Online Appendix
Competition in Pricing Algorithms

Zach Y. Brown
University of Michigan & NBER*

Alexander MacKay
Harvard University†

October 1, 2021

Additional Appendices

Note: Appendices A and B are included with the main text.

*University of Michigan, Department of Economics & NBER. Email: zachb@umich.edu.
†Harvard University, Harvard Business School. Email: amackay@hbs.edu.

Table C1: Price Observations by Website and Brand

Retailer	Allegra	Benadryl	Claritin	Flonase	Nasacort	Xyzal	Zyrtec	Total
A	309,554	219,098	508,768	104,634	66,178	108,854	236,044	1,553,130
B	126,738	58,270	144,098	46,584	12,517	34,177	75,096	497,480
C	89,477	99,608	171,782	80,772	34,633	32,508	90,858	599,638
D	112,273	68,466	128,385	50,130	2,411	47,321	128,123	537,109
E	71,061	47,799	125,171	51,732	38,051	23,185	62,600	419,599
Total	709,103	493,241	1,078,204	333,852	153,790	246,045	592,721	3,606,956

Notes: Count of price observations for the sample period from April 10, 2018 through October 1, 2019.

C Details on High-Frequency Price Data

In this section, we provide further details on the collection of high-frequency price data and product definitions.

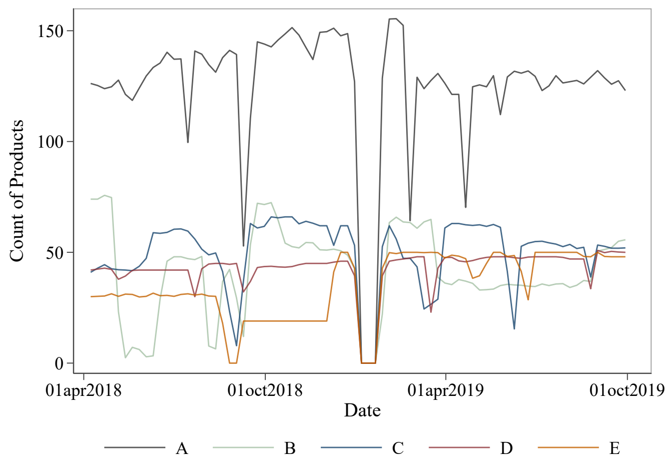
We focus on the main seven brands of allergy drugs. Each of the retained brands specializes in one drug, but they often offer the products in multiple forms (e.g., Liquid Gels, Liquid, or Tablets). Each brand offers many different size options, so there are several products per brands. In addition, most brands offer variants with different amounts of the active drug, targeted for children, 12-hour or 24-hour use. There are also versions of the drug that are combined with a decongestant. These varieties are captured by the *variant* of the drug. Finally, we distinguish products that are sold in a twinpack, so that twinpack of 12 tablets is a different product than a single pack of 24 tablets.[1] When a retailer sells multiple versions of the same product, we select the most popular version by retaining the version that has the greatest number of reviews, on average, in our sample. Retailers *A* and *B* offer significantly more product varieties than the other retailers. This is primarily due to the number of size options offered for each brand.

Due to the technological challenges involved in collecting high-frequency data, there is concern about measurement error. We address this in a few ways. First, we have focused on high-volume brands, helping to ensure the availability of price information. Second, we use supplemental information obtained at the time of our price sample to rule out price changes brought about by a lag in the website. For example, we can see if the description of the product is consistent over time. Third, we impute missing prices by filling in missing prices with the most recently observed price if the gap of missing prices is fewer than six hours. Finally, for the three retailers that do not change prices hourly, we smooth over single-period blips in price that revert back to the earlier price.[2] Table C1 displays the count of observations by brand and retailer.

Figure C1 illustrates the challenge of capturing high-frequency price data over an extended

[1]We drop multipacks that are of greater size than a twinpack, as they are not common across retailers.
[2]Overall, 7.8 percent of the prices are imputed in our analysis sample.

