

Online Appendix for “Sources of limited consideration and market power in e-commerce” *

Michael Sullivan

Department of Economics, University of Western Ontario

June 27, 2024

O.1 Data

This appendix describes the data for books, electronics, and DVD e-commerce used in the paper’s descriptive analyses. The books that I analyze 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 book titles. Across these titles, 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; 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 websites¹ and “other” stores, which includes the other online bookstores with many

*Email address: msulli65@uwo.ca. This project draws on research supported by the Social Sciences and Humanities Research Council.

¹Namely, doubledatbookclub.com, mysteryguild.com, literaryguild.com, and eharlequin.com.

fewer recorded sales than Amazon and Barnes & Noble.² 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.³ 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.

In computing the figures in Table O.11 for books, the price for “book club” is set to the price at the book club with the highest sales for the consumer’s purchased book (or, in the case that the consumer bought from a different book club, the price at this club). I proceed analogously for the “other” store. For categories including eBay as a retailer, I do not include eBay’s price in computing the minimum given that items sold on eBay are often in used condition.

O.2 Within versus across store search

The fact that consumers require brand-specific prescriptions to purchase contact lenses makes the assumption that search occurs for a fixed product across retailers credible in my setting.⁴ This assumption may not hold in other product categories. To illustrate, Table O.1 describes search within Amazon and Barnes & Noble’s online store in search efforts that result in a book purchase. The median number of pages viewed during a visit to Amazon (Barnes & Noble) is 39 (27), and that the median time spent browsing Amazon (Barnes & Noble) is about half an hour (16 minutes). Most consumers do not visit multiple sites before buying a book online; see Table O.10. This suggests that consumers search within sites—possibly over distinct products—while conducting little cross-store search. Consumers search much less within contact lenses sites: as reported by Appendix Table O.12, the median number of pages is 12 and the median duration is 6 minutes (pooling across retailers). This is consistent with brand-specific

²The “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.

³The 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.

⁴This assumption is credible in other settings, including the textbooks setting of Hong and Shum (2006) when students must buy a particular textbook to satisfy a course requirement.

prescriptions that eliminate the rationale for within-store, across-brand search.

Table O.1: Within-site search intensity prior to book purchase

(a) Amazon visits					(b) Barnes & Noble visits				
Measure	Mean	Percentile			Measure	Mean	Percentile		
		25 th	50 th	75 th			25 th	50 th	75 th
# pages	102.0	16	39	105	# pages	71.4	10	27	80
Duration	79.9	10	29	79	Duration	48.5	5	16	62

Note: this table reports summary statistics—the mean, 25th percentile, 50th percentile, and 75th percentile—of the number of pages viewed during a visit to an online bookstore and of the duration of time spent (in minutes) browsing the online bookstore.

O.3 Expressions for search effort outcome probabilities

This appendix provides chains of inequalities relating indirect and reservation utilities for every possible search effort outcome in the model for an arbitrary number of retailers F . Let $\mathcal{F} = \{1, 2, \dots, F\}$. I suppress consumer i , brand j , and search effort t subscripts for notational brevity here as they do not play a role in the analysis. Let $\tilde{f} = \{f_1, f_2, \dots, f_k\}$ be the sequence of stores visited by the consumer, and let f^* be the store from which the consumer makes a purchase. Let $\bar{r} = \max_{f \in \mathcal{F} \setminus \tilde{f}} r_f$ be the greatest reservation utility among all non-searched stores. Let $\bar{u} = \max_{f \in \tilde{f} \setminus \{f^*\}} u_f$ be the indirect utilities of all searched stores excluding the store of ultimate purchase.

Consider first the case in which \tilde{f} includes only a single store: $\tilde{f} = \{f_1\}$. The consumer's choice to visit f_1 and then choose the outside option corresponds to one of the following chains of inequalities:

$$\begin{aligned}
r_{f_1} &\geq u_0 \geq u_{f_1} \vee \bar{r} \\
u_0 &\geq r_{f_1} \geq u_{f_1} \vee \bar{r} \\
u_0 &\geq u_{f_1} \geq r_{f_1} \geq \bar{r}.
\end{aligned} \tag{1}$$

It is possible for the consumer to visit store f_1 when the outside option's indirect utility exceeds f_1 's reservation utility because, by assumption, the consumer must visit at least one store. The first chain of inequalities above corresponds to the case in which the consumer would have searched without a requirement to visit at least one store, but only wishes to visit store f_1 . The next two chains of inequalities correspond to the case in which the consumer would not have searched without this requirement,

and then finds an indirect utility that falls below that of the outside option. Under the article's maintained distributional assumptions, the probability of the first chain of inequalities is

$$\frac{e^{\bar{r}_{f_1}}}{e^{\bar{u}_0} + e^{\bar{u}_{f_1}} + \sum_{g=1}^F e^{\bar{r}_g}} \times \frac{e^{\bar{u}_0}}{e^{\bar{u}_0} + e^{\bar{u}_{f_1}} + \sum_{g \notin \{0, f_1\}}^F e^{\bar{r}_g}} \quad (2)$$

for $\bar{u}_g = u_g - \varepsilon_g$ and $\bar{r}_g = r_g - \eta_g$. The probability of the search effort outcome described above is the sum of the probabilities of the chains of inequalities in (1). I will not explicitly state any more choice probabilities, however, since they follow the same rank-order logit form as (2).

Now consider the case in which i buys from f_1 after visiting f_1 alone. The inequalities corresponding to this outcome are

$$\begin{aligned} r_{f_1} &\geq u_{f_1} \geq u_0 \vee \bar{r} \\ u_{f_1} &\geq r_{f_1} \geq u_0 \vee \bar{r} \\ u_{f_1} &\geq u_0 \geq r_{f_1} \geq \bar{r}. \end{aligned}$$

Next consider the case in which \tilde{f} includes more than one alternative. There are two cases to consider. The first is that in which the consumer buys from $f^* = 0$ (the outside option) or $f^* \in \tilde{f} \setminus \{f_k\}$ (any store except the last visited). The chain of inequalities corresponding to this case is

$$r_{f_1} \geq r_{f_2} \geq \cdots \geq r_{f_{k-1}} \geq r_{f_k} \geq u_{f^*} \geq \bar{u} \vee \bar{r}.$$

Here, the fact that $r_{f_j} \geq r_{f_{j+1}}$ for $j \in \{1, \dots, k-1\}$ is a necessary and sufficient condition for f_j being visited before f_{j+1} under the optimal sequential search strategy (conditioning on the consumer visiting both stores). The fact that $r_{f_k} \geq u_{f^*} \geq \bar{u} \vee \bar{r}$ is a necessary and sufficient condition for the consumer to terminate search after visiting f_k and to purchase from f^* . The fact that $r_{f_j} \geq u_{f^*}$ for all $j \leq k$ ensures that the consumer does not terminate search until visiting store f_k . Last, $u_{f^*} \geq \bar{u}$ ensures that the consumer purchases from f^* rather than from any other seller.

