INSTITUTE OF RETAIL ECONOMICS

IS INTERTEMPORAL PRICE DISCRIMINATION THE CAUSE OF PRICE DISPERSION IN MARKETS WITH LOW SEARCH COSTS?

CHARLIE LINDGREN, SVEN-OLOV DAUNFELDT,
NIKLAS RUDHOLM & SIRIL YELLA

HFI WORKING PAPER No8

Is intertemporal price discrimination the cause of price dispersion in markets with low search costs?

Charlie Lindgren[a,*], Sven-Olov Daunfeldt[a,b], Niklas Rudholm[b] and Siril Yella[a]

[a]Dalarna University, Högskolegatan 2, 791 31, Dalarna, Sweden; [b]Institute of Retail Economics, Regeringsgatan 60, 103 29, Stockholm, Sweden

Abstract: Theories of intertemporal price discrimination imply that prices must be chosen using mixed strategies, with retailers changing their prices randomly over time. Otherwise, consumers will learn which retailer has the lowest price, and eventually, all customers will patronize the lowest price retailer, or all retailers will charge the same price. We test whether price dispersion is explained by intertemporal price discrimination strategies using a dataset of identical products sold through the PriceSpy price comparison website. Our results show that there are clusters of retailers with similar pricing within each cluster but different price levels between clusters even after controlling for retailer heterogeneity. Retailers also remain in the same price cluster over time, suggesting that consumers have ample opportunities to learn which retailers belong to which price cluster. Intertemporal price discrimination is thus unlikely to have caused the observed price dispersion.

Keywords: Clustering; intertemporal price discrimination; online retailing; e-commerce; price comparison websites.

JEL codes: C38, D22, D83, L81.

[*]Corresponding author: clg@du.se.

1. Introduction

In his seminal article, Stigler (1961) pointed out the pervasive price dispersion for homogeneous products sold in well-developed markets. The most widely used models for explaining such price dispersion, even in markets with low search costs, are so-called clearinghouse models. The most widely cited clearinghouse model was presented by Varian (1980), but there have been several followers (Rosenthal, 1980; Narasimhan, 1988), and these models have also been used to explain price dispersion in online markets (Bayliss and Perloff, 2002; Lach, 2002; Baye et al., 2004).

In a clearinghouse model, retailers must simultaneously appeal to two types of customers: shoppers who search and use the available price list to buy from the retailer offering the lowest price and non-shoppers who do not engage in search but learn prices over time as they visit stores' or retailers' webpages. The reasons for not being a shopper can differ. Consumers might have strong preferences for a specific retailer (Rosenthal, 1980; Narasimhan, 1988) or not have access to the clearinghouse price list (Varian, 1980).

For price dispersion to remain in these models, there must be some consumers who are non-shoppers, and prices must be chosen using mixed strategies, with retailers changing their prices randomly (Varian, 1980; Lach, 2002). Otherwise, consumers will eventually learn which retailer has the lowest price, and all customers will either patronize the lowest price retailer, or all retailers will charge the same price.

The use of mixed strategies in these models has some empirically testable implications. First, there can be no grouping of retailers having similar, and thus predictable, price strategies that remain over time. Second, the position of individual retailers within a cross-sectional price distribution will change randomly over time. Therefore, there will be no distinguishable patterns in a transition matrix of prices, and the probability of remaining in the same position in the transition matrix should be low.

Evidence from previous studies is mixed, with some studies rejecting the clearinghouse model (Bayliss and Perloff, 2002) but others supporting it (Lach, 2002; Baye et al., 2004). However, Bayliss and Perloff (2002) and Baye et al. (2004) do not account for retailer heterogeneity in their analysis, while Lach (2002) arbitrarily groups retailers into quartiles when investigating price movements in a transition matrix analysis.

We test the predictions from clearinghouse models using a dataset of identical products sold through the PriceSpy price comparison website. In contrast to previous studies, we account for retailer heterogeneity and use cluster analysis to determine the size and number of retail clusters with similar prices. We show that even under such conditions, pricing does not follow the pattern suggested by theories of intertemporal price discrimination.

