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

Michael Sullivan

Department of Economics, Yale University

May 11, 2023

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 websites[1] and "other" stores, which includes the other online bookstores with many

*Email address: `m.r.sullivan@yale.edu`. This project draws on research supported by the Social Sciences and Humanities Research Council.

[1]Namely, `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.

O.2 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. 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:

$$\begin{aligned} r_{if} &\geq u_{i0} \geq u_{if} \vee \max_g r_{ig} \\ u_{i0} &\geq r_{if} \geq u_{if} \vee \max_g r_{ig} \\ u_{i0} &\geq u_{if} \geq r_{if} \vee \max_g r_{ig}. \end{aligned} \tag{1}$$

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. Under my distributional assumptions, the probability of the first chain of inequalities is

$$\frac{e^{\bar{r}_{if}}}{e^{\bar{u}_{i0}} + e^{\bar{u}_{if}} + \sum_{g=1}^{F} e^{\bar{r}_{ig}}} \times \frac{e^{\bar{u}_{i0}}}{e^{\bar{u}_{i0}} + e^{\bar{u}_{if}} + \sum_{g \notin \{0,f\}}^{F} e^{\bar{r}_{ig}}} \tag{2}$$

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 (1). I

[2] 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`.

[3] 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`.

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 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

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

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

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

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

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

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

O.3 Details of indirect-inference estimation

O.4 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.[4] The

[4]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.

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.

O.5 Optimal weighting matrix

The asymptotic normality of the I-I estimator is ensured by conditions that are standard in the I-I literature.[5] Recall that the I-I estimator is defined by

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

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

The asymptotic normality result for the I-I estimator is

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

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

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

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

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

Letting

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

we can write

$$V_{\hat{\theta}_n^H}(\Omega^*) = \left(\Lambda_0' V_{\hat{\beta}}^{-1}\Lambda_0\right)^{-1}.$$

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.

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 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

$$\mathbb{1}\{t \text{ results in purchase from 1800}\} = \beta_j - \alpha \log\left(\frac{p_{j,1800,t}}{\bar{p}_{jt}}\right) + \varepsilon_t, \tag{3}$$

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.1 provides estimates of (3) 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 (3) 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 (3), where n_j is the number of transactions of brand j in the sample. The between price-sensitivity estimate is larger in

Table O.1: 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

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

I expect that, conditional on a consumer previously purchasing from store f, the prices that the consumer faces and the consumer's unobserved tastes will be correlated. This is because a consumer who buys from a store f despite its high prices will have strong positive tastes for store f to rationalize buying from the store despite its price. To empirically assess this conditional correlation, I regress an indicator for whether a consumer visits stores other than the store f corresponding to the consumer's initial state on the relative price of f at the time that the consumer made the purchase that determined this state. The regression equation is

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

where j is consumer i's brand; p_{jf1} is f's price 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 positive λ_1 estimate would indicate that consumers who bought from a relatively expensive store are less likely to consider purchasing from other stores, which would indicates that these consumers have strong preferences for the store from which they historically bought contact lenses. Appendix Table O.2 provides the regression results. As expected, the estimate of λ_1 is positive.

Table O.2: 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

$$V_f(x) = \max_{p_f \geq 0} \left[\sum_{\zeta \in \mathcal{Z}} \mu(\gamma) \sum_g x_g(\zeta) \sigma_{fg}(p_f, p_{-f}^*(x), \zeta)(p_f - mc) + \beta V_f(Q(x, p_f, p_{-f}^*(x))) \right]. \tag{4}$$

The function Q appearing in (4) 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 (4).

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 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

Table O.3: 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

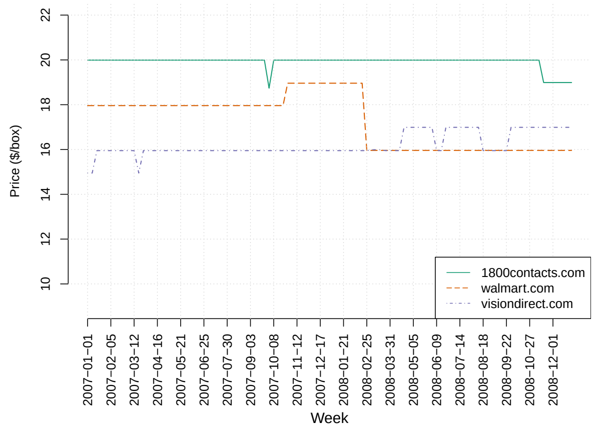
services (COGS) for January 1–June 30, 2007 as reported on 1800's quarterly report.[6] This approach applies 1800 overall markup ratio as defined in the preceding paragraph to a particular product's price to obtain an estimate of that product's marginal cost. Last, I set the discount factor β to 0.95.

Results. Table O.3 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.3 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.

