

# *Price Dispersion Analysis of Homogeneous Products in Online Markets*

Xbox 360 Games

Saumya Gupta  
[h20saugu@du.se](mailto:h20saugu@du.se)

Mian Muhammad Usman Zahid  
[v21miaza@du.se](mailto:v21miaza@du.se)

**Abstract**—Using the pricing data collected from a price comparison website of products belonging to the Xbox 360 category, we tried to identify if price dispersion exists for homogeneous products across multiple online retailers in this market and whether intertemporal price discrimination is the reason behind price dispersion. Price dispersion is substantially present in this market and is increasing with time. Our results show that some retailers are following randomized pricing and trying to undercut each other. Their behaviour reflects the theory of intertemporal price discrimination that suggests retailers should be randomizing their prices over time to avoid price pattern identification by consumers. However, many retailers practice persistent pricing over time and are not worried about consumers learning their pricing patterns. It suggests that intertemporal price discrimination is most likely not the reason for price dispersion for most products. The clearinghouse model is thus not entirely supported by our analysis.

**Keywords**—component; Intertemporal price discrimination; price dispersion; clearinghouse model; hierarchical clustering; time-series data

## I. INTRODUCTION

In his work, Varian (1980) describes price dispersion as multiple retailers selling products with homogeneous characteristics for different prices and then further explains the temporal price dispersion as "each store varying its prices over time". Varian (1980) points out that "the retail markets are characterized by a large degree of price dispersion" which invalidates the "law of one price". This law has been widely accepted as not applicable anymore in multiple studies (Varian 1980; Lach 2002; Baye, Morgan and Scholten, 2004). In their article, Lindgren et al. (2020) quote a 19% standard deviation in prices for identical products sold in the US across 54 markets. A similar study conducted by Lach (2002) that measured four homogeneous goods across retailers in Israel for 48 months concluded that price dispersion exists in the market.

The clearinghouse model explains such price dispersion over time. In his descriptions of the clearinghouse model, Varian (1980) talks about two types of consumers: informed and uninformed. The informed consumers are aware of the price dispersion and go for the low-priced stores compared to the uninformed consumers who shop at random (Varian,

1980). Then, some stores can randomize prices between them to appeal to both.

Retailers who randomize their prices have a better chance of attracting both informed and uninformed consumers (Varian, 1980). Otherwise, informed consumers would identify the trend and always shop at the lowest cost retailer, or all the other retailers would have to reduce their price to meet the competition, and then price dispersion would not exist.

The retailers and regulators, such as competition authorities, would like to know how the market functions to earn greater profits. This report aims to help them do so by answering two questions: (1) Is price dispersion prevalent in the market? If so, we attempt to gauge the extent of it. (2) Is intertemporal price discrimination the cause of this price dispersion in the market as suggested by the clearinghouse model? Hence, we analyze the pricing trends of online retailers for homogeneous products in the market and identify if the retailers are prone to persistently setting their prices in predictable clusters, for example, of a low, medium and high pricing, or are randomizing their prices while competing for consumers, eventually decreasing prices.

This report uses the following structure. After the introduction in section I, section II summarizes the data used for the Xbox 360 category and emphasizes the details of the methods used: (1) to identify price dispersion in the market and (2) to test the clearinghouse model against the provided dataset. It also mentions the materials used in the process. Section III explains the results achieved from the research carried throughout in this report. Lastly, section IV concludes our studies by answering the two questions raised by the retailers and regulators mentioned previously.

## II. METHODOLOGY

For analyzing the functioning of online markets for Xbox 360, the data provided is high frequency (daily) pricing data of 52 retailers in Sweden selling 64 games in the category, for a period of ~54 months (2012 Aug – 2017 Feb), previously collected from PriceSpy by Lindgren et al. (2020) for their studies, in an excel long format. The data is masked, implying we do not have any popularity measure

(click-through rate or the like) to justify price difference across products and therefore, our studies focus on price dispersion across retailers selling the same product. We use ggplot2 (Wickham, 2016) to support our results with figures. We found that even though by 2015, the game releases declined as publishers focussed more on Xbox One after its release in 2013, online retailers kept entering the Xbox 360 market, shown by Fig. 1.

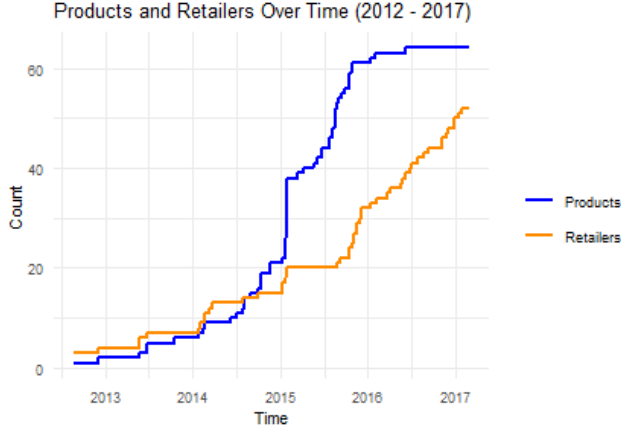


Fig. 1 Product and online retailers over the years in the Xbox 360 market, showing an increase in the number of products and webpages, especially during early 2015.