2. Method

Compared to previous studies (Bayliss and Perloff, 2002; Lach, 2002; Baye et al., 2004), we consider a large set of products (14) that are observed with a higher frequency (daily) and for a longer period (up to 42 months). An important feature of our data is that the products are identical between retailers and that there is no ambiguity in product representation on the price comparison website. However, the average coefficient of variation for the price of these products (Table 1) still reveals substantial price dispersion.

Following Lach (2002), we control for time-invariant retailer heterogeneity using the following model:

$$\log p_{it} = \mu + \alpha_i + \sigma_t + \epsilon_{it} \tag{1}$$

where p_{it} is the CPI-adjusted price of product i on day t, α_i is a retailer-specific fixed effect and σ_t is a day-specific fixed effect. The residual variation ϵ_{it} for each individual retailer represents the percentage deviation of a retailer's price from the geometric mean in the group of retailers when controlling for both retailer and time heterogeneity (Lach, 2002).

Lach (2002) divided the residuals from the estimation of equation (1) into quartiles and then investigated the frequency and likelihood of price movements between these quartiles to determine whether price movements followed the predictions of intertemporal price discrimination models. However, the grouping of prices into quartiles is arbitrary, and both the number and size of retailer price clusters are likely to differ between products. Therefore, we adopt clustering techniques to let the data identify the number and size of retail price clusters instead of arbitrarily dividing the data into quartiles. If there are groups of retailers gathered in well-defined price clusters, these clusters remain over long periods, and the ranking of an individual retailer in a Markovian transition matrix of these clusters is likely to stay the same, indicating that retailers do not follow an intertemporal price discrimination strategy.

There are some characteristics common to retail price series that might confound this analysis. For example, if a retailer changes price in advance of competitors, even by a day, it can influence whether that store belongs to a cluster of retailers or not, even if a similar pricing pattern is otherwise clearly visible. To mitigate this issue, we use the methods of dynamic time warping (Rath & Manmatha, 2003) and piecewise aggregation approximation (Keogh et al., 2001). We then use the "elbow method" (Thorndike, 1953) to determine the appropriate number of clusters and the number of retailers belonging to each cluster. Finally, we determine the degree to which individual retailers transition between clusters using Markovian transition matrices.[1] A pattern of retailer transition probabilities that align with the diagonal of a K x K

[1] These methods are explained in more detail in appendices A and B in the supplemental online material.

transition matrix would indicate that intradistribution mobility is low. The retailer transition probabilities are contingent on the time horizon considered, and following Lach (2002), we use monthly transition matrices.

3. Results

Table 1 shows that there are between 2 and 5 price clusters, depending on the product considered, and that the average amount of time that a retailer spends in a specific cluster ranges from 19 to 246 days. We also present the likelihood of observing an individual retailer along the diagonal, i.e., remaining in the same price cluster as in the month before, in Table 1. In a K x K transition matrix, the likelihood of randomly being observed in the diagonal will equal 1/K, while the likelihood of being observed in the off-diagonal will equal [(K-1) x K)/(K x K)]. Thus, in a 2 x 2 transition matrix, pricing using mixed strategies should give a likelihood for the diagonal of 0.5, while in our data, the likelihood for the diagonal for the two-cluster products ranges from 0.72 to 0.85. For a 3 x 3 transition matrix, we expect to observe the diagonal in 1/3 of all cases (0.33), but the averages instead range from 0.43 to 0.78. For no product do we reach the expected number according to intertemporal price discrimination models.[2]

4. Discussion

We tested the predictions from clearinghouse models using data from a price comparison website. Our results showed that substantial price dispersion remains even after controlling for heterogeneity in retailer offerings and that there are clusters of retailers that maintain persistently high, mid-range, or low prices. One possible explanation is that the share of consumers using price comparison websites in Sweden is still low enough to make it profitable for some retailers to focus on the group of uninformed consumers while also listing their products on the price comparison website.

