DISCUSSION PAPER SERIES

DP16933
(v. 2)

Flexible Demand Estimation with Search Data

Tomomichi Amano, Andrew Rhodes and Stephan Seiler

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

Flexible Demand Estimation with Search Data

Tomomichi Amano, Andrew Rhodes and Stephan Seiler

Discussion Paper DP16933
First Published 20 January 2022
This Revision 25 May 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Flexible Demand Estimation with Search Data

Abstract

Traditional methods for estimating demand are not always well-suited to online markets, where individual products are sold infrequently, unobserved factors such as webpage layout drive substitution, and often only a limited set of product characteristics is observed. We propose a demand model where browsing data—which is abundant in many online settings— is used to infer individual consumers' consideration sets. In our model, the underlying variables which drive consideration can be correlated arbitrarily across products. We estimate the model through a constraint maximization approach, based on the insight that these correlations should rationalize the product-pair co-search frequencies that are observed in the data. In turn, these correlations make it possible to estimate more flexible substitution patterns. We apply the model to data from an online retailer, recover the elasticity matrix, and solve for optimal prices.

JEL Classification: D12, D83, M31

Keywords: Demand estimation, Consideration sets, consumer search

Tomomichi Amano - tamano@hbs.edu
Harvard Business School

Andrew Rhodes - andrew.rhodes@tse-fr.eu
Toulouse School of Economics and CEPR

Stephan Seiler - stephan.a.seiler@gmail.com
Imperial College London and CEPR

Acknowledgements
We thank seminar participants at Boston University, Columbia, Cornell, the European Quant Marketing Seminar (eQMS), Kansai University, Kellogg, Northeastern University, Ohio State, Purdue, UCLA, Washington University St. Louis, and conference participants at the Search and Switching Costs Workshop (2015), IO Fest (2016), Winter Marketing-Economics Summit (2017), Marketing Science (2018), NBER Digitization Summer Institute (2018), Quantitative Marketing and Economics (2018), National Association of Business Economics Conference (2018), and the Toulouse Digital Economics Conference (2021) for useful comments. An earlier version of the paper was circulated under the title “Large-scale Demand Estimation with Search Data”. Please contact Amano (tamano@hbs.edu), Rhodes (andrew.rhodes@tse-fr.eu), or Seiler (stephan.a.seiler@gmail.com) for correspondence. Rhodes acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future program (Investissements d'Avenir, grant ANR-17-EURE-0010). Amano and Seiler did not received external funding for this paper. All errors are our own.

Flexible Demand Estimation with Search Data*

Tomomichi Amano
Harvard Business School

Andrew Rhodes
Toulouse School of Economics & CEPR

Stephan Seiler
Imperial College London & CEPR

This draft: May 25, 2022

Traditional methods for estimating demand are not always well-suited to online markets, where individual products are sold infrequently, unobserved factors such as webpage layout drive substitution, and often only a limited set of product characteristics is observed. We propose a demand model where browsing data—which is abundant in many online settings— is used to infer individual consumers' consideration sets. In our model, the underlying variables which drive consideration can be correlated arbitrarily across products. We estimate the model through a constraint maximization approach, based on the insight that these correlations should rationalize the product-pair co-search frequencies that are observed in the data. In turn, these correlations make it possible to estimate more flexible substitution patterns. We apply the model to data from an online retailer, recover the elasticity matrix, and solve for optimal prices.

Keywords: Demand Estimation, Consideration Sets, Consumer Search

*We thank seminar participants at Boston University, Columbia, Cornell, the European Quant Marketing Seminar (eQMS), Kansai University, Kellogg, Northeastern University, Ohio State, Purdue, UCLA, Washington University St. Louis, and conference participants at the Search and Switching Costs Workshop (2015), IO Fest (2016), Winter Marketing-Economics Summit (2017), Marketing Science (2018), NBER Digitization Summer Institute (2018), Quantitative Marketing and Economics (2018), National Association of Business Economics Conference (2018), and the Toulouse Digital Economics Conference (2021) for useful comments. An earlier version of the paper was circulated under the title "Large-scale Demand Estimation with Search Data". Please contact Amano (tamano@hbs.edu), Rhodes (andrew.rhodes@tse-fr.eu), or Seiler (stephan.a.seiler@gmail.com) for correspondence. Rhodes acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future program (Investissements d'Avenir, grant ANR-17-EURE-0010). Amano and Seiler did not received external funding for this paper. All errors are our own.

1 Introduction

Demand estimation is a cornerstone of quantitative marketing and empirical IO, and the basis for optimal price setting and many other applications. However, online markets pose several challenges for demand estimation. Firstly, purchases are typically sparse, and this makes it difficult to precisely estimate demand parameters. Secondly, important drivers of demand such as webpage layout and product recommendations are not usually observed by researchers. Finally, the number of observed product characteristics is often small relative to the number of available products, which reduces the flexibility of substitution patterns that can be estimated from the data.

In this paper we propose a new and flexible approach to estimating demand in online markets. Our approach leverages data on consumer search (i.e., browsing) behavior, which is typically more abundant than purchase data. These data reveal which products a consumer had in her consideration set, and so provide an additional source of information on substitution patterns. The key idea behind our approach is to directly estimate correlations in product-specific consideration probabilities, rather than explicitly model the factors that drive consideration. In doing this we avoid the risk of mis-specifying the drivers of consideration set formation. We believe this risk is high in online settings, where unobserved factors such as webpage layout and recommendations drive consideration. We use a constrained optimization approach to estimate the large vector of correlations in search between product pairs. Specifically, we exploit the fact that these correlation terms can be estimated by matching predicted consideration probabilities for each pair of products to their empirical counterparts. Because search data is abundant and contains information at the product-pair level, correlations in consideration probabilities can be estimated precisely with sample sizes often available in the context of online markets.

In more detail, our model assumes a two-step process where a consumer first forms her consideration set and then chooses a product from it. Consideration sets are formed based on product-specific inclusion probabilities which are a function of price, product fixed effects, and a set of correlated multi-variate normal errors. Conditional on her consideration set, a consumer chooses which product to purchase based on a standard discrete choice demand model. The correlated error structure in the consideration stage is the key ingredient of our model, because it allows us to flexibly capture co-search patterns between products. We show that these co-search patterns in turn directly influence cross-price elasticities—in particular, substitution between any two products tends to be larger when the two products are searched together more often.

Figure 1 illustrates the role of search data via a simple example with 10 products, denoted by $A, \ldots, J$. Assume focal product A is either searched together with products B, C, D, and E in the top row (illustrated by the solid rectangle) or product F (illustrated by the dashed rectangle). However, product A is never searched together with any other product. Therefore a price increase for any product outside the set of co-searched products (products G, H, I, J) will not affect demand for A, whereas price changes for products B, C, D, E, and F will.[1] Patterns such as the ones

[1]For simplicity, in this example price changes do not alter consideration sets. However in our demand model prices can affect consideration set formation, leading to more complex substitution patterns than those outlined here.

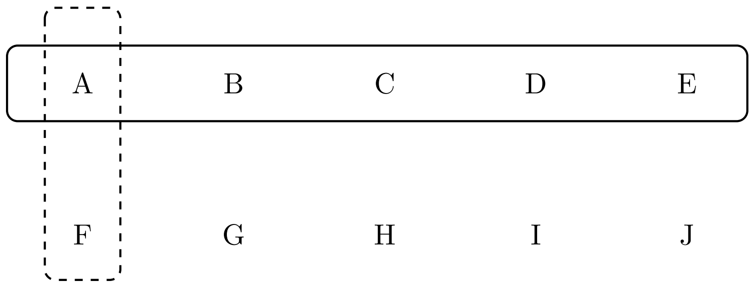

Figure 1: **Illustrative Search Patterns.**

illustrated in Figure 1 could be driven by similarity in product characteristics. For instance, the top and bottom rows could denote respectively high and low quality, and each column could represent a different color. Alternatively, the rows and columns could represent similarity with regards to a characteristic such as product design that is typically unobserved to the researcher or proximity on the webpage. A major advantage of our approach is that we do not need to define the relevant set of characteristics that drive substitutability—because the groupings of substitute products are directly observed in the data. This feature of our modeling approach is particularly valuable in online markets, where the number of products is large relative to the number of observable characteristics, and where other drivers of substitution patterns such as similarity in design or proximity on a webpage are either hard to quantify or are simply not observed by the researcher due to the complexity or novelty of the product.

We apply our estimation approach to search and purchase data from an online retailer selling home-improvement products. We focus on the top 30 products in one popular category and estimate our model based on three months of data that consist of roughly 200,000 product searches and 7,000 purchases by 120,000 users. Using the estimated demand model we compute elasticities and solve for optimal prices. We find that optimal prices differ significantly from current prices—mark-ups are too low for most high-selling products, but too high for some of the products with smaller market shares. Implementing the optimal prices leads to a 23% increase in profits.

Our paper relates to several distinct streams of literature. First, it relates to a literature that employs descriptive methods to uncover and visualize substitution patterns among products. Netzer, Feldman, Goldenberg, and Fresko (2012) use the co-occurrence of products mentioned in online discussion forums, whereas Lee and Bradlow (2011) use customer reviews and the products they mention. The two papers in this realm that are closest to our approach are Ringel and Skiera (2016) and Kim, Albuquerque, and Bronnenberg (2011), which use online search data to analyze competitive market structure and product substitution. However, neither of these papers estimates an elasticity matrix, due to the absence of information on prices and purchases. Instead, these papers provide a visualization of closeness in product space (a “perceptual map”) with an implicit understanding (but no formal derivation) that this visualization informs substitution patterns and hence demand elasticities.[2] Our approach also leverages co-occurrence data, but then also embeds this information into a demand model and uses it to derive an elasticity matrix.[3]

Our paper is also related to models of consideration-set formation and consumer search. Consideration set models are comprised of two separate decision stages: consideration and choice. The consideration stage is typically either modeled as consumers 'passively' becoming aware of products due to external factors such as advertising or product displays (Bronnenberg and Vanhonacker (1996), Mehta, Rajiv, and Srinivasan (2003), Pancras (2010)), or can be understood as a reduced-form approximation of a structural search model. In general, consideration is modeled as a function of observable product characteristics (Andrews and Srinivasan (1995), Bronnenberg and Vanhonacker (1996), Goeree (2008), Draganska and Klapper (2011), Barroso and Llobet (2012), Gaynor, Propper, and Seiler (2016)). Models of consumer search (e.g., Kim, Albuquerque, and Bronnenberg (2010), De Los Santos, Hortacsu, and Wildenbeest (2012), Honka (2014), Chen and Yao (2017), Ursu (2018), Jiang, Chen, Che, and Wang (2021))[4] instead present a unified framework of consumers' utility maximization that rationalizes observed search and purchase patterns. In both types of model, despite different structures and functional forms, observed characteristics determine how often products are considered (or searched) together and hence whether or not they are close substitutes.

