the random-listing rule to be \$104.81 and the total welfare under the proprietary-ranking rule to be \$113.65.[24] Thus, I conclude that welfare was improved by an average of \$8.84 per person when the listings were ordered by the proprietary-ranking rule. Intuitively, this is appealing. While there are valid reasons for concern about the welfare harm that could come from manipulating search-result lists, it is also true that a well-ordered list can improve welfare over a randomly ordered one.

Given our assumption that consumers only look at the first page of search results, price independence will hold reasonably well if, as the price of a good increases, the OTA's listing rule does not remove this good from the first page of search results before most consumers no longer prefer to purchase it. Since the dataset includes ranking information, a simple regression of product ranking on booking price and the other covariates from the earlier demand regression is informative of this relationship. I ran this regression on the bookings that were ordered by the random-ranking rule and then again, separately, on the bookings that were ordered by the proprietary-ranking rule. As expected, the price coefficient in the former case is a weak predictor of ranking. In the latter case, the price coefficient is statistically significant but is not practically significant. For example, in the latter case and for each of the five most popular bookings, I estimate that a three-standard-deviation increase in a good's price increases a product's ranking by less than three positions.[25] This is a small amount, given the first page of search results has space for more than 30 positions. At the same time, the previous section's demand regression predicts that a three-standard-deviation increase in price lowers the quantity demanded for each of the top five products by more than 50%. (See Table 2 for details.) Thus, as the price of a good increases, it is likely to become undesirable before it is removed from the first page of search results, and price independence approximately holds in this environment.

[24] Note that Theorem 2 holds when there is a coefficient on money. This is clear as it is a simple change of variables in the integrals.

[25] If a hotel's ranking increases, it is found farther down the page. A ranking of 1 is the closest to the top of the page and is the first among the search results to be seen by the consumer.

Product	Average Price (Standard Deviation)	Estimated Ranking Change	Estimated Demand Change
A	\$117.76 (\$57.06)	1.16 positions	-82.3%
B	\$71.77 (\$41.03)	.83	-71.0%
C	\$89.55 (\$46.92)	.95	-75.4%
D	\$53.90 (\$33.14)	.67	-62.6%
E	\$107.31 (\$52.04)	1.05	-79.2%

Table 2: *Comparing the Effects of a Price Increase on Predicted Ranking and Predicted Market Share.* This table shows my estimates of the effects of a good's price increase on its own rankings. The values are estimated for a three-standard-deviation price increase from their observed levels. Products A through E above are the same A through E in Table 3 below. That is, these are the five most popular bookings in the Expedia data set. This table also includes my estimates of the effects of a good's three-standard-deviation price increase on its quantity demanded. I estimate that this sharp price increase will only mildly increase its ranking but will decrease its quantity demanded sharply. Thus, price independence holds approximately. Only data from the search impressions that were ranked by the proprietary listing rule were included in this analysis. The data was from 2012 to 2013 and was provided by Expedia.

6.4 Welfare Changes from Removing the Top 5 Products

In this section, I provide estimates of the welfare loss that results from a new listing rule that hides the top five products from the consideration sets. This allowed me to simulate the welfare harm a search platform could cause by suddenly removing certain third-party listings from its search-result lists.[26] The top five products are the products with the largest market share in the sample data. These market shares are listed in Table 3. Together, these firms account for about 20% of the observed bookings. The market is not dominated by any one hotel: even the hotel with the largest market share accounted for less than 6% of total sales in the observed data.

Given my demand estimates from earlier in this section, I calculated welfare using Corollary 2. This calculation amounted to removing product A, then B, then C, et cetera. The results are shown in Table 3. The average welfare lost from the combined removal of all five products is \$20.51 per person. The marginal removal of each additional product reduced

[26]This was a concern that drove the EU Antitrust authorities to fine Alphabet \$2.7 billion in 2017 and there is evidence that Amazon also does this.

consumer welfare by between \$2 and \$6.

Product	Market Share	Estimated Market Share	Estimated Marginal Welfare Loss
A	.05646	.021283	\$2.27
B	.04772	.030695	\$3.37
C	.04261	.043229	\$4.98
D	.04026	.045443	\$5.54
E	.039625	.034399	\$4.34
Total	.2267	.1750	\$20.51

Table 3: *Counterfactual Marginal and Cumulative Welfare Losses from Removing the Top Five Products from the Search-Result Lists.* This table shows my estimates of the marginal welfare loss from removing product A, then B, then C, etc. until E is removed last. Products A through E are the most frequently purchased products in the data set. The table also indicates each product's market share in the data provided by Expedia, as well as the market share predicted by my demand estimation.