Figure C1: Observed Products Over Time



Notes: Figure displays the average daily count of observed products in our sample by week and by retailer. Dips in the data correspond to changes to the retailer website and issues with the researchers' servers. Retailers *A* and *B* offer significantly more product varieties than the other retailers. This is primarily due to the number of size options offered for each brand.

period. Dips in the data correspond to changes to the retailer website and issues with the researchers' servers. We note that we have several periods of many thousands of observations for which we have a consistent sample, and the periods of missing data do not meaningfully affect our results once we account for period fixed effects. We also include specifications using only data from July 1, 2019 through October 1, 2019, which are the most recent three months and for which we have a fairly consistent panel.

D Testing for Differences in Shipping and Distribution Costs

In Section 2, we document substantial differences in prices across retailers for the same products. We then examine how price differences can be generated by asymmetries in pricing technology. Another possible explanation for price differences is differences in supply costs. For online retailers, shipping and distribution costs can be a significant component of costs, in addition to wholesale purchase costs. However, for allergy products, it is reasonable to think that the differences in supply costs across retailers is small. The products are light and small, with each package weighing less than a few ounces. This allows retailers to use standard shipping processes and companies to deliver the products. In addition, wholesale prices are likely similar across retailers, as each retailer sells a large national presence and the brands in our study are large national brands.[3]

Though we do not measure costs directly, we test for variation in shipping and distribution costs by examining how prices vary with within-package quantity for the products in our dataset. Many allergy medications use identical packaging for different quantities of medication. For example, the packaging for 30, 60, and 90 tablets of Zyrtec is identical; other than the label, the only difference is the number of tablets in the (identically-sized) container. Since individual tablets weigh little, these three products should have negligible differences in shipping costs.

We exploit variation in within-package quantity to decompose prices into a fixed component per package and a variable component per unit (tablet, gelcap, etc.). For example, if the Zyrtec packages described above sold for \$22, \$34, and \$46, respectively, we infer that an additional 30 tablets increases the price by \$12, and that the unit price of one additional tablet is \$0.40. We also infer that the fixed component of price is \$10 ($= \$22 - \$12 \times 30$).

From a consumer perspective, the fixed component of price tends to make unit prices lower at larger packages, as the fixed component is spread out over more units. It may also be the case that a retailer charges lower unit prices when selling larger packages. We account for this by allowing the price per unit to decline with quantity. Thus, we model prices according to the following schedule:

$$p_{jr} = a_{jr} + b_{jr}x + c_{jr}x^2 \tag{1}$$

where p is the price of product j at retailer r, a is the fixed component of price, b is the unit price, and x is the within-package quantity. The coefficient c captures quantity discounts on the variable component of price, which we expect to be negative, and e is an error term. Here, product j refers to products with identical characteristics except for the within-package quantity,[4] and p provides the price for a specific product-size combination.

[3]Further, charging different prices to different retailers may be illegal under the Robinson-Patman Act, though enforcement of this Act has not been consistent.

[4]We treat multi-packs as different products, since they might require greater shipping costs.

To evaluate the differences in the fixed component of price across retailers, we divide by x and estimate the following regression equation:

$$\frac{p_{xjr}}{x} = a_r \frac{1}{x} + b_j^{(1)} + b_r^{(2)} + cx + \varepsilon_{xjr} \tag{2}$$

where we allow the fixed component a to vary across retailers. We estimate product-specific unit prices $b_j^{(1)}$ using fixed effects, and we capture differences in unit prices across retailers with $b_r^{(2)}$. The error term ε captures additional variation beyond the differences in means.

Table D1 presents results. Columns (1) and (2) report specifications including product-size combinations that are sold on 3 or more websites, and columns (3) and (4) report specifications including only product-size combinations sold by all five retailers. Column (1) and (3) restrict the fixed component to be the same across retailers. We estimate that the fixed component is approximately \$5.00. The average price for the set of products included products is \$18.67 and \$18.90 in the two specifications, indicating the fixed component of price is 27 percent of the total price on average.

