Price matching and platform pricing

Bottasso, Anna and Marocco, Paolo and Robbiano, Simone

Università degli Studi di Genova, Department of Economics

March 2021

Price Matching and Platform Pricing*

Anna Bottasso‡ Paolo Marocco‡ Simone Robbiano‡†

‡University of Genoa
Department of Economics

April 2022

Abstract

We investigate the effects of Price Matching Guarantees policies adopted by the US NewEgg online platform on prices of a representative sample of consumer electronics products. By applying a Difference-in-Differences identification strategy, we find price reductions of about 4% occurring after the policy implementation. We control for products characteristics recovered from User Generated Contents (products popularity and quality) and perform heterogeneity analysis based on products visibility (Google Search Rank). Estimates suggest that for high (low) visible products prices are higher (lower) during the policy validity period. Results are consistent with the hypothesis that such policies can act as tools for price discrimination.

Keywords. Price Matching Guarantees; Online Sales Platforms; User Generated Contents; Difference-in-Differences.

JEL Classification. L11; L13; L15; L81.

1 Introduction

Online sales platforms have recently gained increasing importance in both retail and wholesale markets.[1] Such markets are characterized by the supply of personalized

*A previous version of this paper was included in the Ph.D. dissertation of Paolo Marocco and Simone Robbiano. The authors thank seminar participants at the 1^{st} JRC Meeting of the NetCIEx (Network on Counterfactual Impact Evaluation), the 61^{st} Conference of the Italian Economic Association and the COMPIE 2021. The usual disclaimer applies. Declarations of interest: none. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

†Corresponding author. *Address*: Via F. Vivaldi, 5 – Darsena – 16126 Genova. *e-mail address:* simone.robbiano@economia.unige.it

[1]The term "online platform" identifies a range of digital services that facilitates interactions between two or more distinct but interdependent sets of users (whether firms or individuals) who interact through the service via the Internet (OECD, 2019). Online sales platforms can operate as online retailers, as a marketplace for third-party sellers or they can offer both services.

services, more convenient delivery schedules and the ability to reach a very high number of consumers. In addition, platforms claim to warrant lower prices with respect to traditional stores through the provision of special offers, promotions, down prices and other price discounting policies. Among these options, online sales platforms often implement Price Matching Guarantees (PMGs) policies, that is the promise to match lower prices offered by competitors.[2]

PMGs policies are surely appealing for customers and can guarantee low prices, increase consumer confidence and brand fidelity; however, the announcement to tie prices to those of competitors can have anti-competitive effects and sustain high prices in certain markets, thus harming consumers' welfare. Indeed, the theoretical literature has analysed PMGs policies along various directions, since they can affect firms and consumers' choices in different ways.

Hay (1981), Salop (1986) and Belton (1987) first suggest that PMGs can sustain collusion in oligopoly models and highlight that such clauses might be considered as threats to punishment for firms that lower cartel prices, thus reducing firms' incentive to deviate from agreements. Several other papers support the pro-collusive argument by extending the basic oligopolistic setting or applying the Hotelling model.[3] However, Hviid and Shaffer (1999) highlight that the presence of hassle costs (costs for activating the guarantee) might undermine possible anti-competitive effects of PMGs.

Some other models suggest that sellers can use PMGs policies as a price discrimination tool. Png and Hirshleifer (1987), Belton (1987), Corts (1996) and Nalca et al. (2010) suggest theoretical models where firms discriminate between different consumers groups, namely "uninformed" customers and "informed" ones. Indeed, if customers differ in terms of information on prices and guarantee terms, willingness to pay, loyalty to a specific retailer or the extent of hassle costs, firms could use the price guarantee to discriminate across them.[4]

Hence, the theoretical literature suggests that PMGs policies can have different effects on prices, according to markets conditions and consumers characteristics. In particular, it seems interesting to investigate such issue for online markets that have received less attention with respect to traditional ones.