7 Conclusion

I have presented several formulas for measuring the changes in consumer welfare that result from an online shopping platform changing the way it lists its search results. Under monotonicity, compensating variation and equivalent variation can be bound with straightforward integrals of aggregate demand. Under quasi-linearity, the exact compensating variation and equivalent variation can be recovered. I have also provided formulas for estimating the counterfactual welfare changes that occur under certain simple listing rule changes. Applications to data that features an OTA's search-result experiment show that ordered listings improve welfare over random listings by an average of \$8.84 per user. They also show that the removal of the five most popular products from search-result lists would lower welfare by an average of \$20.51 per user.

References

Athey, Susan and Glenn Ellison (2011). "Position Auctions with Consumer Search." In: *The Quarterly Journal of Economics* 126.3, pp. 1213–1270. DOI: 10.1093/qje/qjr028. URL: http://dx.doi.org/10.1093/qje/qjr028.

Bhattacharya, Debopam (2015). "Nonparametric Welfare Analysis for Discrete Choice." In: *Econometrica* 83.2, pp. 617–649. DOI: 10.3982/ECTA12574.

Brynjolfsson, Erik, Avinash Collis, et al. (2019). *GDP-B: Accounting for the Value of New and Free Goods in the Digital Economy.* Working Paper 25695. National Bureau of Economic Research. DOI: 10.3386/w25695. URL: http://www.nber.org/papers/w25695.

Brynjolfsson, Erik and JooHee Oh (2012). "The Attention Economy: Measuring the Value of Free Digital Services on the Internet." In: *ICIS*.

Croissant, Yves (2019). *Estimation of multinomial logit models in R: The mlogit Packages.* Version 0.4-2.

Diewert, W. Erwin and Robert Feenstra (2017). *Estimating the Benefits and Costs of New and Disappearing Products.* Microeconomics.ca working papers. Vancouver School of Economics. URL: https://ideas.repec.org/p/ubc/pmicro/tina_marandola-2017-12.html.

Diewert, W. Erwin, Kevin Fox, and Paul Schreyer (2018). *The Digital Economy, New Products and Consumer Welfare.* Economic Statistics Centre of Excellence (ESCoE) Discussion Papers ESCoE DP-2018-16. Economic Statistics Centre of Excellence (ESCoE). URL: https://EconPapers.repec.org/RePEc:nsr:escoed:escoe-dp-2018-16.

Dinerstein, Michael et al. (2017). "Consumer Price Search and Platform Design in Internet Commerce." In: *Working.* http://web.stanford.edu/ leinav/wp/Search.pdf.

Ershov, Daniel (2016). "The Effect of Consumer Search Costs on Entry and Quality in the Mobile App Market." In: *University of Toronto Job Market Paper.* URL: http://www.law.northwestern.edu/research-faculty/searlecenter/events/internet/documents/Ershov_Consumer_Search.pdf.

Feldstein, Martin (2017). "Underestimating the Real Growth of GDP, Personal Income, and Productivity." In: *Journal of Economic Perspectives* 31.2, pp. 145–64. DOI: 10.1257/jep.31.2.145. URL: http://www.aeaweb.org/articles?id=10.1257/jep.31.2.145.

Fradkin, Andrey (2018). "Search, Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb." In: *Working Paper.* URL: https://andreyfradkin.com/assets/SearchMatchingEfficiency.pdf.

Gentzkow, Matthew (2007). "Valuing New Goods in a Model with Complementarity: Online Newspapers." In: *American Economic Review* 97.3, pp. 713–744. URL: https://EconPapers.repec.org/RePEc:aea:aecrev:v:97:y:2007:i:3:p:713-744.

Greenstein, Shane and Ryan McDevitt (2011). "The broadband bonus: Estimating broadband Internet's economic value." In: *Telecommunications Policy* 35.7, pp. 617–632. URL: https://EconPapers.repec.org/RePEc:eee:telpol:v:35:y:2011:i:7:p:617-632.

Groshen, Erica L. et al. (2017). "How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A View from the Trenches." In: *Journal of Economic Perspectives* 31.2, pp. 187–210. DOI: 10.1257/jep.31.2.187. URL: http://www.aeaweb.org/articles?id=10.1257/jep.31.2.187.

Hausman, Jerry (1981). "Exact Consumer's Surplus and Deadweight Loss." In: *American Economic Review* 71.4, pp. 662–76. URL: https://EconPapers.repec.org/RePEc:aea:aecrev:v:71:y:1981:i:4:p:662-76.