Columns (2) and (4) allow the fixed component and unit prices to vary across retailers, while also allowing for a quantity discount. We find modest variation in the fixed component of prices, but greater fixed components do not necessarily correspond to higher prices. For example, we estimate that retailer B and C both have lower fixed components than retailer A, though they have higher prices. Retailer A and D have similar fixed components, between \$4.58 and \$4.95, yet retailer D has significantly higher overall prices (see Table 3). The unit price premium, relative to A, is positive and decreasing with the frequency of price changes, consistent with the findings in Section 2. We estimate a statistically significant per-unit discount for packages with higher quantity.

It is plausible that the fixed component of price represents an upper bound on shipping and distribution costs. With this interpretation, our estimates suggests that differences in these costs across retailers are not able to explain the price differences we observe in the data. The differences are modest, and the least expensive retailers do not have the lowest fixed components of prices. The above results indicate that differences in per-unit prices drive the price differences we observe in the data. This is consistent with price differences arising from larger markups, rather than differences in shipping and distribution costs.

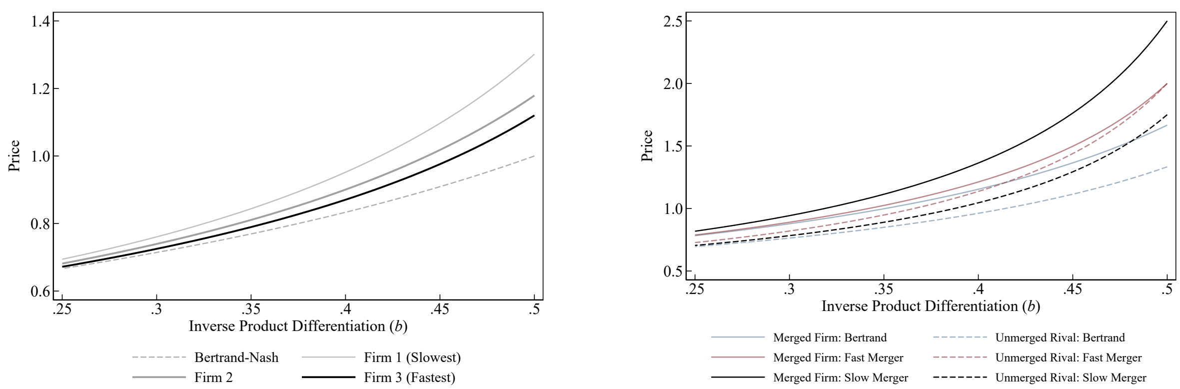
Table D1: Variation in Shipping Costs

	(1)	(2)	(3)	(4)
Fixed Component (All)	5.040***		5.068***	
	(0.165)		(0.192)	
Fixed Component (A)		4.816***		4.595***
		(0.244)		(0.391)
Fixed Component (B)		4.085***		4.244***
		(0.245)		(0.391)
Fixed Component (C)		3.694***		3.310***
		(0.257)		(0.391)
Fixed Component (D)		4.947***		4.585***
		(0.274)		(0.391)
Fixed Component (E)		6.079***		5.907***
		(0.331)		(0.391)
Unit Price Premium (B)	0.027	0.074***	0.062**	0.077**
	(0.020)	(0.025)	(0.025)	(0.032)
Unit Price Premium (C)	0.021	0.090***	0.059**	0.116***
	(0.020)	(0.024)	(0.025)	(0.032)
Unit Price Premium (D)	0.117***	0.119***	0.143***	0.143***
	(0.021)	(0.026)	(0.025)	(0.032)
Unit Price Premium (E)	0.154***	0.106***	0.185***	0.126***
	(0.022)	(0.027)	(0.025)	(0.032)
Quantity Discount		−0.001***		−0.002***
		(0.000)		(0.000)
Product FEs	X	X	X	X
Sold at All Retailers			X	X
Observations	294	294	170	170
R^2	0.967	0.974	0.915	0.933