For all our analysis, we use CPI(consumer price index)-adjusted prices (made available to us) to account for the little inflation during the four years. For result 1, we look at the range and coefficient of variance (abbreviated as CoV) across prices of all products to check for the presence of price dispersion in the market.

The products range from being sold by 2-24 retailers. For result 2, since it is hard to explore patterns among many retailers, we try to find natural groupings using clustering. Before clustering, we z-score the CPI-adjusted prices to avoid dominance by large values in distance computations. We then use hierarchical clustering on time series to cluster many retailers selling the same product. For products sold by a lesser number of retailers ( $< 10$ ), we analyze price trends (time series) without clustering. Following Lindgren et al. (2020), we use dynamic time warping as the distance measure, using dtw (Giorgino, 2009), because it accommodates unequal length time series and is not sensitive to time shifts. To determine the optimum number and size of clusters, we use NbClust (Charrad et al., 2014), which provides up to 30 indices. If clusters undercut each other, it implies retailers are practising intertemporal price discrimination, which is the reason for existing price dispersion.

### III. RESULTS

Fig. 2 (B) shows an increasing range of prices across retailers of a single product, which justifies a similar observation in Fig. 2 (A), where we consider all the products in the category. Increasing CoV of prices supports this and hence suggests increasing price dispersion in the market across the years. Table 1. shows that products with a greater mean CoV are ones with a lesser number of retailers. We first calculate the CoV of CPI-adjusted prices across retailers for each product and time point and then average it over time to calculate the mean CoV. More computations show that more than 50% of products have their mean CoV of prices of greater than 0.3.

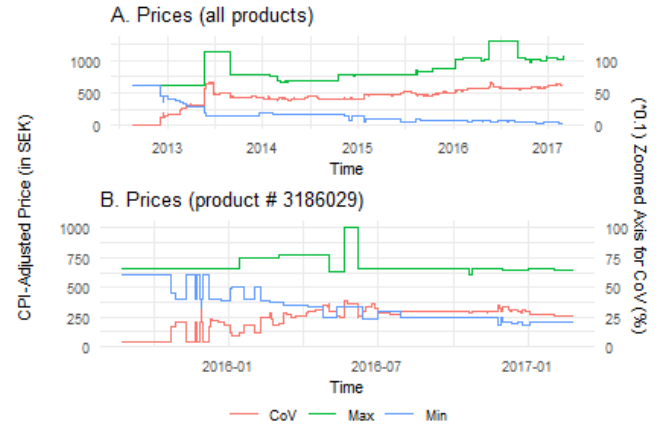


Fig. 2 Max, min and CoV of prices over time, across: (A) all products and (B) product ID 3186029. (Read CoV using the secondary axis.)

TABLE 1. PRICE DISPERSION (AVERAGE OVER FOUR YEARS)

Product ID	Mean CoV*	No. of Retailers
1769910	.77	11
1410410	.64	7
353454	.63	7
917003	.63	13
971856	.52	8
1180267	.52	8
1965068	.50	11
1026418	.50	12
729380	.48	4
2719635	.48	10

\*Coefficient of Variation

During our analysis, we have found an unusual behaviour shown by retailer ID 1260, for product ID 2687705, during the year 2015. The retailer sells the mentioned product

usually for prices ranging from 429 to 466 kr, throughout the year, but abruptly increases it to an exact figure of 4062 kr for a total of 38 days, randomly spread across the year, for reasons unknown, given the limited variables in the data. It is atypical of an online retailer to increase a product price almost ten times within a day. We, hence, decided to remove these 38 data points from our study, assuming them to be outliers. Fig. 2 ignores these data points while computing aggregates across all products.

Fig. 3 shows product-level price analysis. The retailers for the product in Fig. 3 (A) practice consistent pricing strategies of a predictable pattern, whereas retailers for the product in Fig. 3 (B) seem to undercut each other, fiercely competing for shoppers. We found all products in the category oscillate between these two types of results.

Clustering results in 1 – 4 clusters depending on the product. While some products show similar randomization trends as shown for product ID 1341635 in Fig. 4, the maximum of the products shows that retailers are clustered into two levels with little to no overlap, practising consistent pricing over time, similar to the product in Fig. 5.