We view our approach as combining the strengths of the descriptive approaches outlined above as well as models of consideration and search. Similar to the descriptive papers, we let information on co-occurrence in search inform substitution patterns directly without the need to rationalize co-occurrence through similarity in characteristics. However, by embedding the search information into a model of consideration and choice, we are also able to combine it with information on purchases and price variation, which allows us to estimate the elasticity matrix. In terms of model structure, our approach is most closely related to models of consideration set formation, but it differs from the typical approach because it allows for flexible correlations (rather than modeling substitution as a function of observed factors). We are able to estimate the large vector of correlation terms by setting up a constrained optimization approach which leads to a lower computational burden.

Finally, our paper is related to an emerging literature on flexible demand estimation. Smith, Rossi, and Allenby (2019) use a Bayesian approach to flexibly estimate market partitions using supermarket scanner data. Chiong and Shum (2019) use sparse random projection to reduce the dimensionality of the estimation problem. Ruiz, Athey, and Blei (2017) estimate a sequential probabilistic model of basket demand. Smith and Griffin (2021) develop a Bayesian model of demand that allows the direction and rate of shrinkage to depend on a product classification tree.

[2]Kim, Albuquerque, and Bronnenberg (2011) clearly articulate the implied relationship between the perceptual map and substitution patterns (see p.14): "the map can be used to shed some light on substitution patterns. Local subsets of products on the map can be interpreted as stereotypical products or consideration sets that are searched together and, presumably, compete more intensely."

[3]A related application is by Li, Netessine, and Koulayev (2018), who use search data to compute instrumental variables that are used to estimate the relationship between different firms' pricing decisions.

[4]Other papers within the broader literature of consumer search include Seiler (2013), Koulayev (2014), Pires (2016), Honka and Chintagunta (2017), Choi and Mela (2019), Haviv (2021), and Mojir and Sudhir (2021).

A recent working paper by Armona, Lewis, and Zervas (2021) also uses search data to augment purchase data, with the aim of better estimating substitution patterns. Contrary to our unified approach of using search and purchase data Armona, Lewis, and Zervas (2021) estimate a lower dimensional characteristics vector from search data in a first step and then use those characteristics inside a standard random coefficient demand model in a second step. Donnelly, Kanodia, and Morozov (2022) analyze the impact of product rankings in an online retail setting and implement a latent factorization approach within a model of consideration and choice. Dotson, Howell, Brazell, Otter, Lenk, MacEachern, and Allenby (2018) and Dotson, Beltramo, Feit, and Smith (2019) use variables that characterize the similarity between pairs of products and allow the covariance matrix of a probit choice model to depend on these variables. Our approach similarly leverages data at the product-pair level to inform substitution patterns. However, search data enters our model through the formation of consideration sets rather than the utility function.

The remainder of the paper is structured as follows. Section 2 outlines our approach to estimating demand and discusses how it relates to alternative approaches. Section 3 describes the data and provides descriptive statistics. Section 4 applies the model to data from an online retailer and Section 5 derives optimal prices based on the demand estimates. Section 6 concludes.

2 Model Framework

In this section we outline our modeling framework, which is closely related related to the "consideration-then-choice" approach of the consideration set literature. We note that, despite the similarities with this literature, we use the terms "search" and "consideration" interchangeably below when describing the products that consumers inspect before making a purchase.

We assume that consumers form consideration sets based on product-specific "inclusion probabilities". Specifically, we denote by Pr_{ij}^{Search} the probability that consumer i includes product j in her consideration set. We also assume that the outside option is always included in the consideration set. The probability of a specific consideration set s occurring is then given by multiplying the relevant product-specific inclusion (and exclusion) probabilities:

$$Pr_{is}^{Cset} = \prod_{k \in s} \left[Pr_{ik}^{Search} \right] \times \prod_{l \notin s} \left[1 - Pr_{il}^{Search} \right]. \quad (1)$$

We assume that product-specific inclusion probabilities take the form

$$Pr_{ij}^{Search} = Pr(v_{ij} > 0),$$

where v_{ij} is a function that determines consideration set formation and is given by

$$v_{ij} \quad = \quad \bar{v}_{ij} + \nu_{ij} = \gamma_{price} \times price_j + \bar{\gamma}_j + \tilde{\gamma}_{ij} + \nu_{ij}. \quad (2)$$

The sensitivity of consideration to price is captured by γ_{price}, while $\bar{\gamma}_j$ is a vector of product intercepts that is common across consumers. The vector $\tilde{\gamma}_{ij}$ denotes individual-specific parameters that are allowed to be correlated. We assume $\tilde{\gamma}_{ij}$ is drawn from a multi-variate normal distribution centered at zero with the following variance-covariance matrix

$$\Omega = \begin{bmatrix} 1 & \sigma_{12} & \cdots & \sigma_{1J} \\ \sigma_{12} & 1 & & \sigma_{2J} \\ \vdots & & \ddots & \vdots \\ \sigma_{1J} & \sigma_{2J} & \cdots & 1 \end{bmatrix}. \tag{3}$$

We normalize the diagonal terms to one, which serves a function similar to the scale normalization in choice models. A key aspect of our model is that we allow for a fully flexible covariance structure, subject only to (3) constituting a valid (i.e., positive semi-definite) variance-covariance matrix. We show below that with rich search data we can estimate this matrix even in settings with a relatively large number of product pairs. In our empirical application we estimate a model in a setting with 30 products and therefore 435 product pairs. We assume ν_{ij} is logistic and iid across consumers and products, which yields the following expression for the product-specific inclusion probabilities:

$$Pr_{ij}^{Search} = \frac{\exp(\bar{v}_{ij})}{1 + \exp(\bar{v}_{ij})}.$$

Choice conditional on consideration is determined by the utility consumer i obtains when purchasing a product j from her consideration set (or the outside option):

$$\begin{aligned} u_{ij} &= \bar{u}_{ij} + \varepsilon_{ij} = \alpha_{price} \times price_j + \bar{\alpha}_j + \varepsilon_{ij} \\ u_{i0} &= \varepsilon_{i0}, \end{aligned} \tag{4}$$

where α_{price} is the price coefficient, and $\bar{\alpha}_j$ denotes a product fixed effect. We assume ε_{ij} follows an extreme value distribution. All inside goods are contained in one nest with correlation parameter $(1 - \lambda)$, while a second nest contains only the outside option. This error structure is important—since we want to model optimal price setting by a multiproduct retailer, the degree of substitution towards the outside option relative to all inside goods is crucial. We let $Pr_{ij|s}$ denote the probability that consumer i chooses product j from consideration set s, which by the usual logit formula is

$$Pr_{ij|s} = \frac{\exp(\bar{u}_{ij})}{1 + \sum_{k \in s} \exp(\bar{u}_{ik})}.$$

Given the above model of consideration-then-choice, the probability of purchasing product j is

$$Pr_{ij} = \sum_{s \ni j} Pr_{is}^{Cset} \times Pr_{ij|s} = \sum_{s \ni j} \left\{ \prod_{k \in s} \left[\frac{\exp(\bar{v}_{ik})}{1 + \exp(\bar{v}_{ik})} \right] \times \prod_{l \notin s} \left[\frac{1}{1 + \exp(\bar{v}_{il})} \right] \right\} \times \left\{ \frac{\exp(\bar{u}_{ij})}{1 + \sum_{k \in s} \exp(\bar{u}_{ik})} \right\}.$$

Aggregate demand can be computed by integrating the above expression over the distribution of the correlated consideration error terms $\tilde{\gamma}_{ij}$.

We note that many models of consideration set formation in the prior literature take a similar form (e.g. Bronnenberg and Vanhonacker (1996), Swait and Erdem (2007), Goeree (2008)) to our model. However our approach differs from previous work in two important ways: i) we allow for price to impact consideration set formation, and ii) we allow for a flexible correlation structure in the function that determines product-specific inclusion probabilities. As we show in the next two subsections, the latter aspect of our model allows us to flexibly model cross-price elasticities as a function of co-search probabilities. Structural search models that derive search and choice based on an underlying utility function typically allow price to affect search and purchase decisions, but constrain the impact to be driven by a single price coefficient in the utility function. More importantly, search models tend to define utility as a function of observables and therefore do not allow for the type of flexible correlations that we capture in the consideration set formation process. We return to a comparison of our approach with other models of consideration set formation, as well as with search models, in Section 2.5.

2.1 Own- and Cross-Price Effects

We now examine own- and cross-price effects, and show how they are influenced by search patterns.