Notes: Results from OLS regressions in which outcome is log price. Baseline sample in specification (1) includes all major brands of allergy drugs over the period April 10, 2018 to October 1, 2019. Coefficients show price difference relative to unit price of retailer A. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure E1: Differentiation and Equilibrium Prices in Oligopoly

(a) Three Firm Oligopoly　　(b) Post-Merger Prices



Notes: Panel (a) shows simulated prices for the three-firm oligopoly with zero marginal costs. A vertical slice of the plot captures the market equilibrium prices conditional on the product differentiation parameter, b. The dashed lined shows equilibrium prices for Bertrand-Nash competition, which is the same for all firms for any value of b. Solid lines show equilibrium prices for the firms under algorithmic competition when differences in pricing technology are large and the faster firms are assumed to react instantaneously to price changes by slower firms. Panel (b) shows equilibrium prices post-merger for the case in which the faster firms merge (Firm 1 and Firm 2) and the case in which the slower firms merge (Firm 2 and Firm 3). The post-merger prices under Bertrand competition are displayed in light blue.

E Differentiation and Oligopoly Effects

We extend the model considered in Section 5.1 to allow for various levels of differentiation in three-firm oligopoly. We consider a simple differentiated demand system given by

$$q_j = 1 - p_j + b \sum_{k \neq j} p_k \tag{3}$$

where $0 < b \leq \frac{1}{2}$. Assuming marginal cost c, the Bertrand-Nash equilibrium price is $p_j = \frac{1+c}{2-2b}$. The case in which $b = \frac{1}{2}$ and $c = 0$ corresponds to the example in the main text. As in the main text, each firm has technology characterized by (θ_j, γ_j). We assume that $\theta_j = \theta \; \forall j$ and $\gamma_3 > \gamma_2 > \gamma_1 = \theta$. We again assume zero marginal costs.

Figure E1 panel (a) shows the equilibrium prices as a function of b, the differentiation parameter. A market equilibrium is given by a vertical slice for a particular value of b. Consistent with the example in the main text, prices are monotonically decreasing in pricing algorithm frequency and all prices in the pricing algorithm equilibrium are higher than those from the Bertrand-Nash equilibrium. Intuitively, price dispersion is exacerbated when product differentiation is relatively low.

Figure E1 panel (b) shows the post-merger equilibrium prices for different values of b. The merged firm sets the same price for both entities, which are shown by solid lines. The dashed

Table E1: Simulated Equilibrium with Three Firms

(a) Simultaneous Bertrand

	Price			Profit			Merger			
							Price		Profit	
b	1	2	3	1	2	3	1	2/3	1	2/3
0.3	0.71	0.71	0.71	0.51	0.51	0.51	0.76	0.88	0.58	1.08
0.4	0.83	0.83	0.83	0.69	0.69	0.69	0.96	1.15	0.92	1.60
0.5	1.00	1.00	1.00	1.00	1.00	1.00	1.33	1.67	1.78	2.78

(b) Algorithmic Competition

	Price			Profit			Fast Firm Merger				Slow Firm Merger			
							Price		Profit		Price		Profit	
b	1	2	3	1	2	3	1	2/3	1	2/3	1/2	3	1/2	3
0.3	0.76	0.74	0.72	0.52	0.52	0.53	0.82	0.89	0.59	1.11	0.94	0.78	1.07	0.61
0.4	0.95	0.90	0.87	0.72	0.75	0.76	1.14	1.21	0.95	1.76	1.36	1.05	1.54	1.09
0.5	1.30	1.18	1.12	1.10	1.22	1.25	2.00	2.00	2.00	4.00	2.50	1.75	2.08	3.06

Notes: Table displays the simulated price, quantity, and profit under a three firm oligopoly for different values of differentiation, b, when demand is given by equation (3). The first three columns report the outcomes assuming simultaneous Bertrand price-setting behavior, which imply the same price and quantity for all firms. The final nine columns report the outcomes assuming algorithmic competition for each of the three firms, where firm 1 is the firm with the slowest pricing technology and firm 3 is the firm with the fastest pricing technology.

lines show the prices for the single unmerged rival. The equilibrium prices may be compared to the pre-merger prices in panel (a), noting that the y-axis has a different scales.