Our work provides empirical evidence on the effect of platforms' PMGs policies on daily consumer electronics prices observed on the US online market between May

[2]For example, NewEgg's PMG policy states that "if you purchase an item from Newegg.com which is carrying the Price Match Guarantee badge at the time of purchase, then find the exact same item at a lower price by Newegg or a major retailer, just let us know, and we'll send you a Newegg Customer Care Card to cover the difference". See `https://kb.newegg.com/knowledge-base/price-match-guarantee/`.

[3]E.g. Logan and Lutter (1989); Baye and Kovenock (1994); Lu and Wright (2010); Hviid and Shaffer (2010); Pollak et al. (2017); Constantinou and Bernhardt (2018); Cabral et al. (2018).

[4]Similar results can be found also in Edlin (1997). Moreover, some authors (Moorthy and Winter, 2006; Jain and Srivastava, 2000; Moorthy and Zhang, 2006) suggest that PMGs might be a credible signal of low prices, if low cost firms adopt the policy and (high cost) competitors can not match it.

and October 2018. In particular, we focus on the NewEgg platform that exclusively sells consumer electronics products and implements PMGs policies. Given that our identification strategy is based on a comparison of price levels before and after the policy implementation, we have excluded platforms that never stop offering PMGs (like Target). The empirical strategy is based on a Difference-in-Differences (DiD) design where the pool of NewEgg products affected by PMGs policies is considered as the treated sample. Differently from standard practices in studies adopting a DiD approach, we build the control sample with price data for the same products observed on a different platform, namely Amazon, that never offers PMGs to customers. Furthermore, in order to ensure that the counterfactual sample is less likely to be influenced by PMGs policies adopted by NewEgg, we consider data from the Amazon UK platform (instead of Amazon US). Indeed, price observed on Amazon US might not be completely independent from the policy under scrutiny, because of price tracking practices frequently adopted by platforms within the same country.

Estimates provide evidence in favor of an average price reduction of about 4% after the implementation of PMGs policies. Moreover, in order to have a more detailed picture of the issue, we investigate if such effect is heterogeneous and we focus on products features that might affect the outcome of PMGs policies and that can be recovered on online markets through the analysis of Users Generated Contents (UGCs).

Platform data provide evidence on products visibility that can be associated to product-specific search costs; moreover, data on products' popularity and quality are available for each product. The distribution of such characteristics across products is likely to be associated to consumers heterogeneity in terms of information, willingness to search (and to invoke PMGs) and price elasticity and such heterogeneity can affect the impact of PMGs on prices.

Estimates conducted on specific sub-samples based on values assumed by the search index, our proxy for products visibility, suggest that when PMGs are off, products characterised by a low search rank (low visible) experience a price increase of about 2,5%, while for high visible products prices decreases of about 5%. Moreover, since quality and visibility are correlated in our sample, we distinguish low quality and low search rank products from high quality and high search rank ones. Estimates suggest that low quality/visible products experience a price increase of about 3.5% after the policy implementation periods, while for high quality/visible products a price reduction of about 4% is observed.[5]

This study enriches the literature on the effects of PMGs on prices along different lines. First, it extends previous research by analysing very detailed real-time platform data without using price-tracking engines and tools for data extraction from price

[5] It is worth noting that our identification strategy is based on a comparison of price levels before and after the policy implementation.

charts. Second, products characteristics based on platform information and UGCs are employed for the first time in order to study possible heterogeneous effects of PMGs. Third, the DiD identification strategy adopted is based on the construction of a control group with a novel approach that guarantees its independence.

The paper is organised as follows. Section 2 discusses previous empirical literature; Section 3 describes the data extraction process and provides summary statistics; Section 4 explains our identification strategy; Section 5 discusses estimates results and robustness analysis. Section 6 concludes.

2 Empirical Evidence

The empirical literature that studies the effects of PMGs policies on prices focuses on specific markets (tyre, gasoline) and on retailing prices from supermarkets; just one study analyses online markets.

