Sources of limited consideration and market power in e-commerce*

Michael Sullivan

Department of Economics, Yale University

November 28, 2022

Abstract

This paper empirically evaluates the contributions of search frictions and pre-search seller differentiation to limited consumer search and to markups in e-commerce. The internet facilitates consumer learning about product offerings, and it allows firms sell products without physical stores. These conditions seem capable of inducing high consumer awareness and cut-throat price competition. In practice, though, consumers exhibit severely limited consideration in online markets and often pay significantly above the minimum available price for a product. High search costs could explain these facts, as could pre-search seller differentiation: consumers with little aversion to search may not visit a store they believe they are unlikely to purchase from based on information known prior to search. I assess these alternative explanations for limited search using a model of sequential consumer search and retailer price competition. I estimate this model on data describing browsing and transactions in contact lens e-commerce. My approach exploits the panel nature of my data to estimate the extent of state dependence and consumers' persistent unobserved tastes for sellers. I find that pre-search seller differentiation, not search costs, is primarily responsible for limited consideration and market power in contact lens e-commerce.

*Email address: m.r.sullivan@yale.edu. I thank Steven Berry, Phil Haile, Katja Seim, Mitsuru Igami, and Soheil Ghili as well as participants in Yale's IO Prospectus Workshop, at the Young Economist Symposium 2021, and at the 5th Doctoral Workshop on the Economics of Digitization for helpful feedback on this project. This paper draws on research supported by the Social Sciences and Humanities Research Council.

1 Introduction

This paper evaluates the contributions of search frictions and pre-search seller differentiation to limited consumer search and to markups in e-commerce using online browsing data and an empirical sequential search model.

The internet facilitates consumer learning about retailers' product offerings and prices. It also facilitates firms' entry in retail markets given that online retailers do not require costly brick-and-mortar stores. In spite of these conditions, consumers on the internet exhibit severely limited consideration of available sellers and products. Additionally, online markets for minimally differentiated goods often feature significant price dispersion, which suggests the presence of market power. Subsequent sections of my paper and numerous papers in the large literature on e-commerce—e.g., Clay et al. (2001), Clemons et al. (2002), Moraga-González and Wildenbeest (2008), Koulayev (2014), and Jolivet and Turon (2019)—document this finding.

In this paper, I ask why the internet failed to effect expansive consumer knowledge about purchasing opportunities and cut-throat price competition in retail, a plausible outcome in a setting wherein stores sell minimally differentiated products such as particular book titles or contact lens boxes. Indeed, if search were truly costless on the internet and if sellers' product offerings were truly undifferentiated, then consumers would compare a product's prices across all available retailers before selecting a particular retailer from whom to purchase that product, and these industries would feature Bertrand price competition with prices depressed to the level of the lowest marginal cost among potential retailers.

The fact that online retail markets for minimally differentiated products often feature limited consideration and substantial price dispersion could reflect that search frictions remain significant on the internet. Much of the empirical literature on consumer search online has emphasized this explanation. When consumers find it costly to search, retailers can make sales at prices above those of their competitors because consumers are willing to transact prices above the minimum price offered by a retailer to avoid further search.[1] Seller differentiation, however, can also explain limited consideration and price dispersion. Even when the characteristics of the physical product that arrives on a consumer's doorstep do not depend on which retailer the consumer selects, a consumer may differentially value retailers for non-price reasons. Characteristics such as shipping and logistical efficiency, reputation, extent of quality certainty, and customer service quality may vertically differentiate retailers. Consumers who value one-stop shopping—i.e., purchasing a variety of goods at a single store—may additionally value retailers that offer a wide variety of products. Consumers may also differentially value online stores' user interfaces and branding; that is, retailers may be

[1] As noted later, Hortaçsu and Syverson (2004) (mutual funds), Hong and Shum (2006) (online textbook sales), and Moraga-González and Wildenbeest (2008) (computer memory chips) were early papers in the empirical search literature that considered search frictions as an explanation for price dispersion and market power in product markets with little product differentiation. More recently, Jolivet and Turon (2019) use search frictions alongside other aspects of consumer behaviour to justify consumers' choices to purchase listings that are dominated by others in an online marketplace.

horizontally differentiated. A consumer may also be more likely to buy from stores from which that consumer has previously made a purchase because of habit formation, store loyalty, or switching costs. If the consumer has limited knowledge about a seller and its prices, but the consumer's knowledge at the outset of search is enough for the consumer to know it is unlikely that they would want to buy from that seller, then the consumer may not visit the seller even when search costs are negligible.

In this paper, I empirically investigate potential sources of limited consideration and market power in the online direct-to-consumer market for soft contact lenses in the United States. In particular, I assess the contribution of various features of the market to the extent of consumer search and equilibrium markups in contact lens e-commerce. Contact lens e-commerce features a small number of online retailers that market contact lenses purchased from a common set of manufacturers to consumers. It also features consumers who often purchase products above their minimum available online prices; prices for many boxes of contact lenses significantly vary across across the retailers on which I focus in this study. Absent large differences in retailers' costs of providing consumers with a specific box of contact lenses, this price dispersion suggests significant markups charged by at least a subset of retailers.

My motivation for studying contact lenses is two-fold. This setting is well suited for the study of consumer search because consumers require a *brand-specific* prescription to buy contact lenses. A consumer prescribed Acuvue Oasys lenses, for instance, cannot substitute the prescription to buy Acuvue 2 lenses or Freshlook lenses. This allows me to credibly assume that all search occurs across stores for a physically homogeneous product rather than across products and stores simultaneously. Several studies of consumer search online analyze product categories in which it is likely that consumers search both across physical products and across retailers; a consumer searching for books, for example, may visit different stores in pursuit of a particular book title, visit different books' pages at a particular online bookstore, or search across both book titles and bookstores.[2] Thus, in the books setting, the assumption of search exclusively occurring across stores within a product seems implausible. Another favourable aspect of the contact lens setting is that the major contact lens retailers only sell contact lenses and contact lens accessories. Thus, modelling the supply side of the industry does not require accounting for the interdependencies between retailers' operations across different product categories or for consumer preferences for one-stop shopping. The advantages of the contact lens setting for studying search come with the caveat that the applicability of results obtained for the contact lens category to other product categories is is somewhat limited due to features of contact lens e-commerce that confer the advantages described above.

I study contact lens e-commerce using a consumer panel dataset of web browsing and on-

[2] Hong and Shum (2006) and De Los Santos et al. (2012) study search for a specific book title across retailers. Moraga-González and Wildenbeest (2008) study search for a specific computer memory chip product across retailers. Morozov et al. (2021) study search across cosmetics products within a retailer. Jolivet and Turon (2019) study search for a particular CD product within a page of adverts on an e-commerce platform.

line transactions histories. Unlike almost all existing studies of consumer search online, I incorporate a dataset's panel dimension to learn about state dependence–that is, the effect of an agent's previous choice on that agent's contemporaneous choice probabilities—and consumers' persistent heterogeneous preferences for stores. These are forms of seller differentiation that I consider to be important potential drivers of limited consideration and market power in e-commerce.[3] Throughout this paper, I use the term state dependence to refer to the effect of a previous purchase on a consumer's contemporaneous choice probabilities. This effect may be explained by habit formation, store loyalty, switching costs, or some other phenomena.

Understanding limited consumer consideration of available sellers—i.e., why consumers visit few retailers before purchasing—and online price dispersion for minimally differentiated products is important for understanding the nature of competition in e-commerce markets and the efficacy of policies intended to remedy market power in the industry. To illustrate, if search frictions were the primary source of market power in e-commerce, then a policy designed to enhance consumer information may make e-commerce markets more competitive. If search costs were trifling and state dependence were instead primarily responsible for market power online, then this remedy would be ineffective; a policy that promoted switching between online retailers could instead be appropriate.

My first analyses describe limited consideration and price dispersion in US e-commerce market in 2007–2008, the time frame of my study. This analysis reveals that consumer consideration is indeed severely limited in the market for contact lenses even though prices vary significantly across online sellers, which suggests possible gains from search. In 83% of the search efforts for contact lenses in my sample, the consumer visits only one contact lens retailer. Also, the average transaction price for contact lenses is 16% above the minimum price available among the three major retailers for the brand the consumer ultimately purchases. These results are qualitatively similar for other product categories: in the books, iPod music players, PlayStation 3 video consoles, and DVD product categories, I find that most consumers visit only one or two online retailers before making a purchase. Additionally, consumers pay on average 35.3% more for books than if they bought at the minimum available price.

I answer this paper's research questions using a model of sequential consumer search and of retailer price competition. I use this model to determine how search behaviour and equilibrium markups change when consumer preferences counterfactually change. This analysis indicates which aspects of consumer preferences give rise to limited consideration and market power in contact lens e-commerce. In estimating my model, I exploit a one-to-one mapping between search effort outcomes and chains of inequalities relating consumers' preferences for sellers. This mapping is implied by the optimal sequential search strategy of Weitzman (1979). To make the use of this mapping in estimation and simulation computationally practical, I deploy a novel parametric assumption on consumers' search costs. Under this

[3] Two papers that study panels of consumer search efforts are Honka (2014) and Morozov et al. (2021). My paper is novel in estimating a model that features both persistent unobserved tastes and state dependence.

assumption, the probability that a consumer conducts a particular sequence of site visits followed by a particular purchase decision is the sum of products of logit choice probabilities. Additionally, whereas almost all empirical search models have used cross-sectional data variation, I develop techniques for analyzing panel data in the context of a search model. These techniques address an initial conditions problem and an endogeneity problem. The initial conditions problem arises because the distribution of consumer tastes differs across consumers with different initially observed purchases. The endogeneity problem arises because, conditional on a consumer's initially observed purchase, the price charged by a store is correlated with the consumer's unobserved tastes; consumers who buy from a store despite it charging a high price may be more likely to have strong unobserved tastes for that store. As noted above, panel data is especially valuable in my setting because patterns in consumers' search activity across time are informative about state dependence and persistent preferences for specific retailers, which are potential explanations for limited consumer consideration and market power online. My pricing model features retailers who anticipate the long-run dynamic effects of their pricing decisions on the distribution of consumers across purchasing states; this distribution affects sales as long as consumer purchasing behaviour exhibits state dependence.

My first main finding is that the low levels of consumer search observed in my data are primarily justified by state dependence and store differentiation in spite of low estimated search costs. Eliminating state dependence raises the share of consumer search efforts involving a visit to more than one store by over 12 percentage points, whereas eliminating persistent unobserved heterogeneity in consumers' tastes for stores raises this share by almost 50 percentage points. Cutting the median search cost in half only raises the share of search efforts involving a visit to more than one store by 10 percentage points. Note that my estimate of the median search cost is 88 cents per website visit, which is much lower than estimates appearing in the empirical consumer search literature of the cost of searching online for books. Additionally, state dependence and consumers' persistent unobserved tastes for stores give rise to market power (i.e., equilibrium markups) whereas search costs do not significantly contribute to markups at their estimated magnitudes. Indeed, removing state dependence reduces equilibrium markups for one popular brand of contact lenses by over 7% at the two leading online contact lens retailers, and removing persistent unobserved heterogeneity in consumers' tastes for stores reduces equilibrium markups at these retailers by over 40%. Lowering the median search cost by half, meanwhile, barely changes markups. This suggests that seller differentiation, not search costs, explains limited consideration and market power in contact lens e-commerce.

1.1 Related literature

My paper's primary contribution is to compare the roles of search frictions, seller differentiation, and state dependence in explaining limited consideration and market power in e-commerce. Early papers in the empirical consumer search literature—namely Hong and Shum (2006), Hortaçsu and Syverson (2004), and Moraga-González and Wildenbeest (2008)—relied on search frictions to explain price dispersion in markets with seemingly homogeneous goods.[4] Some recent empirical search papers feature models with features other than search frictions—namely, state dependence and persistent unobserved tastes—that may contribute to limited consideration equilibrium price dispersion. Honka (2014) includes the consumer's previous decision as an exogenous utility shifter, thereby introducing state dependence but not persistent unobserved tastes into her search model. Morozov et al. (2021) make the alternative choice of modelling persistent preference heterogeneity but not state dependence. One of my paper's contributions is the incorporation of both state dependence and persistent unobserved heterogeneity in a panel-data study of consumer search. Incorporating both of these features is important because both potentially contribute to limited consideration and market power in online retail markets, and they have different implications for the effects of policies or other changes in the market environment. The panel structure of my data together with the novel empirical methods that I develop in my paper allow me to estimate these two aspects of consumer preferences. Additionally, my paper is similar to Morozov et al. (2021) in that it emphasizes the the relationship between preference heterogeneity and search frictions in determining consumer search and purchasing behaviour. To the best of my knowledge, my paper is the first in the empirical search literature that compares the contribution of search frictions, seller differentiation, and state dependence to limited consideration and markups in e-commerce.

My secondary contribution is the development of empirical techniques for estimating a sequential search model on panel data. I use the characterization of the optimal sequential search strategy provided by Weitzman (1979) to construct a one-to-one mapping between inequalities relating consumer utility measures and search effort outcomes. Additionally, I follow Moraga-González et al. (2022) in (i) inverting an equation defining the reservation utilities involved in the Weitzman (1979) optimal search strategy and (ii) specifying a search cost distribution to obtain closed-form choice probabilities.

My paper situates in the empirical search literature, of which Honka et al. (2019) provide an overview. Several papers in this literature relating to my own are De Los Santos et al. (2012) (e-commerce bookstores), Morozov et al. (2021) (e-commerce for cosmetics), Koulayev (2014) (online hotel booking), Jolivet and Turon (2019) (CDs), and Honka (2014) and Honka and Chintagunta (2017) (automobile insurance). My paper also relates to the empirical literature on inertia in consumer choice in that it (i) estimates the contribution of both state dependence and persistent unobserved heterogeneity in consumer tastes to inertia repeat purchasing across shopping occasions and (ii) assesses the implications of state dependence and persistent heterogeneity for firm pricing. Especially relevant papers in this literature include Dubé et al. (2009) and Dubé et al. (2010). Last, my work relates to a literature on platform design in e-commerce. When seller differentiation is significant, as I find in this paper, platform design plays a important role in matching buyers and sellers. Two relevant

[4] Although Hortaçsu and Syverson (2004) allow for vertical (but not horizontal) differentiation between product offerings, Hong and Shum (2006) and Moraga-González and Wildenbeest (2008) use a model without seller differentiation.

papers on platform design and search are Dinerstein et al. (2018), who study search within eBay and its implications for platform design, and Lee and Musolff (2021) who study seller differentiation on Amazon's Marketplace platform and its role in determining the welfare effects of an algorithm that selects default sellers for products listed on the platform.

1.2 Roadmap

This paper proceeds as follows. Section 2 discusses this study's industrial setting and its data source; Section 3 then conducts descriptive analyses of these data. Sections 4 and 5 present my model of sequential consumer search and my model of retailer price competition, respectively. Section 6 outlines my estimation procedure and informally discusses the identification of my search model. Section 7 reports my parameter estimates and Section 8 describes the counterfactual analyses that I conduct using these estimates. Section 9 concludes.

2 Setting and data

This study's primary data source is the Comscore Web Behavior Panel for the years 2007 and 2008. This dataset includes the online browsing and transactions activities for a large panel of US households.[5] As noted by De Los Santos et al. (2012), the Comscore Web Behavior Panel is representative of online consumers in the United States on various observable demographic variables. The browsing data include a record for each web domain visited by a panelist, and each of these records includes the identifier of the panelist who visited the domain, the visit's time, the visit's duration, and whether the visit is associated with a transaction. The data do not include the list of pages visited by a panelist within a web domain; for example, a record of a panelist visiting `amazon.com` does not reveal which product pages that the consumer visited while browsing Amazon. For each transaction in the data, I observe an identifier for the panelist who conducted the transaction, the name of the purchased product, the unit price of the product, the quantity of the product purchased, the total price of the consumer's shopping basket, the time of the transaction, and the web domain on which the transaction took place.