We begin with cross-price effects. An increase in the price of product k increases the demand for product j through two channels. Firstly, fixing consumers' consideration sets, amongst consideration sets that feature both products j and k, consumers become more likely to choose product j. Secondly, some consumers will drop product k from their consideration set, increasing the probability they buy other products (including product j). Hence one can write

$$\frac{\partial Pr_j}{\partial p_k} = \int \left[\sum_{s \ni j} Pr_{is}^{Cset} \times \frac{\partial Pr_{ij|s}}{\partial p_k} + \sum_{s \ni j} \frac{\partial Pr_{is}^{Cset}}{\partial p_k} \times Pr_{ij|s} \right] dF(\tilde{\gamma}), \tag{5}$$

where the first and second terms in the square-brackets are respectively the first and second channels described above. After some manipulations that we detail in Appendix A, we can rewrite (5) as

$$\frac{\partial Pr_j}{\partial p_k} = \int \sum_{s \ni j,k} Pr_{is}^{Cset} \left[\frac{\partial Pr_{ij|s}}{\partial p_k} - \gamma_{price}(1 - Pr_{ik}^{Search})(Pr_{ij|(s \backslash k)} - Pr_{ij|s}) \right] dF(\tilde{\gamma}). \tag{6}$$

The first term inside square-brackets reflects how a change in p_k affects demand for product j when consideration sets are fixed. For brevity we do not write it out, but $\partial Pr_{ij|s}/\partial p_k$ follows from the standard logit expression. The second term in square-brackets reflects the impact of a change in p_k on demand for product j through the impact that the price change has on consideration sets. Intuitively, as p_k increases some consumers no longer have product k in their consideration set. For those consumers the probability of choosing product j increases by $Pr_{ij|(s \backslash k)} - Pr_{ij|s}$. Importantly, the square-bracketed term is positive (since $\gamma_{price} < 0$) and is multiplied by the probability for consideration sets that involve both products.[5] Therefore $\partial Pr_j/\partial p_k$ is larger, other things equal, when products j and k are searched together more often. Consequently, by allowing for flexible correlations in product-level consideration probabilities, our model allows for flexible substitution patterns.

Similarly the own-price effect is given by

$$\begin{aligned}\frac{\partial Pr_j}{\partial p_j} &= \int \left[\sum_{s\ni j} Pr_{is}^{Cset} \times \frac{\partial Pr_{ij|s}}{\partial p_j} + \sum_{s\ni j} \frac{\partial Pr_{is}^{Cset}}{\partial p_j} \times Pr_{ij|s}\right] dF(\tilde{\gamma}) \\ &= \int \sum_{s\ni j} Pr_{is}^{Cset}\left[\frac{\partial Pr_{ij|s}}{\partial p_j} + \gamma_{price} \times (1 - Pr_{ij}^{Search}) \times Pr_{ij|s}\right] dF(\tilde{\gamma})\end{aligned}$$

which can again be decomposed into a choice effect and a consideration effect.

2.2 Illustrative Example

In this subsection we illustrate the relationship between model parameters, co-search probabilities, and cross-price elasticities using a simple version of the framework outlined above. We assume that consumers can choose from 4 products (and an outside option), and the utility function that determines choice conditional on consideration is given by

$$u_{ij} = 2 - 0.2 \times price_j + \varepsilon_{ij}.$$

The function determining consideration is given by equation (2) from earlier. We set the product intercepts to $\{0.4, 0.3, 0.2, 0.1\}$ to allow for a vertical element in the consideration set formation process, such that some products are more likely to be considered than others. The price coefficient in the consideration stage is set to $\gamma_{price} = -0.2$. We set the variance-covariance matrix governing the distribution of $\tilde{\gamma}_{ij}$ to the values displayed in the top panel of Table 1, thus allowing for a positive correlation in inclusion probabilities between products 1 and 2 as well as between products 3 and 4. We also set the variance of ν_{ij} to a small value in order for consideration to be driven to a large extent by the other components of v_{ij}.[6] We set prices to 1 for all products and evaluate elasticities at those prices.[7]

In the remaining two panels of Table 1 we report search and co-search probabilities as well as the elasticity matrix that result from the data-generating process outlined above. The middle panel displays the product-specific consideration probabilities on the diagonal and joint-search probabilities

[5]To understand why the cross-price effect in the consideration stage depends on Pr_{is}^{Cset}, note that the impact of product k being dropped from a consideration set is more important for products that are frequently considered together with product k.

[6]Specifically, we set the variance of the ν_{ij} to $0.1 \times (\pi^2/6)$.

[7]In this section we are not concerned with inferences and therefore we do not need to induce variation in prices. We merely aim to illustrate search patterns and elasticities generated by the model.

Variance-Covariance Matrix				
	1	2	3	4
1	1	0.7	0	0
2	0.7	1	0	0
3	0	0	1	0.7
4	0	0	0.7	1
Search / Co-Search Probabilities				
	1	2	3	4
1	0.577	0.427	0.289	0.266
2	0.427	0.537	0.269	0.247
3	0.289	0.269	0.501	0.348
4	0.266	0.247	0.348	0.460
Elasticities (of Row Demand w.r.t Column Price)				
	1	2	3	4
1	-1.980	1.033	0.489	0.408
2	1.094	-2.073	0.491	0.423
3	0.586	0.476	-2.185	0.936
4	0.554	0.484	1.114	-2.225
0	0.083	0.074	0.069	0.061

Table 1: **Motivating Example.** Diagonal elements in the search panel show the search frequency for individual products. Off-diagonal elements show the frequency of the row/column-pair of products being searched together. Model parameters are set as follows: utility parameters: $\bar{\alpha}_j = 2 \forall j$, $\alpha_{price} = -0.2$, $\lambda = 0.05$, consideration parameters: $\bar{\gamma}_j = \{0.4, 0.3, 0.2, 0.1\}$, $\gamma_{price} = -0.2$.

of product pairs on the off-diagonal. For product pairs without correlated inclusion probabilities, the co-search probability is simply equal to the product of the product-specific probabilities. For example, the joint search probability of products 1 and 3 is equal to $0.577 \times 0.501 = 0.289$. For product pairs with positively correlated inclusion probabilities, the co-search probability is higher than the product of the product-specific probabilities. For example, the joint search probability of products 1 and 2 is equal to $0.427 > 0.577 \times 0.537$. Differences in the product-specific consideration probabilities displayed on the diagonal reflect differences in the product intercept terms $\bar{\gamma}_j$.

Finally, the comparison of the second and third panels illustrates how the correlations in co-search probabilities affect cross-price elasticities. Consistent with the formula derived in the previous section, demand for product 1 is most strongly affected by the price of product 2 i.e., the product with which it is most frequently considered. The same pattern holds with regards to products 3 and 4, which are close substitutes for each other due to the higher co-search probability. We re-iterate that we see the close relationship between co-search probabilities and elasticities as a key advantage of our approach. By allowing for a flexible relationship in co-search probabilities via correlations in $\tilde{\gamma}_{ij}$, we indirectly allow for a flexible elasticity structure that is partly determined by co-search patterns.

2.3 Correlation in Consideration Set Formation: Microfoundation

As outlined above, we model consideration set formation flexibly, without trying to rationalize why certain pairs of products are more or less likely to be searched together. In this section we lay out a simple microfoundation for these differing co-search probabilities, based on how products are presented on a retailer's webpage.[8]

Consider the setting from our simple example in the previous subsection. Let $\tilde{\gamma}_{ij}$ be a measure of how prominently product j is displayed to consumer i. Suppose there is an algorithm which determines prominence, using a score $\tilde{\gamma}_{ij} = \tilde{\gamma}^1_{ij} + \tilde{\gamma}^2_{ij}$ where $\tilde{\gamma}^1_{ij}$ and $\tilde{\gamma}^2_{ij}$ are independent normally distributed random variables. The first term $\tilde{\gamma}^1_{ij}$ reflects noise in the algorithm; it has mean 0 and variance 0.3. The second term $\tilde{\gamma}^2_{ij}$ reflects the algorithm's attempt to display relevant products to a consumer; it has mean 0 and variance 0.7. Suppose also that the algorithm deems products 1 and 2 to be relevant for some set of customers and products 3 and 4 for a different set of customers —meaning that $\tilde{\gamma}^2_{i1} = \tilde{\gamma}^2_{i2}$ and $\tilde{\gamma}^2_{i3} = \tilde{\gamma}^2_{i4}$. Suppose further that $\tilde{\gamma}^2_{ik}$ and $\tilde{\gamma}^2_{il}$ are uncorrelated, for $k = 1, 2$ and $l = 3, 4$. Such a situation would give rise to a covariance between products 1 and 2 (as well as products 3 and 4) of $Cov(\tilde{\gamma}_{i1}, \tilde{\gamma}_{i2}) = Cov(\tilde{\gamma}^2_{i1}, \tilde{\gamma}^2_{i2}) = 0.7$, whereas the covariance between all other product pairs is zero. Finally, the variance terms are equal to 1 for all products and hence the variance-covariance structure is the same as the one outlined in the previous subsection.

Our example easily extends to the general case of markets with more products and different covariance structures. While we referred above to consideration being driven by an algorithm, one can think more generally of consideration being driven by different aspects of webpage layout. For example, two products might often be considered together because in response to certain search queries they tend to appear together, or because a consumer uses a specific category tag or filtering option that places both products next to each other. Finally, when browsing one product, the webpage might recommend other similar products which will lead to higher co-search probabilities. We think of these various elements of webpage navigation as being captured by the covariance structure of the consideration set formation process. As this discussion illustrates, correlation in consideration can arise for a variety of reasons. Furthermore, many aspects of webpage layout such as product recommendations and the usage of sorting and filtering tools are often not observed by researchers, and would be difficult to model and characterize even if they were observed.

[8] In Goeree (2008) advertising is a key driver of consideration set formation. Webpage layout plays a similar role in our setting. We posit that the presentation of products on the retailer's webpage drives not only how likely a product is to be considered, but also which products it is likely to be considered alongside.

2.4 Estimation

We estimate both stages of the model by maximum likelihood. The two stages can be estimated separately because they do not share any parameters. The conditional choice stage resembles a typical full-information choice model, where the set of available options is given by the observed consideration set. Estimation for this part of the model is standard, and we therefore focus on outlining estimation of the parameters governing the consideration stage, in particular the large vector of covariance terms.

We re-iterate that product-specific inclusion probabilities are given by

$$Pr_{ij}^{Search} = \frac{\exp(\bar{v}_{ij})}{1+\exp(\bar{v}_{ij})}, \text{ where } \bar{v}_{ij} = \gamma_{price} \times price_j + \bar{\gamma}_j + \tilde{\gamma}_{ij} + \nu_{ij},$$

and the (log) likelihood function is equal to

$$LL_{Cset} = \sum_i \sum_s \log(Pr_{is}^{Cset}) \times y_{is},$$

where y_{is} is equal to one for the consideration set that is observed for consumer i in the data, and Pr_{is}^{Cset} is the product of product-specific inclusion and exclusion probabilities as described earlier.