Arbatskaya et al. (2000) recover daily price quotes from the tyre industry advertisements from 61 US Sunday newspapers observed for three months in 1996. Authors find weak evidence of anti-competitive effects of PMGs and show that an increase in the number of firms implementing the policy leads to a 10% increase in prices.[6] Cabral et al. (2018) focus instead on daily pricing policies adopted by the Shell network of gas stations in Germany in 2015.[7] Leveraging on gas stations localization and consumers demographics as sources of identification, they suggest that PMGs can be a collusion enacting policy. Gas station prices have been analyzed also by Byrne and De Roos (2019) for Australia by means of a detailed 15 years time series dataset. Authors argue that the majority of gas stations prices follow a weekly cycle and that dominant firms can use PMGs to coordinate market prices and reduce price competition. Similar results can be found in Chilet (2018), who analyses pricing policies of three big retail pharmacy chains in Chile, observed over the period 2006-2008. The author follows an identification strategy based on the estimation of a demand model, in which quantity sold is a function of price differentials with competitors, around the time period where collusive price increases occurred.

Hess and Gerstner (1991) analyse instead the effect of PMGs on prices by collecting weekly data of 114 goods sold in several US supermarkets and grocery stores, from 1984 to 1986 and provide evidence in favor of higher prices (about 1-2%) when the guarantee is introduced. Different results are provided by Moorthy and Winter (2006), who analyse prices of several products sold by 46 Canadian retailers in 2002 and assume the existence of informed and uninformed consumers. Authors argue that the adoption of PMGs policies might be interpreted as a way to signal lower prices to

[6] The same authors in Arbatskaya et al. (2006) confirm their results by analysing the same data with a different approach.

[7] See also Wilhelm (2016).

uninformed customers and suggest that PMGs are mainly adopted by low cost/low service chain stores.[8] Similar results can be found in Chung et al. (2016) for three leading hypermarkets in Korea.

Finally, Zhuo (2017) focuses on online platforms and collect US price data from online price trackers for 150 products offered on Amazon in 2012. The author observes prices during and after the implementation of PMGs policies by two big-box stores (Target and Best Buy) targeted specifically on Amazon prices; by applying DiD and RDD methods, the author suggests that prices increase by about six percentage points during the period of validity of the policy. Moreover, the analysis highlights an heterogeneous impact of PMGs, with larger price increases for initially lower-priced goods.[9]

3 Data

3.1 Data Extraction

In order to study the impact of PMGs on prices, we focus on the online consumer electronics market, since it is one of the most widespread sector on online retailing and is often affected by such pricing policies. In particular, electronic products are search goods, whose quality can be evaluated before the purchase: the advent of online markets has made this process much cheaper and faster and is most likely to affect the impact of such policies, whose outcome depends, among other factors, by the level of search and hassle costs. Moreover, electronic goods are barely affected by seasonal effects, so that prices signals are more stable over time.[10]

Among different online retailing platforms, we choose to focus on NewEgg, a leading online US retailer of consumer electronics products, that implements PMGs policies and switch them on and off over time on selected items. Given that our identification strategy is based on the comparison of prices before and after a policy's switch off, we do not consider platforms that apply PMGs to wide groups of products continuously over time (i.e. Target, among others). In particular, NewEgg communicates the period of validity of the price guarantee by means of a label that appears on the specific product online page; the customer who discovers the PMG badge has 14 calendar days of time to find the same "title" at a lower price from US competitors

[8] Authors suggest that firms offering higher prices do not find convenient to apply PMGs as it would imply devolving their pricing decisions to low price competitors (Moorthy and Winter, 2006).

[9] Some other authors analyse the impact of price-beating guarantees, that are less widespread policies with similar terms as price matching ones (in price beating guarantees refund exceeds the price difference). Studies that refer to these policies argue that, with respect to price matching guarantees, they might be serving different purposes in practice and likely be effective in enhancing competition. Experimental literature also focuses on the effect of price matching and price-beating guarantees: however, experimental results lack the complexity of real interactions between sellers and consumers.