Acknowledgments

Research funding from the Swedish Retail and Wholesale Council, grant number 2018:773, is gratefully acknowledged. The authors would also like to thank Mark Dougherty and the participants at the Microdata analysis seminar (December 6, 2019) for their valuable comments and suggestions.

[2] Since the results are contingent on the time horizon considered, we also present results using biweekly and bimonthly transition matrices in appendix B in the online material. Except for the bimonthly result for the Apple iPhone, we observe the above expected probabilities along the diagonal of the transition matrices for all products. Additionally, for a direct comparison with Lach (2002), we present results for retailer prices divided into quartiles in appendix C in the online material. This analysis also shows higher than expected probabilities along the diagonal of the transition matrices.

Table 1: Cluster characteristics and transition matrix results.

			Cluster sizes					Cluster rank characteristics			Monthly transition matrix characteristics	
Category	Product	Mean CoV*	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Total days sold	Frequency of shifts between clusters	Average days within a cluster	Average diagonal	Expected value diagonal
Cellphones	Apple iPhone 5s 16GB	0.72	13	24	13			787	29	26	0.43	0.33
Consoles	Sony PlayStation 4 Pro 1TB	0.54	8	5				544	6	78	0.85	0.50
Headphones	Bose QuietComfort 35	1.46	16	10	62	5		1623	14	108	0.67	0.25
Laptops	Apple MacBook Pro	0.62	8	6	11	6	3	719	31	22	0.44	0.20
Mobile Speakers	Sonos Play:1	1.44	6	21				1231	4	246	0.72	0.50
Nintendo 3DS	Pokémon Sun	0.29	4	4	6			728	3	182	0.58	0.33
Nintendo Wii U	tLoZ: BotW	0.28	7	1	7			422	6	60	0.78	0.33
PC	Battlefield 1	1.28	2	8	4	3		422	4	84	0.61	0.25
PlayStation 3	Grand Theft Auto V	0.64	8	5	3			1274	13	91	0.52	0.33
PlayStation 4	FIFA 17	0.26	8	13	2			301	16	18	0.59	0.33
Tablets	nVidia Shield Tablet K1 16GB	0.83	9	13	6			638	13	46	0.52	0.33
TV	Samsung UE55KS7005	1.02	6	1	1	1	1	270	13	19	0.59	0.20
Xbox 360	Grand Theft Auto V	0.66	6	3				1274	9	127	0.79	0.50
Xbox One	Grand Theft Auto V	0.34	9	4				483	7	60	0.81	0.50

*Coefficient of Variation.

References

Baye, M. R., Morgan, J., & Scholten, P. (2004). Temporal price dispersion: Evidence from an online consumer electronics market. *Journal of Interactive Marketing*, 18(4), 101–115.

Bayliss, K & Perloff, J.M. (2002). Price dispersion on the internet: Good firms and bad firms. *Review of Industrial Organization*, 21(3), 305–324.

Keogh, E., Chakrabarti, K., Pazzani, M., & Mehrotra, S. (2001). Dimensionality reduction for fast similarity search in large time series databases. *Knowledge and Information Systems*, 3(3), 263–286.

Lach, S. (2002). Existence and persistence of price dispersion: An empirical analysis. *Review of Economics and Statistics,* 84(3), 433–444.

Narasimhan, C. (1988). Competitive promotional strategies. *The Journal of Business*, 61(4), 427–449.

Rath, T. M., & Manmatha, R. (2003). Word image matching using dynamic time warping. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2*.

Rosenthal, R. W. (1980). A model in which an increase in the number of sellers leads to a higher price. *Econometrica*, 48(6), 1575–1579.

Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy*, 69(3), 213–225.

Thorndike, R. L. (1953). Who belongs in the family. *Psychometrika*, 18(4), 267–276.

Varian, H. R. (1980). A model of sales. *The American Economic Review*, 70(4), 651–659.

Appendix A: Hierarchical clustering and cluster evaluation.