The contact lens transactions analyzed in this paper occur at the three major contact lens retailers in the data, which collectively account for about 95% of observed contact lens transactions in the Comscore data during my sample. These retailers are 1-800 Contacts (1800), Vision Direct (VD), and Walmart (WM). Contact lens e-commerce was sizeable by 2007; 1-800 Contacts made net sales of $125 million in the first half of 2007. As is clear from Table 1, the retailers specializing in contact lens sales—1800 and VD—have much higher sales volumes than WM. Both 1800 and VD almost exclusively sold contact lenses in the sample period; their other offerings were contact lens solutions and other contact lens accessories.

[5] About 92 000 households are included in the 2007 panel, and about 58 000 are included in the 2008 panel.

Although I focus on contact lens e-commerce in this paper, I also analyze online shopping for books, iPod music players, Playstation 3 (PS3) video game consoles, and DVDs. I choose these categories because they contain products for which I observe numerous purchases. The books that I include in the analysis are those for which I observe sales in the Comscore data and that were either (i) a *New York Times* best-seller in either fiction or non-fiction for at least one week in 2007 or 2008 or (ii) one of Amazon's top selling books of 2007. This yields 26 books for which I observe 1696 transactions. The iPod category includes the iPod Shuffle (1GB) and iPod Nano (4GB) as products, and the PlayStation 3 (PS3) category includes the 40GB, 60GB, and 80GB versions of the PS3 as products. The DVD products that I study are the standard edition DVDs for *Ratatouille* and for the first three films in the *Pirates of the Caribbean* series as products; these were among the best-selling DVDs of 2007–2008. I observe 355 iPod purchases, 89 PS3 purchases, and 250 DVD purchases. The four online stores that I analyze in the books category are `amazon.com`, `barnesandnoble.com`, and two composite stores: the "book club" store, which includes various book club websites[6] and "other" stores, which includes the other online bookstores with many fewer recorded sales than Amazon and Barnes & Noble.[7] My scheme of combining several websites follows De Los Santos et al. (2012), who also study Amazon, Barnes & Noble, book clubs, and other stores as four online retailers. For each DVD, iPod, and PS3 product, I include all stores for which I observe a sale of the product.[8] The average transaction prices for each category are: \$16.55 for books, \$13.33 for DVDs, \$137.86 for iPods, and \$517.90 for PS3s.

For each online retailer and each brand of contact lens in the data, I construct a daily time series of the retailer's prices for that brand. My construction of price time series assumes that prices remain constant between observed transactions. Thus, if I observed a transaction for Acuvue 2 at 1800 on July 7, 2007 for \$20.00 and an observation for the same brand at 1800 on July 10, 2007 for \$21.00, I would assume that the price on July 8 and July 9 of 2007 was \$20.00 and that the price from July 10 until the date of the next observed transaction of Acuvue 2 at 1800 was \$21.00. Although this procedure introduces some measurement error into my price variables, the magnitude of the error is likely to be small because my sample size is reasonably large. As reported by Table 2, the mean gap between transactions for top brands (and, consequently, updates of their prices on a site) is generally under two weeks. Also, there is little intertemporal variation in brands' prices for a particular store relative to variation in prices across brands and across stores; see Appendix Table 17 for a description of the price variation in my transactions data.

The prices in my data do not include shipping fees, although 1800 and VD both waived shipping fees for sufficiently large purchases during the period that I study; 1800, for ex-

[6]Namely, `doubledatbookclub.com`, `mysteryguild.com`, `literaryguild.com`, and `eharlequin.com`.

[7]The "other" store includes `walmart.com`, `abebooks.com`, `zooba.com`, `overstock.com`, `booksamillion.com`, `alibris.com`, `borders.com`, `target.com`, and `booksite.com`, `costco.com`, `indigo.ca`, `powells.com`, `bestbuy.com`, `buy.com`, `mytowersafe.com`, and `monstercommercesites.com`.

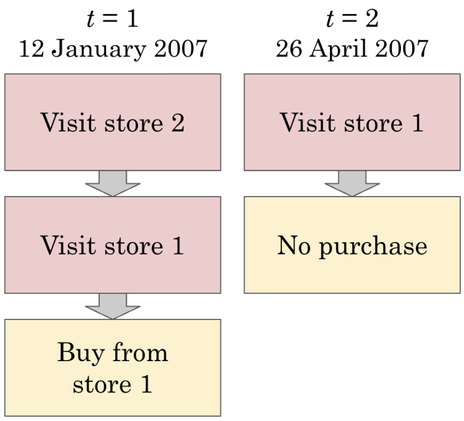
[8]The stores for which I observe iPod purchases are `amazon.com`, `apple.com`, `bestbuy.com`, and `circuit.com`. The stores for which I observe PS3 purchases are `bestbuy.com`, `sonystyle.com`, `buy.com`, `walmart.com`, `amazon.com`, `toysrus.com`, `circuitcity.com`, and `sears.com`. The stores for which I observe DVD purchases are `amazon.com`, `buy.com`, `ebay.com`, `bestbuy.com`, `overstock.com`, and `barnesannoble.com`.

ample, offered free shipping on orders over \$50. Additionally, contact lens manufacturers often offered rebates for contact lens purchases; consumers could receive rebates by sending information on their purchased boxes to the manufacturer, who would then send funds to the consumer. Since these rebates were offered by manufacturers of contact lenses and not their retailers, they should not affect the appeal of buying from one retailer compared to another.

I use the Comscore data to construct a panel of search efforts; each search effort is an ordered sequence of visits to stores and an associated purchase decision, where the available alternatives for the purchase decision are the visited stores and the outside option of purchasing from none of the online stores. Figure 1 illustrates a panel of two possible search efforts. I construct the search effort for each transaction by determining all contact lens retailers that the consumer visited in the days prior to making the transaction. In particular, I include all visits to 1800 or VM in the K days prior to the transaction and all visits to WM in the $K' \leq K$ days prior to the transaction. In my preferred specification, I set $K = 14$ and $K' = 2$. I assess search behaviour under alternative values of K and K' in Section 3, and I find that consumer search efforts are insensitive to the choice of K and K'. The reason for using a shorter time window for WM is that consumers may visit Walmart for purposes that are not related to search for contact lenses, and setting a shorter time window for Walmart is likely to exclude visits unrelated to contact lenses. I also construct a search effort for each visit to 1800 or VM that does not result in a transaction. In doing so, I search for visits to other contact lens retailers within R days (1800 and VM) or R' days (WM) of this visit, and I associate these nearby visits with the same search effort as the initial visit. In my preferred specification, I set $R = 7$ and $R' = 2$, although I find that patterns of search are insensitive to my choices of R and R'. I proceed to add visits that are within R (1800 and VM) or R' (WM) days of visits that have already been added to the search effort, and I continue to iteratively add visits to a search effort until no more visits are added in an iteration. In both the case in which the search effort involves a transaction and in which it does not, I identify the chronologically first, second, and third visits in the search effort, and I use this ordering in my empirical analysis. In the other e-commerce categories that I study (books, iPods, PS3s, and DVDs), most of the large retailers sell products across multiple categories, and consumers are not limited by prescriptions to buy a particular product within a category. This makes the assumption that a visit not resulting in a transaction represents a search effort for a particular product untenable. I therefore do not construct search efforts around visits without transactions in categories other than contact lenses.

In the United States, optometrists and ophthalmologists prescribe contact lenses to their patients after administering eye exams and contact lens fittings for these patients. A contact lens prescription specifies the prescribed brand of contact lenses, various parameters of the prescription (e.g., diameter and power), and an expiration date. Prescriptions are typically valid for one to two years, and there is no limit on the quantity of contact lens boxes that a consumer can purchase with a particular prescription if the consumer makes these purchases before the prescription expires. I infer the prescription of consumers in my sample based

Figure 1: Illustration of search efforts



on the brand of contact lens that these consumers buy. When I see a consumer buy a different brand than the consumer purchased in his previous search effort, I assume that the consumer's prescription has changed to this newly purchased brand and the consumer holds this prescription alone (not in addition to his previous one) until the consumer's next purchase. Over 85% of consumers in my sample never switch between brands.

My empirical analysis relies on the panel structure of my data to study the role of state dependence in guiding search and purchase behaviour. To facilitate the treatment of consumers' previous purchases as observable variables, I drop from my estimation sample each consumer's search efforts made before and including the search effort in which the consumer first made an online purchase. This reduces the number of transactions from 1956 to 1160.

3 Descriptive analysis

3.1 Overview of data

This section provides evidence of limited consideration in contact lens e-commerce and more generally describes the data that I analyze throughout the paper. To begin, Table 1 provides the number of transactions at each of the three retailers studied in this paper as well as their average transaction prices relative to the price at 1800. Note that 1800 and VD sell many more contact lenses than WM, and that 1800 has the highest sales despite having the highest average prices.

Table 2 describes the transactions in my sample. First, panel (a) reports the number of observed transactions for the brands with the most transactions. Note that there are 42 distinct contact lens brands in my sample. Panel (b) provides several quantiles of prices and transaction quantities in my sample. The interquartile range of transaction prices is \$19.99 to \$38.99. The median number of boxes purchased in a transaction is two, which is not surprising given that someone prescribed contact lenses will generally have a different

Table 1: Sales and prices by store (2007–2008)

Store	Transactions	Average relative price
1800contacts.com	849	1
visiondirect.com	416	0.85
walmart.com	70	0.94

Note: the average relative price column reports the average ratio of the store's price to 1800's price across transactions in the 2007–2008 sample.

prescription strength for each eye and will therefore need to purchase a distinct box for each eye.

Table 3a reports how often consumers in my sample search for and purchase contact lenses. Contact lens consumers make, on average, two and a half search efforts and less than two transactions, with some consumers making many more search efforts and transactions. I exclude consumers' first search efforts from my estimation sample so that I can infer consumer's lagged purchase for each included search effort. This is helpful because my model features state dependence. Dropping consumers' first search efforts leaves 1160 search efforts by 494 unique consumers in my estimation sample. Table 3b describes the distribution of the time between transactions in weeks across pairs of adjacent purchases made by the same consumer. The gap between purchases exhibits moderate variation, and centres around 14 weeks.

Table 4a displays the share of contact lens search efforts involving one, two, and three store visits. The "Baseline" column provides results for search efforts constructed by including visits to 1800 or VD up to 14 days before a purchase and including Walmart visits up to 2 days before a purchase in a search effort using the algorithm described in Section 2. The "2 days before" column only includes visits made up to two days before a purchase (or, in the case of a search effort without a transaction, two days within another visit; $K = K' = R = R' = 2$). Table 4a reveals that consumer consideration of available sellers in my sample is severely limited; a full 83% of search sessions involve a visit to only one store under my baseline data construction. The table also shows that search efforts are insensitive to the choice of the parameters used in constructing search efforts. Tables 4d, 4d, and 4d report the distribution of the number of visited stores across search efforts for other product categories under varying definitions of a search effort. In particular, each table shows the distribution of the number of store visits in a search effort when visits at different number of days before a transaction are included in that transaction's associated search effort. For all categories except PS3s, over 75% of search efforts involve a visit to only one or two stores under the "five days before" definition of a search effort. Note that the two sorts of products for which consumers conduct the most search, PS3s and iPods, are the two most expensive sorts of products: recall that the average transaction prices for PS3s and iPods are \$517.90 and \$137.86 compared to \$16.55 for books and \$13.33 for DVDs. Consumers may search more for PS3s and iPods because they perceive the stakes of search for these products to be higher than for books, DVDs, and contact lenses. A comparison of Tables 4d, 4d, and 4d

Table 2: Description of transactions in contact lens data

(a) Transactions by brand

Brand	# Transactions	Weeks b/w transactions Mean	Median
Acuvue 2	188	0.56	0.0
Acuvue Advance	145	0.72	1.0
Acuvue Oasys	129	0.81	1.0
Acuvue Advance for Astigmatism	95	1.05	1.0
Biomedics	57	1.79	2.0
Freshlook Colorblends	56	1.89	1.0
Acuvue 2 Colors	51	1.90	1.5
Soflens 66 Toric	48	1.91	1.0
Focus Night & Day	46	2.29	1.0
O2 Optix	46	2.22	2.0
Other brands	474		
Total	1335		

(b) Quantiles of transaction prices and quantities

Price quantile: 0.10	16.99
Price quantile: 0.25	19.99
Price quantile: 0.50	29.95
Price quantile: 0.75	38.99
Price quantile: 0.90	49.99
Quantity quantile: 0.25	1
Quantity quantile: 0.50	2
Quantity quantile: 0.75	4

Table 3: Description of consumers in contact lens data

(a) Consumer search efforts and transactions

	Mean	Quantiles 0.25	0.5	0.75	0.95
N. search efforts	2.47	1	2	3	6
N. transactions	1.65	1	1	2	4
# consumers = 793					

(b) Number of weeks between transactions

Mean	Quantiles 0.25	0.5	0.75
17.45	7.00	14.00	24.00

Table 4: Share of search efforts by number of visited stores

(a) Contact lenses

# of store visits	Share of sessions Baseline	2 days before
1	0.83	0.84
2	0.16	0.15
3	0.01	0.01

(b) Other categories: 2 days before

# of store visits	Share of sessions Books	iPod	PS3	DVD
1	0.79	0.48	0.38	0.60
2	0.18	0.34	0.35	0.30
3	0.02	0.16	0.26	0.10
4+	0.00	0.02	0.01	0.00

(c) Other categories: 5 days before

# of store visits	Share of sessions Books	iPod	PS3	DVD
1	0.74	0.40	0.32	0.49
2	0.22	0.37	0.27	0.34
3	0.03	0.21	0.33	0.16
4+	0.00	0.02	0.08	0.00

(d) Other categories: 14 days before

# of store visits	Share of sessions Books	iPod	PS3	DVD
1	0.66	0.27	0.22	0.35
2	0.28	0.39	0.27	0.42
3	0.06	0.29	0.36	0.22
4+	0.00	0.05	0.16	0.02

suggests that consumers' search efforts are not overly sensitive to the definition of a search effort.

Consumers visit few stores despite the possibility of spending less on contact lenses by visiting and purchasing from other stores. Indeed, Table 5a shows that 70% of transactions involve purchasing a brand of contact lenses from a store that sells that brand above the minimum price offered among the three major retailers. The magnitude of spending in excess of these minimum prices is significant — consumers pay, on average, 16.3% above the minimum available price for their brand. Table 5b provides analogous analysis for the other e-commerce categories that I consider. In all categories except the PS3 category, most consumers pay above the minimum available price for a product, with the average overpayment ranging from $2.10 in the books category to $6.73 in the PS3 category.

Table 6 characterizes the persistence of consumer search and purchase decisions. In partic-

Table 5: Transactions above minimum available price

(a) Contact lenses

	Value
Share of transactions above min price	0.70
Average payment over min price ($)	4.31
Average payment over min price (%)	16.3

(b) Other categories

	Books	iPod	PS3	DVD
Share of transactions above min price	0.70	0.68	0.18	0.51
Average payment over min price ($)	2.10	6.19	6.73	1.83
Average payment over min price (%)	35.29	4.07	1.38	13.30

Notes: This table reports, for each of the contact lenses, books, iPod, PS3, and DVD product categories, the (i) share of transactions made above the minimum available price among retailers included in the analysis; (ii) the average difference of the transaction price and the minimum available price, and (iii) the average relative difference in percentage terms of the transaction price over the minimum available price. For the books category, the price for the "book club" store is set to the price at the book club with the highest sales in the sample for the book purchased by the consumer. In the case that the consumer bought from a different book club, the price is the minimum of the price at the book club with the highest sales and the book club from which the consumer made a purchase. I proceed analogously for the "other" store. For categories for which I include eBay as a retailer, I do not include eBay's price in computing the minimum. This is because eBay is qualitatively different from other retailers as an auction platform often used for the sale of used goods.