The second case to consider is that in which the consumer purchases from f_k , the final store visited. Above, it was not possible for the indirect utility u_{f_j} of seller f_j with

$j < k$ to exceed the reservation utility of any visited store: in that case, the consumer would have stopped to purchase the good rather than continue searching until f_k . In the second case described above, though, this is possible; as an implication, there are multiple chains of inequalities corresponding to outcomes covered by the second case. These chains of inequalities are

$$\begin{aligned}
r_{f_1} &\geq r_{f_2} \geq \cdots \geq r_{f_{k-1}} \geq r_{f_k} \geq u_{f_k} \geq \bar{u} \vee \bar{r} \\
r_{f_1} &\geq r_{f_2} \geq \cdots \geq r_{f_{k-1}} \geq u_{f_k} \geq r_{f_k} \geq \bar{u} \vee \bar{r} \\
r_{f_1} &\geq r_{f_2} \geq \cdots \geq u_{f_k} \geq r_{f_{k-1}} \geq r_{f_k} \geq \bar{u} \vee \bar{r} \\
&\vdots \\
r_{f_1} &\geq u_{f_k} \geq \cdots \geq r_{f_{k-1}} \geq r_{f_k} \geq \bar{u} \vee \bar{r} \\
u_{f_k} &\geq r_{f_1} \geq \cdots \geq r_{f_{k-1}} \geq r_{f_k} \geq \bar{u} \vee \bar{r}.
\end{aligned}$$

An alternative and perhaps simpler characterization of these inequalities is

$$\begin{aligned}
r_{f_1} &\geq r_{f_2} \geq \cdots \geq r_{f_{k-1}} \geq r_{f_k} \geq \bar{u} \vee \bar{r} \\
&\text{AND} \\
u_{f_k} &\geq \bar{u} \vee \bar{r}.
\end{aligned}$$

The first chain of inequalities provides a sufficient and necessary condition for the consumer to search until f_k in the order given by \tilde{f} . The second inequality (after “AND”) is a necessary and sufficient condition for the consumer to purchase from f_k rather than any other retailer and to terminate search after visiting f_k .

O.4 Structure of regressions underlying the I-I estimator

Let $Y_n = \{y_i\}_{i=1}^n$ 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 .⁵ The

⁵This 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

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

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

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 j th regression. Similarly, each $x_{it,j}$ is some vector-valued transformation of x_{it} that is used as the regressor vector in the j th 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.

O.5 Optimal weighting matrix and inference

The asymptotic normality of the I-I estimator is ensured by conditions that are standard in the I-I literature.⁶ Recall that the I-I estimator is defined by

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

structural model, for instance, excludes the consumer's state.

⁶See Gouriéroux et al. (1993) for details.

The asymptotic normality result for the I-I estimator is

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

where

$$V_{\hat{\theta}_n^H}(\Omega) = (B_0' \Omega B_0)^{-1} B_0' \Omega \Gamma_0^{-1} V_{\hat{\beta}} \Gamma_0^{-1} \Omega B_0 (B_0' \Omega B_0)^{-1} \quad (3)$$

for

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

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 $n \rightarrow \infty$ 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 $\hat{\beta}_n$ parameters under a given vector of structural parameters θ . Last, $\beta_0 = b(\theta_0)$.

The optimal weighting matrix Ω^* is

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

A practical problem arises in calculating an empirical analogue of the optimal weighting matrix when $V_{\hat{\beta}}$ is singular or close to singular. Due to this problem, I instead use an approximately optimal weighting matrix

$$\Omega^{\text{approx}} = \Gamma_0 \left(V_{\hat{\beta}} + \epsilon I \right)^{-1} \Gamma_0,$$

where ϵ is a small number (I use $\epsilon = 0.1$) and I is the identity matrix. Note that this approximation does not affect the estimator's consistency, as the I-I estimator is consistent for any positive semi-definite weighting matrix.

I estimate the weighting matrix Ω^{approx} and asymptotic variance $V_{\hat{\theta}_n^H}(\Omega)$ by replacing population objects appearing in expressions above with their sample analogues. My estimation of the asymptotic variance is based on (3). Estimation of B_0 follows from estimation of Γ_0 and

$$\Lambda_0 = \Gamma_0 B_0 = -\frac{\partial^2 Q}{\partial \beta \partial \theta'}(\beta_0; \theta_0),$$

which implies $B_0 = \Gamma_0^{-1} \Lambda_0$. The estimates $\hat{\Gamma}_0$ and $\hat{\Lambda}_0$ of Γ_0 and Λ_0 , respectively, are based on second- and cross-partial numerical derivatives of Q_n with respect to β and θ as evaluated at $(\hat{\beta}_n, \hat{\theta})$. My estimate $\hat{V}_{\hat{\beta}}$ of $V_{\hat{\beta}}$ is the sample variance of $s_{i0} - (1/H) \sum_h s_{ih}$ as evaluated at $(\hat{\beta}, \hat{\theta})$. Estimating the weighting matrix Ω^{approx} requires a preliminary consistent estimate $\hat{\theta}$ of θ_0 ; I use an estimate $\hat{\theta}$ obtained upon setting the weighting matrix equal to the identity matrix. When I later estimate Γ_0 , Λ_0 , and $V_{\hat{\beta}}$ as inputs in the estimation of the asymptotic variance $V_{\hat{\theta}_n^H}(\Omega^{\text{approx}})$, I set $\hat{\theta}$ equal to the estimate obtained under my estimate $\hat{\Omega}^{\text{approx}}$ of Ω^{approx} (see below). The estimate of B_0 that I use in estimating the asymptotic variance $\hat{V}_{\hat{\theta}_n^H}(\Omega^{\text{approx}})$ is $\hat{B}_0 = \hat{\Gamma}_0^{-1} \hat{\Lambda}_0$. Last, I obtain estimates of Ω^{approx} and $V_{\hat{\theta}_n^H}(\Omega^{\text{approx}})$ by substituting the estimates of the Γ_0 , B_0 , and $V_{\hat{\beta}}$ matrices described above in for the true values of these matrices:

$$\begin{aligned} \hat{\Omega}^{\text{approx}} &= \hat{\Gamma}_0 \left(\hat{V}_{\hat{\beta}} + \epsilon I \right)^{-1} \hat{\Gamma}_0 \\ \hat{V}_{\hat{\theta}_n^H}(\Omega^{\text{approx}}) &= (\hat{B}_0' \hat{\Omega}^{\text{approx}} \hat{B}_0)^{-1} \hat{B}_0' \hat{\Omega}^{\text{approx}} \hat{\Gamma}_0^{-1} \hat{V}_{\hat{\beta}} \hat{\Gamma}_0^{-1} \hat{\Omega}^{\text{approx}} \hat{B}_0 (\hat{B}_0' \hat{\Omega}^{\text{approx}} \hat{B}_0)^{-1} \end{aligned}$$

O.6 Responses to price variation

To assess the relative contributions of cross-brand and intertemporal price variation to the price coefficient estimates in Table 6, 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

$$\mathbb{1}\{t \text{ results in purchase from 1800}\} = \beta_j - \alpha \log \left(\frac{p_{j,1800,t}}{\bar{p}_{jt}} \right) + \varepsilon_t, \quad (4)$$

Table O.2: 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

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 O.2 provides estimates of (4) 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 (4) 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 (4), 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.

O.7 Conditional dependence of store tastes and prices

Consider a consumer i who makes multiple search efforts across time and whose initial observed purchase is from store f_i^{initial} . I expect that, conditional on f_i^{initial} , the prices that consumer i faces and the consumer's unobserved tastes γ_i will be correlated. This is because a consumer who buys from a store f_i^{initial} despite its high prices requires strong positive tastes for store f_i^{initial} to rationalize buying from f_i^{initial} .