Product-specific terms $\bar{\gamma}_j$ play the role of matching product-specific inclusion probabilities to their empirical counterparts, i.e. the frequency at which a given product is considered (regardless of other products in the consideration set). Similarly, the covariance terms governing the distribution of $\tilde{\gamma}_{ij}$ play the role of matching pair-specific joint-consideration probabilities to their empirical counterparts. Therefore both sets of parameters can be obtained by solving the following system of equations

$$\begin{aligned} s_j &= \int Pr_{ij}^{Search} dF(\tilde{\gamma}_{ij}) \quad \forall j \in J \qquad (7) \\ s_{jk} &= \int \left[Pr_{ij}^{Search} \times Pr_{ik}^{Search} \right] dF(\tilde{\gamma}_{ij}) \quad \forall j \in J \text{ and } \forall k > j, \end{aligned}$$

where the terms on the left-hand sides of the expressions denote, respectively, search frequencies for each product, and co-search frequencies for every unordered pair of products. These equations can be used to solve for J product-specific terms $\bar{\gamma}_j$ and for $J(J-1)/2$ covariance terms σ_{jk}.

The role of product intercepts in the consideration stage is analogous to the role of product fixed effects in discrete choice demand models, which match predicted market shares to their empirical counterparts. The product-specific terms $\bar{\gamma}_j$ that enter search probabilities similarly match predicted search frequencies to those observed in the data. Given a set of product-specific parameters, we can apply a similar logic to the covariance parameters. To build intuition, consider the case where the covariance between two products' inclusion probabilities is equal to zero. In this scenario the joint-consideration frequency s_{jk} is equal to the product of the product-specific frequencies $s_{jk} = s_j \times s_k$. Hence if the observed joint-consideration frequency is larger than the product of the product-specific frequencies, a positive covariance between the two products' inclusion probabilities is required to rationalize what is observed in the data. The joint-consideration probabilities for a given pair only depend on the covariance term specific to the pair (as well as the product-specific $\bar{\gamma}_j$ terms which are determined by the product-specific search frequencies). Hence, each covariance term is equal to the value that matches the predicted joint-consideration probability for a product pair to its empirical counterpart.

We set up estimation as an optimization problem where the conditions in (7) act as constraints. This formulation mirrors the idea of estimating product fixed effects through a constraint optimization procedure in a standard choice model (Dubé, Fox, and Su (2012)). We apply the same idea to both individual products' search shares as well as co-search frequencies.

2.5 Discussion & Relationship to Other Approaches

Our approach closely follows the structure of several papers in the consideration set literature such as Bronnenberg and Vanhonacker (1996), Swait and Erdem (2007), Goeree (2008), and Barroso and Llobet (2012). The models proposed in these papers all have product-specific inclusion probabilities that result in the consideration set probability expression in equation (1).[9] Choice conditional on consideration is then modeled as a function of a product-specific utility function, where consumers are assumed to choose only from the subset of products they consider. The approach we propose in this paper shares all of these features with the prior literature, but departs from it in two important ways: i) we allow price to influence consideration set formation, and ii) we model correlations in product-specific inclusion probabilities by estimating the covariance matrix of the correlated shocks $\tilde{\gamma}_{ij}$ which enter each product's inclusion probability. Allowing price to impact consideration is appropriate in our setting because consumers observe the price of a product before visiting its product detail page. Correlations in inclusion probabilities allow us to estimate cross-price elasticities flexibly, and in a way that is directly informed by co-search patterns. To the best of our knowledge this type of flexibility with regards to correlation in consideration probabilities has not been allowed for in similar models.

We are able to allow for such flexibility because the typical dataset of online browsing and purchase behavior contains rich information on co-search patterns between products. We show in the next section that in our data the average co-search frequency across all product pairs is larger than the purchase frequency of the average product. Moreover, the insight that correlations in product-specific inclusion probabilities can be estimated by imposing that estimated co-search probabilities match their empirical counterparts allows us to estimate a large set of covariance terms without a large increase in computational burden.

[9] A notable exception is Draganska and Klapper (2011) who directly model the probability that each consideration set occurs rather than specifying product-specific inclusion probabilities.

Flexible Correlations and Unobserved Drivers of Consideration

The flexibility of our approach with regards to modeling co-search patterns is important because it allows us to capture the influence of unobserved factors on search patterns. As outlined in Section 2.3, the role of webpage layout is likely to be an important factor that influences which products are searched together (either because certain products are displayed together, or because viewing one product causes another similar product to be recommended). By estimating correlations flexibly across products—rather than modeling consideration as a function of observed product characteristics—we are able to account for the influence of external factors such as webpage layout that are not captured in most standard search and consideration models.[10]

It is easy to extend our approach to a setting where the researcher has some information on how products are presented. For example, it is straightforward to include variables such as product rankings in the function that drives consideration. Alternatively, one could estimate the model outlined above and then, in the second step, project the estimated product intercepts ($\bar{\gamma}_j$) and correlation parameters (σ_{jk}) of the consideration stage on observed aspects of the webpage such as products rankings and product recommendations. We focus on a model without observable factors that impact consideration because many elements of webpage design (such as recommendation or joint-presentation of two products on the homepage of a retailer) are typically not observed by researchers.

Because our approach does not directly model the drivers of co-search, we need to be careful when conducting counterfactuals that might alter the underlying drivers of consideration and therefore the consideration probabilities. For example, a change in webpage design such as a change in the ranking algorithm would likely lead to changes in the parameters of the consideration stage. As we explained above, if information on rankings and other relevant factors is available, then the impact of these factors and a specific change in the algorithm could be modeled. The set of counterfactuals that can be analyzed is therefore not constrained by the nature of our modeling approach, but rather by the availability of data on relevant drivers of consideration. In our empirical application, we focus on estimating price elasticities and study optimal price setting by the retailer. Pricing counterfactuals do not depend on other drivers of consideration and can therefore be conducted without explicitly modeling the drivers of consideration.

Consideration Sets versus Structural Search Model

An alternative approach to modeling consideration and choice would be a structural search model, where both search and choice are driven by the same underlying utility function. We opt for the two stage "consideration-then-choice" framework for several reasons. In the typical structural search model, selection into different consideration sets will depend on consumer preferences.[11] However, a

[10]Because utility is usually defined at the product level, factors that guide a consumer from one product to another (e.g., recommendations) are hard to include, because they are specific to pairs of products.

[11]For example, if consumers tend to search products with similar price levels, a search model will rationalize this behavior through heterogeneity in preferences over price. Such heterogeneity would induce high (low) price sensitivity consumers to search sets of products with low (high) prices.

search model will not be able to rationalize a situation where two products do not share any observable characteristics and yet are frequently searched together. Our approach, by contrast, directly estimates correlations in product-specific inclusion probabilities and can therefore capture co-search patterns that are not driven by similarity in observed product characteristics. Our modeling approach is therefore preferable in a setting where factors that drive co-search (unobserved product characteristics and external factors such as webpage layout) are unobserved to the researcher. A second benefit of our approach is computational. Structural search models are typically computationally burdensome, and thus are usually estimated for small numbers of products and without flexible correlations in utility between products. Scaling a structural model of search to a larger assortment size would be computationally more demanding than our approach.

A key assumption that distinguishes the consideration set approach from a structural search model is the assumption that consideration and choice can be modeled as two separate stages. This assumption is arguably more appropriate in settings where external factors such as webpage layout are important drivers of consideration. The typical structural search model à la Weitzman (1979) instead assumes that consumers are fully informed about the characteristics of all products except for an idiosyncratic "match value", and actively decide which products to consider. We believe that a passive model of consideration set formation is likely to be more appropriate in many online retail settings where consumers rely on ranking algorithms and recommendations and where consumers have relatively little information about products before visiting the product detail page.

3 Data and Descriptive Statistics

We estimate the model using data from an online retailer that sells home-improvement products (hardwood flooring, tiles, etc.). We focus on a single product category, which is one of the largest sold by the retailer, and estimate demand for the top 30 highest-selling products in that category.[12] We observe the entire history of consumers' search and purchase behavior during a 13-week period from April 20 2016 to July 16 2016. The final dataset contains 201,363 searches and 7,264 purchases (basket additions) by 118,992 distinct users. A consumer is considered to have searched a product if she accessed the product description page. We treat basket additions as the choice outcome in the demand model, because the retailer did not store purchase information in a way that was easily accessible to us. (Basket additions were, however, tracked as part of the browsing data.)[13] We assume the rate of conversion from basket additions to purchases does not vary across products and is unaffected by price. Under this assumption the conversion rate simply scales up demand, and hence has no effect on estimated elasticities or optimal prices. Table 2 provides some descriptive statistics. Roughly 6.14% of all search sessions in this category end in a "purchase." On average a search session contains 1.68 products.

Each product belongs to one of four different brands and we observe the weekly price for each product that is posted on the webpage. In addition to prices, we also observe whether a product

[12]The top 30 products in terms of market share account for roughly 60% of total sales in the category.
[13]For simplicity we refer to basket additions as "purchases" going forward.

Panel A:

	Purchase Probability	6.14%	
	Average # of Searches	1.68	
			Percentage
	Number of Searches	1	65.39
		2	18.13
		3	8.04
		4	3.91
		5	2.13
		6	1.16
		7	0.57
		8	0.32
		9	0.16
		≥10	0.19

Panel B:

			Product-pair Co-search Patterns		
	Average Number of Purchases	Average Number of Searches	Number of Product Pairs	Average Number of Searches	Number of Zeros
Top 10 products	458	10,744	45	1271	0
Top 20 products	314	8,404	190	646	0
Top 30 products	242	6,684	435	389	0

Table 2: **Descriptive Statistics: Purchase & Search Patterns.**

is "on deal," i.e., is on sale in a given week and is highlighted as such on the website. We also observe five physical product characteristics (two of them are discrete and three are continuous) that we are not allowed to disclose. Finally, we have data on the star rating of each product based on customer reviews. The review score is reported on a scale from 1 to 5.[14]

In the bottom panel of Table 2, we document search and purchase patterns at the product level and describe the distribution of consideration sets that occur in the data. The first two columns report the average number of purchases and searches separately for the top 10, 20, and 30 products in the assortment by market share. As mentioned earlier, we find that products are searched significantly more frequently than they are purchased. The ratio of searches to purchases is roughly 30 to 1 for the top 30 products, for example. We also document that most pairs of products are searched together (possibly with other products) relatively frequently. Specifically, despite our relatively short time horizon of 13 weeks, amongst the the top 30 products for example, we observe an average number of joint searches of 389 across all product pairs, and all 435 possible product pairs are searched at least once. This large amount of co-search data is important, because our approach estimates a full set of covariance terms that rationalizes the frequency of co-search

[14]Although these ratings could vary over time, given our short sample period, there is minimal variation.

between pairs of products. We posit that a situation with sparse purchases but relatively rich search data is common in other online settings as well, and hence our approach will likely be applicable in other similar contexts.