Table 6: Share of search efforts with the same choice as in the previous search effort

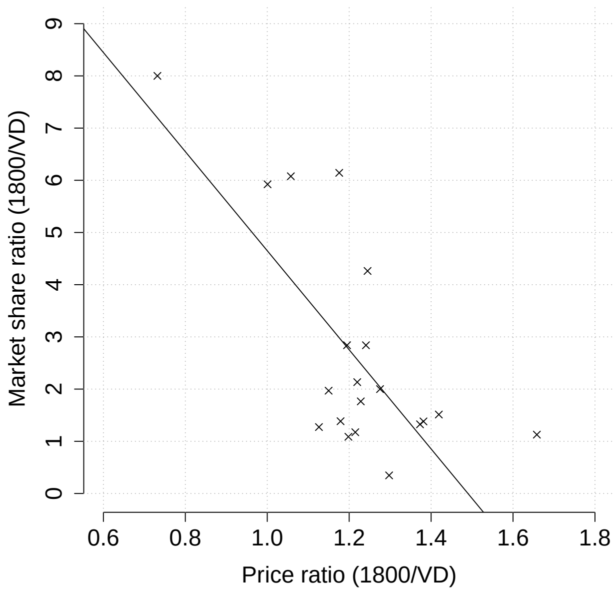
	First visit	Purchase choice
Share	0.85	0.49
N	1163	1163

ular, it shows the share of consumers who choose to visit the same store first in consecutive search efforts. The table also reports the share of consumers making the same purchase decision (i.e., the decision to buy from a particular store or to not buy from any online store) in consecutive search efforts. Last, Table 6 reports the share of transactions that are from the same online retailer as the consumer's previous online transaction from an online retailer. Table 6 shows that search and purchasing behaviour exhibits a high degree of inertia in my sample, with most consumers choosing the same online store for their first visit across search efforts. Additionally, an overwhelming majority of consecutive online purchases take place at the same retailer.

3.2 Prices, browsing, and purchasing

I now turn to the role of prices in directing consumers' browsing and purchasing behaviour. As Table 1 shows, 1800 boasts the highest sales despite charging the highest average prices. This positive relationship between prices and quantities suggests price endogeneity, i.e., that sellers set prices based on consumers' perceptions of their qualities. My solution to the price

Figure 2: Prices and intrabrand market shares at 1800 and VD



Note: Each marker represents a brand of contact lenses in my sample. "Market share ratio (1800/VD)" provides the number of transactions at 1800 for a particular brand divided by the number of transactions at VD for that brand. "Price ratio (1800/VD)" provides the average price of a particular brand at 1800 across days in my sample divided by the analogous quantity for VD. The plot includes the best 20 selling brands in the sample. The plot also displays a line of best fit obtained by regressing the market share ratio on the price ratio for these brands.

endogeneity problem involves exploiting cross-brand differences in stores' relative prices. The idea is that, if stores' quality differences equally affect their sales of all brands, then the extent to which a store's relatively expensive brands sell relatively fewer units will identify consumers' sensitivity to price. Figure 2, which plots 1800's sales relative to VD's against its price relative to VD's for the top 20 brands in my sample, illustrates this idea. This figure shows that, even though 1800 boasts the highest overall sales among online contact lens retailers while charging the highest prices on average, the brands for which 1800 charges especially high prices relative to those of its primary competitor VD are those for which 1800's sales are especially low sales relative to those of VD.

To exploit between-brand variation in prices and quantities to estimate consumers' price sensitivity, I use store fixed effects in my specification of consumers' indirect utilities. I assess the suitability of this approach by running descriptive multinomial logit regressions of consumers' purchasing and browsing decisions on the major online retailers' prices with and without store fixed effects. An additional purpose of these regressions is the determination of whether prices guide consumer search. The finding that prices predict which stores consumers visit would suggest that consumers have some knowledge of stores' prices prior to conducting search.

The estimating equation for the multinomial logit regressions is

$$u_{ift} = q_{ft} - \alpha p_{ift} + \varepsilon_{ift} \qquad f \in \{1800, \mathrm{WM}, \mathrm{VD}\}, \tag{1}$$

where $y_{it} = \arg\max_f u_{ift}$ is, depending on the regression, either the store from which the consumer purchases contact lenses or or the store that the consumer visits first in a search effort. In equation (1), i indexes consumers, t indexes search efforts, and p_{ft} is the price charged by retailer f for i's prescribed brand at time period t. Additionally, ε_{ift} is an unobservable type 1 extreme value random variable taken to be independently distributed across i, f, and t. When I do not include store fixed effects, I impose $q_{ft} = \bar{q}$ and estimate the $\bar{q}$ parameter. When I do include store fixed effects, I set $q_{ft} = q_{f\tau}$, where τ indicates the half-year (e.g. first half of 2007, second half of 2008, etc.) in which search effort t takes place. I estimate the regressions in which the outcome y_{it} is consumer i's purchase decision on a dataset of all search efforts that end in a transaction. I use a disjoint dataset for the regressions in which the outcome y_{it} is consumer i's first-visited store in search effort t: this is the dataset of all search efforts that do not end in a transaction.

Table 7: Descriptive multinomial regression estimates

Specification 1: $q_{ft} = \bar{q} \quad \forall f, t$

	Purchase	First visit
α	-0.033	-0.067
	(0.005)	(0.011)
Implied elasticity	-1.11	-2.25
	(0.18)	(0.36)

Specification 2: seller/half-year fixed effects

	Purchase	First visit
α	0.023	0.025
	(0.004)	(0.015)
Implied elasticity	0.78	0.83
	(0.13)	(0.49)

Notes: Standard errors are reported in parentheses. The "Implied elasticity" is the average own-price elasticity at 1-800 Contacts, where the average is taken across transactions.

Table 7 provides the results of the descriptive logit regressions. Note that, when we do not use seller/time period fixed effects, we do not obtain the expected positive sign for our estimate of price sensitivity. That is, we estimate that consumers become more likely to purchase from a seller when it charges a higher price. This problem is resolved by the introduction of seller/half-year fixed effects. Additionally, the consumer's choice of which store to visit first responds to stores' prices in a similar way as the consumer's purchase choice. This suggests that consumers have some knowledge of stores' prices before conducting search, and thus consumers may search over stores' non-price characteristics (e.g., whether a brand is immediately available in a particular prescription strength, shipping times, or store marketing). The implied elasticities in Specification 2 of Table 7 fall below one, which is inconsistent with profit maximization with non-negative marginal costs. I attribute this to misspecification of the simple logit demand model; estimates from my panel search model imply reasonable elasticities.

The Specification 2 estimates in Table 7 could reflect consumers' responses to cross-brand price differences (i.e., consumers are less likely to buy from a store with a relatively high

Table 8: Between-store and within-store price sensitivities

	α		
	OLS	Between	Within
Estimate	0.31	0.48	0.40
Std. Error	0.13	0.20	0.22

price for their brand on average across time) or intertemporal price variation (i.e., consumers become less likely to buy from a store when the relative price of their brand at that store increases). To assess the relative contributions of cross-brand and intertemporal price variation to the price coefficient estimates in Table 7, I run between and within (fixed-effects) regressions of consumers' purchase decisions on prices. The cross-sectional units of my panel are brands, and the time units are transactions ordered by time. The estimating equation upon which my regressions are based is

$$\mathbb{1}\{t \text{ results in purchase from 1800}\} = \beta_j - \alpha \log\left(\frac{p_{j,1800,t}}{\bar{p}_{jt}}\right) + \varepsilon_t, \tag{2}$$

where j is the prescribed brand of the consumer making transaction t, $p_{j,1800,t}$ is 1800's price for this brand at the time of transaction t, and $\bar{p}_{jt}$ is the average price of brand j across retailers at the time of transaction t.

Table 8 provides estimates of (2) obtained via ordinary least squares (OLS), the between estimator, and the within/fixed-effects estimator. The between estimator is computed by regressing each brand's cross-transaction average of the outcome variable on that brand's cross-transaction average of the regressor in a specification of (2) with $\beta_j = \beta_0$ for all brands j. The within estimator is instead computed by applying the within transform $x_{jt} \mapsto x_{jt} - (1/n_j)\sum_\tau x_{j\tau}$ to each of the outcome variable and the regressor before conducting the regression in (2), where n_j is the number of transactions of brand j in the sample. The between price-sensitivity estimate is larger in absolute value and is more statistically significant than the within estimate, although the difference between the magnitudes of these estimates is small and the within estimate is almost statistically significant at the usual 0.05 level. This suggests that the relationship between purchase decisions and prices in my sample owes to responses to both differences in stores' relative prices across brands and to responses to stores' price changes across time.

Appendix Table 17 characterizes the variation in prices across brands, across stores, and across time. To summarize, the variation in prices across brands is greater than the variation across stores for a particular brand, and variation across time in a brand's price at a particular store is the smallest of the three types of price variation. Appendix Figure 5, which plots price time series for the most popular six brands of contact lenses, conveys this point as well: there are significant differences in prices across brands and stores, although prices typically stay fixed for a brand/store pair for many weeks at a time. This does not reflect the fact that I only update prices upon observing a transaction for a brand, as the prices of the most popular brands in my sample for which I frequently update prices tend to stay fixed for many weeks at a time.

4 Model of consumer search

This section outlines the model of consumer search that I use in my empirical analysis. In the model, consumers search for contact lenses across F online retailers at different occasions across time. Each consumer i has a prescription for a particular brand j of contact lenses. The consumer makes search efforts $t \in \{1, \ldots, T_i\}$ at exogenously determined calendar times. In each search effort, the consumer determines which online retailers $f \in \mathcal{F} = \{1, \ldots, F\}$ to visit. Each online retailer f charges a price p_{jft} for brand j during a given search effort t. The consumer additionally chooses a store f from which to purchase under the constraint that it is only possible to purchase from visited stores. The consumer can also choose to purchase from the outside option, which I denote by $f = 0$ and which represents the possibility of purchasing contact lenses offline or not purchasing contact lenses at all.

Search is costly for the consumer, who incurs a search cost of κ_{ift} for visiting store f in search effort t. In each search effort, the consumer conducts directed sequential search according to the optimal strategy characterized by Weitzman (1979). Consumer i's utility from purchasing from store f at time t is

$$\text{(Online)} \quad u_{ijft} = q_f - \alpha p_{jft} + \phi h_{ift} + \gamma_{if} + \varepsilon_{ift} \tag{3}$$

$$\text{(Offline)} \quad u_{ij0t} = \varepsilon_{i0t}, \tag{4}$$

where q_f is the quality of store f; γ_{if} is persistent component of consumer i's idiosyncratic taste for f; ε_{ift} is the time-varying component of consumer i's idiosyncratic taste for f. Additionally, h_{ift} is an indicator for whether the consumer purchased from f in the previous search effort $t-1$. I will refer to $h_{it} = \{h_{ift}\}_{f \in \mathcal{F}}$ as consumer i's *state* at t throughout this paper. Additionally, α and ϕ are model parameters that governs consumers' price sensitivity and the extent of state dependence, respectively.

There are several interpretations of state dependence in my setting. The ϕh_{ift} term in consumers' indirect utilities, for instance, could be explained by habit formation, switching costs, or store loyalty. In my setting, I expect state dependence to owe in some part to the fact that consumers must send their prescriptions to online retailers before purchasing contact lenses. Once the consumer has sent a recent prescription to an online retailer as a part of a purchase, then that consumer can make future purchases from the retailer without the effort of sending in a prescription again. Similarly, a consumer who has uploaded billing and delivery information to an online retailer can make future purchases without incurring the effort of uploading this information again. Both of these features of contact lens e-commerce suggest that $\phi > 0$.

I assume that, prior to search, the consumer knows each term appearing in u_{ijft} except ε_{ift}, which I call i's *match value* with f. Additionally, I assume that the consumer knows u_{i0t}

prior to search. Note that the first assumption implies that the consumer knows each store's price before visiting that store. Section 4.1 justifies this assumption in my setting. I also assume that consumers are myopic in that they do not anticipate the effects of their choices in a given search effort on their payoffs in future search efforts. This is a common assumption in the literature on state dependence in consumer choice; it is invoked, for example, by Dubé et al. (2010). Future payoffs depend on current choices in my setting because the ϕh_{ift} term in (3) gives rise to state dependence.

The optimal sequential search strategy of Weitzman (1979) involves sorting alternatives in descending order by a *reservation utility* index and then searching the stores in this order until obtaining an indirect utility higher than the maximum reservation utility among unsearched alternatives. Consumer i's reservation utility r_{ift} for store f in search effort t, is defined by

$$\kappa_{ift} = \int_{r_{ift}}^{\infty} (u - r_{ift}) dF_{ift}(u), \tag{5}$$

where F_{ift} is the distribution of u_{ift} conditional on everything except ε_{ift}. Note that r_{ift} is the quantity that makes the consumer indifferent between (i) enjoying a payoff of r_{ift} without further search and (ii) visiting store f before enjoying a payoff equal to whichever of u_{ift} and r_{ift} is greater.

In my model, the reservation utilities can be written as

$$r_{ift} = q_f + \gamma_{if} - \alpha p_{ft} + \phi h_{ift} + \Gamma_0^{-1}(\kappa_{ift}), \tag{6}$$

for

$$\Gamma_0(\kappa) = \int_{\kappa}^{\infty} (u - \kappa) dF_0(u),$$

where F_0 is the distribution of the ε_{ift} match values; this is the type 1 extreme value distribution in my empirical analysis. Note that Γ_0 (and its inverse) are strictly decreasing functions, which means that a store's reservation utility is decreasing in the search cost associated with a visit to that store. Kim et al. (2010) and Moraga-González et al. (2022) similarly invert equations defining reservation utilities to obtain expressions resembling my equation (5) that express reservation utilities as a sum of the parts of the indirect utilities known prior to search and of a decreasing function of the search cost.

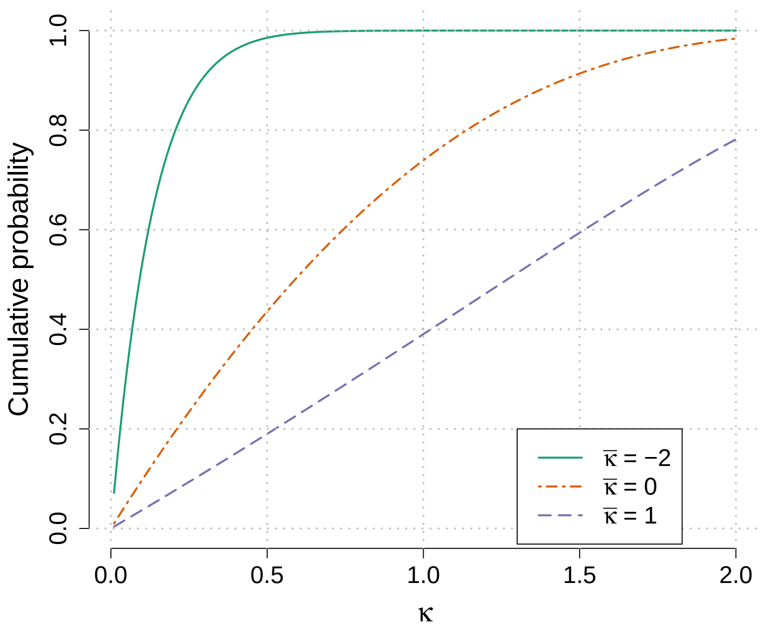
There is a convenient parametric distribution of the search costs κ_{ift} that yields tractable choice probabilities for consumers' sequences of visits and purchases. Suppose that $\kappa_{ift} \sim F_\kappa(\cdot; \bar{\kappa})$ independently of all else, where

$$F_\kappa(\kappa; \bar{\kappa}) = 1 - \exp\left\{- \exp\left\{-\Gamma_0^{-1}(\kappa) - \bar{\kappa}\right\}\right\}. \tag{7}$$

Then, we can express equation (6) as

$$r_{ift} = q_f + \gamma_{if} - \alpha p_{ft} + \phi h_{ift} - \bar{\kappa} + \eta_{ift},$$

Figure 3: Illustration of the search cost distribution function



where the η_{ift} are mutually independent (across i, f, and t) random variables following a type 1 extreme value distribution. Note that the $\bar{\kappa}$ parameter has a positive relationship with both the mean and variance of the search cost distribution. Figure 3 plots F_κ for various values of $\bar{\kappa}$.