To empirically assess this conditional correlation, I regress an indicator for whether a consumer visits stores other than the store f_i^{initial} on the relative price of f_i^{initial} at the time that the consumer made the purchase that determined this state. The regression

equation is

$$\mathbb{1}\{i \text{ visits store other than } f_i^{\text{initial}} \text{ in } t\} = \lambda_0 + \lambda_1 (p_{j, f_i^{\text{initial}}, 1} / \bar{p}_{j1}) + \epsilon_{it}$$

where j is consumer i 's brand; $p_{j, f_i^{\text{initial}}, 1}$ is the price of f_i^{initial} when i first purchased lenses in the sample; and \bar{p}_{j1} is the mean price of j across 1800, WM, and VD at the time i first purchased lenses in the sample. I run the regression on a dataset including all search efforts excluding those of consumers' first purchases. A negative λ_1 estimate would indicate that consumers who bought from a relatively expensive store are less likely to consider other stores, which would indicate that these consumers have strong preferences for the store from which they historically bought contact lenses. Appendix Table O.3 provides the regression results. As expected, the estimate of λ_1 is negative.

Table O.3: 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.

O.8 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 heterogeneity $\zeta_i = (\gamma_i, \alpha_i)$ 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} . The set of types ζ_i is

$$\mathcal{Z} = \{(\gamma, \alpha) : \gamma \in \mathcal{G}, \alpha \in \{\alpha_0, \alpha_0 + \alpha_1\}\}.$$

Recall that α_0 is the price sensitivity of low-income consumers and that α_1 is the price sensitivity of high-income consumers. Let $x_{f\tau}(\zeta)$ denote the share of consumers of type $\zeta \in \mathcal{Z}$ whose previous purchase in time τ was made at store f , let \mathcal{F} be the collection of all competing online retailers, and let $x_\tau = \{x_{f\tau}(\zeta) : f \in \mathcal{F}, \zeta \in \mathcal{Z}\}$. 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

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

where β is a discount factor shared by all competing firms, $\mu(\zeta)$ is the share of consumers of type ζ , 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

$$V_f(x) = \max_{p_f \geq 0} \left[\sum_{\zeta \in \mathcal{Z}} \mu(\zeta) \sum_g x_g(\zeta) \sigma_{fg}(p_f, p_{-f}^*(x), \zeta) (p_f - mc) + \beta V_f(Q(x, p_f, p_{-f}^*(x))) \right]. \quad (5)$$

The function Q appearing in (5) is the state transition function, which provides the next period's state given the contemporary state x and prices p . The state transition is deterministically determined by consumer choice probabilities conditional on type ζ_i , 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 (5).

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 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 times the share of consumers with price sensitivity α as the corresponding population shares $\mu(\gamma, \alpha)$ for support points γ . Additionally, I use $K = 2$ 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 Toric—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.⁷ 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 O.4 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 O.4 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.

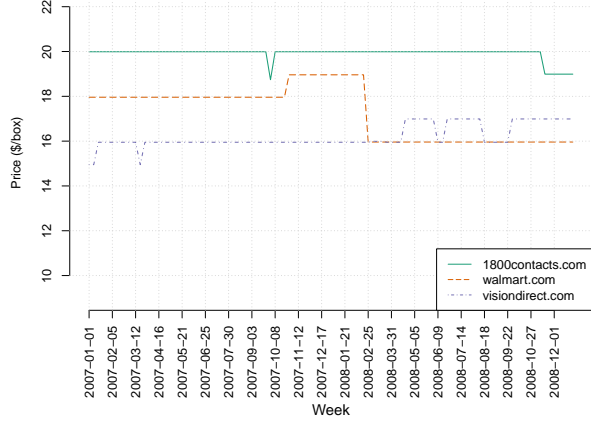
⁷Net sales and COGS were \$125,202,000 and \$73,962,000, respectively, in this time period. The ratio of these values is 1.69.

Table O.4: Percentage changes in markups from dynamic pricing model

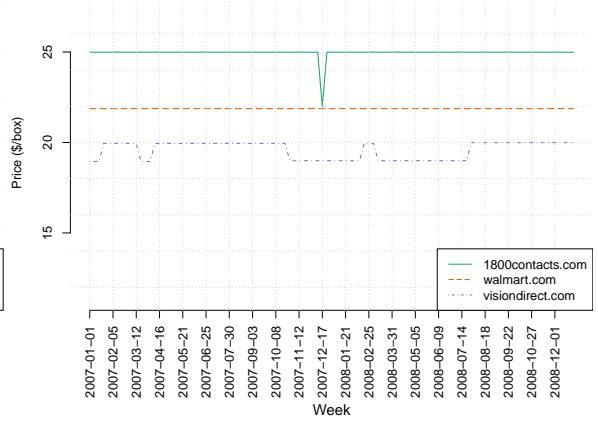
Store	Low search costs	No state dependence	No persistent unobs.	No price insensitive
1800	-0.5	-0.5	-15.9	-4.9
VD	-1.2	-3.1	-31.5	-6.2