3.1 Price Variation

Prices in our data vary both across products (at a given point in time) and over time (for a given product). Time-series variation in prices is driven either by temporary deals or by changes in the regular price implemented by the retailer's data-analytics team. Deals tend to be accompanied by other changes such as more salient display (a colored price tag alerting the consumer to the deal) and more prominent placement on the webpage. As a result, consumers' reaction to deals is likely to be attributable not just to the price change, but also to other elements that change alongside it. We therefore control for deal status and estimate the impact of price on consideration and conditional choice entirely from within-product changes in regular price.

We have ample within-product price variation. During our 13-week sample, the average product changes price 1.9 times. All products experience a price change at least once, and the magnitude of the average price change is equal to 11.3% of the average product price over the 13 weeks. During the whole period, 13 products are "on deal" at least once. Conditional on having deal status at least once, the median product has it for five out of the 13 weeks.

We treat regular price changes as exogenous. Based on conversations with the company, this assumption is reasonable because most price changes were part of an attempt to induce price variation in order to understand how responsive demand is to such changes. Of course, some price changes could be triggered by changes in demand that the firm is trying to adapt to. However, within the short time frame of our data (13 weeks), we think that large changes in product-level demand are unlikely.

3.2 Browsing Data and Product Characteristics

Next, we examine the relationship between the search data and product characteristics. First, we show that similarity in characteristics helps predict the likelihood of two products appearing in the same consideration set. Hence, the search data allow us to capture product similarity that is driven by observed characteristics. Second, however, we find that a large part of the variation in search patterns remains unexplained by flexible measures of similarity in observed characteristics.

In more detail, we measure the closeness of products in characteristic space by using price, brand identity, the consumer review star rating and the physical characteristics that we mentioned earlier. For all discrete variables, we define a dummy that is equal to 1 if the variable takes the same value for both products j and j'. (For example, one of the regressors is a dummy that takes the value of 1 if two products belong to the same brand.) For all continuous variables, we compute the absolute value of the difference between the two products, and then to facilitate comparisons, we normalize by the variable's standard deviation.

Dependent Variable	(1) # Joint Searches	(2) Joint Search Ratio	(3) Log # Joint Searches	(4) # Joint Searches	(5) Joint Search Ratio	(6) Log # Joint Searches
Mean	391.6	0.093	5.259	391.6	0.093	5.259
S.D.	586.9	0.103	1.257	586.9	0.103	1.257
Same Brand	322.8***	0.048***	0.447***			
	(53.2)	(0.009)	(0.099)			
Price Difference	-55.3	-0.033***	0.097			
	(42.7)	(0.007)	(0.080)			
Star Rating	88.8	0.005	0.196			
Difference	(63.9)	(0.011)	(0.119)			
Same Char. 1	566.2**	0.217***	0.543			
(Discrete)	(226.2)	(0.039)	(0.423)			
Same Char. 2	242.7**	0.031*	0.985***			
(Discrete)	(100.5)	(0.017)	(0.188)			
Diff. Char. 3	-114.5***	-0.009	-0.457***			
(Continuous)	(42.8)	(0.007)	(0.080)			
Diff. Char. 4	-111.0***	-0.010**	-0.135***			
(Continuous)	(25.4)	(0.004)	(0.047)			
Diff. Char. 5	-112.1***	-0.010**	-0.592***			
(Continuous)	(26.2)	(0.005)	(0.049)			
Similarity				339.0***	0.055***	0.501***
Score				(23.0)	(0.004)	(0.055)
Products	30	30	30	30	30	30
Observations	435	435	435	435	435	435
R-squared	0.286	0.310	0.456	0.334	0.286	0.159

Table 3: **Determinants of Co-search of Product Pairs.** The joint-search ratio is defined as [# Searches (j, j')]/[# Searches $(j)\times$# Searches (j')]. Characteristics 1-2 are discrete variables. Regressors are defined as a dummy equal to 1 if the characteristic has the same value for both products. Characteristics 3-5 are continuous. Regressors are defined as the absolute difference between characteristics. All continuous variables (price difference, consumer review star rating difference, and characteristics 3-5) are standardized.

In column (1) of Table 3 we regress the number of times a pair of products (j, j') was searched together during the entire sample period on the measures of product closeness that we just described. Most characteristics have a significant impact and the coefficients have the expected sign. Sharing the same discrete characteristic increases joint search, whereas a larger difference in any continuous characteristic lowers joint search. Some coefficients are relatively large in magnitude, for example, belonging to the same brand increases the number of joint searches by over half a standard deviation. The r-squared is equal to 0.286 and therefore a substantial part of the variation in search patterns is not explained by closeness in observed product characteristics.

The remaining columns of Table 3 probe the robustness of this result. In column (2) the dependent variable is

$$\frac{\sum_i \mathbf{1}((j,j') \in s_i)}{(\sum_i \mathbf{1}(j \in s_i)) \times (\sum_i \mathbf{1}(j' \in s_i))}, \tag{8}$$

where, for example, $\mathbf{1}(j \in s_i)$ takes the value of 1 if product j is contained in consumer i's consideration set. In other words, the dependent variable is the number of consumers who searched (j, j') together, divided by the product of the number of consumers who searched j and j', respectively. This metric adjusts for the fact that products that are searched more often will automatically have higher joint search with any other product. In column (3), we use a logarithmic transformation of the number of joint searches as the dependent variable. The r-squared is higher in both specifications, especially in the log-specification where it is equal to 0.456. In columns (4) to (6), we run the same set of regressions but use as the regressor a product similarity score that is computed internally by the firm and is based on the characteristics we have already used. Directionally these regressions confirm that product similarity predicts more frequent co-search of products. In unreported regressions, we also probe robustness to removing outliers and including higher-order terms of all covariates, and find the results do not change qualitatively.

In summary, these regressions suggest that co-search correlates with similarity in product characteristics, which is likely driven by how products are presented on the webpage. Products with similar physical characteristics will be grouped together in specific sub-categories and might be recommended on each other's product detail pages (for example, as part of a "consumer who looked at this product also considered ..." list of recommendations). Products that are similarly priced are likely to appear next to each other in a list of search results, especially if a consumer sorts products explicitly by price.[15] While there is slight variation in the statistical significance of specific coefficients across specifications, most coefficients indicate more frequent co-search for products with more similar characteristics. Furthermore, we find that observed characteristics only partly explain observed co-search patterns and there are thus aspects of the way products are presented on the webpage that cannot be explained by observed product similarity. Therefore an important advantage of our approach is that we directly use information on joint-search patterns and leverage them to estimate cross-elasticities.

4 Estimation and Results

We outlined the general structure of the demand model in Section 2. We briefly re-cap the relevant equations here and describe a few small modifications that are required to adapt the model to our

[15]We do not observe any aspects of webpage layout such as search result lists or recommendations, but we take the correlation of co-search frequencies with product similarity as evidence that webpage layout is important.

empirical setting.

With regards to consideration set formation, we slightly modify equation (2) from earlier to include deal status:[16]

$$v_{ijt} = \bar{v}_{ijt} + \nu_{ijt} = \gamma_{price} \times price_{jt} + \gamma_{deal} \times deal_{jt} + \bar{\gamma}_j + \tilde{\gamma}_{ij} + \nu_{ijt}.$$

The parameters to be estimated are the coefficients on price and deal status, a vector of product fixed effects, and the correlation terms that govern the distribution of $\tilde{\gamma}_{ij}$ denoted by $\Omega = \{\sigma_{12}, \sigma_{13}, ..., \sigma_{J-1,J}\}$. We denote the full set of parameters in the consideration stage by $\theta_{cset} = \{\gamma_{price}, \gamma_{deal}, \bar{\gamma}_1, ..., \bar{\gamma}_J, \sigma_{12}, \sigma_{13}, ..., \sigma_{J-1,J}\}$. We assume ν_{ijt} is iid and follows a logistic distribution. We set the variance of ν_{ijt} to a small value in order to minimize the role of the logistic error relative to the joint normal errors $\tilde{\gamma}_{ij}$ that drive co-search patterns.[17]

With regards to conditional choice, we also modify equation (4) to include deal status:

$$\begin{aligned} u_{ijt} &= \bar{u}_{ijt} + \varepsilon_{ijt} = \alpha_{price} \times price_{jt} + \alpha_{deal} \times deal_{jt} + \bar{\alpha}_j + \varepsilon_{ijt} \\ u_{i0t} &= \varepsilon_{i0t}. \end{aligned}$$

We assume an error structure where all inside goods are contained in one nest with correlation parameter $(1-\lambda)$, and the outside option is contained in another nest. The full vector of parameters to be estimated is given by $\theta_{choice} = \{\alpha_{price}, \alpha_{deal}, \bar{\alpha}_1, ..., \bar{\alpha}_J, \lambda\}$.

The two stages of the model yield the following probability of choosing consideration set s and the conditional probability of choosing product j from consideration set s:

$$\begin{aligned} Pr_{ist} &= \prod_{k \in s} \left[\frac{\exp(\bar{v}_{ikt})}{1 + \exp(\bar{v}_{ikt})} \right] \times \prod_{l \notin s} \left[\frac{1}{1 + \exp(\bar{v}_{ilt})} \right] \\ Pr_{ijt|s} &= \frac{\exp(\bar{u}_{ijt})}{1 + \sum_{k \in s} \exp(\bar{u}_{ikt})}. \end{aligned}$$

Both stages of the model are estimated by maximum likelihood subject to constraints based on purchase, search, and co-search shares as described in Section 2.4. Because the two stages do not share any parameters, they are estimated separately.