The search cost distribution proposed above is one of two features of my model that gives rise to tractable choice probabilities. The other is a collection of inequalities implied by Weitzman (1979)'s optimal sequential search strategy. This strategy provides a bijective mapping between (i) inequalities relating reservation utilities r_{ift} and indirect utilities u_{ift} and (ii) outcomes of consumer search efforts (i.e., search and purchase decisions). Given my distributional assumptions on ε_{ift} and κ_{ift}, these inequalities yield rank-ordered logit probabilities of outcomes. To illustrate, suppose that a consumer visits a store f and f' before buying from store f. This sequence of visits implies that the consumer's highest reservation utility is that for store f. It also implies that the reservation utility for store f' exceeds the indirect utility for store f; otherwise, the consumer would have terminated search after visiting f to buy from that store. Analogous reasoning establishes that the reservation utility for store f' exceeds the reservation utility for the outside option. Since the consumer purchases from store f, the indirect utility of store f must exceed the indirect utilities of store f' and of the outside option in addition to the reservation utilities of all stores other than f and f'. The reasoning above is summarized by the following chain of inequalities (in which I suppress the search effort subscript t):[9]

$$r_{if} \geq r_{if'} \geq u_{if} \geq u_{i0} \vee u_{if'} \vee \max_{g \in \mathcal{F} \setminus \{f, f'\}} r_{ig}.$$

[9]Note that $\vee$ is the maximum operator, i.e. $a \vee b = \max\{a, b\}$.

Given my distributional assumptions, the probability of the consumer's search outcome is

$$\frac{e^{\bar{r}_{if}}}{\sum_{g=1}^{F} e^{\bar{r}_{ig}} + e^{\bar{u}_{i0}} + e^{\bar{u}_{if}} + e^{\bar{u}_{if'}}} \times \frac{e^{\bar{r}_{if'}}}{\sum_{g \neq f} e^{\bar{r}_{ig}} + e^{\bar{u}_{i0}} + e^{\bar{u}_{if}} + e^{\bar{u}_{if'}}} \\ \times \frac{e^{\bar{u}_{if}}}{\sum_{g \notin \{f,f'\}} e^{\bar{r}_{ig}} + e^{\bar{u}_{i0}} + e^{\bar{u}_{if}} + e^{\bar{u}_{if'}}}, \tag{8}$$

where $\bar{u}_{ig} = u_{ig} - \varepsilon_{ig}$ and $\bar{r}_{ig} = r_{ig} - \eta_{ig}$. Appendix B provides the chains of inequalities and choice probabilities corresponding to other search effort outcomes.

The choice probabilities in (8) are straightforward to compute. Without using either the search cost distribution (7) or the chains of inequalities implied by the Weitzman (1979) strategy, computing choice probabilities would require, for a given draw of unobservables κ_{ift} and ε_{ift}, the inversion of a function defined by an integral (i.e., Γ_0) to compute reservation utilities. It would then require the sequential solution of the consumer's search problem by comparing reservation utilities and indirect utilities revealed by search at each step in the consumer's search effort. Last, it would require integration over κ_{ift} and ε_{ift} in order to obtain the probabilities of the various search effort outcomes. By contrast, my choice probabilities have a convenient closed form. Note that the mapping between chains of inequalities involving reservation and indirect utilities and search effort outcomes reduces the burden of computing choice probabilities without a parametric assumption on the search cost distribution F_κ. The application of this mapping does not even require an assumption that the search costs are identically and independently distributed. Other empirical papers that have exploited rankings of indirect and reservation utilities in estimating and analyzing a sequential search model include Moraga-González et al. (2022) and Morozov et al. (2021). Moraga-González et al. (2022) use a result from the theoretical literature on sequential search (Armstrong 2017 and Choi et al. 2018) that choosing an alternative by following the Weitzman (1979) optimal search strategy is equivalent to choosing the alternative that maximizes the minimum of an alternative's reservation utility and its indirect utility. Morozov et al. (2021) specify separate inequalities for (i) the order of visits, (ii) the consumer's continuation and stopping decisions, and (iii) the consumer's purchase decision, and pool these inequalities together to characterize the probability with which a consumer makes a particular sequence of visits and a particular purchase. The primary difference between the approach of Morozov et al. (2021) and my own is that I specify inequalities characterizing all stages of the consumer's search effort that naturally give rise to rank-ordered logit choice probabilities, whereas Morozov et al. (2021) specify multiple inequalities that do not together give rise to tractable closed-form choice probabilities.

The search cost distributions used in the empirical consumer search literature are not typically chosen to achieve tractable choice probabilities in the same way that I have chosen my search cost distribution. Several papers use a log-normal distribution for search costs, e.g., Kim et al. (2010) and Morozov et al. (2021). The paper whose approach most closely relates to my own is Moraga-González et al. (2022), whose authors derive a search cost distribution that ensures a random variable determining their choice probabilities has a type 1 extreme value distribution. This considerably simplifies computation in their context.

4.1 Justification of assuming search over match value

The assumption of known prices and search over match values is common in the literature on consumer search; see, e.g., Kim et al. (2010) and Moraga-González et al. (2022). The assumption is justified in my context for several reasons. First, the descriptive multinomial logit regressions from Section 3.2 suggest that consumers respond to prices in choosing which stores to visit even when they do not make purchases. This is compatible with the consumer choosing which store to visit based on the consumer's knowledge of the prices that each store charges for that consumer's prescribed brand of contact lenses. A question that naturally arises is whether there is any reason for consumers to know stores' prices for their contact lens brand before search. Recall also that prices exhibit relatively little intertemporal variation (see Appendix Table 17) and that I drop consumers' first search efforts from my estimation sample. This means that all consumers in my sample have previous search experience. One explanation for consumer knowledge of prices is that consumers obtained information about stores' prices from previous search efforts.

Another justification of my assumption that consumers search over match values ε_{ift} is that there are several sources of non-price variation in stores' contact lens offerings that the consumer must conduct search to reveal. One important non-price characteristic that consumers may search over is shipping time. Contact lenses vary not only by brand but by other parameters of the consumer's prescription; these include base curve, power, sphere, diameter, cylinder, axis, and addition. Importantly, contact lens' prices vary by brand but not by these other parameters. Whether or not a particular retailer has the consumer's exact specification in stock for that consumer's prescribed brand determines the store's shipping time for the consumer's order and hence the consumer's valuation of ordering from that store. Additionally, online contact lens retailers frequently update their websites to highlight different brands, and this level of variation in brand/site-specific promotion may induce idiosyncratic variation in consumers' valuation of sites across search efforts.

The alternative assumption that consumers conduct search to learn stores' prices faces several problems relating to the specification of consumers' beliefs about prices. One common approach in the empirical search literature is to select a parametric distribution for prices, to estimate this distribution using observed prices, and then to assume that consumers' beliefs follow this estimated price distribution.[10] My setting features relatively little intertemporal price variation for particular brand/store pairs, which means that each estimated brand/store-specific price distributions will be concentrated around the store's mean price for a given brand. In practice, the assumption that these distributions represent consumers' beliefs will therefore be similar to assuming that consumers believe that the mean

[10] See Mehta et al. (2003), Hong and Shum (2006), Moraga-González and Wildenbeest (2008), and Honka (2014).

price (which is similar to the price at any given point in time) is the current price. Thus, I do not expect the approach of estimating price distributions to substantially differ from my approach in terms of its implications for consumer knowledge of prices.

4.2 Probabilities of sequences of search efforts

Consumers' search efforts at different calendar times are related by state dependence and persistent store tastes. In this section, I provide an expression for the probability of a consumer's sequence of search efforts across time. Let $y_i = \{y_{it}\}_{t=1}^{T_i}$, where y_{it} denotes consumer i's search/purchase choices in search effort t. Similarly let $p_i = \{p_{it}\}_{t=1}^{T_i}$, where p_{it} denotes the prices of consumer i's brand at search effort t across all stores f. Next, let h_{i1} denote consumer i's initial state, i.e., a vector encoding the store of the consumer's previous purchase. I will also use θ to denote an arbitrary model parameter vector belonging to the parameter space Θ, and I will use θ_0 to denote the true parameter vector under which we assume the estimation sample was generated.

The model outlined in the preceding sections provides conditional probabilities of search effort outcomes under given model parameters θ; I will denote these probabilities by

$$P(y_{it}|p_{it}, h_{it}, \gamma_i; \theta).$$

Then, the overall probability of consumer i's sequence of search efforts conditional on the observables x_i and h_{i1} is

$$\Pr(y_i \mid p_i, h_{i1}; \theta) = \int P(y_i|p_i, h_{i1}, \gamma_i; \theta) dG(\gamma_i|p_i, h_{i1}; \theta),$$

where G is the distribution of γ_i conditional on p_i and h_{i1}.

Integrating over the conditional distribution of γ_i raises two econometric problems. The first is the standard initial condition problem: the distribution of γ_i conditional on p_i and h_{i1} will depend on h_{i1} because h_{i1} reflects consumers' past choices, which depended on γ_i. Thus, we cannot drop h_{i1} from the conditioning set. The second problem, which I call the endogeneity problem, relates to the dependence of γ_i and prices p_i conditional on h_{i1}. To understand this dependence, suppose that store f sold two brands of contact lenses and that its price for the first brand was high relative to other stores whereas its price for the second brand was relatively low. In that case, consumers with a prescription for the first brand who buy at f will tend to have favourable tastes for the store (i.e., high γ_{if} values) in order to justify buying from f despite its high price for their brand. Similarly, consumers with prescriptions for the second brand who buy from f will tend to have low tastes for the brand because some consumers will buy from f despite their distaste for the store to take advantage of its low price. My model generally implies that, conditional on the initial state, the prices that prevailed when the consumer made an initial purchase and that consumer's persistent unobserved tastes for stores will be correlated. Given that prices are highly persistent across time, I expect the same correlation to hold for the prices that the consumer faces in later search efforts. Thus, independence of γ_i and p_i conditional on h_{i1} is not plausible. Appendix A presents empirical evidence that consumers who have previously purchased contact lenses from a seller who charges a relatively high price for that consumer's prescribed brand have especially strong tastes for that seller.

The two problems above prevent me from making the simplifying assumption that $G(\gamma_i|p_i, h_{i1}; \theta)$ depends neither on the initial state nor on prices. As discussed in the subsequent paragraph, specifying a parametric model of γ_i's distribution conditional on p_i and h_{i1} allows the researcher to address this problem without generating the computational problems associated with alternative approaches. I assume

$$\gamma_{if} \mid (p_i, h_{i1}) \sim \begin{cases} N\left(\lambda \tilde{p}_{jf}, \sigma_\gamma^2\right), & h_{if1} = 1 \\ N\left(\Gamma_{fg}, \sigma_\gamma^2\right), & h_{ig1} = 1 \end{cases}$$

where g designates an online seller other than f; λ, Γ_{fg}, and σ_γ^2 are parameters; and $\tilde{p}_{if}$ is the relative price of consumer i's brand at f at i's first observed purchase:

$$\tilde{p}_{if} = \frac{p_{if1} - \frac{1}{F}\sum_{g=1}^{F} p_{ig1}}{\frac{1}{F}\sum_{g=1}^{F} p_{ig1}}.$$

The parameter λ governs the extent to which consumers who initially buy from f despite f charging a high price for their brand have more favourable tastes for f. The parameter Γ_{fg} governs the tastes for store f of consumers who initially buy from store g. Last, the parameter σ_γ^2 governs variability in consumers' persistent idiosyncratic tastes for sellers.

My approach to modelling the conditional distribution of γ_i is based on commonly used approaches to the initial condition problem and endogeneity problems in panel data settings. First, my assignment of a parametric distribution to unobserved heterogeneity γ_i conditional on the initial state h_{i1} follows Wooldridge (2005).[11] Second, modelling the dependence of γ_i on prices conditional on the initial state follows the correlated random effects (CRE) approach commonly used to account for endogeneity in panel data models. CRE approaches involve explicitly modelling the dependence of an individual i's persistent unobserved heterogeneity on the regressors observed for that individual across time.[12]

Morozov et al. (2021) similarly assume a normal distribution for persistent unobserved hetero-

[11] As discussed by Wooldridge (2005), the primary alternative to this approach is to specify the distribution of the initial state conditional on the unobserved heterogeneity. One could obtain this distribution by computing the steady-state distribution of the initial state for a consumer with a particular value of γ_i after selecting an assumption on the transition of stores' prices (e.g., that they follow a Markov chain). This approach is far more computationally burdensome than the Wooldridge (2005)-based approach that I use.

[12] In the widely used CRE model of Chamberlain (1980), the conditional expectation of the unobserved heterogeneity is a linear function of the explanatory variables for each time period. Mundlak (1978) proposes a closely related approach that is also widely used. Wooldridge (2010) uses the term "correlated random effects" to refer to both Chamberlain (1980)'s and Mundlak (1978)'s approaches. Like Chamberlain (1980), I specify a parametric form for the expectation of unobserved heterogeneity conditional on explanatory variables, although I use a nonlinear function of the explanatory variables that captures the influence of prices on the consumer's decision at the time the initial state is determined.

geneity in their panel model of consumer search online. The difference between my approach and theirs is that I model conditional distributions of persistent unobserved heterogeneity whereas they model the unconditional distribution of persistent unobserved heterogeneity. The unconditional distribution is enough in their setting because they do not face the initial conditions and endogeneity problems that I face on account of the fact that their model does not feature state dependence.

5 Price competition

To analyze the effects of search costs and seller differentiation on markups, I use a model of static pricing with a Nash equilibrium solution concept. This model is static in that each of its retailers sets one price for each of its brands that does not dynamically evolve. The model does, however, capture the long-run responses of consumer states to sellers' prices by incorporating a demand system that accounts for these responses. An alternative to the analysis of Nash equilibria of a static pricing game is the analysis of Markov perfect equilibria (MPE) of a dynamic pricing game wherein sellers adjust prices over time in response to changes in payoff relevant state variables. In my setting, these state variables are the shares of consumers of each γ_i type who previous purchased from each seller. The benefit of a static pricing model is that it is straightforward to find Nash equilibria in the the model without simplifying it, whereas simplifications are required to lower the computational burden of solving for an MPE to a feasible level. In particular, the state space is infinite dimensional in a dynamic model with consumer preferences as specified in Section 4, which precludes the computation of MPE without adjustments to the model that reduce the state space. One primary advantage of a dynamic pricing model is that it captures the effects of contemporaneous price changes on future sales in a realistic manner; the static model that I propose, however, captures these effects by virtue of the fact that sellers in the model set prices to maximize profits against a demand system that captures responses of consumer states to prices. See Appendix D for explication of, and results from, the dynamic pricing model. I find qualitatively similar results for the static and dynamic pricing models.