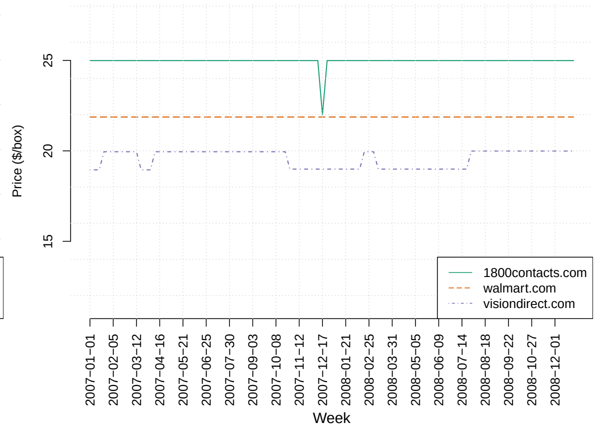
[6] Net sales and COGS were \$125,202,000 and \$73,962,000, respectively, in this time period. The ratio of these values is 1.69.

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

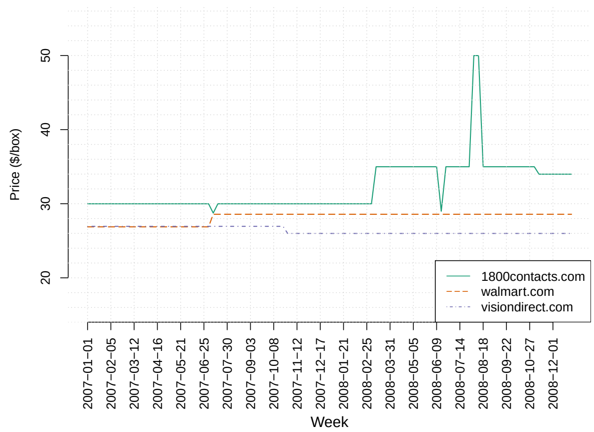
(a) Acuvue 2



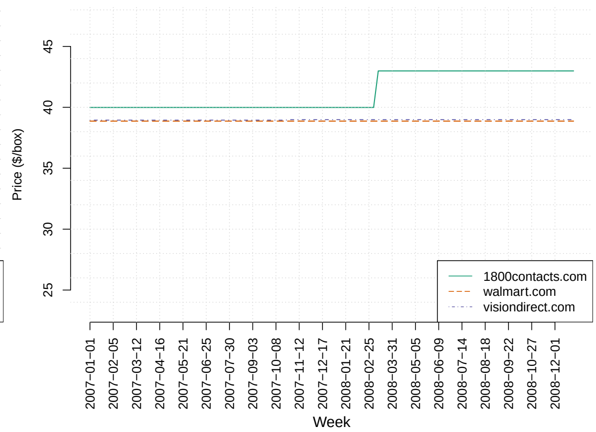
(b) Acuvue Advance



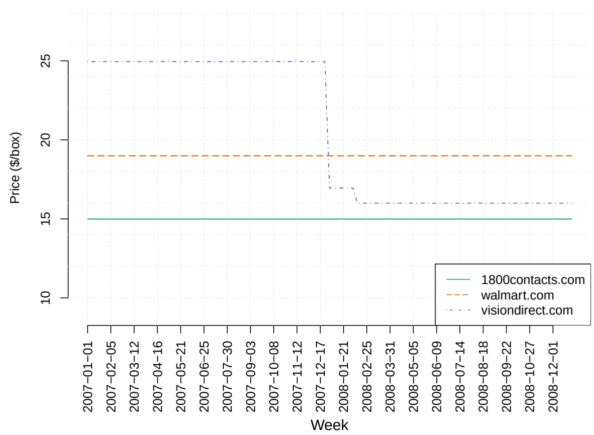
(c) Acuvue Oasys



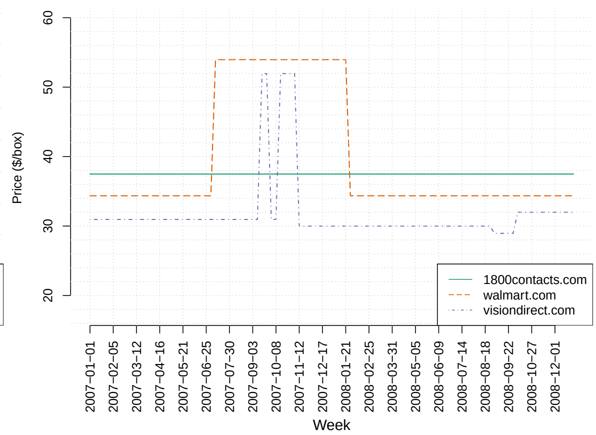
(d) Acuvue Advance



(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.

Table O.4: 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.5: Descriptive multinomial regression estimates (books)

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

	Purchase	First visit
α	0.143	0.114
	(0.011)	(0.011)
Average elasticity	1.171	0.789
	(0.092)	(0.077)

Specification 2: seller/half-year fixed effects

	Purchase	First visit
α	0.147	0.136
	(0.013)	(0.015)
Average elasticity	1.203	0.943
	(0.107)	(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.