Consideration Set Formation		
	Coeff.	S.E.
Price	-0.956***	0.010
Deal Dummy	-0.004*	0.002
Product Fixed Effects	Yes	
Product-Pair Covariance Terms	Yes	
Conditional Choice		
	Coeff.	S.E.
Price	-1.449***	0.281
Deal Dummy	0.104***	0.038
Within-Nest Correlation	0.797***	0.018
Product Fixed Effects	Yes	
# Observations	118,992	

Table 4: **Estimation Results: Price and Deal Coefficients and Nesting Parameter.** The table reports estimation results for all parameters except fixed effects (in both stages) and correlation parameters (in the consideration stage).

4.1 Estimation Results

The top panel of Table 4 reports results for the consideration model excluding product fixed effects and covariance terms, which we return to in more detail below. The impact of price has the expected sign—a product is more likely to be considered when its price is lower. We find that deal status is not significant at the 5% level, and the point estimate of the deal coefficient is small in magnitude. The bottom panel of Table 4 reports results for the conditional choice model, where we find a negative price coefficient and a positive deal coefficient. We obtain an estimated within-nest correlation parameter of $\lambda = 0.797$.[18] Thus, taste shocks for the inside good are correlated, which leads to more substitution among inside goods relative to substitution towards the outside option.

A key innovation of our approach is that it uses covariance terms in the consideration stage to flexibly model correlations in search behavior. To illustrate our approach, Table 5 reports covariance terms together with co-search probabilities and implied elasticities for the top 10 products in terms of sales. (The structure mirrors that of Table 1 from earlier, where we presented the same three

[16]We also add a time subscript t which denotes a week. Prices vary at the weekly level and consumers that arrive at different points in time therefore face a different vector of prices.

[17]The presence of the ν_{ijt} term imposes an upper bound on the degree of correlation in search, because these errors are independent across products. We include ν_{ijt} to obtain analytical expressions for search probabilities, but minimize the role of the iid errors by setting the variance of ν_{ij} to $1/15 \times (\pi^2/6)$. We experimented with different values of the variance term, and found that estimation took longer when the variance was smaller. We therefore settled on a variance that is as large as possible, while still small enough to rationalize observed co-search patterns in the data.

[18]Note that if $\lambda = 1$ then the choice stage becomes a standard (i.e., non-nested) logit model.

	1	2	3	4	5	6	7	8	9	10
Variance-Covariance Matrix										
1	1	-0.111	0.523	-0.025	0.423	0.074	-0.298	0.344	0.256	0.695
2	-0.111	1	-0.179	0.668	-0.142	-0.302	0.046	-0.135	0.330	-0.086
3	0.523	-0.179	1	-0.110	0.525	0.178	-0.223	0.478	0.094	0.685
4	-0.025	0.668	-0.110	1	-0.189	-0.251	0.130	-0.209	0.509	0.068
5	0.423	-0.142	0.525	-0.189	1	0.066	-0.327	0.703	-0.012	0.459
6	0.074	-0.302	0.178	-0.251	0.066	1	-0.170	-0.059	0.061	0.073
7	-0.298	0.046	-0.223	0.130	-0.327	-0.170	1	-0.254	-0.211	-0.221
8	0.344	-0.135	0.478	-0.209	0.703	-0.059	-0.254	1	-0.069	0.441
9	0.256	0.330	0.094	0.509	-0.012	0.061	-0.211	-0.069	1	0.328
10	0.695	-0.086	0.685	0.068	0.459	0.073	-0.221	0.441	0.328	1
Search / Co-Search Probabilities										
	1	2	3	4	5	6	7	8	9	10
1	0.148	0.008	0.040	0.008	0.026	0.012	0.007	0.018	0.010	0.036
2	0.008	0.086	0.006	0.022	0.004	0.002	0.012	0.003	0.007	0.005
3	0.040	0.006	0.120	0.005	0.027	0.014	0.009	0.021	0.006	0.031
4	0.008	0.022	0.005	0.065	0.003	0.002	0.011	0.002	0.009	0.005
5	0.026	0.004	0.027	0.003	0.084	0.008	0.004	0.024	0.003	0.016
6	0.012	0.002	0.014	0.002	0.008	0.086	0.006	0.004	0.004	0.007
7	0.007	0.012	0.009	0.011	0.004	0.006	0.143	0.004	0.003	0.005
8	0.018	0.003	0.021	0.002	0.024	0.004	0.004	0.068	0.002	0.014
9	0.010	0.007	0.006	0.009	0.003	0.004	0.003	0.002	0.038	0.007
10	0.036	0.005	0.031	0.005	0.016	0.007	0.005	0.014	0.007	0.068
Elasticities (of Row Demand w.r.t Column Price)										
	1	2	3	4	5	6	7	8	9	10
1	-1.800	0.025	0.049	0.021	0.030	0.026	0.017	0.021	0.021	0.034
2	0.024	-2.584	0.013	0.082	0.009	0.009	0.028	0.007	0.023	0.007
3	0.079	0.022	-2.376	0.018	0.041	0.037	0.024	0.031	0.016	0.040
4	0.028	0.106	0.014	-2.719	0.007	0.010	0.032	0.005	0.035	0.011
5	0.070	0.024	0.058	0.014	-2.230	0.028	0.017	0.052	0.011	0.026
6	0.046	0.019	0.040	0.014	0.020	-3.299	0.032	0.010	0.018	0.014
7	0.030	0.047	0.025	0.041	0.010	0.030	-4.575	0.010	0.008	0.009
8	0.065	0.024	0.060	0.012	0.067	0.020	0.021	-2.563	0.009	0.028
9	0.052	0.061	0.024	0.071	0.013	0.025	0.014	0.008	-2.776	0.022
10	0.106	0.024	0.078	0.029	0.036	0.028	0.021	0.031	0.029	-2.504

Table 5: **Estimation Results: Consideration Stage Covariances, Co-Search Probabilities, and Elasticities for the Top 10 Products.** Diagonal elements in the co-search panel show the search frequency for individual products, while off-diagonal elements show the frequency of the row/column-pair of products being searched together.

constructs in the context of an illustrative example.) Looking at the top panel of Table 5 we find relatively large heterogeneity in the covariance terms—the largest covariance term is 0.703 (for products 5 and 8) and the smallest covariance term is -0.327 (for products 5 and 7). As discussed earlier, these covariance patterns are informed by how the co-search frequency of a product pair relates to the search frequencies of the products that comprise the pair. For example, using the middle panel of Table 5, if search probabilities for products 1 and 3 were independent, their co-search probability would be equal to $0.148 \times 0.120 = 0.018$. Since the observed co-search probability for this product pair in the data is 0.040, a large positive covariance term is required to rationalize the degree of co-search. In the case of products 1 and 2, the co-search probability under independent search is equal to $0.148 \times 0.086 = 0.013$, which is larger than the probability observed in the data, and hence leads to a negative estimated covariance term.

The bottom panel of Table 5 reports elasticity estimates. As discussed in Section 2.1, cross-price elasticities are partly determined by co-search frequencies. This relationship between elasticities and search patterns is visible in our estimates. For example, products 1 and 3 are co-searched more frequently than products 1 and 2, and demand for product 1 is more strongly affected by the price of product 3 than by the price of product 2. As we document in more detail below, most products have a small set of strong competitors. For example, demand for product 2 is most strongly affected by the price of product 4, while cross-price elasticities with regards to other products are all substantially lower. As expected, product 4 is the product that is most frequently searched together with product 2.

In summary, Table 5 highlights the main innovation of our approach: we rationalize observed co-search patterns in the data through flexible co-variance terms in the consideration process, which in turn leads to cross-price elasticities that are reflective of co-search patterns.

4.2 Elasticities

In Table 6 we report summary statistics of the distribution of cross-elasticities. The average cross-price elasticity across all product pairs is equal to 0.019 and the median is equal to 0.012. These numbers are relatively small because any given product is likely to have few close substitutes, and will therefore have relatively small cross-price elasticities with most products in the assortment. To illustrate this pattern, for each product we compute the highest cross-price elasticity with any other product in the assortment. We find that the average (across products) maximum cross-price elasticity is 0.068, which is more than three-times larger than the average elasticity. We also report the distributions of the 10th largest, 20th largest, and the smallest cross-price elasticity. These elasticities decline relatively rapidly, again suggesting most products have a small set of important substitutes. We also report the distribution of own-price elasticities in Table 6 and find an average elasticity equal to -3.651.

Finally, we report the distribution of cross-price elasticities of the outside option with respect to price changes for each of the 30 products. We find that these cross-price elasticities are small and similar in magnitude to the smallest cross-price elasticities of each product with the other products

	Mean	25th Perc.	Median	75th Perc.
All Cross-price Elasticities	0.019	0.007	0.012	0.024
Max Cross-price Elasticity	0.068	0.036	0.060	0.096
10th Largest Cross-price Elasticity	0.019	0.011	0.014	0.024
20th Largest Cross-price Elasticity	0.011	0.005	0.009	0.016
Min Cross-price Elasticity	0.006	0.003	0.004	0.008
Own-price Elasticity	-3.651	-4.722	-3.481	-2.621
Elasticity vis-à-vis Outside Option	0.005	0.004	0.003	0.007

Table 6: **Cross- and Own-Price Elasticities.** The unit of observation in each row is a product, expect for the first row which reports the distribution of elasticities across all product pairs.

in the assortment. This pattern of the outside and inside good cross-elasticities suggests that substitution among the 30 products is quantitatively important, and a large share of the demand decrease when increasing the price of an individual product is due to substitution to other products rather than substitution to the outside option.

4.3 Role of the Consideration and Choice Stages

Apart from modeling correlation in search probabilities, the second innovation of our approach relative to the previous literature on consideration set formation is the inclusion of price in the consideration stage.[19] We believe that it is natural for price to influence consideration in an online context. This is because price influences how prominently a given product is displayed on the webpage, and because price is visible to consumers when viewing a list of search queries.