— (1996). "Valuation of New Goods under Perfect and Imperfect Competition." In: *The Economics of New Goods.* National Bureau of Economic Research, Inc, pp. 207–248. URL: https://EconPapers.repec.org/RePEc:nbr:nberch:6068.

Honka, Elisabeth (2014). “Quantifying Search and Switching Costs in the US auto insurance industry.” In: *RAND Journal of Economics* 45.4, pp. 847–884.
Honka, Elisabeth, Ali Hortaçsu, and Maria Ana Vitorino (2017). “Advertising, consumer Awareness and Choice: Evidence from the U.S. Banking Industry.” In: *RAND Journal of Economics* 48.3, pp. 611–646.
Khan, Lina M. (2017). “Amazon’s Antitrust Paradox.” In: *Yale Law Yournal* 126.3, pp. 564–907. URL: http://www.yalelawjournal.org/article/amazons-antitrust-paradox.
Lewis, Gregory and Albert Wang (2013). “Who benefits from improved search in platform markets.” In: *SSRN*. URL: http://ssrn.com/abstract=2249816.
Liu, Xudong et al. (2013). “Combination of Diverse Ranking Models for Personalized Expedia Hotel Searches.” In: *CoRR* abs/1311.7679. URL: http://dblp.uni-trier.de/db/journals/corr/corr1311.html#LiuXZYPLSW13.
Moraga-Gonzalez, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest (2017). “Nonsequential Search Equilibrium with Search Cost Heterogeneity.” In: *International Journal of Industrial Organization* 50, pp. 392–414.
Nevo, Aviv (2003). “New products, quality changes, and welfare measures computed from estimated demand systems.” English (US). In: *Review of Economics and Statistics* 85.2, pp. 266–275. ISSN: 0034-6535. DOI: 10.1162/003465303765299792.
Petrin, Amil (2002). “Quantifying the Benefits of New Products: The Case of the Minivan.” In: *Journal of Political Economy* 110.4, pp. 105–130.
Quan, Thomas W. and Kevin R. Williams (2018). “Product variety, across-market demand heterogeneity, and the value of online retail.” In: *The RAND Journal of Economics* 49.4, pp. 877–913. DOI: 10.1111/1756-2171.12255. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/1756-2171.12255.
R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: http://www.R-project.org/.
Seiler, Stephan and Song Yao (2017). “The Impact of Advertising along hte Conversion Funnel.” In: *Quantitative Marketing and Economics* 15.3, pp. 241–278.
Small, Kenneth A. and Harvey S. Rosen (1981). “Applied Welfare Economics with Discrete Choice Models.” In: *Econometrica* 49.1, pp. 105–130. URL: http://www.jstor.org/stable/1911129.
Syverson, Chad (2017). “Challenges to Mismeasurement Explanations for the US Productivity Slowdown.” In: *Journal of Economic Perspectives* 31.2, pp. 165–86. DOI: 10.1257/jep.31.2.165. URL: http://www.aeaweb.org/articles?id=10.1257/jep.31.2.165.
Ursu, Raluca M. (2017). “The Power of Rankings.” In: *Working Paper*. May 6, 2017 version.

Appendix A Proofs

A.1 Proof of Theorem 1

Fix an arbitrary consumer (y, η, ζ) and consumer listing rules A and B. For this proof, I define the following terms. Let

$$f^{\star} := \arg\max_{j \in \mathcal{C}_A} u_j(y - p_j^m, \eta)$$

and let

$$s^{\star} := \arg\max_{j \in \mathcal{C}_B} u_j(y - p_j^m + S^{CV}, \eta).$$

Finally, note that by monotonicity for every $\alpha \in \{ A, B \}$ there is some $\bar{p}_\alpha$ such that

$$\max_{j \in \mathcal{C}_\alpha} u_j(y - p_j^m, \eta) = u_{s^\star}(y - \bar{p}_\alpha, \eta),$$

so

$$u_{s^\star}(y - p_{s^\star}^m + S^{CV}, \eta) = u_{f^\star}(y - p_{f^\star}^m, \eta) = u_{s^\star}(y - \bar{p}_A, \eta),$$

and therefore $p_{s^\star}^m - \bar{p}_A = S^{CV}$ whenever $f^\star \in \mathcal{R}$ (so utility is raised by the platform change) and 0 otherwise.

Now, I prove Equation (8). To that end, I first show that the individual inequalities in Equation (15) below hold for our arbitrary consumer. That is,

$$\begin{aligned} -\int_{p_k^m}^{\infty} q_k(y, p_k, p_{-k}^m, B, \eta, \zeta) dp_k &\geq S^{CV} \\ &\geq -\sum_{j \in \mathcal{R}} \lim_{p_k \to \infty \forall k \in \mathcal{R} \setminus \{ j \}} \int_{p_j^m}^{\infty} q_j(y, p_j, (p_{\mathcal{R} \setminus \{ j \}}, p_{-(R,j)}^m), B, \eta, \zeta) dp_j, \end{aligned} \tag{15}$$

for all $k \in \mathcal{R}$.