Figure O.1: Prices of contact lenses across stores, brands, and time

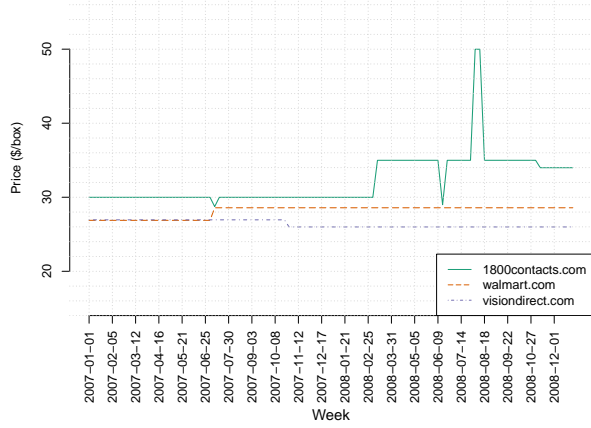
(a) Acuvue 2



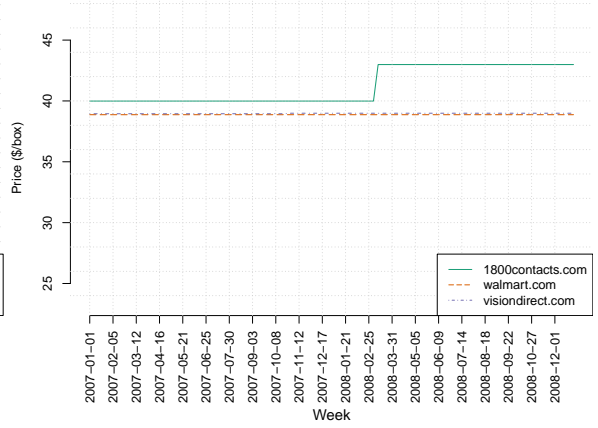
(b) Acuvue Advance



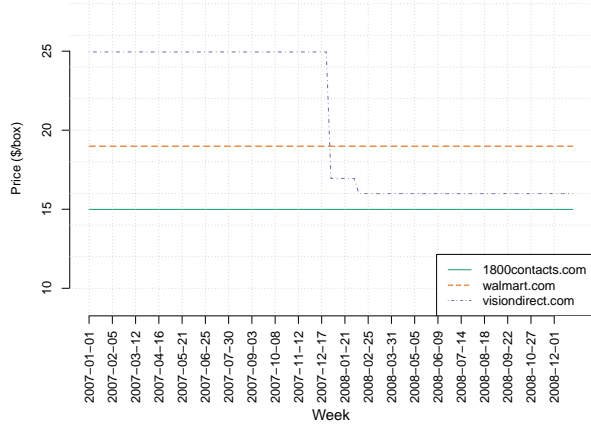
(c) Acuvue Oasys



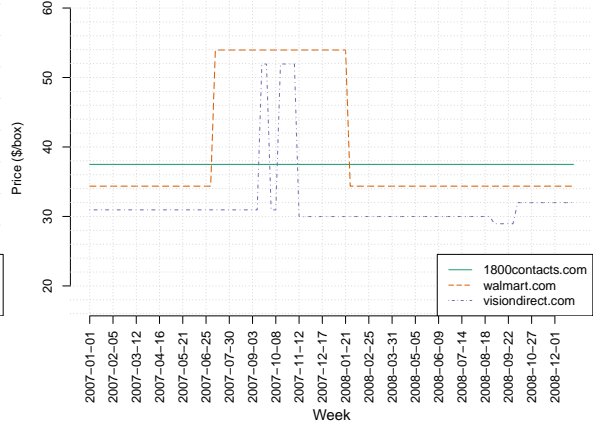
(d) Acuvue Advance



(e) Biomedics



(f) Freshlook Colorblends



O.9 Results for books

This appendix provides results of my analysis for the book category. I follow De Los Santos et al. (2012) by specifying the set of retailers as (i) Amazon, (ii) Barnes & Noble (“B&N”), (iii) book clubs considered together (“Club”), (iv) and various other book retailers considered together (“Other”). See Appendix O.1 for a list of retailers included in the groupings (iii) and (iv). In computing pricing equilibria for books, I drop “Club” and “Other” as retailers, and study price competition between Amazon and Barnes & Noble. I do not compute dynamic pricing equilibria for the books category.

The auxiliary statistics that I use in estimating the model for the books category are similar to those used in estimating the model for contact lenses. They are:

- (i) the share of search efforts including a visit to each retailer;
- (ii) the share of search efforts including a visit to more than one store;
- (iii) the “Inertia” statistics described in Appendix B;
- (iv) the “Role of lagged price” statistics described in Appendix B, but with Amazon taking the place of 1800;
- (v) the “Price sensitivity” statistics described in Appendix B;
- (vi) the “Cross-visiting behaviour” statistics described in Appendix B;
- (vii) the “Dependence of tastes and prices conditional on the initial state” statistics described in Appendix B; and
- (viii) the “Price sensitivity heterogeneity” statistics described in Appendix B.

Table O.5: Sales and prices by store (books, 2007–2008)

Store	Transactions	Average relative price
amazon.com	884	1.00
barnesandnoble.com	281	1.17
Book clubs	369	1.15
Other bookstores	162	0.98

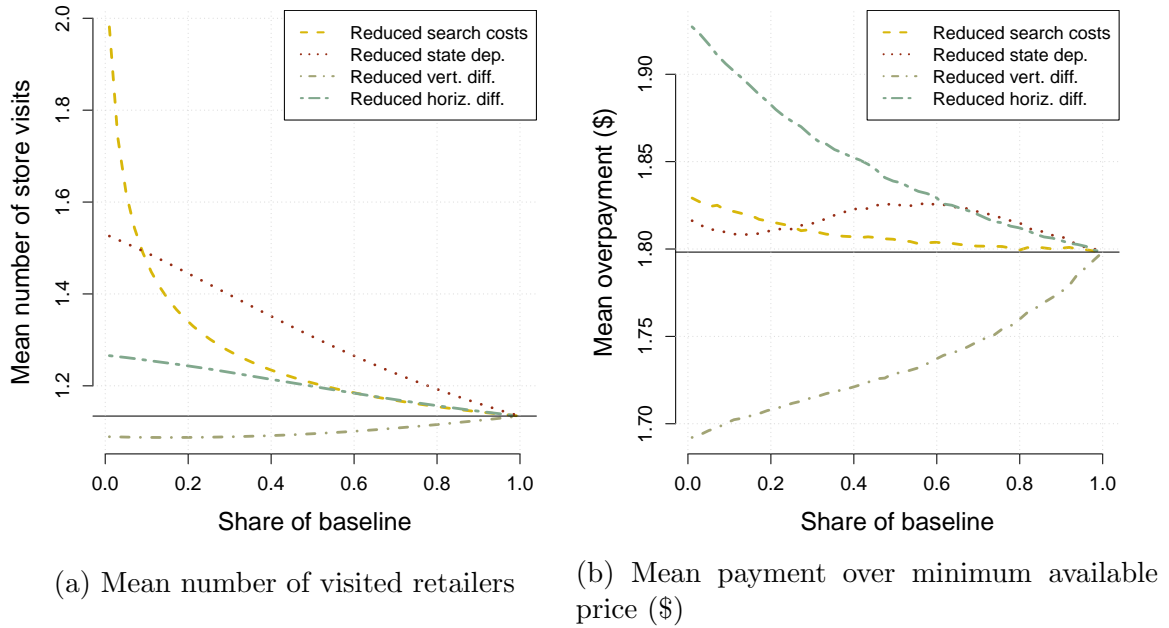
Note: the average relative price column reports the average ratio of the store’s price to the price at `amazon.com` across transactions in the 2007–2008 sample.

Table O.6: Descriptive multinomial regression estimates (books)

Specification 1: $q_{ft} = \bar{q} \quad \forall f, t$			Specification 2: seller/half-year fixed effects		
	Purchase	First visit		Purchase	First visit
α	0.143 (0.011)	0.114 (0.011)	α	0.147 (0.013)	0.136 (0.015)
Average elasticity	1.171 (0.092)	0.789 (0.077)	Average elasticity	1.203 (0.107)	0.943 (0.103)

Notes: The table reports maximum likelihood estimates of (1) for the books category. Standard errors are reported in parentheses. The “Average elasticity” is the average elasticity taken across transactions.

Figure O.2: Counterfactual search patterns (books)



Notes: the figure plots outcomes of counterfactual search efforts when each of search costs, state dependence, vertical differentiation, and horizontal differentiation are reduced from their baseline extent to zero in the manner described in the main text. The plotted quantities are averages over 5000 simulated search-effort histories (i.e., sequences of distinct search efforts over time) for each consumer in the estimation sample.

Table O.7: Parameter estimates (books)

Parameter	Estimate	SE
q_{Amazon}	3.099	0.577
$q_{\text{B\&N}}$	0.845	0.223
q_{Club}	-1.715	0.548
ϕ	3.576	0.607
α_0	0.199	0.058
α_1	-0.025	0.049
$\bar{\kappa}$	-1.307	0.342
Γ_{Amazon}	-2.695	0.706
$\Gamma_{\text{B\&N}}$	-1.781	0.667
Γ_{Club}	-0.941	0.402
Γ_{Other}	-1.406	0.538
σ_γ^2	0.540	0.262
λ	-0.688	0.606
Median search cost (utils)	0.18	0.06
Median search cost (\$)	0.90	0.50

Note: The “Estimate” columns provide point estimates obtained from the indirect inference estimator outlined in Section 6 whereas the “SE” columns report standard errors. In estimating the model on the books data, I specify that, for each f , $\Gamma_{fg} = \Gamma_f \quad \forall g$. Thus, Γ_f is the mean value of γ_i among consumers with initial state h_{i1} given by $h_{if1} \neq 1$.

Table O.8: Model fit and counterfactual search patterns (books)

Spec.	Share visiting one store only	Mean # of visits	Share paying > min. price	Mean over- payment (\$)
Observed	0.81	1.20	0.57	1.90
	-	-	-	-
Baseline	0.89	1.12	0.60	1.97
	(0.35)	(1.16)	(0.04)	(0.18)
Low search costs	0.84	1.19	0.61	1.99
	(0.35)	(1.16)	(0.04)	(0.18)
No state dep.	0.58	1.51	0.63	1.96
	(0.35)	(1.16)	(0.04)	(0.18)
No vertical diff.	0.91	1.10	0.59	1.89
	(0.35)	(1.16)	(0.04)	(0.18)
No horiz. diff.	0.75	1.28	0.67	2.23
	(0.35)	(1.16)	(0.04)	(0.18)
Logit only	0.00	4.00	0.49	1.47
	(0.35)	(1.16)	(0.04)	(0.18)

Notes: see the notes for Tables 9 and O.16.