[10] Our sample covers the period May - October 2018 and does not include important dates like Thanksgiving or Christmas.

belonging to a declared list.[11]

In order to build the treated sample we have identified all NewEgg electronics products affected by PMGs on May 2018. We have observed such products (100) from May 2018 to October 2018 and we have collected price data, PMGs information and other Users Generated Contents (UGCs).[12] The control sample has been built by selecting the same products sold on other platforms (Amazon US and UK) that never offer PMGs policies. However, crucially for our identification strategy, we have assumed that PMGs implemented by NewEgg might affect products that are sold in US, so that prices of identical products sold by Amazon UK are less likely to be affected and well represent a counterfactual sample, mimicking what would have happened to prices of treated products in the absence of PMGs. This approach has implied a reduction in the number of observed products, so that the final sample includes 58 products belonging to 19 sub-categories (computer hardware, tablet and computers, mobile phones, printers and scanners, PC accessories, speakers for domotics, screens and audio devices).[13] It is worth noting that we have collected data of products sold directly from NewEgg and Amazon and we did not consider third-party sellers products.

The retrieving of sample data has been a challenging task. Given the absence of ready-made and easy-to-use repositories on price data, we have developed an ad-hoc scraping program (in Python language) able to protect the scraping process from unpredictable changes of the page and capable to recover the data without stressing the site, thus limiting the risk of interruptions due to firewalls. In particular, the scraping process has been supported by several alert tools signalling periodical changes of the internal page structure, given that platforms frequently change the deep structure of the page, in a not visible way by the human reader but in a way that affects the program code and the scraping process. The process of data collection has required the subscription to the Amazon Web Service (AWS) cloud in order to use virtual servers where installing and launching the daily loop process. The scraping code allowed us to navigate among product pages, select the field tags, get the data and save on a server disk. Each scraping session run about 20 minutes every day.

In addition to information on product prices and PMGs, we also have collected some product characteristics available exclusively on online sales platforms. In particular, we have recovered some UGCs, like the number of reviews and products rating given by customers. The absolute number of reviews is a dynamic information which represents a sort of popularity index, since it is proportional to the product

[11] With "title" we refer to a product with the same brand and model number. NewEgg, after checking the validity of the claim, sends a Customer Care Card to refund the price difference (Source: `https://kb.newegg.com/knowledge-base/price-match-guarantee/`).

[12] The average number of treatments occurred in our sample is about 2.38, thus suggesting that, on average, the policy is applied twice to each product during the sample period.

[13] In the Appendix B we provide a detailed list of selected products (Tables B.1 and B.2).

market diffusion. Starting from this information, we have built the relative number of reviews as the ratio between the latter and the amount of reviews received from the most popular product in the same sub-category. This normalized index, that ranges from zero to one, shows the relative popularity of the product with respect to other items of the same sub-category. Another interesting information is the products rating (stars) provided by consumers. We consider the number of stars gained by each product, ranging from zero (low quality) to five (high quality), as a proxy of product quality. Finally, we have built a product specific search rank as a proxy of the time spent on the search engine to discover the specific web-page of a certain product. In particular, for each product we have launched, at the beginning of the sample period, the Google query composed by the sentence ("*product name*" AND "*platform name*") and we have recovered its ranking position.[14] Such position has been normalized in order to interpret the search index as the probability to find the product in first ranked positions of Google. It is worth noting that, although products analysed are sold by Amazon and NewEgg in different countries, information on some of considered UGCs maintain their consistency across countries. This property is typical of consumer electronics goods that have a standardized nature. However, we adopt a country-specific search index by launching the Google search engine with specific country settings.[15]

3.2 Descriptive Statistics

The analyzed sample includes 9028 daily price observations for 58 products observed on NewEgg and Amazon UK platforms, from May 2018 until October 2018. Table 1 shows summary statistics on prices and selected product characteristics for the overall sample and for treated and control ones. Prices show a large variability, being the average for the overall sample $240.43 and the standard deviation $283.53. Average prices for the control sample are lower than those observed over the treated sample; however, such pattern does not represent an issue for our identification strategy as long as the parallel trend assumption is satisfied (see Section 5.2).