Analysis was conducted in R (R Core Team, 2019). We apply Piecewise Aggregate Approximation (PAA) using the jmotif package (Keogh et al., 2001) to reduce dimensionality of our residual price series. This method entails averaging over equal-sized segments of the price series and is useful in similarity searches, in our main case 30-day windows were chosen. The commonly used and package default option of the Euclidean distance is used in our clustering, and we scale our residual time series as is commonly advised in time series data mining (Keogh & Kasetty, 2003). When calculating the distance matrix we use Dynamic Time Warping (Sakoe & Chiba, 1978) using the dtw package (Giorgino, 2009) for aligning our retailer heterogeneity adjusted price series, the reason being that if a retailer reduces the price in advance of competitors, even by a single day, it will have a disproportionally large effect on whether that retailer belongs to the cluster of retailers, even if the price patterns are otherwise visibly similar. The resulting distance matrices were then analysed using the hierarchical clustering algorithm which is part of base R (R Core Team, 2019).

A fundamental question in clustering techniques is how to choose the number of clusters, and we choose to use the "elbow method" (Thorndike, 1953), where in our case the height of the dendrogram in the hierarchical clustering described above is used as the entropy measure. The elbow is found when the height of the dendrogram branches sharply reduces, at this point the number of clusters are selected. To find the elbow, we use a novel stopping rule by calculating the point of sharpest decline in height of the dendrogram, e.g. the elbow itself. This is done by iteratively calculating the delta of heights between cluster solutions k and $k + 1$, and thereafter calculating the difference between these delta values. These in turn represents the first and second order difference which is then subtracted to give a numerical indication of the strength of the elbow. The cluster number K is then chosen according to the largest numerical strength.

Appendix B: Transition matrices between clusters.

To assess retailer transition between clusters, we create Markovian transition matrices using the markovchain package (Spedicato, 2017). A Markovian K x K transition matrix has elements which contain probabilities of transitioning from one state to another, in our case for an individual retailer to transition between retailer price clusters. To distinguish whether a particular retailer transitions from one cluster to another, we needed to create a boundary between the cluster centroids, our choice was the straightforward midpoint between these centroids. That is, if a retailer belonging to a certain cluster passes the midpoint boundary across to another cluster this would register as a transition between these two clusters in the corresponding transition matrix.

The assumption of time-homogeneity in transition matrices should always be subject to scrutiny. In this paper, we use the χ^2 test designed by Anderson and Goodman (1957) to investigate whether transition probabilities are constant. Rejecting the null hypothesis of time-homogeneity would therefore entail that transition probabilities are not constant. We find that the only products that fail the test at the 5 % level is FIFA 17 when considering the monthly and biweekly transition matrices, and the Xbox One version of Grand Theft Auto for the biweekly transition matrix, implying that we have evidence that the assumption of time-homogeneity holds for the majority of our data.

We provide results for K x K transition matrices for monthly transitions in the main article, while also providing the same results for biweekly and bimonthly transitions in Table B.1 below, the former being measured on Mondays every odd week while the latter is measured on the 1st of each odd month. In a K x K transition matrix, the likelihood of randomly being observed in the diagonal will equal 1/K, which is calculated and presented in the last column of Table B.1. Comparing the biweekly and bimonthly diagonal probabilities to the expected ones, the results show that, except for the bimonthly result for the Apple iPhone, we observe above expected probabilities along the diagonal of the transition matrices for all products.

Table B.2: Cluster characteristics and transition matrix results with biweekly and bimonthly transition.