In order to quantify the importance of price sensitivity in the two stages, we calculate elasticities for each stage separately while holding prices in the other stage fixed.[20] We find that the average own-price elasticity in the choice stage is -1.741, and the average own-price elasticity in the consideration stage is -1.946. Since both stages of the model are similarly reactive to price changes, modeling the influence of price in the consideration stage is important. In Table A1 in the appendix, we report the cross-price elasticity matrix for each stage of the model as well as the total elasticity matrix for the top 10 products. We find that, similar to own-price elasticities, the cross-price elasticities in the choice and consideration stages are of similar magnitude.

Finally, we note that our model accounts for different frequencies of product-level search and

[19] As we discussed earlier, structural search models tend to allow for price to influence both search and choice, but constrain the impact of price to be determined by a single price coefficient in the utility function.

[20] That is, to calculate the consideration stage price elasticity, we recompute demand when changing the price for a given product by a small amount in the consideration stage, but we hold all prices in the choice stage constant. We obtain elasticities in the the choice stage by only changing price in the choice part of the model.

purchase behavior via a separate set of product fixed effects in both stages.[21] Our modeling approach can therefore capture niche products that are searched rarely but purchased frequently conditional on search, as well as mainstream products that are frequently searched but might have lower conversion probabilities. Interestingly, we find a very weak negative correlation (-0.02) between the fixed effects in the two stages of the model. This suggests that products which are searched more often, presumably due to more salient display on the webpage, are not necessarily the same products that are likely to be purchased conditional on search.[22]

5 Optimal Prices

We next solve for the optimal prices implied by our model and compare them with the retailer's actual prices. Among the 30 products used in estimation, 3 of them are "marketplace" products which the retailer itself does not stock. The retailer takes a commission of 25% of total sales revenue from these products, but does not control their prices. When computing optimal prices, we take the prices of marketplace products as given but we account for the commissions that the retailer earns by selling them. For the remaining products we have information on wholesale costs, which allows us to calculate optimal mark-ups and profits for each product. In order to preserve the anonymity of the retailer, we rescale wholesale and retail prices so that the retailer's current price for product 1 is equal to 1 dollar. We similarly rescale profits such that profits under current prices are equal to 100 dollars.

Mark-ups under current prices are relatively heterogeneous. For the most popular products mark-ups are typically around 15-30%, whereas for lower-selling products substantially larger mark-ups of up to 90% are observed. The heterogeneity in mark-ups suggests that the retailer is not using a simple constant mark-up rule (so-called cost-plus pricing) to set prices. From conversations with the retailer we know that current price setting is based on a quantitative analysis of price elasticities. The retailer did not disclose the exact analysis underpinning their pricing, but we know that current price setting does not take into account substitution across products.

There are substantial differences between the retailer's actual prices and the optimal prices derived from our demand model. On average optimal mark-ups are higher—46%, relative to 35% under the retailer's actual prices—and also less heterogeneous. However, price differences are not uniform across the assortment. Our demand model suggests higher prices for the top 19 selling products. In some cases the differences are large, such as for the top-selling product, where the retailer's actual mark-up is only 23% but the optimal mark-up is 73%. For most of the smaller products optimal prices are lower than current prices, especially for the products with the smallest market shares. We find that profits increase by 23% when switching from current prices to the optimal prices.

[21] Jiang, Chen, Che, and Wang (2021) similarly allow for product-specific factors that impact search and choice separately.

[22] Ursu (2018) finds that product rankings influence search behavior, but not purchase probabilities conditional on search. Product-specific factors that drive search (such as rankings) but not purchases could be the reason for the lack of correlation between product intercepts in the two stages of our model.

	Marketplace Product	Wholesale Price	Current Price	Current Mark-up	Optimal Price	Optimal Mark-up
Product 1		0.81	1.00	23.2%	1.40	73.0%
Product 2		1.12	1.33	18.3%	1.69	50.0%
Product 3		0.92	1.21	31.3%	1.50	62.8%
Product 4		1.12	1.33	18.3%	1.65	47.2%
Product 5		0.92	1.10	19.2%	1.47	59.3%
Product 6		1.28	1.68	31.0%	1.86	45.1%
Product 7		2.01	2.51	24.6%	2.62	30.2%
Product 8		0.92	1.22	32.1%	1.47	59.3%
Product 9		1.12	1.29	14.5%	1.65	47.2%
Product 10		0.92	1.18	28.3%	1.45	57.7%
Product 11		1.86	2.35	26.5%	2.42	30.3%
Product 12	Yes		2.01		2.01	
Product 13		1.20	1.43	19.4%	1.72	42.7%
Product 14		2.01	2.47	22.9%	2.58	27.9%
Product 15		1.12	1.33	18.3%	1.65	47.2%
Product 16	Yes		2.33		2.33	
Product 17		0.92	1.21	31.3%	1.45	57.6%
Product 18		1.00	1.12	11.7%	1.51	51.6%
Product 19		2.01	2.47	22.8%	2.59	28.7%
Product 20		1.11	2.03	82.9%	1.69	51.9%
Product 21	Yes		1.54		1.54	
Product 22		0.86	1.39	62.5%	1.34	56.2%
Product 23		2.01	2.51	24.4%	2.58	27.9%
Product 24		1.39	2.16	55.0%	1.94	39.2%
Product 25		0.86	1.47	71.6%	1.34	56.4%
Product 26		1.48	2.11	42.2%	2.03	36.9%
Product 27		1.15	2.18	90.1%	1.71	49.0%
Product 28		1.39	2.16	55.0%	1.92	38.1%
Product 29		1.39	1.74	25.1%	1.91	37.2%
Product 30		1.54	2.20	42.6%	2.08	34.3%
Profit			100		123	

Table 7: **Current and Optimal Prices.** Products are ranked by decreasing sales. Wholesale and retail prices are re-scaled such that the current retail price for product 1 is equal to \$1. Profits are re-scaled such that profits under current prices are equal to \$100. Prices for marketplace products are kept constant when optimizing the remaining prices.

6 Conclusion

In this paper we propose a demand model that uses consumer search data to flexibly estimate substitution patterns. Our approach is based on a model of consideration set formation that allows product-specific search probabilities to be correlated in a fully flexible fashion. We show that in a typical online retail setting, search data is sufficiently rich to precisely estimate the correlation matrix driving co-search behavior. In order to estimate the large vector of correlation parameters, we employ a constrained optimization approach based on the idea that product-pair correlations play the role of matching co-search probabilities to their empirical counterparts. We also show that co-search patterns correlate with similarity in observed product characteristics which drive substitution patterns in typical models of demand. However, observed characteristics only partially explain search patterns, suggesting that unobserved factors such as webpage layout also drive co-search patterns. Modeling co-search flexibly (rather than as a function of observed characteristics) allows us to capture such unobserved drivers of consideration set formation.

We believe our approach is particularly useful in online markets, where purchases are sparse, demand may be driven by factors such as page layout that are hard to record, and the number of products is large relative to the number of observed product characteristics. We apply our approach to data from an online retailer, and show that our model generates optimal prices that differ substantially from current prices and lead to a 23% increase in profits.

References

Andrews, R. L., and T. C. Srinivasan (1995): "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32(1), 30–41.

Armona, L., G. Lewis, and G. Zervas (2021): "Learning Product Characteristics and Consumer Preferences from Search Data," Working Paper.

Barroso, A., and G. Llobet (2012): "Advertising and Consumer Awareness of New, Differentiated Products," *Journal of Marketing Research*, 49(6), 773–792.

Bronnenberg, B. J., and W. R. Vanhonacker (1996): "Limited Choice Sets, Local Price Response and Implied Measures of Price Competition," *Journal of Marketing Research*, 33(2), 163–173.

Chen, Y., and S. Yao (2017): "Sequential Search with Refinement: Model and Application with Click-stream Data," *Management Science*, 63(12), 4345–4365.

Chiong, K. X., and M. Shum (2019): "Random Projection Estimation of Discrete-Choice Models with Large Choice Sets," *Management Science*, 65(1), 256–271.

Choi, H., and C. Mela (2019): "Monetizing Online Marketplaces," *Marketing Science*, 38(6), 913–1084.

De Los Santos, B. I., A. Hortacsu, and M. Wildenbeest (2012): "Testing Models of Consumer Search using Data on Web Browsing and Purchasing Behavior," *American Economic Review*, 102(6), 2955–2980.

Donnelly, R., A. Kanodia, and I. Morozov (2022): "Welfare Effects of Personalized Rankings," Working Paper.

Dotson, J. P., M. A. Beltramo, E. M. Feit, and R. C. Smith (2019): "Modeling the Effect of Images on Product Choices," Working Paper.

Dotson, J. P., J. R. Howell, J. D. Brazell, T. Otter, P. J. Lenk, S. MacEachern, and G. M. Allenby (2018): "A Probit Model with Structured Covariance for Similarity Effects and Source of Volume Calculations," *Journal of Marketing Research*, 55(1), 35–47.

Draganska, M., and D. Klapper (2011): "Choice Set Heterogeneity and The Role of Advertising: An Analysis with Micro and Macro Data," *Journal of Marketing Research*, 48(4), 653–669.

Dubé, J.-P., J. T. Fox, and C.-L. Su (2012): "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, 80(5), 2231 – 2267.

Gaynor, M., C. Propper, and S. Seiler (2016): "Free to Choose? Reform, Choice and Consideration Sets in the English National Health Service," *American Economic Review*, 106(11), 3521–3557.

Goeree, M. S. (2008): "Limited Information and Advertising in the US Personal Computer Industry," *Econometrica*, 76(5), 1017–1074.

Haviv, A. (2021): "Does Purchase Without Search Explain Counter-Cyclic Pricing?," *Marketing Science*, 41(2), 211–432.

Honka, E. (2014): "Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry," *RAND Journal of Economics*, 45(4), 847–884.

Honka, E., and P. Chintagunta (2017): "Simultaneous or Sequential? Search Strategies in the US Auto Insurance Industry," *Marketing Science*, 36(1), 21–42, 2017.