The general patterns remain similar to the example in the main text. Algorithmic competition exacerbates the price effects of mergers. The prices for the merged firm under algorithmic competition are higher than those under Bertrand competition. The unmerged rival also has a strong incentive to increase price. For lower values of differentiation (high values of b), the resulting price for the unmerged rival is greater than the Bertrand-Nash price for merging firms.

The effects on market average prices are similar whether or not the firm with the fastest technology is one of the merging firms. However, the post-merger patterns of price dispersion depend on the pricing technology of the merging firms. When slower firms merge, price dispersion across firms is exacerbated, but a merger between faster firms can yield lower price dispersion.

Table E1 reports pre-merger and post-merger prices and profits for specific values of b.

F Details of Spatial Differentiation Model

We model demand through the lens of spatial differentiation. Each consumer is located between two firms; these two firms represent each consumer's first and second choice at equilibrium prices. Consumers vary in their proximity to each firm, therefore the "travel" costs associated with each firm varies across consumers. In our setting, travel costs represent psychological costs and hassle costs of visiting each website. This may roughly be interpreted as search costs, though we provide no formal connection.

The model is a generalization of the Hotelling (1929) line. Unlike the circle model of Salop (1979), firms compete with all other firms, not just their closest neighbors. In this way, the model is related to the pyramid model of von Ungern-Sternberg (1991) and the spokes model of Chen and Riordan (2007). Unlike previous models, our approach allows for the mass of consumers on each segment to be different, including the mass of consumers on segments that link to an outside option. This feature is important since it allows for flexible substitution patterns that could explain differences in prices across retailers. This is also an advantage over models of vertical differentiation, such as the logit model, which restrict the horizontal substitution patterns to be symmetric across firms.

Each firm j lies in a $(J-1)$-dimensional space. A mass of consumers μ_{jk} lie along the line segment connecting j to k.[5] The distance between each firm is 1 unit. Each firm sells a single product, which consumers value at $v_j > 0$, and each firm chooses a price p_j. Each firm also has a mass of consumers on a line segment of distance D_0 connecting to an outside option ($j = 0$), with $p_0 = 0$ and $v_0 = 0$. Consumers lie on these segment with density μ_{j0} and mass $\mu_{j0} D_0$. D_0 may be arbitrarily large, so that the firm never captures the full segment. Figure F1 provides a visual representation of the demand system for the case of three firms.

Each consumer i is indexed by its location and bears a travel cost τd_{ij} for traveling a distance d_{ij} to firm j to purchase its product. A consumer along segment jk will choose j if $u_{ij} > u_{ik}$, or

$$(v_j - p_j) - (v_k - p_k) > \tau(d_{ij} - d_{ik}). \tag{4}$$

That is, the consumer will prefer j to k if the added value of product j is greater than the additional travel cost of visiting firm j. The consumer also has the option to stay home and get $u_{i0} = 0$, which he will do if $u_{ij} < 0$ and $u_{ik} < 0$.

Consumers are distributed along each line segment connecting j to k according to a distribution F_{jk} with support $[0, 1]$. We assume that the distribution is symmetric about the midpoint of the segment. Symmetry implies $F_{jk} = F_{kj}$, so the direction of the connection is arbitrary. We also assume that the same distribution is applied to all segments: $F_{jk} = F$, though this could easily be relaxed. Demand along each segment can then be characterized by the distribution function F.