Table O.9: Percentage changes in markups from static pricing model (books)

Panel A: Point estimates (%)

Store	Low search costs	No state dependence	No persistent unobs.	No price sens. het.
Amazon	0.7	17.1	-32.5	-4.0
Barnes	1.1	28.8	-83.6	-3.5

Panel B: Standard errors

Store	Low search costs	No state dependence	No persistent unobs.	No price sens. het.
Amazon	0.5	6.9	26.3	2.2
Barnes	1.3	21.8	8.1	2.2

Note: see the notes for Table O.9. This table presents estimates of percentage changes in markups for the book title “*T*” is for *Trespass*.

O.10 Additional results

Table O.10: Share of search efforts by number of visited stores (other categories)

(a) 2 days before					
# of 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	

(b) 5 days before					
# of 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	

(c) 14 days before					
# of 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	

Table O.11: Transactions above minimum available price (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 the (i) share of transactions made above the minimum available price (MAP); (ii) the average difference of paid price and the MAP, and (iii) the average relative difference of the transaction price over the MAP.

Table O.12: Within-site search intensity prior to contact lens purchase

(a) # of pages					(b) Duration (minutes)				
Store	Mean	Percentile			Store	Mean	Percentile		
		25 th	50 th	75 th			25 th	50 th	75 th
All	15.9	5	12	21	All	10.8	2	6	15
1800	16.1	5	12	21	1800	10.6	2	6	14
WM	16.7	4	10	20	WM	10.6	1	4	15
VD	15.0	5	11	20	VD	11.5	2	7	16

Note: this table reports summary statistics—the mean, 25th percentile, 50th percentile, and 75th percentile—of the number of pages viewed during a visit to an online contact lens retailer and of the duration of time spent (in minutes) browsing the retailer’s website. The “All” row gives results for all of the three major contact lens retailers pooled together.

Table O.13: Explanatory power of consumer characteristics

Specification	R^2_{McF}	R^2_{pseudo}
Demographic variables	0.16	0.15
Demographic and web use variables	0.23	0.21

Note: this table reports measures of fit from multinomial logistic regressions in which retailer of purchase (1800, WM, or VD) is the outcome and consumer characteristics are the regressors. For the first row of the table, the included consumer characteristics are: (i) indicators for the consumer having a university degree and for a “missing” value of the data’s educational attainment variable, (ii) indicators for age groups (30s, 40s, and 50+), (iii) indicators for household income (\$25-75k and over \$75k), (iv) indicators for racial groups (black, Asian, and other non-white), (v) indicators for household size (2, 3, and 4 or greater), (vi) an indicator for the presence of children in the household, (vii) an indicator for broadband internet, (viii) an indicator for Hispanic ethnicity, and (ix) indicators for census region of residence. The regression also includes year and brand fixed effects. For the second row of the table, variables characterizing the consumer’s internet usage are included in addition to the demographic variables listed above. Each of these variables reports the number of internet browsing sessions in which the consumer visited a website in a particular category; the categories are adult, advert, career, finance, gaming, government, information, malware, media, portal, retail, social media, video, weather, web service, dating, internet/wireless companies, news, sports, travel, downloads, and directories. See Saruya and Sullivan (2023) for a description of how these variables were constructed. R^2_{McF} reports values of McFadden’s R^2 for the regressions, whereas R^2_{pseudo} reports values of a pseudo- R^2 defined as follows:

$$R^2_{\text{pseudo}} = 1 - \frac{\sum_i \sum_f (y_{if} - \hat{y}_{if})^2}{\sum_i \sum_f (y_{if} - \bar{y}_f)^2},$$

where y_{if} is an indicator for observation i involving a purchase from retailer f , \hat{y}_{if} is the model-predicted probability of a purchase from retailer f , and \bar{y}_f is the sample average of y_{if} .

Table O.14: Elasticity estimates for Acuvue Toric

(a) Point estimates				(b) Standard errors			
Share	Price			Share	Price		
	1800	WM	VD		1800	WM	VD
1800	-2.14	0.18	0.08	1800	0.66	0.07	0.03
WM	4.41	-9.69	0.78	WM	1.29	2.52	0.57
VD	0.23	0.10	-2.14	VD	0.07	0.05	0.35

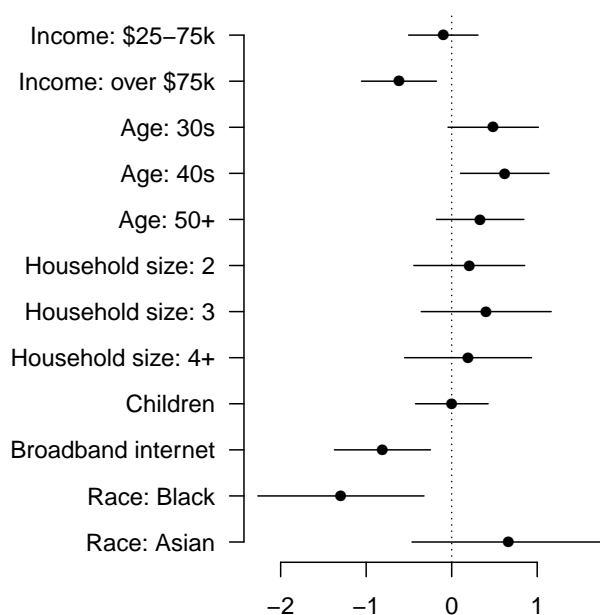
Note: Each entry corresponds to the elasticity of long-run demand at the store indicated by the entry’s row with respect to the price indicated by the entry’s column. Standard errors computed using the parametric bootstrap with 100 replicates.

Table O.15: Decomposition of price variation

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

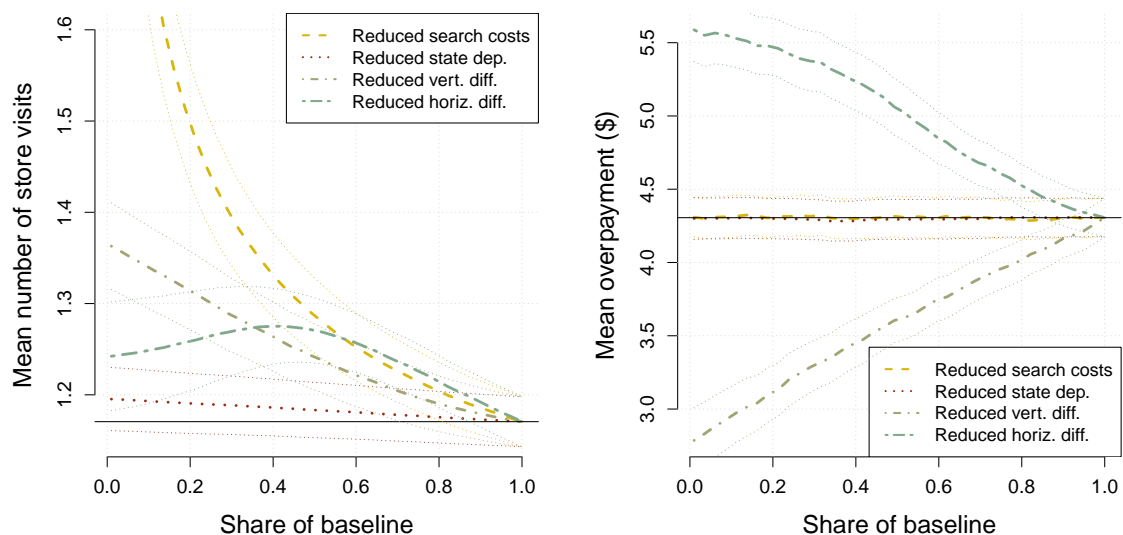
Notes: “Interbrand” provides the cross-brand standard deviation of brands’ average transaction prices. “Interstore” provides the average standard deviation of a brand’s price across stores, where the average is taken over transactions in the sample. “Intertemporal” provides the average standard deviation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