Another important issue is related to the representativeness of our sample. Figure 1 represents the distribution of products by price classes (10). The histogram shows that 78% of the products belong to the first two price decile, with price ranging between 14$ and 287$, and 17% to the third to fifth decile (prices between 316$ and 774$); the remaining products are ranked from the sixth to the tenth decile, with price ranging between 886$ and 1565$. This pattern closely matches a typical price distribution observed in several markets (Coad, 2009), often characterized by a large amount of

[14] It is worth noting that such query can provide not only the specific product page but also a similar product page or a bucket of products that includes the specific object of the search. We rank only the product's specific web page.

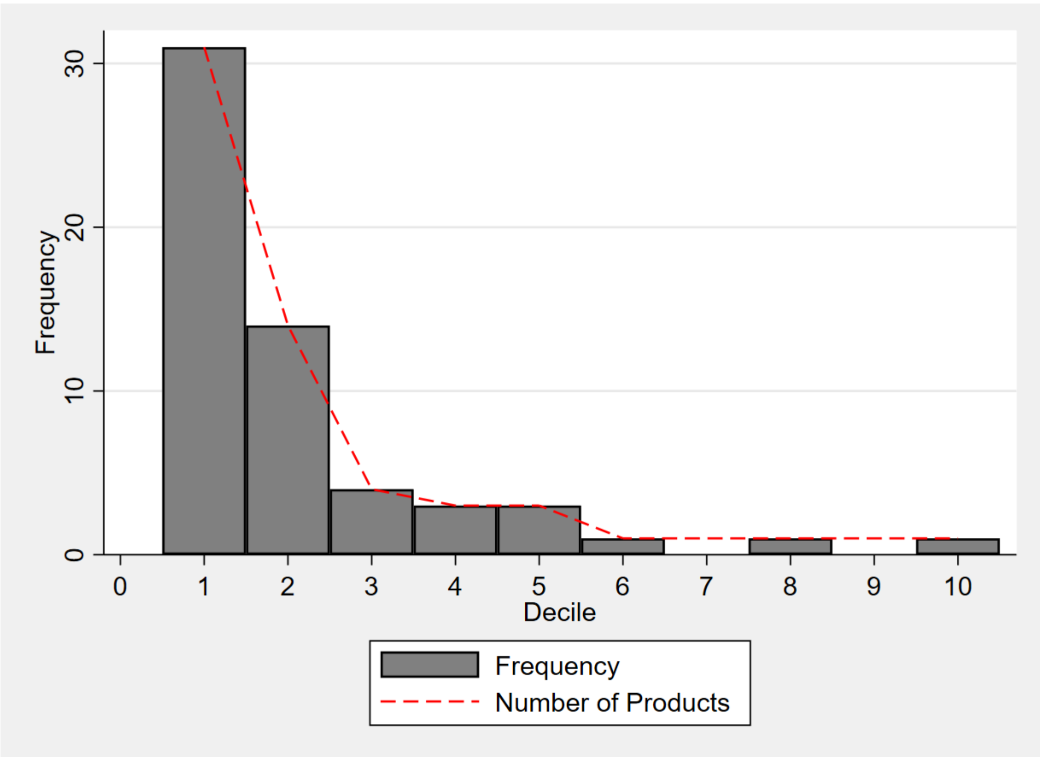
[15] Amazon UK prices have been converted into dollars at the daily exchange rate.

Table 1: Summary Statistics.

Variables	Full Sample	Treated Sample	Control Sample
		NewEgg	Amazon UK
Provider Price ($)	240.43	253.15	227.72
	(283.53)	(302.39)	(262.74)
Product Popularity (0-1)	0.23	0.20	0.26
	(0.27)	(0.23)	(0.30)
Search Rank (0-1)	0.75	0.64	0.85
	(0.30)	(0.36)	(0.17)
Rating (0-5 stars)	4.14	4.15	4.14
	(0.68)	(0.83)	(0.48)

Notes: The treatment of interest is the PMGs policies' switching off, so that prices observed on the NewEgg platform represent the treated sample, while the control sample has been built by recovering price data for the same products observed on NewEgg but sold on the Amazon UK platform. Product popularity is built like the ratio between the absolute number of product's reviews and the amount of ones received from the most popular product in the same sub-category. Products ratings provided by consumers range from zero (low quality) to five (high quality). The search rank is a proxy of the time spent on search engines to discover the web-page of a certain product. The sample period includes 175 days.