[5]Demand can be represented by a graph. The graph is complete if $\mu_{jk} > 0$ for all $\{j, k\}$.

Noting that $d_{ik} = 1 - d_{ij}$ for a consumer on segment jk, a consumer on this segment will choose j if $u_{ij} > u_{ik}$ and if $u_{ij} \geq 0$, i.e., $\frac{1}{2}+\frac{1}{2\tau}\left((v_j - p_j) - (v_k - p_k)\right) > d_{ij}$ and $\frac{1}{\tau}\left(v_j - p_j\right) \geq d_{ij}$. Firm j receives customers for which d_{ij} satisfies both conditions. Therefore, firm j receives a quantity of $\mu_{jk}F(y_{jk})$ from line segment jk, where

$$y_{jk} = \min\left\{\frac{1}{2} + \frac{1}{2\tau}\left((v_j - p_j) - (v_k - p_k)\right), \frac{1}{\tau}\left(v_j - p_j\right)\right\}. \tag{5}$$

For the outside segments, $y_{j0} = \frac{1}{D_0}\frac{1}{\tau}\left(v_j - p_j\right)$, as these segments have length D_0 instead of 1. The parameter D_0 can also be interpreted as the relative travel cost of choosing the outside option relative to an inside good, as the model has an isomorphic parameterization with outside travel costs $\tilde{\tau}_0 = D_0\tau$.

Overall, quantities are given by

$$q_j = \sum_{k \neq j} \mu_{jk}F(y_{jk}). \tag{6}$$

The flexibility in substitution patterns from this relatively parsimonious model comes primarily through the mass of consumers on each segment $\{\mu_{jk}\}$ and the choice of distribution F. In equilibrium, the consumers $\{\mu_{j0}\}$ that have no next-best alternative other than the outside option are also important in determining substitution patterns.

We introduce some terminology to facility discussion of the model. When $\max(u_{ij}, u_{ik}) \geq 0$ for all i on segment jk and $y_{jk} < 1$, the segment is *contested*.[6] When some consumers prefer to stay home, rather than purchase, the segment is *uncontested*. If segment jk is uncontested, there is no consumer indifferent between j and k, so those firms have local monopoly power over a portion of consumers on that segment. That is, a change in the price of firm k does not affect demand for firm j at the margin. When all segments between firms (the "inside" segments) are contested, we say the market is *covered*. For a covered market, all consumers on inside segments purchase.

For our calibration exercise, we assume that consumer locations are distributed uniformly within each segment. We also assume that the products are homogeneous (but for the travel costs), so that $v_j = v$ for all j except for the outside option, for which $v_0 = 0$. Finally, we assume that consumer valuations are sufficiently high that all consumers on the inside segments purchase a product.[7] Demand for retailer j is equal to

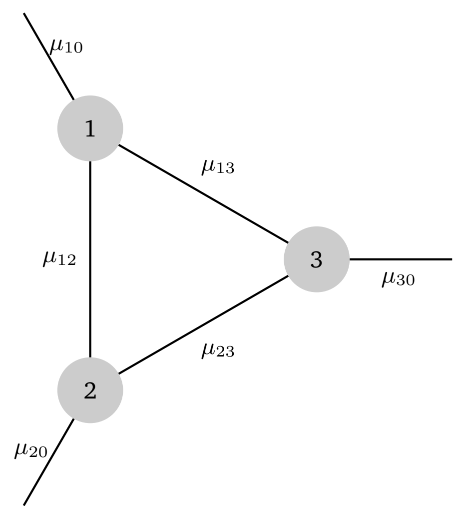
$$q_j = \sum_{k \neq j, 0} \mu_{jk}\left(\frac{1}{2} - \frac{1}{2\tau}\left(p_j - p_k\right)\right) + \mu_{j0}\frac{1}{\tau}\left(v - p_j\right). \tag{7}$$

Rearranging terms yields equation 25 in the main text.