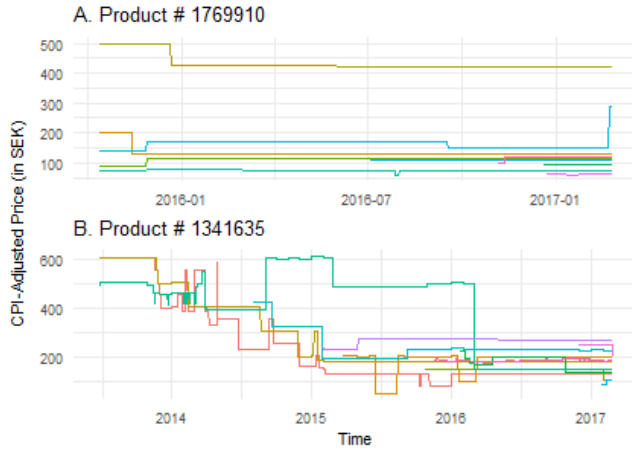


Fig. 3 Pricing trends across all retailers for the product ID: (A) 1769910 and (B) 1341635. Different retailers in different colours.

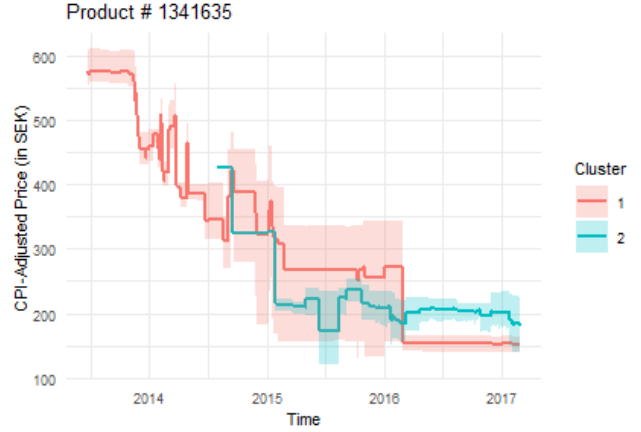


Fig. 4 Clustered retailers for product ID 1341635 with 25% - 75% interquartile range (IQR) ribbons.

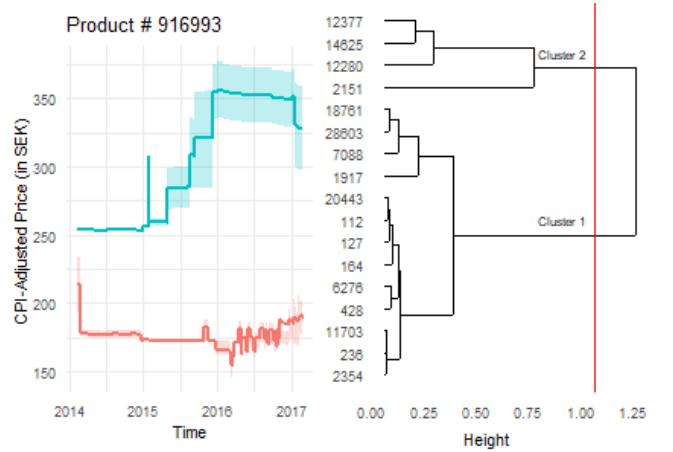


Fig. 5 Clustered retailers for product ID 916993 (left) Cluster dendrogram (right). The red line cuts the tree into two clusters.

#### IV. CONCLUSION

We investigated the predictions of the clearinghouse model on the dataset of homogenous goods obtained from the price comparison website. We looked at a single category of products (Xbox 360) and explored the data accounting for four years. In congruence with the expectations of the mentioned model, we have found price dispersion in the market. We notice that it is also increasing over the years. Regarding the cause of this price dispersion, some retailers are maintaining consistent pricing across time for the same product, whereas other retailers are randomizing their prices to undercut each other and to remain unpredictable in their pricing trends. There is a mixed pricing strategy in this market, suggesting we cannot hold intertemporal price discrimination entirely responsible for the price dispersion present in the market. Clustering results that show many retailers grouped into two non-overlapping clusters of consistently high and low

pricing further verify this. Based on our findings, we can conclude that the market has mixed behaviour where some retailers randomize their prices while others have consistent pricing over time, therefore, rejecting the clearinghouse model.

# REFERENCES

- [1] Baye, M. R., Morgan, J., Scholten, P., 2004. Temporal price dispersion: Evidence from an online consumer electronics market. *Journal of Interactive Marketing*, 18(4), pp. 101-115.
- [2] Charrad, M., Ghazzali, N., Boiteau, V. & Niknafs, A., 2014. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, 61(6), pp. 1-36.
- [3] Giorgino, T., 2009. Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package. *Journal of Statistical Software*, 31(7), pp. 1-24.
- [4] Lach, S., 2002. Existence and Persistence of Price Dispersion: An Empirical Analysis. *The Review of Economics and Statistics*, 84(3), pp. 1-12.
- [5] Lindgren, C., Daunfeldt, S.-O., Rudholm, N. & Yella, S., 2020. Is intertemporal price discrimination the cause of price dispersion in markets with low search?. *Applied Economics Letters*.
- [6] Varian, H., 1980. A Model of Sales. *The American Economic Review*, 70(4), p. 1-9.
- [7] Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. 2nd ed. New York: Springer-Verlag.