Figure O.3: Selected estimates from regression of purchase choice on consumer characteristics



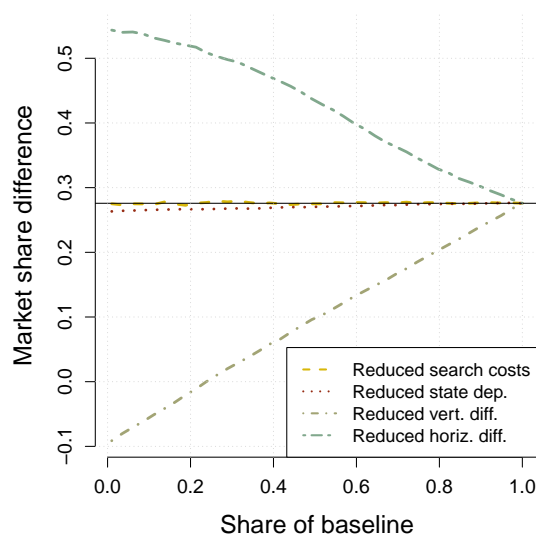
Notes: this figure reports selected estimates and their corresponding 95% confidence intervals from the regression described in the notes of Table O.13 that includes the web use variables.

Figure O.4: Counterfactual search patterns with pointwise confidence intervals



(a) Mean number of visited retailers

(b) Mean payment over minimum available price (\$)



(c) Difference in market share between 1800 and VD

Notes: the dotted bands around each curve provide 95% confidence intervals as computed from a parametric bootstrap with 100 replicates.

Table O.16: Model fit and counterfactual search patterns

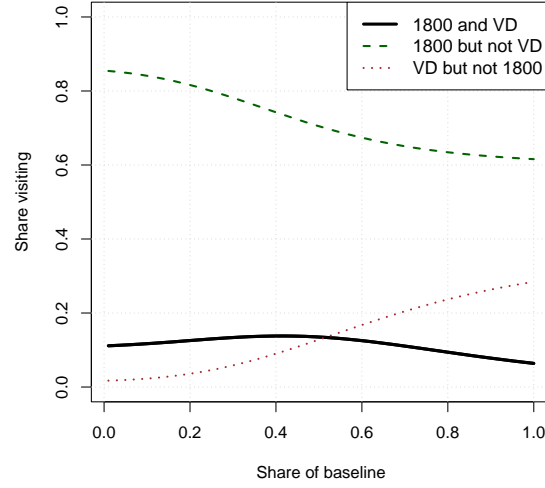
Spec.	Share visiting one store only	Mean # of visits	Share buying from...			Share paying > min. price	Mean over- payment (\$)
Observed	0.82	1.20	0.61	0.36	0.22	0.66	3.95
	-	-	-	-	-	-	-
Baseline	0.84	1.18	0.55	0.34	0.20	0.67	4.13
	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.17)
Low search costs	0.74	1.30	0.55	0.34	0.20	0.67	4.08
	(0.02)	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)	(0.18)
No state dep.	0.80	1.22	0.55	0.33	0.19	0.67	4.09
	(0.03)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)	(0.17)
No vertical diff.	0.73	1.30	0.55	0.20	0.29	0.53	2.89
	(0.04)	(0.05)	(0.02)	(0.01)	(0.03)	(0.03)	(0.22)
No horiz. diff.	0.61	1.48	0.55	0.50	0.03	0.88	5.47
	(0.05)	(0.07)	(0.02)	(0.05)	(0.05)	(0.07)	(0.65)
Logit only	0.00	3.00	0.55	0.13	0.25	0.52	2.19
	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)	(0.02)	(0.14)

Notes: the counterfactual preference changes considered 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 mean store taste $Q_f = 0$ for each store f to eliminate mean quality differences between stores;⁸
- (iv) No horizontal differentiation: set $\gamma_{if} = \mathbb{E}[\gamma_{if}]$ for all consumers i and retailers f ; and
- (v) Logit only: eliminate search costs, state dependence, vertical differentiation, and horizontal differentiation so that only prices and ε_{ijft} shocks differentiate retailers.

Altering consumer preferences changes the probability that a consumer buys from any online store. To focus on the effects of qualitative preference changes rather than those of changes in the magnitude of tastes for e-commerce, I add a compensating constant q^\dagger to each consumer's indirect utility for every online store to ensure the outside good's share is constant across counterfactuals. The value of q^\dagger differs across counterfactuals. I compute standard errors using a parametric bootstrap with 100 replicates.

Figure O.5: Changes in mean number of visited retailers as horizontal differentiation is reduced



Notes: this figure plots (i) the share of consumers visiting both 1800 and VD; (ii) the share of consumers visiting 1800 but not VD; and (iii) the share of consumers visiting VD but not 1800 for each extent of horizontal differentiation between zero and 100% of the baseline estimates. The main text provides details on the implementation of reductions in horizontal differentiation. Consumers visiting any of the combinations of retailers included in the figure may also possibly visit Walmart (e.g., 1800 and not VD includes consumers who visit 1800 and Walmart but not VD).

Table O.17: Characteristics of consumers who switch between brands

Coefficient	Estimate	SE
Intercept	1.148	0.103
Household size: 2	-0.024	0.094
Household size: 3	-0.103	0.109
Household size: 4	0.080	0.109
Household size: 5	-0.032	0.115
Household size: 6	-0.031	0.129
Age: 30s	0.078	0.075
Age: 40s	0.105	0.072
Age: 50 and over	0.070	0.071
Children in HH?	-0.044	0.058
R^2	0.023	

Notes: this table provides estimates from a regression of an indicator for whether a consumer ever switched brands of contact lenses on various consumer characteristics, including: (i) the size of the consumer's household, (ii) the age group of the head of the consumer's household, and (iii) an indicator for whether there is a child in the consumer's household. The number of observations is 494. The "SE" provides classical asymptotic standard errors.

Table O.18: Brand prices before and after switching

Variable	Mean	SE
Before	30.62	0.89
After	31.48	0.91
Difference	0.86	1.28
N	136	-

Notes: this table provides the mean prices of consumers' prescribed contact lens brands before and after a change of brand. In particular, it provides mean prices in a dataset of brand changes as defined as events in which the consumer purchases a brand of contact lenses different than the brand that the consumer previously purchased. The price before purchase is the mean price across retailers of the consumer's initial brand at the time of the consumer's search effort just before the change in brand. The price after purchase is the mean price across retailers of the consumer's new brand at the time of the search effort in which the consumer switches to this new brand. The number of observations is 136. The "SE" provides classical asymptotic standard errors.

O.11 Sensitivity of parameter estimates to auxiliary statistics

In this appendix, I assess the sensitivity of the structural model parameter estimates to the auxiliary statistics used in I-I estimation. This assessment involves re-estimating the model under alternative values of the auxiliary statistics and determining how parameter estimates obtained under these alternative values differ from the baseline parameter estimates.