The primary challenging in specifying a static pricing game is in specifying a demand system that takes account of state dependence. The demand system that I propose, which I call *long-run demand*, represents consumer choice under the long-run stationary distribution of states corresponding to a particular vector of prices. This demand involves the *long-run state probabilities* $\{\rho_f(p,\gamma_i)\}_{f=1}^F$, which I define as the solutions of the system of linear equations

$$\rho_f(p,\gamma_i) = \sum_g \sigma_{fg}(p,\gamma_i)\rho_g(p,\gamma_i) \qquad \forall f, \tag{9}$$

where $\sigma_{fg}(p,\gamma_i)$ is the probability with which a consumer with state $h_{igt} = 1$ buys from store f given prices p. Note that the right-hand side of (9) is the overall probability of a consumer belonging to state f after a search effort when the probability that consumer belongs to state g prior to the search effort is $\rho_g(p,\gamma_i)$. Thus, the condition (9) imposes that a type-γ_i

consumer's probability of belonging to state f does not change after an additional search effort. The long-run market share for store f is then

$$\sigma_f^L(p) := \int \sum_g \rho_g(p, \gamma_i)\sigma_{fg}(p, \gamma_i) dH(\gamma_i)$$

where H is the unconditional distribution of γ_i.

6 Estimation

6.1 Indirect inference

I estimate the model using an indirect inference (I-I) estimator of the sort detailed by Gouriéroux et al. (1993).[13] This approach involves (i) computing auxiliary statistics $\hat{\beta}_n$ on my estimation sample; (ii) simulating outcomes y_i conditional on (p_i, h_{i1}) and a trial parameter value θ using my search model; and (iii) computing the same statistics on the simulated data as I initially computed on the estimation sample, letting $\tilde{\beta}_n(\theta)$ denote the statistics computed on the simulated data. The I-I estimator is a value $\hat{\theta}$ that minimizes a measure of the distance between $\hat{\beta}_n$ and $\tilde{\beta}_n(\hat{\theta})$.

To be precise, my I-I estimator is defined by

$$\hat{\theta}_n = \arg\min_{\theta \in \Theta} (\hat{\beta}_n - \tilde{\beta}_n^H(\theta))' \hat{\Omega}_n (\hat{\beta}_n - \tilde{\beta}_n^H(\theta))$$

where $\hat{\beta}_n$ are ordinary least squares (OLS) estimators computed on my estimation sample and $\tilde{\beta}_n^H(\theta)$ are the same OLS estimators computed on a dataset of outcomes simulated under θ conditional on the x_i and h_{i1} values of observations in my estimation sample, outcomes simulated H times for each (p_i, h_{i1}) in my estimation sample.[14] Additionally, Θ is the parameter space and $\hat{\Omega}_n$ is a weighting matrix selected so that $\hat{\Omega}_n$ converges in probability to a positive definite weighting matrix Ω as n grows large. Appendix C provides additional details on the I-I estimator that I deploy and the regressions used in computing the I-I criterion function. It also discusses the form of the asymptotically optimal weighting matrix and my procedure for estimating this matrix, which involves a preliminary consistent estimator of θ_0 obtained by setting the weighting matrix $\hat{\Omega}_n$ to the the identity matrix.

[13] The primary reason that I use an I-I estimator instead of a maximum likelihood estimator (MLE) is that MLEs tend to exhibit poor finite-sample performance in discrete-choice settings with many low probability potential outcomes. This is because the likelihood function highly penalizes observations to which the model assigns a near-zero probability of occurrence, which makes the MLE highly sensitive to low probability observations. If my model under a particular parameter vector, for instance, assigned a probability near zero to a particular search outcome that was observed in my data with a probability of 1%, then the likelihood for this parameter vector would be highly negative. The MLE estimate would be selected in large part to justify low probability observations in the data, and thus it would be very sensitive to which low probability observations are realized. Other papers that similarly justify the use of I-I or moment-based estimators include Krasnokutskaya and Seim (2011), Pakes et al. (2007), and Collard-Wexler (2013).

[14] Thus, the sample size of the dataset on which I run the regressions yielding $\tilde{\beta}_n^H$ is H times the sample size of the dataset on which I run the regressions yielding $\hat{\beta}_n$.

I describe the regression coefficients included in $\hat{\beta}_n$ and the structural parameters that they are included to target in Appendix C. Many of these coefficients are simple sample averages obtained by regressing a variable on a vector of ones. To briefly summarize the regression coefficients:

(i) *Stores' visit shares*: shares of search efforts with a visit to each store.

(ii) *Consideration set size*: shares of search efforts wherein the consumer visited all available stores.

(iii) *Inertia*: regressions of indicators for whether a consumer visited a store on lagged purchases.

(iv) *Role of lagged price*: regressions of an indicator for buying from 1800 on the contemporaneous and lagged price of the consumer's brand at 1800.

(v) *Price sensitivity*: regression of indicators for purchasing at stores on stores' prices.

(vi) *Cross-visiting behaviour*: shares of consumers in various states who visit each store.

(vii) *Dependence of tastes and prices conditional on initial state*: regressions of indicators for whether the consumer visited a particular store on the ratio of the store's price to the average price across stores.

Appendix Table 15 presents these statistics' values in my estimation sample and provides some additional details on their computation.

6.2 Identification

I now informally discuss the identification of my model's parameters. First, as noted in Section 3.2, my general approach to the problem of price endogeneity is to use cross-brand variation in stores' relative prices to learn about consumers' price sensitivity. In my model, this means assuming that that store quality q_f varies by store but not by brand in order to identify the α parameter. This assumption could be partially relaxed; we could allow store qualities to vary in additional dimensions, e.g., time period, manufacturer, whether the lenses are spherical versus toric or transparent versus coloured, etc. What is essential is that there remains residual variation in price within the units for which we use fixed effects.

Second, we may worry that state dependence and persistent unobserved tastes are not separately identified as explanations for inertia in consumer choice. These explanations for inertia, though, imply different patterns of switching behaviour. Although both state dependence and persistent unobserved tastes promote consistency in consumer choice across the consumer's purchasing occasions, they have different implications for the nature of switching when it does occur. For a consumer who purchases from two distinct stores within three purchasing occasions, for example, strong switching costs would make it more likely that the consumer would switch once from the originally selected store, whereas a model with strong persistent store tastes would assign higher probability to the consumer switching from the originally visited store and then switching back. Additionally, a model with stronger persistent store tastes features a greater correlation between contemporaneous choice and choice two or more purchasing occasions ago conditional on the choice in the previous purchasing occasion than a model in which state dependence primarily explains inertia. This is because, conditional on the choice made last period, the choice made two periods ago correlates with the consumer's persistent unobserved tastes, which is a driver of contemporaneous choice. This motivates my inclusion of a regression of the consumer's contemporaneous choice on choices from the previous purchase occasion and the purchase occasion before that in my regression. As reported by Appendix Table 15, the estimated coefficients for previous purchase and purchase two search efforts in the past are 0.495 and 0.392, respectively. The fact that the second lag of the consumer's purchase decision affects the contemporaneous purchase decision conditional on the previous purchase suggests that a role for persistent idiosyncratic tastes in consumer behaviour.

Another explanation for the separate identification, which Dubé et al. (2010) invoke in their study of inertia in consumers' grocery store purchasing, involves variation in covariates. Consider a consumer who is initially observed making a transaction from a store f. Suppose that the next time the consumer makes a search effort, store f raises its price and the consumer responds by purchasing from store g instead. Last, suppose that the third and final time that we observe the consumer make a transaction, store f's price returns to its original level. If there is a high degree of state dependence, the consumer is likely to purchase from store g again because this was the store that the consumer most recently purchased from. If the consumer's initial purchase from f stemmed from favourable, persistent tastes for store f, then we would expect her to instead purchase from store f in the final search effort as long as state dependence is not too strong. Thus, state dependence and persistent idiosyncratic tastes for stores imply different predictions concerning the nature of consumer switching. This insight motivates my inclusion of the "Role of lagged price" regression among my I-I auxiliary statistics. As reported by Appendix Table 15, I estimate the lag of 1800's price to have a coefficient that is close to zero and slightly positive in a regression with an indicator for purchase at 1800 as the dependent variable and with the contemporaneous price at 1800 as another regressor. This suggests a limited role for state dependence in consumer preferences.

I conclude my discussion of identification by considering the identification of the parameter $\bar{\kappa}$ governing search costs. Given that each of search costs, state dependence, and persistent unobserved tastes for stores tend to limit the number of stores that consumers visit, it may seem unclear how the magnitude of search costs is separately identified by these latter two aspects of consumer preferences. The separate identification stems from the fact that state dependence and persistent unobserved tastes induce dynamics in consumer behaviour that iid search costs unique to a purchasing occasion do not induce. Search costs would induce dynamic patterns in consumer behaviour if they were serially correlated, although these patterns would differ from those induced by either state dependence or persistent unobserved tastes. I would not expect serially correlated search costs, for example, to give rise to the sort of uninterrupted steaks of purchases at the same store that I would expect from a model with state dependence. Similarly, serial correlation in search costs would not make consumers highly attached to particular sellers, whereas persistent unobserved tastes have this effect.

7 Parameter estimates

This section presents and discusses the parameter estimates yielded by the indirect inference estimator outlined in the preceding section. Before discussing my estimates, I provide some details of my estimation procedure. I minimize the indirect inference criterion function using the genetic algorithm, which is a global optimization algorithm. I simulate each consumer $H = 50$ in computing the indirect inference criterion function. Last, I de-mean the prices that I enter in consumers' indirect utilities and reservation utilities using the average price across stores for the brand and calendar time corresponding to the search effort in question. Without performing this de-meaning procedure, the model would mechanically predict a larger probability of choosing the outside option for brands that are more expensive on average. I similarly apply this de-meaning procedure in my counterfactual simulations, in which I hold fixed the average prices used in demeaning under counterfactual changes in stores' prices.

Table 9 presents estimates of my model's key parameters. To make the estimates of the search-cost distribution parameter $\bar{\kappa}$ more readily interpretable, I include the median search cost implied by these estimates in both utils and dollars in the bottom two rows of Table 9. The median search cost in dollar terms is only \$0.88, which is low compared to the median transaction price in my sample of about \$30. My estimates suggest, however, that state dependence and heterogeneous tastes for sellers as reflected in γ_i exercise significant influence on consumer decisions. Indeed, having previously purchased from a store raises a consumer's valuation of the store by almost \$12. Additionally, the standard deviation of γ_i conditional on initial state and prices is about 1.23, or about \$12 in dollar terms.

Table 10 provides estimates of $q_f + \mathbb{E}[\gamma_{if}]$—which I call f's *mean store taste*— implied by my choice model estimates for each store f. In line with 1800 boasting higher sales than its rivals despite higher prices, 1800 has a higher mean store taste than WM and VD.

My search cost estimates fall substantially below some others in the empirical literature on consumer search. Hong and Shum (2006), for instance, find median search costs for textbooks between \$2.32 and \$29.40. De Los Santos et al. (2012), who use the Comscore Web Behavior Database to analyze search for books, find average search costs of \$4.14. Although this comparison is somewhat limited by the fact that my contact lens setting differs from the book setting, my results suggest that high estimated search costs in the empirical search literature may reflect a failure to account for forms of seller differentiation that limit consumer consideration.

Table 9: Selected parameter estimates

Parameter	Estimate	SE
q_{1800}	0.040	0.050
q_{WM}	-0.349	0.160
q_{VD}	0.240	0.231
ϕ	1.157	0.246
α	0.103	0.033
$\bar{\kappa}$	-2.012	0.379
$\Gamma_{1800,VD}$	-2.078	0.309
$\Gamma_{VD,1800}$	-5.416	0.654
σ_γ^2	1.508	0.361
λ	6.597	1.333
Median search cost (utils)	0.091	0.038
Median search cost ($)	0.881	0.534

Note: The "Estimate" columns provide point estimates obtained from the indirect inference estimator outlined in 6 whereas the "SE" columns report the estimates' standard errors. Additionally, Γ_{fg} is the mean value of γ_i among consumers with initial state h_{i1} given by $h_{ig1} = 1$.

Table 10: Estimates of mean store tastes

Store f	Mean taste for store f $q_f + \mathbb{E}\gamma_{if}$
1800	-0.65
WM	-3.03
VD	-3.68

Table 11 reports estimates of own- and cross-price elasticities for the popular Acuvue Advance for Astigmatism brand computed using each of the static notions of demand described by Section 5. In particular, each entry corresponds to the elasticity of demand for 1-Day Acuvue at the store indicated by the entry's row with respect to the price indicated by the entry's column. For the two leading stores, 1800 and VD, demand is more elastic in the long run than in the short run. This is to be expected given that an increase in a store's price will eventually lead fewer consumers to belong to that store's state, which amplifies the short-run effect of the price increase on quantity sold under the presence of state dependence. Also, the cross-elasticities for these stores are substantially larger in the long run.

8 Counterfactuals

In this section, I conduct counterfactual analyses intended to assess the sources of limited consideration and market power in contact lens e-commerce.

Table 11: Elasticity estimates for Acuvue Advance for Astigmatism

Panel A: Point estimates

Short-run demand

	Price		
Share	1800	WM	VD
1800	-1.62	0.01	0.01
WM	0.10	-7.72	4.16
VD	0.00	0.09	-1.55

Long-run demand

	Price		
Share	1800	WM	VD
1800	-2.52	0.19	0.19
WM	1.38	-8.23	4.90
VD	0.28	0.97	-2.12

Panel B: Standard errors

Short-run demand

	Price		
Share	1800	WM	VD
1800	0.86	0.03	0.49
WM	1.77	0.89	0.94
VD	1.24	0.10	0.72

Long-run demand

	Price		
Share	1800	WM	VD
1800	1.58	1.39	0.62
WM	1.36	1.90	1.64
VD	1.63	1.33	1.25

Note: Standard errors computed using the parametric bootstrap with 100 bootstrap replicates.

8.1 Sources of limited consideration

My assessment of the sources of limited consideration in online search for contact lenses involves simulating search efforts under counterfactual preference parameters. I consider an aspect of consumer preferences to be a driver of limited consideration if it exerts significant influence on the extent of consumer consideration (i.e., on how many stores a consumer visits). Note, though, that the exercise considered by this section involves simulating search efforts conditional on observed prices rather than simulating search efforts in a pricing equilibrium computed under counterfactual model parameters. Thus, the exercise addresses the question of why consumers exhibit limited consideration in response to the prices they face in the data rather than the question of why consideration is limited in a pricing equilibrium. To produce the simulated datasets discussed throughout this section, I simulate each consumer's history of search efforts 50 times; in each simulation, I draw outcomes conditional on that consumer's prescribed brand, the prices faced by that consumer, and the consumer's initial state. In order to condition on the initial state, I drop all search efforts before and including that in which I first observe the consumer make an online purchase.

The counterfactual consumer preferences that I consider are

(i) Low search costs: reduce $\bar{\kappa}$ so that the median search cost equals one half of the median search cost under the estimated value of $\bar{\kappa}$;

(ii) No state dependence: set $\phi = 0$;

(iii) No vertical differentiation: set $q_f + \mathbb{E}[\gamma_{if}] = 0$ for each store f to eliminate mean quality differences between stores;[15]

[15] I compute $\mathbb{E}[\gamma_{if}]$ by first integrating over each consumer i's estimated distribution of γ_{if} conditional on

(iv) No persistent unobserved store tastes: set $\gamma_{if} = 0$ for all consumers i and online retailers f; and

(v) Logit only: eliminate search costs, state dependence, vertical differentiation, and persistent unobserved store tastes. Under these counterfactual consumer preferences, only prices and the ε_{ijft} unobservable differentiate retailers from the consumer's perspective.