Table O.6: 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.7: Model fit and counterfactual search patterns (books)

Spec.	Share visiting one store only	Mean # of visits	Share buying from... any	1800	VD	Visit order	Share paying > > min. price	Mean over-payment ($)
Observed	0.815	1.196	1.000	0.401	0.446	0.585	0.426	1.36
	-	-	-	-	-	-	-	-
Baseline	0.890	1.122	1.000	0.434	0.412	0.377	0.460	1.41
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)
Low search costs	0.839	1.187	1.000	0.436	0.411	0.352	0.462	1.43
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)
No state dep.	0.585	1.509	1.000	0.561	0.331	0.482	0.491	1.42
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)
No vertical diff.	0.907	1.104	1.000	0.408	0.412	0.464	0.446	1.22
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)
No persistent unobs.	0.759	1.266	1.000	0.596	0.307	0.400	0.545	1.88
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)
No price sens. het.	0.890	1.122	1.000	0.435	0.411	0.379	0.461	1.42
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)
No search	0.000	4.000	1.000	0.466	0.404	1.000	0.460	1.40
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)
Logit only (comp.)	0.000	4.000	1.000	0.297	0.280	1.000	0.405	-0.45
	(0.350)	(1.163)	(0.000)	(0.083)	(0.054)	(0.245)	(0.037)	(0.62)

Notes: see the notes for Tables 10 and O.12.

Table O.8: 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 loyalty	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.8. This table presents estimates of percentage changes in markups for the book title *"T" is for Trespass*.

O.10 Additional results

Table O.9: Within-site search intensity prior to contact lens purchase

(a) # of pages

Store	Mean	Percentile 25th	50th	75th
All	15.9	5	12	21
1800	16.1	5	12	21
WM	16.7	4	10	20
VD	15.0	5	11	20

(b) Duration (minutes)

Store	Mean	Percentile 25th	50th	75th
All	10.8	2	6	15
1800	10.6	2	6	14
WM	10.6	1	4	15
VD	11.5	2	7	16

Note: this table reports summary statistics—the mean, 25^{th} percentile, 50^{th} percentile, and 75^{th} 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.

Figure O.2: Role of search costs in limiting consumer search

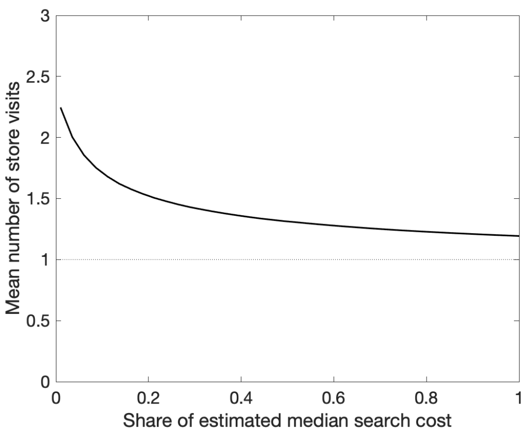


Table O.10: Elasticity estimates for Acuvue Advance for Astigmatism

(a) Point estimates

Share	Price 1800	WM	VD
1800	-2.14	0.18	0.08
WM	4.41	-9.69	0.78
VD	0.23	0.10	-2.14

(b) Standard errors

Share	Price 1800	WM	VD
1800	0.66	0.07	0.03
WM	1.29	2.52	0.57
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.11: 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.

Table O.12: Model fit and counterfactual search patterns (full results)

Spec.	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	0.843	1.170	0.730	0.504	0.210	0.418	0.713	4.36
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
Low search costs	0.747	1.286	0.724	0.499	0.206	0.389	0.715	4.36
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
Low search costs (comp.)	0.745	1.289	0.730	0.502	0.208	0.393	0.714	4.36
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No state dep.	0.821	1.194	0.677	0.470	0.189	0.425	0.717	4.38
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No state dep. (comp.)	0.802	1.217	0.730	0.502	0.206	0.417	0.714	4.34
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No vertical diff.	0.596	1.502	0.893	0.352	0.370	0.623	0.583	3.04
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No vertical diff. (comp.)	0.669	1.396	0.730	0.266	0.341	0.608	0.548	2.81
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No persistent unobs.	0.562	1.470	0.434	0.249	0.185	0.528	0.570	2.83
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No persistent unobs. (comp.)	0.387	1.672	0.729	0.415	0.312	0.536	0.587	2.94
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No price sens. het.	0.841	1.172	0.729	0.494	0.217	0.428	0.701	4.19
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No price sens. het. (comp.)	0.840	1.173	0.730	0.495	0.218	0.427	0.701	4.19
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
No search	0.000	3.000	0.710	0.478	0.204	1.000	0.705	4.19
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)
Logit only (comp.)	0.000	3.000	0.730	0.161	0.352	1.000	0.481	1.83
	(0.283)	(0.608)	(0.094)	(0.110)	(0.064)	(0.198)	(0.077)	(0.80)

Notes: This table expands upon Table 10 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. “No search” 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.

Bibliography

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.

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.