To see Equation (15), note that since $\mathcal{C}_A \subseteq \mathcal{C}_B \subseteq \mathcal{C}_A \cup \mathcal{R}$, either $S^{CV} = 0$ or both $S^{CV} < 0$ and $s^\star \in \mathcal{R} \setminus \mathcal{C}_A$. The fact that demand is positive everywhere and that the integrals are over a positive interval shows that the second inequality holds when $S^{CV} = 0$. When $S^{CV} = 0$, it must be the case that for each good $k \in \mathcal{R}$, either this consumer has revealed that she does not prefer the good at market prices—or any prices above market prices, given monotonicity—or she still does not have the good in her consideration set $\mathcal{C}_B$. Either way, the first integral in Equation (15) must be zero for this consumer for each good $k \in \mathcal{R}$ and therefore the first inequality in Equation (15) holds. Thus, it only remains to show that the inequalities in Equation (15) hold in the case where $S^{CV} < 0$.

To see the second inequality in Equation (15) in the case where $S^{CV} < 0$, note that

$$\begin{aligned} S^{CV} &= p_{s^\star}^m - \bar{p}_A && \text{(from above)} \\ &= -\int_{p_{s^\star}^m}^{\infty} \lim_{p_k \to \infty \forall k \in \mathcal{R} \setminus \{ s^\star \}} q_{s^\star}(y, p, p_{-s^\star}^m, B, \eta, \zeta) dp \\ &= -\lim_{p_k \to \infty \forall k \in \mathcal{R} \setminus \{ s^\star \}} \int_{p_{s^\star}^m}^{\infty} q_{s^\star}(y, p, p_{-s^\star}^m, B, \eta, \zeta) dp && \text{(Monotonic Convergence Theorem)} \\ &\geq -\sum_{j \in \mathcal{R}} \lim_{p_k \to \infty \forall k \in R \setminus \{ j \}} \int_{p_j^m}^{\infty} q_j(y, p_j, (p_{\mathcal{R} \setminus \{ j \}}, p_{-(\mathcal{R},j)}^m), B, \eta, \zeta) dp_j. \end{aligned}$$

Next, for the first inequality in Equation (15), note that

$$-\int_{p_k^m}^{\infty} q_k(y, p_k, p_{-k}^m, B, \eta, \zeta) dp_k = \begin{cases} 0 \text{ if } k \neq s^\star \\ -\bar{p}_B + p_{s^\star}^m \text{ if } k = s^\star. \end{cases}$$

Finally note that $\bar{p}_B \leq \bar{p}_A$ since $\mathcal{C}_A \subseteq \mathcal{C}_B$, so

$$p_{s^\star}^m - \bar{p}_B \geq p_{s^\star}^m - \bar{p}_A.$$

which finishes the verification of the inequalities in Equation (15). Adding the pricing limits (as in Equation (8)) to the integral on the left-hand side of Equation (15) reduces the area by switching $\bar{p}_B$ to $\bar{p}_A$ in the above statements. The limits can pass through the integral by the monotonic convergence theorem.

Finally, to get Equation (8) from Equation (15), all we need to do is integrate each part over η and ζ, pass the limit through the integral by the monotone convergence theorem (which is possible by monotonicity) and finally switch the integral order by Tonnelli's theorem (which is possible since demand is nonnegative). Since the left-hand inequality holds for all products in $\mathcal{R}$, even after integrating over all η and ζ, we can take the area-maximizing integral for our bound.

A.2 Welfare Theorems Under Monotonicity in the Case of the Removal of Listings from Search Results

The following theorem bounds the equivalent variation in the case of the removal of a collection of goods from the search-result list.

Theorem 4 (Probable Removal of Goods)**.** *Suppose the listing rule changes exogenously from A to B in a way that makes each consumer unable to consider goods in the collection* $\mathcal{R}$.[27] *Then, the average equivalent variation of this change in listing rule is such that*

$$\max_{j \in \mathcal{R}} \int_{p_j^m}^{\infty} Q_j(y, p_j, p_{-j}^m, A) dp_j \leq \mu^{EV}$$
$$\leq \sum_{j \in \mathcal{R}} \lim_{p_k \to \infty \forall k \in R \setminus \{j\}} \int_{p_j^m}^{\infty} Q_j(y, p_j, (p_{\mathcal{R}\setminus\{j\}}, p_{-\mathcal{R}}^m), A) dp_j. \quad (16)$$

The proof of this result is analogous to the proof of Theorem 1 above.