Counterfactually altering a consumer's preferences changes that consumer's inclusive value of the online stores considered together, and therefore changes the probability that a consumer buys from any of the online stores. State dependence, for example, is a major driver of a consumer's decision to purchase from any online store. Removing store loyalty therefore substantially decreases the volume of online transactions. Thus, the effects of the counterfactual preference changes described above would reflect both a qualitative change in consumer preferences and a change in the magnitude of consumers' tastes for online retail. To focus on the effect of various qualitative changes in consumer preferences, I make an additional adjustment to consumers' preferences in each counterfactual. In particular, I add a compensating constant $q^{\dagger}$ to each consumer's indirect utility for every online store to ensure that the outside good's share is constant across the counterfactuals. The value of $q^{\dagger}$ differs across counterfactuals. Appendix Table 18 provides results for counterfactual preference changes in which this compensating factor is not included.

consumer i's initial state and the prices of that consumer i faces while searching, and then integrating over the distribution of consumers in my sample.

Table 12: Model fit and counterfactual search patterns

Specification	Share visiting one store only	Mean # of visits	Share buying from... any	1800	VD	Visit order	Share paying > min. price	Mean over-payment ($)
Observed	0.819	1.196	0.610	0.364	0.220	0.496	0.660	3.95
Baseline	0.823	1.193	0.733	0.479	0.214	0.419	0.717	4.45
Low search costs (comp.)	0.724	1.314	0.733	0.476	0.214	0.417	0.716	4.42
No state dep. (comp.)	0.705	1.334	0.733	0.472	0.204	0.459	0.720	4.42
No vertical diff. (comp.)	0.710	1.339	0.733	0.297	0.325	0.602	0.581	3.22
No persistent unobs. (comp.)	0.330	1.978	0.733	0.203	0.432	0.431	0.442	1.73
Logit only (comp.)	0.000	3.000	0.733	0.179	0.326	1.000	0.531	2.27

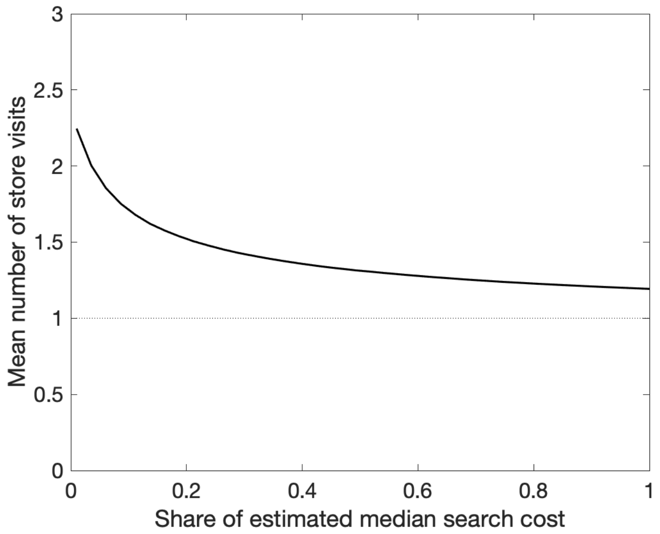
Table 12 characterizes consumer search in the estimation sample, in search efforts simulated under the parameter estimates, and in search efforts simulated under counterfactual parameters. For a version of the table with additional counterfactual preferences and with standard errors, see Appendix Table 18. The "comp." label appearing in parentheses for some of the specifications indicates that the value of buying from each of the stores has been adjusted by a compensating factor $q^{\dagger}$ as described in the preceding paragraph. The "Share visiting one store only" column provides the share of search efforts involving a visit to only one of the three online stores; the "Mean # of visits" column provides the average number of visits in a search effort; and the "Share buying from" columns report the shares of search efforts resolving in a purchase from either any store or from one of the two leading stores, 1800 and VD. Next, the "Visit order" column reports the share of search efforts involving a visit to each of 1800 and VD in which 1800 is visited first. The final two columns characterize the extent to which consumers pay above the minimum available price for contact lenses: "Share paying over min. price" provides the share of search efforts involving the purchase of a contact lens brand at a price above the minimum price available among the three retailers. Last, "Mean overpayment ($)" reports the mean difference between the price at which the consumer purchased contact lenses and the minimum available price for the consumer's brand across search efforts ending in online transactions.

A comparison of the first two rows provides an evaluation of model fit; in general, the model's predictions closely match the data. A comparison of the second row with the remaining rows characterizes the sources of limited consideration under the model. Of the aspects of consumer preferences that I consider, persistent unobserved heterogeneity plays the largest role in explaining why consumers exhibit limited consideration; the share of search efforts involving a visit to more than one store rises from about 18% to 67% upon the elimination of consumers' persistent unobserved tastes that horizontally differentiate sellers. Additionally, the extent to which consumers overpay for contact lenses decreases in this counterfactual. These results together suggest that the consumer's preference for purchasing from sellers that they idiosyncratically prefer explains why the consumer avoids visiting other stores even when they offer lower prices. Eliminating state dependence also expands consumer consideration, although it does not meaningfully decrease the amount that consumers overpay for contact lenses.

Eliminating vertical differentiation only modestly expands the extent of consumer consideration. This is because it leads some consumers who previously visited only 1800 to also consider VD. Given that 1800 is estimated to be the vertically superior store in terms of mean store tastes $q_f + \mathbb{E}[\gamma_{if}]$, the elimination of 1800's mean quality advantage over its less expensive competitor VD leads more consumers to consider and ultimately purchase from VD. This decreases the average overpayment. Therefore, we can conclude that consumers' overpayment for contact lenses partially reflects superior quality offered by more expensive stores that justifies the overpayment, and thus that overpayment for contact lenses does not necessarily reflect consumer inattention.

Search costs play a smaller role in limiting consumer search, and reducing search cost has a

Figure 4: Role of search costs in limiting consumer search



negligible effect on the extent that consumers overpay for contact lenses. Figure 4 shows how the number of stores that consumers visit changes as search costs are reduced. In particular, it shows the relationship between the mean number of visits in a search effort and the median search cost as a fraction of its estimated level. The median search cost must fall below about 20% of its estimated level for the average number of visits in a search effort to exceed one and a half.

8.2 Sources of market power

I assess the sources of market power in online contact lens retail by recomputing pricing equilibria for a particular brand of contact lenses after changing consumer preferences and then computing the change in equilibrium markups from the baseline equilibria to the counterfactual equilibria. The first assessment that I conduct uses a model of static pricing with a Bertrand-Nash equilibrium solution concept and the long-run demand system that I defined in Section 5. That is, I assume that each store f's equilibrium prices p_f maximize its long-run profits

$$\Pi_f(p) = (p_f - mc_f)\sigma_f^L(p)$$

given the equilibrium prices of f's competitors. Two elements of the Π_f function are unknown: the long-run demand function σ_f^L and the marginal costs mc_f. In order to compute firm profits in practice, I use the estimate of σ_f^L derived from my search model estimates, and estimates of marginal costs mc_f obtained by solving firms' first-order conditions for profit maximizations under the observed prices and my demand estimates σ_f^L. Throughout this section, I focus on price competition within a single brand of contact lenses. The brand for which I present my counterfactual results is Acuvue Advance for Astigmatism, which is one of the most popular brands in my sample.

Table 13: Percentage changes in markups from static pricing model

Panel A: Point estimates (%)

Store	Low search costs	No state dependence	No persistent unobs.
1800	-0.8	-7.4	-41.3
WM	1.0	-1.3	-57.5
VD	-1.0	-7.8	-77.5

Panel B: Standard errors

Store	Low search costs	No loyalty	No persistent . unobs.
1800	0.2	1.4	5.1
WM	1.3	2.2	5.8
VD	0.2	1.4	1.7

Note: This table presents estimates of percentage changes in markups for Acuvue Advance for Astigmatism under counterfactual consumer preferences relative to markups in a pricing equilibrium computed at the estimated model parameters. The standard errors were computed using a parametric bootstrap with 100 bootstrap draws.

The changes in consumer preferences that I consider are similar to those considered in Section 8.1. They are:

(i) Low search costs: reduce $\bar{\kappa}$ so that the median search cost equals one half of the median search cost under the estimated value of $\bar{\kappa}$;

(ii) No state dependence: set $\phi = 0$; and

(iii) No persistent unobserved store tastes: set $\gamma_{if} = \mathbb{E}\gamma_{if}$ for each consumer i and each store f.

I do not add any compensating constant $q^{\dagger}$ to consumers' utilities in my pricing counterfactuals.

Table 13 reports percentage changes in equilibrium markups relative to the baseline estimated consumer preferences for each set of counterfactual preferences. Reducing search costs does little to change equilibrium markups, implying that price dispersion for physically identical goods sold online is not a consequence of search frictions providing sellers with market power. Instead, Table 13 suggests that the sources of market power online are store differentiation and, to a lesser extent, state dependence. Indeed, eliminating persistent unobserved tastes reduces markups by over 40% at each retailer, and eliminating state dependence decreases markups by over 7% at each of the two largest retailers (1800 and VD).

The contribution of state dependence to market power has implications for business practices that make switching between stores difficult. As discussed earlier, there are several interpretations of state dependence in my setting. These include habit formation, switching costs, and the convenience of not having to re-enter prescription, billing, and delivery information. Given that eliminating state dependence would decrease payment decrease equilibrium markups, it may benefit consumers to reduce the extent of state dependence. An example of a policy that would accomplish this is the introduction of an intermediary service to which the user uploads prescription, billing, and delivery information. This service could then share the user's information with any online retailer with which the consumer seeks to make a transaction without the user having to separately input this information for each online retailer. This proposed service resembles e-commerce platforms like Amazon and eBay that provide an interface through which a consumer can deal with many retailers.

Search costs in contact lens e-commerce are small and do not meaningfully contribute either to limited consideration or market power in the industry. Given that searching across contact lens retail sites is qualitatively similar to searching across sites operating in other product categories, I expect the same conclusion to hold for e-commerce more broadly. Additionally, searching within a site for products seems less difficult than searching across sites since it does not require navigating to sites via search engines or URL entry. Thus, I expect that the costs of searching within Amazon or eBay, for example, are lower than the search costs that I estimate in this paper. As such, remedies to market power in the industry that aim to make search easier, e.g., by introducing comparison tools or by increasing retailers' transparency about their product offerings, are unlikely to meaningfully lower prices or otherwise improve the consumer experience in online retail.

9 Conclusion

This paper applied a model of consumer search to a panel dataset describing consumers' browsing and purchasing behaviour in contact lens e-commerce. The paper's first primary contribution is its development of a tractable empirical framework for studying sequential search models. This framework exploits a property of the Weitzman (1979) search strategy and, optionally, a convenient set of parametric assumptions to simplify the computation of probabilities of particular search outcomes. Additionally, my framework can be used to learn about state dependence and persistent unobserved heterogeneity in a search setting from panel data; these are aspects of consumer preferences that have not been simultaneously accounted for in previous empirical studies of consumer search. The paper's other primary contribution is in drawing substantial conclusions about limited consideration and market power in e-commerce. My analysis suggests that various forms of seller differentiation play a much larger role than search frictions in accounting for these phenomenon.

Although the contact lens setting is suitable for studying search, the applicability of my study's conclusions to other e-commerce settings is somewhat limited due to differences between contact lens e-commerce and other online retail industries. The fact that consumers must enter their prescriptions when they initially buy contact lenses from a particular retailer, for instance, may make state dependence higher in contact lens e-commerce than in other e-commerce product categories. Contact lens retailers, whether online or offline, resemble pharmacies in that they require consumers to present prescriptions from medical professionals before purchasing products and in that their product offerings are limited to a specific product category (e.g., contact lenses or prescription drugs). Last, my methodological framework for studying contact lens e-commerce may be straightforwardly applied to other e-commerce industries, although the assumption that consumer search for a fixed product across stores seems less plausible than in the contact lens setting.

Bibliography

Armstrong, Mark. 2017. "Ordered Consumer Search." *Journal of the European Economic Association* 15 (5): 989–1024. 10.1093/jeea/jvx017.

Chamberlain, Gary. 1980. "Analysis of covariance with qualitative data." *Review of Economic Studies* 47 (1): 225–238.

Choi, Michael, Anovia Yifan Dai, and Kyungmin Kim. 2018. "Consumer Search and Price Competition." *Econometrica* 86 (4): 1257–1281.

Clay, Karen, Ramayya Krishnan, and Eric Wolff. 2001. "Prices and price dispersion on the web: evidence from the online book industry." *Journal of Industrial Economics* 49 (4): 521–539.

Clemons, Eric K., Il-Horn Hann, and Lorin M. Hitt. 2002. "Price dispersion and differentiation in online travel: an empirical investigation." *Management Science* 48 (4): 534–549.

Collard-Wexler, Allan. 2013. "Demand fluctuations in the ready-mix concrete industry." *Econometrica* 81 (3): 1003–1037.

De Los Santos, Babur, Ali Hortaçsu, and Matthijs R. Wildenbeest. 2012. "Testing models of consumer search using data on web browsing and purchasing behavior." *American Economic Review* 102 (6): 2955–2980.

Dinerstein, Michael, Liran Einav, Jonathan Levin, and Neel Sundaresan. 2018. "Consumer Price Search and Platform Design in Internet Commerce." *American Economic Review* 108 (7): 1820–1859.

Dubé, Jean-Pierre, Günter Hitsch, and Peter E. Rossi. 2009. "Do switching costs make markets less competitive?" *Journal of Marketing Research* 46 (4): 435–445.

Dubé, Jean-Pierre, Günter Hitsch, and Peter E. Rossi. 2010. "State dependence and alternative explanations for consumer inertia." *RAND Journal of Economics* 41 (3): 417–445.

Gouriéroux, Christian, Alain Monfort, and Eric Renault. 1993. "Indirect inference." *Journal of Applied Econometrics* 8 (S1): S85–S118.

Hong, Han, and Matthew Shum. 2006. "Using price distributions to estimate search costs." *RAND Journal of Economics* 37 257–275.

Honka, Elisabeth. 2014. “Quantifying search and switching costs in the US auto insurance industry.” *RAND Journal of Economics* 45 (4): 847–884.

Honka, Elisabeth, and Pradeep Chintagunta. 2017. “Simultaneous or sequential? Search strategies in the U.S. auto insurance industry.” *Marketing Science* 36 (1): 21–42.

Honka, Elisabeth, Ali Hortaçsu, and Matthijs R. Wildenbeest. 2019. “Empirical search and consideration sets.” *Handbook of the Economics of Marketing* 1 193–257.

Hortaçsu, Ali, and Chad Syverson. 2004. “Product differentiation, search costs, and competition in the mutual fund industry: a case study of S&P 500 index funds.” *Quarterly Journal of Economics* 119 (2): 403–456.

Jolivet, Grégory, and Hélène Turon. 2019. “Consumer search costs and preferences on the internet.” *Review of Economic Studies* 86 (3): 1258–1300.

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

Koulayev, Sergei. 2014. “Search for differentiated products: identification and estimation.” *RAND Journal of Economics* 45 (3): 553–575.

Krasnokutskaya, Elena, and Katja Seim. 2011. “Bid preference programs and participation in highway procurement auctions.” *American Economic Review* 101 (6): 2653–2686.

Lee, Kwok Hao, and Leon Musolff. 2021. “Entry into two-sided markets shaped by platform-guided search.” Unpublished working paper.

Mehta, Nitin, Surendra Rajiv, and Kannan Srinivasan. 2003. “Price uncertainty and consumer search: a structural model of consideration set formation.” *Marketing Science* 22 (1): 58–84.

Moraga-González, José Luis, Zsolt Sándor, and Matthijs R. Wildenbeest. 2022. “Consumer search and prices in the automobile market.” *Review of Economics Studies*, Forthcoming.

Moraga-González, José Luis, and Matthijs R. Wildenbeest. 2008. “Maximum likelihood estimation of search costs.” *European Economic Review* 52 (5): 820–848.

Morozov, Ilya, Stephan Seiler, Xiaojing Dong, and Liwen Hou. 2021. “Estimation of preference heterogeneity in markets with costly search.” *Marketing Science* 40 (5): 871–899.