Category	Product	Mean CoV*	Total days sold	Biweekly transition matrix Average diagonal	Bimonthly transition matrix Average diagonal	Expected value diagonal
Cellphones	Apple iPhone 5s 16GB	0.72	787	0.45	0.25	0.33
Consoles	Sony PlayStation 4 Pro 1TB	0.54	544	0.86	0.71	0.50
Headphones	Bose QuietComfort 35	1.46	1623	0.59	0.51	0.25
Laptops	Apple MacBook Pro	0.62	719	0.69	0.42	0.20
Mobile Speakers	Sonos Play:1	1.44	1231	0.88	0.74	0.50
Nintendo 3DS	Pokémon Sun	0.29	728	0.67	0.58	0.33
Nintendo Wii U	tLoZ: BotW	0.28	422	0.86	0.74	0.33
PC	Battlefield 1	1.28	422	0.66	0.52	0.25
PlayStation 3	Grand Theft Auto V	0.64	1274	0.71	0.59	0.33
PlayStation 4	FIFA 17	0.26	301	0.72	0.43	0.33
Tablets	nVidia Shield Tablet K1 16GB	0.83	638	0.65	0.43	0.33
TV	Samsung UE55KS7005	1.02	270	0.73	-[#]	0.20
Xbox 360	Grand Theft Auto V	0.66	1274	0.85	0.68	0.50
Xbox One	Grand Theft Auto V	0.34	483	0.84	0.75	0.50

*Coefficient of Variation. [#]This probability was not calculated due to the short time period of only 270 days, or 4 bimonthly transitions.

Appendix C: Transition matrices between quartiles.

Our study adopts clustering techniques to determine the number of groups of retailers with similar pricing (price clusters), and then calculate transition matrices using the number of clusters found, instead of arbitrarily dividing prices into quartiles as in Lach (2002). However, for a direct comparison to Lach (2002), we present results for the heterogeneity adjusted retailer prices divided into quartiles in Table C.1. We compare at the lowest frequency of data that Lach (2002) has available, which is the one-month horizon, and our data is divided into monthly intervals by using the first day of the months as subset.

In a 4 x 4 transition matrix, pricing using mixed strategies should give a likelihood for the diagonal of 0.25, while in our data the likelihood for the diagonal for the quartile separated retailer prices range between 0.47 and 0.86. As such, the results from the main analysis remain, even if using the quartile analysis of Lach (2002).

Table C.1: Cluster characteristics and transition matrix results with quartiles.

Category	Product	Mean CoV*	Total days sold	Average diagonal	Expected value diagonal
Cellphones	Apple iPhone 5s 16GB	0.72	787	0.58	0.25
Consoles	Sony PlayStation 4 Pro 1TB	0.54	544	0.55	0.25
Headphones	Bose QuietComfort 35	1.46	1623	0.80	0.25
Laptops	Apple MacBook Pro	0.62	719	0.66	0.25
Mobile Speakers	Sonos Play:1	1.44	1231	0.77	0.25
Nintendo 3DS	Pokémon Sun	0.29	728	0.80	0.25
Nintendo Wii U	tLoZ: BotW	0.28	422	0.86	0.25
PC	Battlefield 1	1.28	422	0.64	0.25
PlayStation 3	Grand Theft Auto V	0.64	1274	0.61	0.25
PlayStation 4	FIFA 17	0.26	301	0.47	0.25
Tablets	nVidia Shield Tablet K1 16GB	0.83	638	0.45	0.25
TV	Samsung UE55KS7005	1.02	270	0.47	0.25
Xbox 360	Grand Theft Auto V	0.66	1274	0.60	0.25
Xbox One	Grand Theft Auto V	0.34	483	0.53	0.25

References

Anderson, T.W., & Goodman, L.A. (1957). Statistical inference about Markov chains. *The Annals of Mathematical Statistics*, 26(1), 89–110.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: The {dtw} Package. *Journal of Statistical Software*, 31(7), 1–24.

Keogh, E., Chakrabarti, K., Pazzani, M., & Mehrotra, S. (2001). Dimensionality reduction for fast similarity search in large time series databases. *Knowledge and Information Systems*, 3(3), 263–286.

Keogh, E., & Kasetty, S. (2003). On the need for time series data mining benchmarks: a survey and empirical demonstration. *Data Mining and Knowledge Discovery*, 7(4), 349–371.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. Retrieved from https://www.r-project.org/

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43–49.

Spedicato, G.A. (2017). Discrete time Markov chains with R. *The R Journal*, 9(2), 84–104.

Thorndike, R. L. (1953). Who belongs in the family. *Psychometrika*, 18(4), 267–276.