A.3 Proof of Corollary 1 and Corollary 3

In the case of the addition of a single product that was previously unconsidered or the complete removal of a single product from the search-result lists, the bounds in Theorem 1 and Theorem 4, respectively collapse into an equality. The results of Corollary 1 and Corollary 3 build on this result.

First, I state Corollary 3 and then I prove both.

[27]Specifically, $\mathcal{C}_B \cap \mathcal{R} = \emptyset$ and $\mathcal{C}_B \subseteq \mathcal{C}_A \subseteq \mathcal{C}_B \cup \mathcal{R}$.

Corollary 3 (Probable Removal of a Single Good)**.** *Suppose the listing rule changes exogenously from A to B in a way that makes each consumer less likely to consider product 1.*[28] *Then, the average equivalent variation of this change in listing rule is*

$$\mu^{EV} = \int_{p_1^m}^{\infty} \Delta Q_1^m(y, p_1, p_{-1}^m) dp_1. \tag{17}$$

A.3.1 Proof of Corollary 3

Fix a consumer (y, η, ζ). There are two cases. In case 1, the consumer makes the same good choice under A as under B. In case 2, her choice changes. In case 1, S^{EV} must be zero, and we see

$$\int_{p_1^m}^{\infty} q_1^m(y, p, \eta, \zeta, A) dp - \int_{p_1^m}^{\infty} q_1^m(y, p, \eta, \zeta, B) dp = 0$$

since if 1 is her choice then it must be that $1 \in \mathcal{C}_A, \mathcal{C}_B$ and if 1 is not her choice under A, then it mustn't be under B as well and both integrals are 0. This finishes the proof for case 1.

In case 2, our fixed consumer must have purchased good 1 at time $t = 0$. We know from Theorem 4 that

$$S^{EV} = \int_{p_1^m}^{\infty} q_1^m(y, p, \eta, \zeta, A) dp.$$

Moreover, for this consumer, since $1 \notin \mathcal{C}_B$ (otherwise she would have purchased it),

$$0 = \int_{p_1^m}^{\infty} q_1^m(y, p, \eta, \zeta, B) dp.$$

Thus, for this case as well,

$$S^{EV} = \int_{p_1^m}^{\infty} q_1^m(y, p, \eta, \zeta, A) dp - \int_{p_1^m}^{\infty} q_1^m(y, p, \eta, \zeta, B) dp.$$

Aggregating over consumers and cases and then switching the integration order by Tonelli's theorem yields the desired result.

A.3.2 Proof of Corollary 1

This proof is very similar to the proof for Corollary 3 above and is therefore omitted.

[28] Specifically, I mean for every (y, η, ζ), $\mathcal{C}_B$ is either equal to $\mathcal{C}_A$ or $\mathcal{C}_A \setminus \{ 1 \}$.

A.4 Proof of Theorem 2

Fix an arbitrary consumer (y, η, ζ) and let ω_α be defined by

$$\omega_\alpha := \lim_{p_2^m,\dots,p_J^m \to \infty} \int_{p_1^m}^{\infty} q_1(y, p, p_{-1}^m, \eta, \zeta, \alpha) dp + \lim_{p_3^m,\dots,p_J^m \to \infty} \int_{p_2^m}^{\infty} q_2(y, p, p_{-2}^m, \eta, \zeta, \alpha) dp \tag{18}$$

$$+ \cdots + \int_{p_J^m}^{\infty} q_J(y, p, p_{-J}^m, \eta, \zeta, \alpha) dp. \tag{19}$$

By quasi-linearity and Tonelli's Theorem, it suffices to prove that

$$\omega_\alpha = \max_{j \in \mathcal{C}_\alpha} \left[y - p_j^m + \tilde{U}_j(\eta) \right] - u_0(y, \eta) \tag{20}$$

because then

$$\omega_0 - \omega_1 = \max_{j \in \mathcal{C}_A} \left[y - p_j^m + \tilde{U}_j(\eta) \right] - \max_{j \in \mathcal{C}_B} \left[y - p_j^m + \tilde{U}_j(\eta) \right] = S^W$$