Mundlak, Yair. 1978. “On the pooling of time series and cross section data.” *Econometrica* 46 (1): 69–85.

Pakes, Ariel, Michael Ostrovsky, and Steven Berry. 2007. “Simple estimators for the parameters of discrete dynamic games (with entry/exit examples).” *RAND Journal of Economics* 38 (2): 373–399.

Weitzman, Martin L. 1979. "Optimal search for the best alternative." *Econometrica* 47 (3): 641–654.

Wooldridge, Jeffrey M. 2005. "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity." *Journal of Applied Econometrics* 20 (1): 39–54.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* Volume 1. The MIT Press, , 2nd edition.

Appendices

A Conditional dependence of store tastes and prices

In Section 4, I note that the prices that a consumer encounters and that consumer's tastes for stores will be correlated conditional on the consumer's initial state. The reason is that, conditional on the consumer having previously purchased from store f, higher prices at store f for the consumer's brand of contact lenses will be associated with stronger tastes for store f. This is because strong tastes for f are required to justify the consumer's history of purchasing from f despite its high prices.

I now consider empirical evidence for the conditional correlation described in the preceding paragraph. In particular, I consider the regression of an indicator for whether a consumer visits stores other than the store f of corresponding to the consumer's initial state on the relative price of f at the time that the consumer made the purchase that determined his initial state. To be clear, the regression equation is

$$\mathbb{1}\{i \text{ visits store other than } f \text{ in } t\} = \lambda_0 + \lambda_1 \left(p_{jf1}/\bar{p}_{j1}\right) + \epsilon_{it}$$

where j is consumer i's prescribed brand; p_{jf1} is f's price when i first purchased contact lenses in my sample; and $\bar{p}_{j1}$ is the mean price of j across 1800, WM, and VD at the time i first purchased contact lenses in my sample. I run the ordinary least squares regression on a dataset including all search efforts observed after their consumers' first purchases. I expect my estimate of λ_1 to be positive, as this would indicate that consumers with a history of purchasing from a relatively expensive store are less likely to even consider purchasing from other stores; this indicates that these consumers have strong preferences for the store from which they have historically bought contact lenses.

Appendix Table 14 provides the regression results. As expected, the estimate of λ_1 is positive. I take this as evidence of a positive correlation between store f's price and consumer i's tastes for store f conditional on the consumer having previously purchased from store f.

Table 14: Results for regression assessing conditional dependence of prices and store tastes

Parameter	Estimate	SE
Intercept	0.434	0.112
Slope	-0.227	0.109

Notes: the "SE" column provides asymptotic standard errors.

B Expressions for search effort outcome probabilities

This appendix provides chains of inequalities relating indirect and reservation utilities for every possible search effort outcome in my model. As explained in Section 4, I use these inequalities in computing conditional choice probabilities. Throughout this appendix, I suppress the brand j and search effort t subscripts.

First, consider the case in which consumer i visits only store f and then chooses the outside option. This corresponds to one of the following chains of inequalities:

$$\begin{aligned} r_{if} &\geq u_{i0} \geq u_{if} \vee \max_g r_{ig} \\ u_{i0} &\geq r_{if} \geq u_{if} \vee \max_g r_{ig} \\ u_{i0} &\geq u_{if} \geq r_{if} \vee \max_g r_{ig}. \end{aligned} \tag{10}$$

It is possible for the consumer to visit store f when the outside option's indirect utility exceeds f's reservation utility because, by assumption, the consumer must visit at least one store in a search effort. Under the distributional assumptions outlined in Section 4, the probability of the first chain of inequalities is

$$\frac{e^{\bar{r}_{if}}}{e^{\bar{u}_{i0}} + e^{\bar{u}_{if}} + \sum_{g=1}^{F} e^{\bar{r}_{ig}}} \times \frac{e^{\bar{u}_{i0}}}{e^{\bar{u}_{i0}} + e^{\bar{u}_{if}} + \sum_{g \notin \{0,f\}}^{F} e^{\bar{r}_{ig}}} \tag{11}$$

for $\bar{u}_{ig} = u_{ig} - \varepsilon_{ig}$ and $\bar{r}_{ig} = r_{ig} - \eta_{ig}$. The probability of the search effort outcome described above is the sum of the probabilities of the chains of inequalities in (10). I will not explicitly state any more choice probabilities, however, since they follow the same rank-order logit form as (11).

Now consider the case in which i buys from f after visiting f alone. The inequalities inducing this outcome are

$$\begin{aligned} r_{if} &\geq u_{if} \geq u_{i0} \vee \max_g r_{ig} \\ u_{if} &\geq r_{if} \geq u_{i0} \vee \max_g r_{ig} \\ u_{if} &\geq u_{i0} \geq r_{if} \vee \max_g r_{ig}. \end{aligned}$$

Now consider the case in which i visits f_1 and f_2 in that order, but does not buy from either firm. The inequality leading to this outcome is

$$r_{if_1} \geq r_{if_2} \geq u_{i0} \geq u_{i1} \vee u_{i2} \vee \max_{g \notin \{f_1, f_2\}} r_{ig}.$$

Now consider the case in which i visits f_1 and f_2 before buying from f_1. The inequality leading to this outcome is

$$r_{if_1} \geq r_{if_2} \geq u_{if_1} \geq u_{i0} \vee u_{if_2} \vee \max_{g \notin \{f_1, f_2\}} r_{ig}$$

Now consider the case in which i visits f_1 and f_2 before buying from f_2. The inequalities leading to this outcome are

$$r_{if_1} \geq r_{if_2} \geq u_{if_2} \geq u_{i0} \vee u_{if_1} \vee \max_{g \notin \{f_1, f_2\}} r_{ig}$$

$$r_{if_1} \geq u_{if_2} \geq r_{if_2} \geq u_{i0} \vee u_{if_1} \vee \max_{g \notin \{f_1, f_2\}} r_{ig}$$

$$u_{if_2} \geq r_{if_1} \geq r_{if_2} \geq u_{i0} \vee u_{if_1} \vee \max_{g \notin \{f_1, f_2\}} r_{ig}.$$

Now consider the case in which i visits f_1, f_2, and f_3 (in that order) but does not buy from any seller. The inequality leading to this outcome is

$$r_{if_1} \geq r_{if_2} \geq r_{if_3} \geq u_{i0} \geq \max_{1 \leq j \leq 3} u_{if_j} \vee \max_{g \notin \{f_1, f_2, f_3\}} r_{ig}.$$

Now consider the case in which i visits f_1, f_2, and f_3 (in that order) and buys from firm f_1. The inequalities leading to this outcome are

$$r_{if_1} \geq r_{if_2} \geq r_{if_3} \geq u_{if_1} \geq u_{i0} \vee \max_{2 \leq j \leq 3} u_{if_j} \vee \max_{g \notin \{f_1, f_2, f_3\}} r_{ig}.$$

Now consider the case in which i visits f_1, f_2, and f_3 (in that order) and buys from firm f_2. The inequalities leading to this outcome are

$$r_{if_1} \geq r_{if_2} \geq r_{if_3} \geq u_{if_2} \geq u_{i0} \vee \max_{j \in \{1,3\}} u_{if_j} \vee \max_{g \notin \{f_1, f_2, f_3\}} r_{ig}.$$

Now consider the case in which i visits f_1, f_2, and f_3 (in that order) and buys from firm f_3. The inequalities leading to this outcome are

$$r_{if_1} \geq r_{if_2} \geq r_{if_3} \geq u_{if_3} \geq u_{i0} \vee \max_{j \in \{1,2\}} u_{if_j} \vee \max_{g \notin \{f_1, f_2, f_3\}} r_{ig}$$

$$r_{if_1} \geq r_{if_2} \geq u_{if_3} \geq r_{if_3} \geq u_{i0} \vee \max_{j \in \{1,2\}} u_{if_j} \vee \max_{g \notin \{f_1, f_2, f_3\}} r_{ig}$$

$$r_{if_1} \geq u_{if_3} \geq r_{if_2} \geq r_{if_3} \geq u_{i0} \vee \max_{j \in \{1,2\}} u_{if_j} \vee \max_{g \notin \{f_1, f_2, f_3\}} r_{ig}$$

$$u_{if_3} \geq r_{if_1} \geq r_{if_2} \geq r_{if_3} \geq u_{i0} \vee \max_{j \in \{1,2\}} u_{if_j} \vee \max_{g \notin \{f_1, f_2, f_3\}} r_{ig}.$$

C Details of indirect-inference estimation

C.1 Structure of regressions underlying the I-I estimator

Let $Y_n = \{y_{it}\}_{i=1}^{n}$ denote the collection of search effort outcomes in the estimation sample, where $y_i = \{y_{it}\}_{t=1}^{T_i}$ and y_{it} is a vector of search outcomes for consumer i in search effort t (i.e., the sequence of stores that consumer i visited in search effort t and consumer i's purchase decision in search effort t). Next, let $X_n = \{x_i\}_{i=1}^{n}$ denote the collection of explanatory variables in the estimation sample, where $x_i = \{x_{it}\}_{t=1}^{T_i}$ and x_{it} is a vector including the prices for consumer i's prescribed brand of contact lenses during search effort t as well as the consumer's state during search effort t.[16] The statistic $\hat{\beta}_n$ is the value of β minimizing the criterion function

$$Q_n(Y_n, X_n, \beta) = \frac{1}{n}\sum_{i=1}^{n} g(y_i, x_i, \beta).$$

where

$$g(y_i, x_i, \beta) = \sum_{j=1}^{J}\sum_{t=1}^{T_i} w_{ijt}(y_{it,j} - x'_{it,j}\beta_k)^2.$$

Under this form of the g function, the value of β minimizing the auxiliary criterion function is the vector obtained by stacking J weighted least squares estimators, each computed on a dataset of search efforts. Each j corresponds to a distinct regression, and each $y_{it,j}$ is some scalar-valued transformation of y_{it} that is used as the dependent variable in the jth regression. Similarly, each $x_{it,j}$ is some vector-valued transformation of x_{it} that is used as the regressor vector in the jth regression. The weights w_{ijt} will generally depend on the data (y_i, x_i).

Consider, for the sake of illustration, the regression j corresponding to the share of search efforts in which a consumer in state $h_{ift} = 1$ visits store g. In this case, $y_{it,j}$ is an indicator for whether consumer i visited store g in search effort f, $x_{it,j} = 1$, and w_{ijt} is an indicator for whether consumer i's state at search effort t was $h_{ift} = 1$.

The auxiliary model statistics computed on data that are simulated under structural model parameter θ are defined by

$$\tilde{\beta}_n^H(\theta) = \arg\min_{\beta\in B} Q_{nH}(\tilde{Y}_n^H(\theta), \tilde{X}_n^H, \beta).$$

Here, H is the number of simulates, $\tilde{Y}_n^H(\theta)$ are outcome variables simulated under θ conditional on $\tilde{X}_n^H$, and $\tilde{X}_n^H$ is constructed by repeating X_n H times.

[16]This is a minor abuse of notation, since I use y_i and x_i to signify subtly different random elements in the main structural model and in the auxiliary model. The x_i appearing in my exposition of the structural model, for instance, excludes the consumer's state.

C.2 Optimal weighting matrix

The asymptotic normality of the I-I estimator is ensured by conditions that are standard in the I-I literature.[17] Recall that the I-I estimator is defined by

$$\hat{\theta}_n^H(\Omega) = \arg\min_{\theta\in\Theta}(\hat{\beta} - \tilde{\beta}_n^H(\theta))'\hat{\Omega}_n(\hat{\beta} - \tilde{\beta}_n^H(\theta)).$$

The asymptotic normality result for the I-I estimator is

$$\sqrt{n}(\hat{\theta}_n^H(\Omega) - \theta_0) \to_d N\left(0, V_{\hat{\theta}_n^H}(\Omega)\right)$$

where

$$V_{\hat{\theta}_n^H}(\Omega) = (B_0'\Omega B_0)^{-1}B_0'\Omega\Gamma_0^{-1}V_{\hat{\beta}}\Gamma_0^{-1}\Omega B_0(B_0'\Omega B_0)^{-1}$$

for

$$\begin{aligned}
V_{\hat{\beta}} &= \mathrm{Var}\left(s_{i0} - \frac{1}{H}\sum_{h=1}^{H} s_{ih}\right) \\
s_{ih} &= \begin{cases} \frac{\partial g}{\partial \beta}(y_i, x_i, \beta_0), & h = 0, \\ \frac{\partial g}{\partial \beta}(\tilde{y}_i^h(\theta_0), x_i, \beta_0), & h \in \{1, \ldots, H\} \end{cases} \\
\Gamma_0 &= \frac{\partial^2 Q}{\partial\beta\partial\beta}(\beta_0; \theta_0) \\
B_0 &= \frac{\partial b}{\partial \theta}(\theta_0).
\end{aligned}$$

In the definitions above, $\tilde{y}_i^h(\theta_0)$ are search effort outcomes simulated under model parameters θ_0 and $Q(\beta;\theta)$ is the population criterion function, i.e., the uniform probability limit of $Q_n(Y_n, X_n, \beta)$ as $n \to \infty$ when (Y_n, X_n) are generated under the model with structural parameter θ. Also, the binding function

$$b(\theta) = \arg\min_{\beta\in B} Q(\beta;\theta)$$

is the probability limit of the $\hat{\beta}$ parameters under a given vector of structural parameters θ. Last, $\beta_0 = b(\theta_0)$.

The optimal weighting matrix Ω^* is

$$\Omega^* = \Gamma_0 V_{\hat{\beta}}^{-1}\Gamma_0,$$

which yields

$$V_{\hat{\theta}_n^H}(\Omega^*) = \left(B_0'\Gamma_0 V_{\hat{\beta}}^{-1}\Gamma_0 B_0\right)^{-1}.$$

[17]See Gouriéroux et al. (1993) for details.

I estimate the optimal weighting matrix and asymptotic variance of my estimator by replacing population objects appearing in expressions above with their sample analogues. Additionally, as is standard in the estimation of optimal weighting matrices in generalized method of moments and I-I estimators, I replace the true value of the structural parameter θ_0 with $\hat{\theta}_n^H(I)$ in the expression for the optimal weighting matrix when estimating this weighting matrix; here, I is the identity matrix.

C.3 Auxiliary statistics

The following list describes the auxiliary statistics that I include in $\hat{\beta}$ in estimating my model.

(i) *Stores' visit shares*: For each store f, I compute the mean across search efforts of an indicator for whether the consumer visited store f in search effort t. These statistics are intended to target the estimation of stores' qualities q_f and the search cost parameter $\bar{\kappa}$.

(ii) *Consideration set size*: I compute the mean across search efforts of an indicator for whether the consumer visited all available stores in search effort t. This statistic is intended to target the estimation of the search cost parameter $\bar{\kappa}$.

(iii) *Inertia*: I regress indicator I regress an indicator for whether consumer i's search effort t involved a visit to store f on store indicators, an indicator for whether the consumer purchased from store f in search effort $t-1$, and an indicator for whether the consumer purchased from store f in search effort $t-2$. The dataset used for running this regression includes three observations for each search effort, one corresponding to each of the stores. I include all observations for which t exceeds three in the regression. This statistic is intended to target the estimation of the state dependence parameter ϕ and the parameters governing the distribution of consumers' persistent tastes γ_i.

(iv) *Role of lagged price*: I regress an indicator for whether a search effort t ended in a transaction at 1800 on the price of the consumer's brand at 1800 during search effort t and the price of that brand at 1800 during the consumer's previous search effort $t-1$. This statistic is intended to target the estimation of the state dependence parameter ϕ and the parameters governing the distribution of consumers' persistent store tastes γ_i.