[6]When $y_{jk} \geq 1$, the segment is *dominated* by j.

[7]In a slight abuse of notation, we omit the arrival rate of consumers $m(t)$.

Figure F1: Spatial Differentiation Model with Three Firms

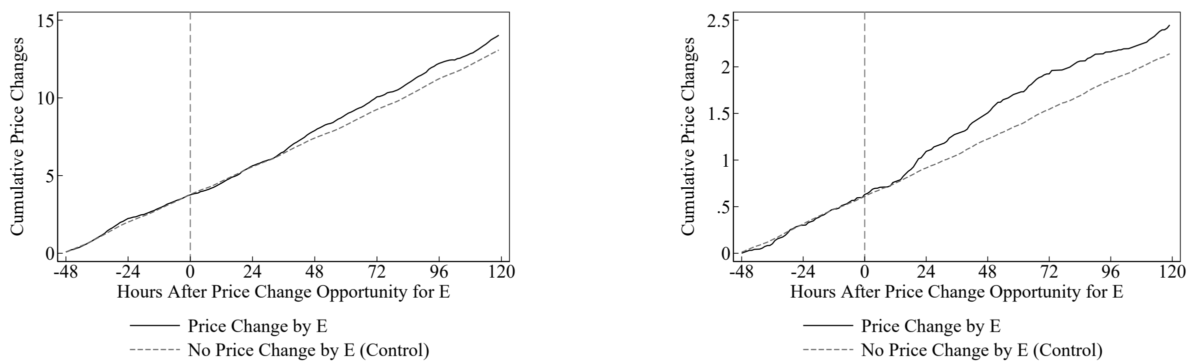


Notes: Example of demand for three firms with an outside option. The mass of consumers along each segment is given by μ_{jk}. The segments with mass μ_{10}, μ_{20}, and μ_{30} represent consumers whose next-best alternative to the linked firm is the outside option.

G Additional Tables and Figures

Figure G1: Price Changes by Fastest Retailers in Response to a Price Change by Retailer *E*

(a) Response by Retailer *A* (b) Response by Retailer *B*



Notes: Figure displays the cumulative price changes of high-frequency retailers *A* and *B* in response to a price change occurring at retailer *E*. The solid line displays the cumulative price change when retailer *E* changes a price of the same product in that week. The dashed line plots the cumulative price changes when the product at retailer *E* does not have a price change. The pre-period differences are netted out so that the difference is zero at period 0.

Figure G1 shows the reaction of high-frequency firms (retailers *A* and *B*) to price changes by low-frequency retailer *E*. The charts imply that the high-frequency firms respond to a price change by Retailer *E* within about 48 hours.

Table G1: Measures of Retailer Market Share

Retailer	Share of Online Personal Care	Google Search		
		"Retailer name"	"Retailer name" + Allergy	Mean
A	0.338	0.427	0.188	0.307
B	0.252	0.311	0.263	0.287
C	0.084	0.139	0.123	0.131
D	0.119	0.062	0.188	0.125
E	0.207	0.061	0.237	0.149

Notes: Share of personal care category reflect 2019 revenue figures from ecommerceDB.com. This includes online sales of medical, pharmaceutical, and cosmetic products for each of the retailers, including sales through mobile channels. Google search figures refer to the searches over the sample period as a share of total searches for all of the five retailers. Google search data are obtained from Google Trends (trends.google.com).