and

$$\Omega_0 - \Omega_1 = \int \omega_0 dF - \int \omega_1 dF = \mu^W.$$

Thus, to show Equation (20), first note that

$$\begin{aligned}
&\lim_{p_2^m,\dots,p_J^m \to \infty} \int_{p_1^m}^{\infty} q_1(y, p, p_{-1}^m, \eta, \zeta, \alpha) dp \\
&\quad = \lim_{p_2^m,\dots,p_J^m \to \infty} \int_{p_1^m}^{\infty} \mathbb{1}(1 \in \mathcal{C}_\alpha) \cdot \mathbb{1}\left[-p_1 + \tilde{U}_1(\eta) > \max_{j \neq 1, j \in \mathcal{C}_\alpha} -p_j + \tilde{U}_j(\eta) \right] dp_1 \\
&\quad = \mathbb{1}(1 \in \mathcal{C}_\alpha) \cdot \int_{p_1^m}^{\infty} \mathbb{1}\left[-p_1 + \tilde{U}_1(\eta) > \tilde{U}_0(\eta) \right] dp_1 \qquad \text{(Monotone Convergence Thm)} \\
&\quad = \begin{cases} -p_1^m + \tilde{U}_1(\eta) - \tilde{U}_0(\eta) & \text{if } \tilde{U}_1(\eta) - p_1^m > \tilde{U}_0(\eta) \text{ and } 1 \in \mathcal{C}_\alpha \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\lim_{p_3^m,\dots,p_J^m \to \infty} \int_{p_2^m}^{\infty} q_2(y, p, p_{-2}^m, \eta, \zeta, \alpha) dp \\
&\qquad = \begin{cases} -p_2^m + \tilde{U}_2(\eta) - (\max_{j \in \mathcal{C}_\alpha \cap \{0,1\}} -p_j^m + \tilde{U}_j(\eta)) \\ \qquad \text{if } -p_2^m + \tilde{U}_2(\eta) > \max_{j \in \mathcal{C}_\alpha \cap \{0,1\}} -p_j^m + \tilde{U}_j(\eta) \text{ and } 2 \in \mathcal{C}_\alpha \\ 0 \text{ otherwise.} \end{cases}
\end{aligned}$$

Continuing this pattern and putting this all together, we see

$$
\begin{aligned}
\omega_\alpha = & \lim_{p_2^m,\ldots,p_J^m \to \infty} \int_{p_1^m}^{\infty} q_1(y, p, p_{-1}^m, \eta, \zeta, \alpha) dp + \lim_{p_3^m,\ldots,p_J^m \to \infty} \int_{p_2^m}^{\infty} q_2(y, p, p_{-2}^m, \eta, \zeta, \alpha) dp \qquad (21) \\
& + \cdots + \int_{p_J^m}^{\infty} q_J(y, p, p_{-J}^m, \eta, \zeta, \alpha) dp \\
= & \mathbb{1}\left(1 = \arg \max_{\mathcal{C}_\alpha \cap \{0,1\}} \tilde{U}_j(\eta) - p_j^m\right) \left[\max_{j \in \mathcal{C}_\alpha \cap \{0,1\}} \tilde{U}_j(\eta) - p_j^m - \left(\max_{j \in \mathcal{C}_\alpha \cap \{0\}} \tilde{U}_j(\eta) - p_j^m\right)\right] \qquad (22) \\
& + \mathbb{1}\left(2 = \arg \max_{\mathcal{C}_\alpha \cap \{0,1,2\}} \tilde{U}_j(\eta) - p_j^m\right) \left[\max_{j \in \mathcal{C}_\alpha \cap \{0,1,2\}} \tilde{U}_j(\eta) - p_j^m - \left(\max_{j \in \mathcal{C}_\alpha \cap \{0,1\}} \tilde{U}_j(\eta) - p_j^m\right)\right] \\
& + \cdots + \mathbb{1}\left(J = \arg \max_{\mathcal{C}_\alpha} \tilde{U}_j(\eta) - p_j^m\right) \left[\max_{j \in \mathcal{C}_\alpha} \tilde{U}_j(\eta) - p_j^m - \left(\max_{j \in \mathcal{C}_\alpha \setminus \{J\}} \tilde{U}_j(\eta) - p_j^m\right)\right].
\end{aligned}
$$

I now conclude the proof of Equation (20) using Equation (21) with induction on $J \in \mathbb{N}$.