Jiang, Z., T. Chen, H. Che, and Y. Wang (2021): "Consumer Search and Purchase: An Empirical Investigation of Retargeting Based on Consumer Online Behaviors," *Marketing Science*, 40(2), 219–240.

Kim, J. B., P. Albuquerque, and B. J. Bronnenberg (2010): "Online Demand under Limited Consumer Search," *Marketing Science*, 29(6), 1001–1023.

Kim, J. B., P. Albuquerque, and B. J. Bronnenberg (2011): "Mapping Online Consumer Search," *Journal of Marketing Research*, 48(1), 13–27.

Koulayev, S. (2014): "Search for Differentiated Products: Identification and Estimation," *RAND Journal of Economics*, 45(3), 553–575.

Lee, T. Y., and E. T. Bradlow (2011): "Automated Marketing Research Using Online Customer Reviews," *Journal of Marketing Research*, 48(5), 881–894.

Li, J., S. Netessine, and S. Koulayev (2018): "Price to Compete ... with Many: How to Identify Price Competition in High-Dimensional Space," *Management Science*, 64(9), 4118–4136.

Mehta, N., S. Rajiv, and K. Srinivasan (2003): "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science*, 22(1), 58–84.

Mojir, N., and K. Sudhir (2021): "A Model of Multi-Pass Search: Price Search Across Stores and Time," *Management Science*, 67(4), 2126–2150.

Netzer, O., R. Feldman, J. Goldenberg, and M. Fresko (2012): "Mine Your Own Business: Market-Structure Surveillance Through Text Mining," *Marketing Science*, 31(3), 521–543.

Pancras, J. (2010): "A Framework to Determine the Value of Consumer Consideration Set Information for Firm Pricing Strategies," *Computational Economics*, 35(3), 269–300.

Pires, T. (2016): "Costly Search and Consideration Sets in Storable Goods Markets," *Quantitative Marketing and Economics*, 14(3), 157–193.

Ringel, D. M., and B. Skiera (2016): "Visualizing Asymmetric Competition Among More Than 1,000 Products Using Big Search Data," *Marketing Science*, 35(3), 511–534.

Ruiz, F. J. R., S. Athey, and D. M. Blei (2017): "SHOPPER: A Probabilistic Model of Consumer Choice with Substitutes and Complements," *Annals of Applied Statistics*, 14(1), 1 – 27.

Seiler, S. (2013): "The Impact of Search Costs on Consumer Behavior: A Dynamic Approach," *Quantitative Marketing and Economics*, 11(2), 155–203.

Smith, A. N., and J. E. Griffin (2021): "Shrinkage Priors for High-Dimensional Demand Estimation," Working Paper.

Smith, A. N., P. E. Rossi, and G. M. Allenby (2019): "Inference for Product Competition and Separable Demand," *Marketing Science*, 38(4), 690–710.

Swait, J., and T. Erdem (2007): "Brand Effects on Choice and Choice Set Formation Under Uncertainty," *Marketing Science*, 26(5), 679–697.

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.

Weitzman, M. L. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47(3), 641–654.

A Cross-Price Effect Derivation

In this section, we provide additional details on how the cross-price derivative formula presented in Section 2.1 is derived. As a reminder, the derivative of the purchase probability of product j with respect to the price of product $k \neq j$ is given by:

$$\frac{\partial Pr_j}{\partial p_k} = \int \sum_{s \ni j,k} Pr_{is}^{Cset} \left[\frac{\partial Pr_{ij|s}}{\partial p_k} - \gamma_{price}(1 - Pr_{ik}^{Search})(Pr_{ij|(s \setminus k)} - Pr_{ij|s}) \right] dF(\tilde{\gamma})$$

In order to derive the expression above, it is useful to re-write the summation over the relevant consideration sets (those which contain product j) for a consumer i as follows:

$$Pr_{ij} = \sum_{s \ni j} Pr_{is}^{Cset} \times Pr_{ij|s} = \sum_{s \ni j,k} Pr_{is}^{Cset} Pr_{ij|s} + \sum_{(s:k \notin s) \ni j} Pr_{is}^{Cset} Pr_{ij|s}.$$

This expression decomposes the probability of choosing product j into two terms—one involves consideration sets that include product k, and the other involves consideration sets that do not include product k.

Next, we take the derivative with respect to p_k for a given consumer:

$$\begin{aligned}
\frac{\partial Pr_{ij}}{\partial p_k} &= \sum_{s \ni j,k} Pr_{is}^{Cset} \frac{\partial Pr_{ij|s}}{\partial p_k} + \sum_{s \ni j,k} \frac{\partial Pr_{is}^{Cset}}{\partial p_k} Pr_{ij|s} + \sum_{(s:k \notin s) \ni j} \frac{\partial Pr_{is}^{Cset}}{\partial p_k} Pr_{ij|s} \\
&= \sum_{s \ni j,k} Pr_{is}^{Cset} \frac{\partial Pr_{ij|s}}{\partial p_k} + \gamma_{price} \sum_{s \ni j,k} Pr_{is}^{Cset}(1 - Pr_{ik}^{search}) Pr_{ij|s} - \gamma_{price} \sum_{(s:k \notin s) \ni j} Pr_{is}^{Cset} Pr_{ik}^{search} Pr_{ij|s} \\
&= \sum_{s \ni j,k} Pr_{is}^{Cset} \frac{\partial Pr_{ij|s}}{\partial p_k} - \gamma_{price} \sum_{s \ni j,k} Pr_{is}^{Cset}(1 - Pr_{ik}^{search})(Pr_{ij|(s \setminus k)} - Pr_{ij|s}) \\
&= \sum_{s \ni j,k} Pr_{is}^{Cset} \left[\frac{\partial Pr_{ij|s}}{\partial p_k} - \gamma_{price}(1 - Pr_{ik}^{search})(Pr_{ij|(s \setminus k)} - Pr_{ij|s}) \right]
\end{aligned}$$

where the second line follows because p_k only enters Pr_{is}^{Cset} via the product specific inclusion probability Pr_{ik}^{search} and the derivative of the inclusion probability with respect to p_k is given by $\gamma_{price} \times Pr_{ik}^{search} \times (1 - Pr_{ik}^{search})$. The derivative of the exclusion probability is given by the same term multiplied by (-1). The third line follows by changing the summation over consideration sets in the third term in the second line to also include product k.

We note that the final expression is positive because both terms are positive, since $\gamma_{price} < 0$ and $Pr_{ij|(s \setminus k)} - Pr_{ij|s} > 0$. The expression in the text is then obtained by integrating over $\tilde{\gamma}$.

B Additional Tables

Consideration Stage Elasticity										
	1	2	3	4	5	6	7	8	9	10
1	-0.845	0.014	0.022	0.012	0.014	0.013	0.009	0.010	0.011	0.016
2	0.012	-1.345	0.007	0.040	0.005	0.005	0.012	0.004	0.012	0.004
3	0.034	0.013	-1.166	0.011	0.020	0.018	0.012	0.015	0.009	0.019
4	0.013	0.046	0.007	-1.453	0.003	0.006	0.013	0.003	0.017	0.006
5	0.031	0.014	0.026	0.009	-1.127	0.014	0.009	0.024	0.006	0.013
6	0.024	0.013	0.021	0.009	0.011	-1.660	0.016	0.006	0.011	0.008
7	0.019	0.029	0.016	0.026	0.007	0.018	-2.096	0.006	0.006	0.005
8	0.030	0.014	0.028	0.008	0.031	0.011	0.011	-1.331	0.005	0.014
9	0.023	0.030	0.012	0.035	0.007	0.012	0.007	0.004	-1.531	0.011
10	0.039	0.013	0.031	0.017	0.017	0.014	0.010	0.015	0.016	-1.282
Conditional Choice Stage Elasticity										
	1	2	3	4	5	6	7	8	9	10
1	-0.963	0.011	0.028	0.009	0.016	0.013	0.009	0.011	0.009	0.019
2	0.012	-1.255	0.006	0.042	0.004	0.004	0.016	0.003	0.011	0.003
3	0.045	0.009	-1.225	0.007	0.021	0.019	0.013	0.016	0.007	0.021
4	0.014	0.061	0.007	-1.284	0.003	0.004	0.019	0.002	0.018	0.005
5	0.040	0.010	0.033	0.005	-1.115	0.014	0.008	0.028	0.005	0.014
6	0.022	0.006	0.019	0.005	0.009	-1.666	0.016	0.005	0.007	0.006
7	0.011	0.018	0.010	0.015	0.004	0.012	-2.532	0.004	0.003	0.004
8	0.036	0.010	0.032	0.005	0.037	0.010	0.011	-1.249	0.004	0.015
9	0.030	0.032	0.013	0.037	0.006	0.013	0.007	0.004	-1.264	0.011
10	0.067	0.011	0.047	0.012	0.020	0.014	0.011	0.016	0.013	-1.238
Total Elasticity (Both Stages)										
	1	2	3	4	5	6	7	8	9	10
1	-1.800	0.025	0.049	0.021	0.030	0.026	0.017	0.021	0.021	0.034
2	0.024	-2.584	0.013	0.082	0.009	0.009	0.028	0.007	0.023	0.007
3	0.079	0.022	-2.376	0.018	0.041	0.037	0.024	0.031	0.016	0.040
4	0.028	0.106	0.014	-2.719	0.007	0.010	0.032	0.005	0.035	0.011
5	0.070	0.024	0.058	0.014	-2.230	0.028	0.017	0.052	0.011	0.026
6	0.046	0.019	0.040	0.014	0.020	-3.299	0.032	0.010	0.018	0.014
7	0.030	0.047	0.025	0.041	0.010	0.030	-4.575	0.010	0.008	0.009
8	0.065	0.024	0.060	0.012	0.067	0.020	0.021	-2.563	0.009	0.028
9	0.052	0.061	0.024	0.071	0.013	0.025	0.014	0.008	-2.776	0.022
10	0.106	0.024	0.078	0.029	0.036	0.028	0.021	0.031	0.029	-2.504

Table A1: **Elasticity Decomposition: Consideration, Choice Stage, and Total Elasticities for the Top 10 Products.** The top two panels report elasticities when changing price only in the consideration or conditional choice stage respectively.