First, I discuss how the availability of data on both sales and browsing behaviour permits the separate identification of search costs and parameters affecting purchase utility. Figure O.6 displays changes in the estimated retailer-specific indirect utility intercepts q_f and search cost parameters $\bar{\kappa}_f$ when, for each $f \in \{1800, \text{WM}, \text{VD}\}$, (i) the share of search efforts involving a visit to f increases by 5 percentage points (p.p.) and (ii) the share of search efforts involving a purchase from f increases by 5 p.p. The first notable pattern in the results is that the q_f indirect utility intercepts are sensitive to both the share of consumers visiting f and the share of consumers purchasing from f , which indicates that these auxiliary statistics help pin down the estimates of the indirect utility intercepts. Another pattern of note is that the search cost parameters $\bar{\kappa}_f$ are primarily sensitive to the extent to which consumers visit f rather than purchase from f : raising the share of search efforts involving a visit to each of 1800, WM, and VD reduces the corresponding store's $\bar{\kappa}_f$ parameter, although these parameters do not significantly fall when the store's share of purchases rises. The fact that browsing shares and purchasing shares differentially influence the q_f and $\bar{\kappa}_f$ estimates reflects the argument for the separate identification of indirect utility and search cost parameters in Section 6.2 of the main text.

Next consider estimation of the price coefficient α_0 . Figure O.7 plots changes in the estimate of α_0 due to changes in the values of various auxiliary statistics used in estimation; the table notes provide details on these changes. To summarize, the price coefficient is sensitive to the auxiliary statistics that are related to price sensitivity that I included to target α_0 .

I last discuss estimation of the state dependence parameter ϕ . Figure O.8 plots changes in the estimate of ϕ due to changes in the values of various auxiliary statistics used in estimation. See the table notes for details. I choose the magnitude and direction of the statistic changes for interpretability and comparability. The first change is in raising

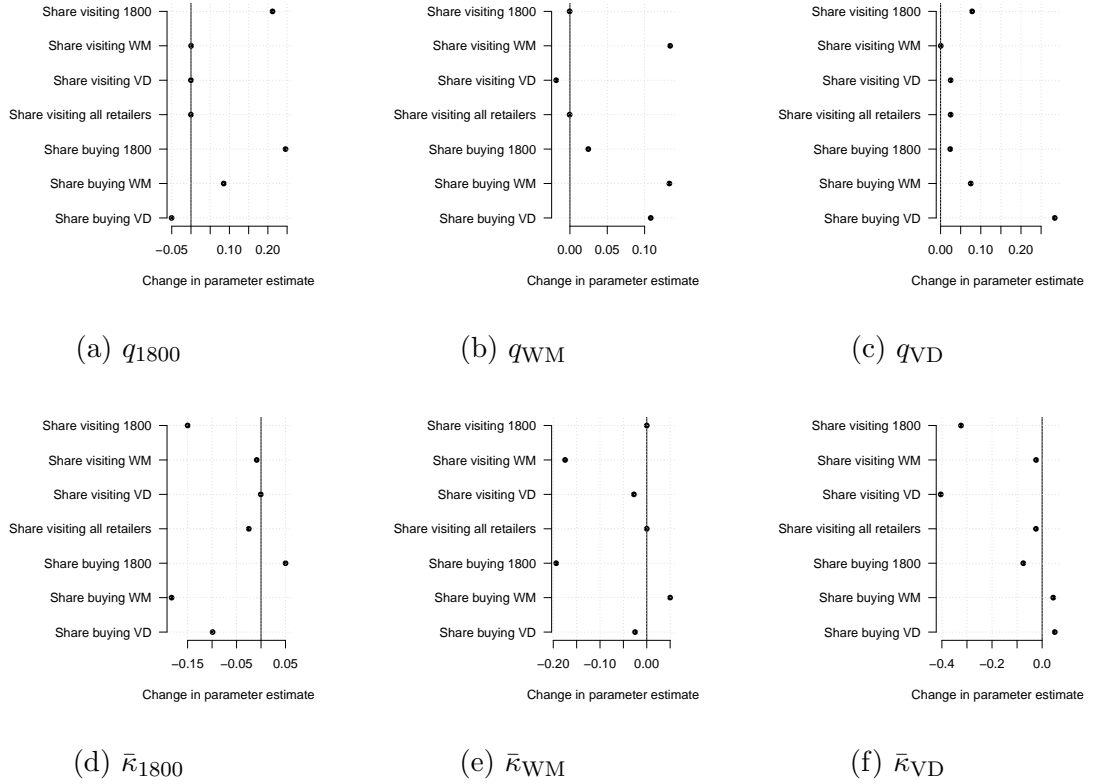
the share of consumers who visit all three retailers by 1 p.p. Since state dependence reduces the extent of search, imposing more expansive search reduces the estimated extent of state dependence. Second, I reduce the share of search efforts featuring the same first-visited store as the associated consumer’s previous search effort (“Inertia share”). Reducing inertia in searching naturally reduces the estimated extent of state dependence. Next, I raise the coefficients of the “role of lagged price” regressions: making consumers less sensitive to lagged prices reduces the extent of estimated state dependence, as expected. The “Inertia reg.: shift 2nd lag to 1st” row provides the change in the estimate of ϕ when the coefficient on the first lag of purchase in the “Inertia regression” is raised from 0.495 to 0.888 and the coefficient on the second lag is reduced from 0.392 to 0 (i.e., the first-lag coefficient is set equal to the sum of the estimated first- and second-lag coefficients whereas the second-lag coefficient is set to zero). Making consumers more sensitive to their immediately previous decisions (which drive state dependence in the model) rather than decisions further back in time (which also depend on unobserved heterogeneity in the model) raises the estimated extent of state dependence. Last, changing the share of consumers that visit a store f among those previously purchased from a store $g \neq f$ reduces the estimated extent of state dependence. This is because state dependence limits a consumer’s tastes for stores other than the store from which the consumer previously made a purchase.

O.12 Robustness to treatment of Walmart

In the article’s primary analysis, Walmart is treated differently than 1800 and VD, the retailers that primarily sell contact lenses: as described by Appendix A, I require a visit to Walmart to be nearer in time to a contact lens purchase or a visit to another contact lens retailer than I require for the other retailers. In this appendix, I consider the robustness of the model estimates to instead using the same cutoffs for including a visit to Walmart in a search effort as I use for the other retailers — i.e., to setting $K' = K = 14$ and $R' = R = 2$. Changing the search effort definition in this manner raises the share of search efforts including a visit to Walmart from 14.5% to 25.1%. In addition, I consider robustness of the parameter estimates to the inclusion of Walmart in the sample.