Note that for $J = 1$, $\mathcal{C}_\alpha \subseteq \{0, 1\}$ and

$$
\begin{aligned}
\omega_\alpha &= \mathbb{1}(1 = \arg \max_{\mathcal{C}_\alpha}) \left[\tilde{U}_1(\eta) - p_1^m - \tilde{U}_0(\eta)\right] \\
&= \max_{j \in \mathcal{C}_\alpha} \left[U_j(\eta) - p_j^m\right] - \tilde{U}_0(\eta),
\end{aligned}
$$

which proves the base case. Now suppose this holds for the collection of goods $\{0, 1, \ldots, K\}$. Then, for $J = K + 1$,

$$
\begin{aligned}
\omega_\alpha = & \left[\max_{j \in \mathcal{C}_\alpha \setminus \{J\}} \tilde{U}_j(\eta) - p_j^m\right] - \tilde{U}_0(\eta) \qquad \text{(by inductive hypothesis)} \\
& + \mathbb{1}\left(J = \arg \max_{\mathcal{C}_\alpha} \tilde{U}_j(\eta) - p_j^m\right) \left[\max_{j \in \mathcal{C}_\alpha} \tilde{U}_j(\eta) - p_j^m - \left(\max_{j \in \mathcal{C}_\alpha \setminus \{J\}} \tilde{U}_j(\eta) - p_j^m\right)\right] \\
= & \max_{j \in \mathcal{C}_\alpha} \left[\tilde{U}_j(\eta) - p_j^m\right] - \tilde{U}_0(\eta),
\end{aligned}
$$

which concludes the proof.

A.5 Proof of Theorem 3

Fix a consumer (y, η, ζ). I start by showing

$$
S^W = \int_{p_1^m}^{\infty} q_1^m(y, p, \eta, \zeta, A) dp - \int_{p_M^m}^{\infty} q_M^m(y, p, \eta, \zeta, B) dp
$$

In the cases where only good 1 exits this consumer's consideration set, or only good M enters this consumer's consideration set, or her purchase behavior does not change, the result is clear from Corollary 1 and Corollary 3. This leaves only the case where the consumer purchases good 1 under A and purchases good M under B. In this case, $\mathcal{C}_A \setminus \{1\} = \mathcal{C}_B \setminus \{M\}$

Thus,

$$
\begin{aligned}
&\int_{p_1^m}^{\infty} q_1(y,p,p_{-1}^m,\eta,\zeta,A)dp - \int_{p_M^0}^{\infty} q_M^m(y,p,p_{-M}^m,\eta,\zeta,B)dp \\
&\qquad = u_1(y-p_1^m,\eta) - \max_{j\in\mathcal{C}_A\setminus\{1\}} u_j(y-p_1^m,\eta) - \left[u_M(y-p_M^m,\eta) - \max_{j\in\mathcal{C}_B\setminus\{M\}} u_j(y-p_j,\eta)\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Corollary 1, Corollary 3 and quasi-linearity)} \\
&\qquad = u_1(y-p_1^m,\eta) - u_M(y-p_M^m,\eta) \\
&\qquad = S^W.
\end{aligned}
$$

Extending the results from S^W to μ^W proceeds exactly as in the rest of the proofs.

A.6 Proof of Corollary 2 and Analogous Result for Multiple Product Removal from Search-Result Lists

A.6.1 Proof of Corollary 2

This is a corollary of Theorem 2. As discussed in Section A.4, the ordering of the goods does not matter. So, let the goods $1,\ldots,J = 1,2\ldots,r_1,\ldots,r_R$. That is, good $J = r_R$, good $J-1 = r_{R-1}$, $\ldots$, $J-R+1 = r_1$. Then

$$
\begin{aligned}
\Omega_\alpha = &\lim_{p_2,\ldots,p_J\to\infty}\int_{p_1^m}^{\infty} Q_1(y,p,p_{-1},\alpha)dp + \cdots + \\
&\qquad \lim_{p_{r_1},\ldots,p_{r_R}\to\infty}\int_{p_{r_1-1}^m}^{\infty} Q_{r_1-1}(y,p,(p_1^m,\ldots,p_{r_1-2}^m,p_{r_1},\ldots,p_{r_R}),\alpha)dp + \Gamma_\alpha(\mathcal{R}).
\end{aligned}
$$

Since

$$
\begin{aligned}
\Lambda_\alpha := &\lim_{p_2,\ldots,p_J\to\infty}\int_{p_1^m}^{\infty} Q_1(y,p,p_{-1},\alpha)dp + \cdots + \\
&\qquad \lim_{p_{r_1},\ldots,p_{r_R}\to\infty}\int_{p_{r_1-1}^m}^{\infty} Q_{r_1-1}(y,p,(p_1^m,\ldots,p_{r_1-2}^m,p_{r_1},\ldots,p_{r_R}),\alpha)dp
\end{aligned}
$$

is invariant to the inclusion of products $\{r_1,\ldots,r_R\}$ in consideration sets or not—the price limits make Λ_α independent of them—we see $\Lambda_A = \Lambda_B$ and, thus

$$
\Omega_A - \Omega_B = \Gamma_A(\mathcal{R}) - \Gamma_B(\mathcal{R}),
$$

which completes the proof.