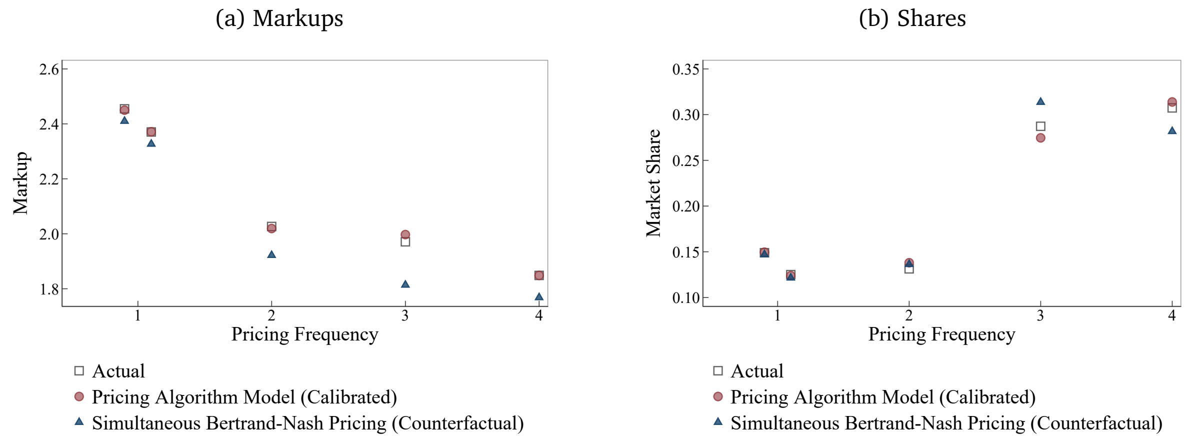
Table G1 provides measures of aggregate shares for the retailers in our data. We calibrate our model to Google search shares, using the mean of search shares for the retailer name and search shares for the retailer name along with the word "allergy." We cross-check these shares against revenue shares provided by ecommerceDB.com. The measures of online revenue shares are obtained for the category of personal care, which includes all medical, pharmaceutical, and cosmetic products. Four of our retailers are in the top five for the personal care category by revenue, and all are in the top ten. The other retailers in the top ten have a focus on cosmetics.

Table G2: Calibrated Segment Weights

		Retailer k					
		A	B	C	D	E	Outside
Retailer j	A	0.00	11.44	2.10	0.54	0.54	0.00
	B	11.44	0.00	2.10	0.54	0.54	1.74
	C	2.10	2.10	0.00	0.54	0.54	1.45
	D	0.54	0.54	0.54	0.00	0.54	3.07
	E	0.54	0.54	0.54	0.54	0.00	3.99

Notes: Row j column k shows the mass of customers on the segment between retailer j and k (μ_{jk}). The weights are symmetric; for convenience, they are displayed twice ($\mu_{jk} = \mu_{kj}$), representing the perspective of each firm. The outside segment weights represent the share of customers captured from the outside segments at the equilibrium prices.

Figure G2: Calibration Fit for Markups and Shares



Notes: Figure displays the markups (panel (a)) and the relative shares (panel (b)) plotted against the pricing frequency of each retailer. Frequency is normalized to the relative sequence. The black squares indicate the data, and the red dots are the fitted prices from a calibration exercise. The relative prices are obtained from the estimated coefficients in specification (1) of Table 3. The markup level is pinned down by the calibrated model. The green triangles display the counterfactual simultaneous Bertrand markups at the calibrated parameters and the corresponding shares.

The fit of the calibration exercise is displayed in Figure G2. In panel (a), squares indicate the relative prices in the data; these prices are translated to markups based on the calibrated model. The x-axis displays the pricing frequency in terms of the relative sequence. The red dots indicate the markups from the calibrated model. Likewise, the black squares in panel (b) represent observed shares, and the red dots indicated the predicted shares from the model. Our eight-parameter model fits prices and shares quite well.

References

CHEN, Y. AND M. H. RIORDAN (2007): "Price and Variety in the Spokes Model," *The Economic Journal*, 117, 897–921.

HOTELLING, H. (1929): "Stability in Competition," *The Economic Journal*, 39, 41.

SALOP, S. C. (1979): "Monopolistic Competition with Outside Goods," *The Bell Journal of Economics*, 10, 141–156.

VON UNGERN-STERNBERG, T. (1991): "Monopolistic Competition on the Pyramid," *The Journal of Industrial Economics*, 355–368.