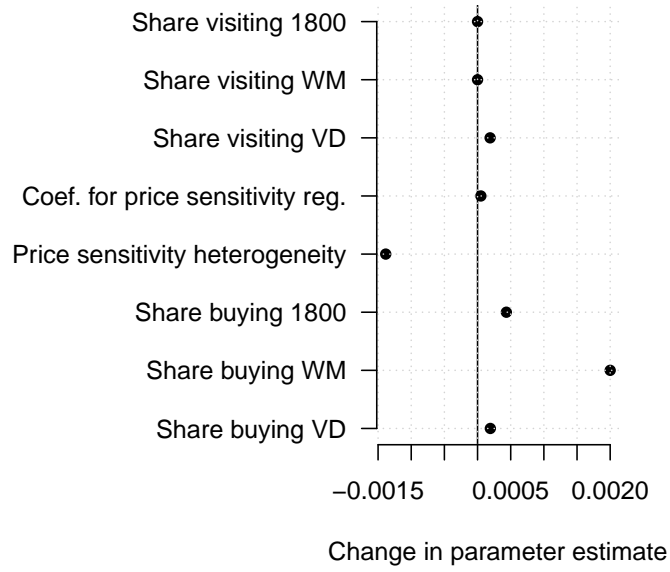
Table O.19 juxtaposes results obtained under the alternative cutoffs described above and obtained upon dropping Walmart from the sample to the baseline parameter

Figure O.6: Sensitivity of indirect utility and search cost parameters to auxiliary statistics



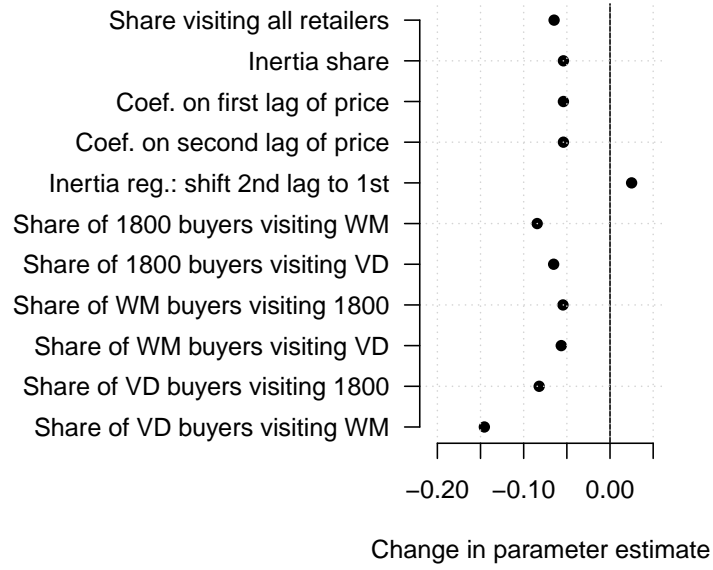
Notes: this figure provides changes in the values of the indirect utility intercept parameters q_f and in the search cost parameters $\bar{\kappa}_f$ under the following changes in the I-I auxiliary statistics $\hat{\beta}_n$ used in estimation: (i) increasing the share of search efforts involving a visit to 1800 by 5 percentage points (p.p.); (ii) increasing the share of search efforts involving a visit to WM by 5 p.p.; (iii) increasing the share of search efforts involving a visit to VD by 5 p.p.; (iv) increasing the share of search efforts involving a purchase from 1800 by 5 percentage points (p.p.); (v) increasing the share of search efforts involving a purchase from WM by 5 p.p.; and (vi) increasing the share of search efforts involving a purchase from VD by 5 p.p.

Figure O.7: Sensitivity of price coefficient α_0 to auxiliary statistics



Notes: this figure provides changes in the values of the price coefficient parameter α_0 under the following changes in the I-I auxiliary statistics $\hat{\beta}_n$ used in estimation: (i) increasing the share of search efforts involving a visit to 1800 by 5 percentage points (p.p.); (ii) increasing the share of search efforts involving a visit to WM by 5 p.p.; (iii) increasing the share of search efforts involving a visit to VD by 5 p.p.; (iv) reducing the slope coefficient in the “price coefficient” regression described in Appendix B by 0.05, from -0.155 to -0.205 (here, I lower the slope coefficient auxiliary statistic because a lower value of this statistic corresponds to greater price sensitivity whereas a greater value of α_0 corresponds to greater price sensitivity in the structural model); (v) raising the “price sensitivity heterogeneity” auxiliary statistic described in Appendix B by 0.1, from 0.045 to 0.145; (iv) increasing the share of search efforts involving a purchase from 1800 by 5 percentage points (p.p.); (v) increasing the share of search efforts involving a purchase from WM by 5 p.p.; and (vi) increasing the share of search efforts involving a purchase from VD by 5 p.p.

Figure O.8: Sensitivity of state dependence parameter ϕ to auxiliary statistics



Notes: this figure provides changes in the values of the state dependence parameter ϕ under the following changes in the I-I auxiliary statistics $\hat{\beta}_n$ used in estimation: (i) increasing the share of search efforts including a visit to all retailers by 1 percentage point (p.p.); (ii) reducing the inertia share (i.e., the share of search efforts with the same first-visited store as the associated consumer's previous search effort) by 10 p.p.; (iii) raising the coefficient on the first lag of price in the "role of lagged price regression" by 0.5; (iv) raising the coefficient on the second lag of price in the "role of lagged price regression" by 0.5; (v) setting the coefficient on the first lag of purchase in the "Inertia regression" equal to the sum of the estimated coefficients for the first and second lags and setting the coefficient on the second lag to zero; (vi) raising the share of consumers who previously bought from store f that visit store g in a search effort (i.e., the "cross-visiting" shares) by 5 p.p. each, one at a time.

estimates discussed in the main text. Each of these results is obtained using the same weighting matrix as used in computing the baseline estimates. In addition, I drop the Walmart-specific parameters and all Walmart-specific auxiliary statistics from the estimation procedure. See the notes of Table O.19 for details. The table shows that the model estimates are largely robust to the treatment of Walmart.

Bibliography

- De Los Santos, Babur, Ali Hortacısu, 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.
- 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.
- Gouriéroux, Christian, Alain Monfort, and Eric Renault.** 1993. “Indirect inference.” *Journal of Applied Econometrics* 8 (S1): S85–S118.
- Saruya, Hiroki, and Michael Sullivan.** 2023. “Cross-channel competition and complementarities in US retail.” Unpublished working paper.

Table O.19: Robustness of parameter estimates to treatment of Walmart

Parameter	Specification		
	Baseline	Alt. cutoffs	No WM
q_{1800}	-0.335	-0.328	-0.412
q_{WM}	-2.234	-2.002	-
q_{VD}	0.300	0.325	0.350
ϕ	0.493	0.448	0.493
α_0	0.110	0.107	0.110
α_1	-0.084	-0.096	-0.063
$\bar{\kappa}_{1800}$	-2.711	-2.762	-3.065
$\bar{\kappa}_{WM}$	-1.887	-2.112	-
$\bar{\kappa}_{VD}$	-1.546	-1.504	-1.596
Γ_{12}	-3.547	-2.987	-
Γ_{13}	-3.257	-3.266	-3.257
Γ_{21}	-1.233	-0.973	-
Γ_{23}	-1.109	-0.852	-
Γ_{31}	-5.574	-5.546	-5.641
Γ_{32}	-4.097	-2.793	-
σ_γ^2	1.298	1.298	0.863
λ	3.986	4.068	3.809
Med. search cost: 1800	0.414	0.403	0.292
Med. search cost: WM	0.930	0.764	-
Med. search cost: VD	1.294	1.380	1.234

Notes: this table contains (i) the baseline parameter estimates described in the main text, (ii) parameter estimates obtained upon setting $K' = K = 14$ and $R' = R = 2$ (“Alt. cutoffs”; see Appendix A), and (iii) parameter estimates obtained upon dropping Walmart from the analysis (“No WM”). In estimating the “No WM” specification I drop the Walmart-specific parameters q_{WM} , $\bar{\kappa}_{WM}$, and—for all retailers f other than Walmart—the parameters $\Gamma_{f,WM}$ and $\Gamma_{WM,f}$. I similarly drop Walmart-specific auxiliary statistics from the estimation procedure. These are: the share of search efforts with a visit to Walmart, the intercept for Walmart in the inertia regression, all cross-visiting shares involving Walmart, and the share of search efforts with a purchase from Walmart. The table also displays median search costs (in dollars) for each retailer under each specification.