(v) *Price sensitivity*: I regress an indicator for whether a search effort t ended in a transaction at store f on store indicators and the price of the consumer's brand at store f. The dataset used for running this regression includes three observations for each search effort, one corresponding to each of the stores. This statistic is intended to target the estimation of the price sensitivity parameter α and the store qualities q_f.

(vi) *Cross-visiting behaviour*: For each pair of distinct stores (f, g), I compute the mean across search efforts t in which the consumer's state is given by $h_{igt} = 1$ of an indicator for whether the search effort involved a visit to store f. This statistic is intended to target the estimation of the parameters Γ_{fg}, which govern the mean tastes of consumers for each store conditional on each initial state.

(vii) *Dependence of tastes and prices conditional on initial state*: I regress an indicator for whether consumer i visited store g in search effort t on the ratio of the price of the consumer's brand at the store f for which $h_{if1} = 1$ to the average price of the consumer's brand across the three major stores. In doing so, I use the prices from the time at which the consumer made his first observed purchase. I use each store $g \neq f$ in the regressions. This statistic is intended to target the estimation of λ.

Table 15: Auxiliary model statistics computed on estimation sample

Statistic	Value	SE
Share visiting 1800	0.688	0.014
Share visiting WM	0.145	0.010
Share visiting VD	0.360	0.014
Share visiting every store	0.013	0.003
Inertia: indicator for 1800	0.309	0.011
Inertia: indicator for VD	0.115	0.010
Inertia: indicator for WM	0.149	0.011
Inertia: purchased from store last search effort	0.495	0.017
Inertia: purchased from store two search efforts ago	0.392	0.018
Role of lagged price: slope for current price	-0.351	0.252
Role of lagged price: slope for lagged price	0.023	0.240
Price sensitivity: indicator for WM	0.181	0.071
Price sensitivity: indicator for VD	0.377	0.064
Price sensitivity: slope	-0.155	0.070
Cross-visiting behaviour: share of 1800 buyers visiting WM	0.116	0.009
Cross-visiting behaviour: share of 1800 buyers visiting VD	0.033	0.005
Cross-visiting behaviour: share of WM buyers visiting 1800	0.308	0.014
Cross-visiting behaviour: share of WM buyers visiting VD	0.128	0.010
Cross-visiting behaviour: share of VD buyers visiting 1800	0.193	0.012
Cross-visiting behaviour: share of VD buyers visiting WM	0.124	0.010
Dep. of tastes and prices cond. on initial state: slope	-0.302	0.098

Notes: See Section 6 for a description of the various auxiliary model statistics. The "SE" column reports classical asymptotic standard errors computed under an assumption of homoskedasticity. I do not use the estimated coefficient for the indicator for 1800 in the price sensitivity regression as an auxiliary statistic in my indirect inference estimation. I similarly do not use the intercept estimated in the "Dependence of tastes and prices conditional on initial state" regression.

D Dynamic pricing model

In addition to the analysis of static pricing in the main text, I additionally study online retailers' pricing in a dynamic framework. My approach to studying dynamic pricing in a setting with state dependence follows that of Dubé et al. (2009), who provide additional information on the properties of the general dynamic pricing model that their paper proposes and that I amend to my setting in this paper.

I analyze a model of online retailers' dynamic pricing using a Markov perfect equilibrium (MPE) solution concept. In the MPE that I consider, firms' pricing strategies maximize their payoffs subject to the constraint that their strategies condition only on information relevant to contemporaneous payoffs. This information includes the share of consumers with each value of unobserved heterogeneity γ_i that belong to each state (i.e., whose previous purchase was from each seller). It is not computationally feasible to find an MPE in a setting in which γ_i is continuously distributed; therefore, I compute MPE in a simplified version of the model in which γ_i takes on one of K support points in $\mathcal{G}$. Let $x_{f\tau}(\gamma)$ denote the share of consumers of type $\gamma \in \mathcal{G}$ whose previous purchase in time τ was made at store f, let $\mathcal{F}$ be the collection of all competing online retailers, and let $x_\tau = \{x_{f\tau}(\gamma) : f \in \mathcal{F}, \gamma \in \mathcal{G}\}$. Following the standard terminology used in dynamic programming, I refer to x_τ as the *state* at risk of causing confusion with the consumer's state h_i as defined in Section 4.

Firm f's payoffs in my dynamic pricing model are the firm's present discounted profits. When players use strategies $p^* : x_\tau \mapsto p_f$, these payoffs are

$$\sum_{\tau=0}^{\infty} \beta^\tau \sum_{\gamma \in \mathcal{G}} \mu(\gamma) \sum_g x_g(\gamma) \sigma_{fg}(p^*(x_\tau), \gamma)(p_f^*(x_\tau) - mc),$$

where β is a discount factor shared by all competing firms, $\mu(\gamma)$ is the share of consumers of type γ, and mc is firm f's marginal cost of providing a consumer with a box of contact lenses. I assume that firms share a marginal cost mc.

The Bellman equation associated with firm f's dynamic programming problem is

$$V_f(x) = \max_{p_f \geq 0} \left[\sum_{\gamma \in \mathcal{G}} \mu(\gamma) \sum_g x_g(\gamma) \sigma_{fg}(p_f, p_{-f}^*(x), \gamma)(p_f - mc) + \beta V_f(Q(x, p_f, p_{-f}^*(x))) \right]. \tag{12}$$

The function Q appearing in (12) is the state transition function, which provides the next period's state given the contemporary state x and firms' prices p. The state transition is deterministically determined by consumers' choice probabilities conditional on their type γ_i, their state h_i, and prices p. A MPE is a pricing strategy function $p^* : x \mapsto p$ and an associated value function V_f for each firm f that solves the Bellman equation (12).

Implementation. To limit the size of the state space of the dynamic programming problem that I solve in finding equilibria, I remove Walmart from the market in computing these equilibria. Thus, I consider competition between the two largest online contact lens retailers: 1800 and VD. Solving for equilibria of the dynamic pricing game requires a finitely supported distribution of unobserved heterogeneity γ_i, a marginal cost mc, and a discount factor β. To obtain a finitely supported distribution of γ_i, I follow Dubé et al. (2009) in clustering consumers into a finite number of types. My clustering procedure involves (i) taking

Table 16: Percentage changes in markups from dynamic pricing model

Store	Low search costs	No state dependence	No persistent unobs.
1800	-1.7	-0.6	-22.4
VD	-0.6	-6.1	-29.9

2000 draws from my estimated unconditional distribution of γ_i and (ii) performing K-means clustering on these draws. I use the cluster centroids as the members of γ_i's support, and I use the share of observations in each cluster as the corresponding population shares $\mu(\gamma)$ of the support points γ. Additionally, I use $K = 3$ clusters. I use information from 1-800 Contacts's quarterly report for the second quarter of 2007 to obtain a marginal cost mc. In particular, I divide the price of Acuvue Advance for Astigmatism—which is the brand on which I focus in my analysis of online retailers' pricing—at 1800 in the first week of 2007 by the ratio of net sales to costs of goods and services (COGS) for January 1–June 30, 2007 as reported on 1800's quarterly report.[18] This approach applies 1800 overall markup ratio as defined in the preceding paragraph to a particular product's price to obtain an estimate of that product's marginal cost. Last, I set the discount factor β to 0.95.

Results. Table 16 provides the results of the analysis. In particular, it provides percentage changes in steady-state markups under counterfactual consumer preferences. Following Dubé et al. (2009), I compute steady-state markups by simulating an equilibrium price path from an arbitrary initial state until firms' prices converge. The initial state that I use is one in which no consumers are loyal to any online store. These results reported by Table 16 largely accord with those obtained using a static pricing model: equilibrium markups are largely unaffected by a reduction in search costs, but markedly decrease upon an elimination of persistent unobserved heterogeneity that horizontally differentiates sellers and, to a lesser extent, upon an elimination of state dependence.

[18]Net sales and COGS were $125,202,000 and $73,962,000, respectively, in this time period. The ratio of these values is 1.69.

E Supplemental tables and figures

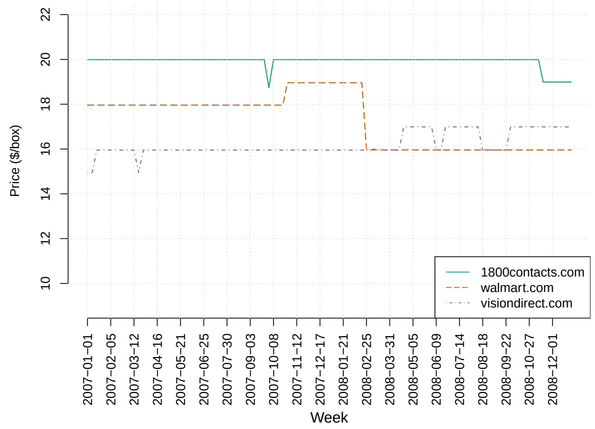
Table 17: Decomposition of price variation

Type of variation	Std. dev
Interbrand	12.26
Interstore	3.91
Intertemporal	1.15

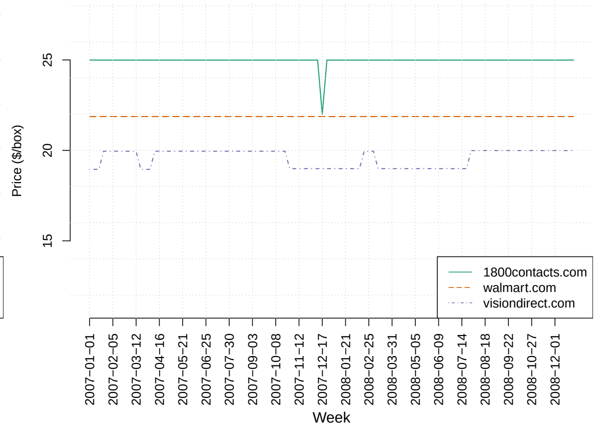
Notes: "Interbrand" provides the cross-brand standard deviation of brands' average transaction prices. "Interstore" provides the average standard deviation of a brand's price across stores, where the average is taken over transactions in the sample. "Intertemporal" provides the average standard deviation of a particular brand's price at a particular store, where the average is taken across both brands and stores.

Figure 5: Prices of contact lenses across stores, brands, and time

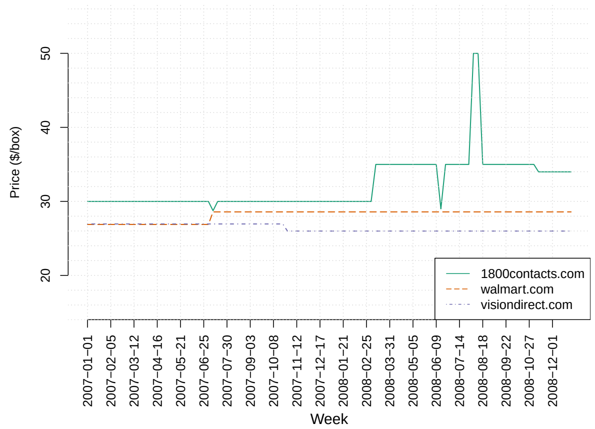
(a) Acuvue 2



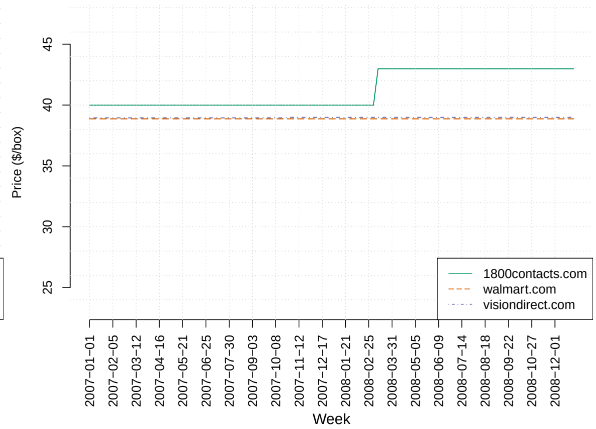
(b) Acuvue Advance



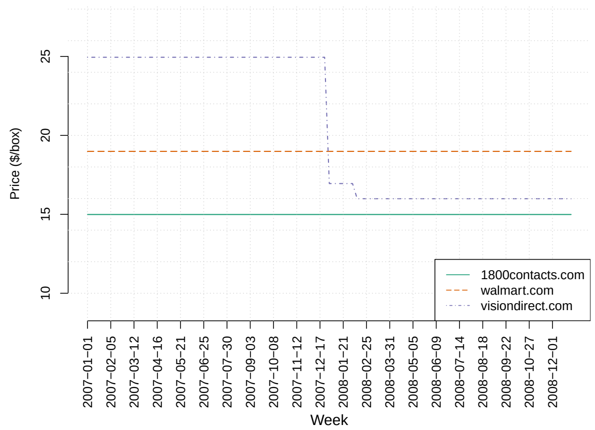
(c) Acuvue Oasys



(d) Acuvue Advance



(e) Biomedics



(f) Freshlook Colorblends

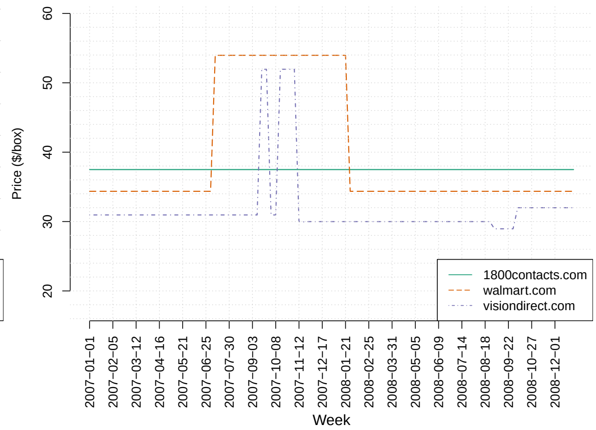


Table 18: Model fit and counterfactual search patterns: full results

Spec.	Share visiting one store only	Mean # of visits	Share buying from...			Visit order	Share paying > > min. price	Mean over-payment ($)
			any	1800	VD			
Observed	0.819	1.196	0.610	0.364	0.220	0.496	0.660	3.95
	-	-	-	-	-	-	-	-
Baseline	0.823	1.193	0.733	0.479	0.214	0.419	0.717	4.45
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
Low search costs	0.726	1.312	0.727	0.473	0.212	0.417	0.716	4.43
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
Low search costs (comp.)	0.724	1.314	0.733	0.476	0.214	0.417	0.716	4.42
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
No state dep.	0.768	1.257	0.594	0.388	0.166	0.455	0.722	4.46
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
No state dep. (comp.)	0.705	1.334	0.733	0.472	0.204	0.459	0.720	4.42
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
No vertical diff.	0.644	1.435	0.928	0.409	0.361	0.635	0.617	3.45
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
No vertical diff. (comp.)	0.710	1.339	0.733	0.297	0.325	0.602	0.581	3.22
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
No persistent unobs.	0.661	1.428	0.231	0.058	0.148	0.464	0.387	1.49
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
No persistent unobs. (comp.)	0.330	1.978	0.733	0.203	0.432	0.431	0.442	1.73
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
No search	0.000	3.000	0.716	0.462	0.206	1.000	0.717	4.42
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)
Logit only (comp.)	0.000	3.000	0.733	0.179	0.326	1.000	0.531	2.27
	(0.285)	(0.635)	(0.157)	(0.135)	(0.083)	(0.205)	(0.113)	(1.09)

Notes: This table expands upon Table 12 by adding rows corresponding to additional counterfactual parameters and also by including standard errors obtained by a parametric bootstrap with 100 replicates. The rows "Low search costs," "No state dependence," and "No persistent unobs." all report results for the counterfactual discussed in Section 8 with the exception that no adjustment is made to the value of the outside option to ensure that the share purchasing from any store is held fixed in the counterfactual. The "No search" row reports results for a counterfactual in which consumer i knows each ε_{ijft} without searching and is able to purchase from any store without having visited that store.