A.6.2 Multiple Product Removal

Suppose now that the collection of products $\mathcal{R}$ is removed from search-result listings. Under quasi-linearity, this simply requires reversing the sign on the results in Corollary 2. The only difference is that the roles of A and B are switched.[29] For completeness, I state the results below.

Corollary 4. *Let $\mathcal{R} = \{r_1, \dots, r_R\}$. Suppose that a change in the listing rule from A to B increases the probability that each of the products in $\mathcal{R}$ is considered. Then, under quasi-linearity*

$$\mu^W = \Gamma_A(\mathcal{R}) - \Gamma_B(\mathcal{R}).$$

[29]Indeed, renaming A and B above for their reversed roles gives the proof for Corollary 4.

Appendix B Additional Figures

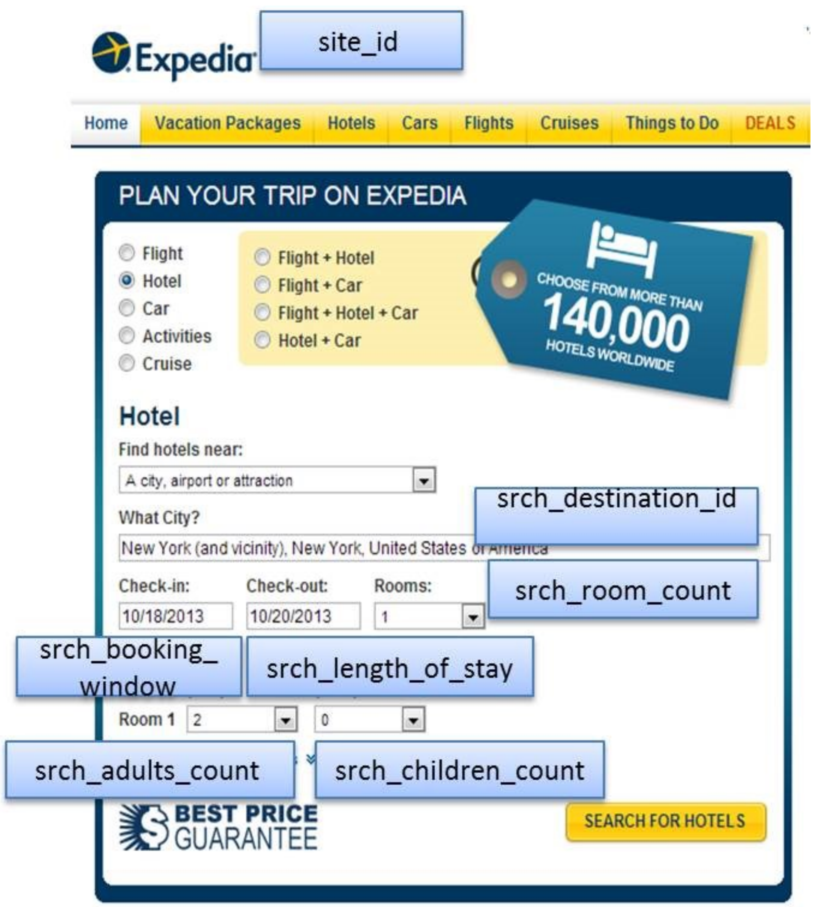
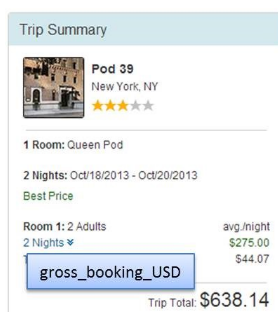


Figure 4: *2012 and 2013 Home Page of Expedia.com.* This is the first page encountered by users of *Expedia.com*. Users select their travel destination, the number of rooms they wish to book, the number of days they wish to spend at the destination, and the number of adults and children who will be staying in the room. All of this information, highlighted by the blue boxes, was collected by Expedia and included in the data set they provided. This figure was provided by Expedia.

Figure 5: *A Search Listing and the Final Booking Page on Expedia.com in 2012 and 2013.*

(a) *A Typical Search Listing on Expedia.com in 2012 and 2013.* This picture shows a typical listing *Expedia.com* users would have observed during the period in which the data was collected. Each search impression contains between 1 and 34 of these listings. This figure was provided by Expedia.



(b) *A Typical Booking Page on Expedia.com in 2012 and 2013.* This is an example of the page a consumer would encounter when finalizing her booking on *Expedia.com* between 2012 and 2013. The consumer would have encountered this page after she clicked on a listing, such as the one shown above in Figure 5a. This figure was provided by Expedia.