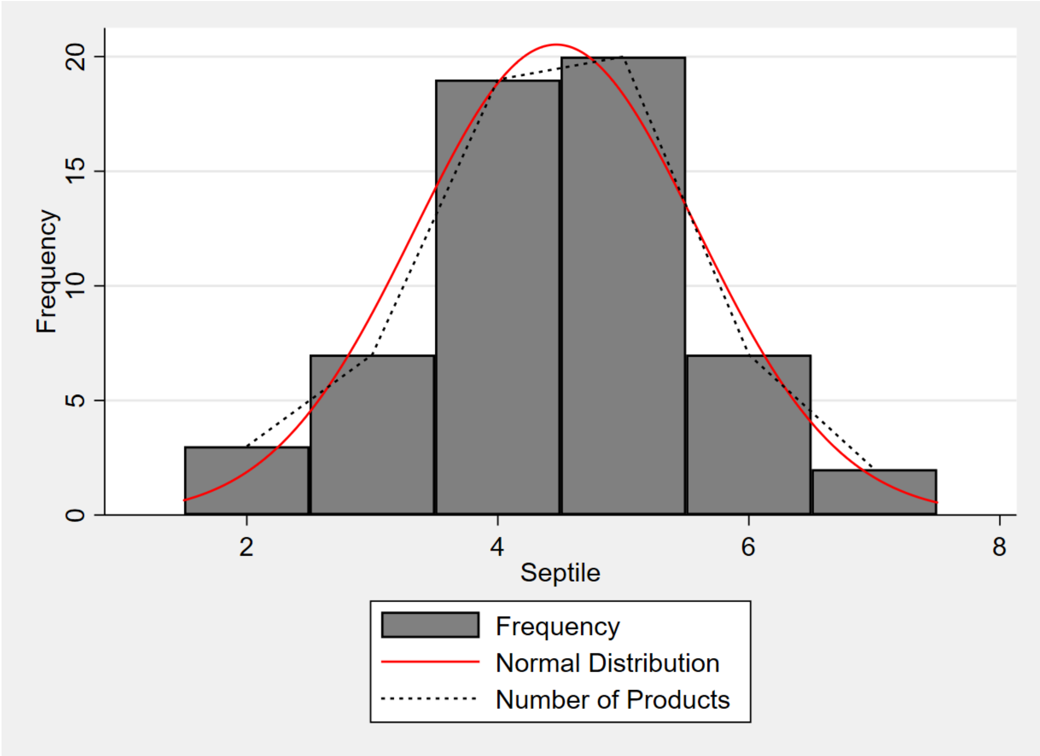
Figure 1: Products Distribution by Price Classes.



Notes: The figure provides the distribution of products by price classes (10). Author's elaboration.

low cost accessories and few luxury goods. Furthermore, calculating the log-price distribution (Figure 2) and mapping the integer part of this value on the x-axis, we obtain a septile-partition. By plotting the distribution of products by log-price classes we obtain a distribution that resembles the Normal one. Such result is in line with those obtained by Coad (2009).

Figure 2: Products Distribution by Log-Price Classes.



Notes: The figure is obtained by calculating the log-price distribution and mapping the integer part of this value on the x-axis. A septile-partition is shown. Author's elaboration.

4 Identification Strategy

We identify the causal effect of PMGs on prices by comparing price levels before and after the switch off of the policies for a treated sample of products sold by NewEgg, to the prices average change for the same products sold by Amazon UK (the control group) in the same period. Indeed, and crucially for our identification strategy, PMGs implemented by NewEgg only affect products that are sold in US, thereby naturally creating a treatment and a control group; identical products sold by Amazon UK (that never offers price warranties) are less likely to be affected by the policy and well represent a counterfactual sample mimicking what would have happened to prices of treated products in the absence of PMGs.[16] This framework provides a quasi-natural experiment that allows us to study the causal impact of PMGs on prices through a DiD research design.

This identification approach requires the estimation of the following panel FE model:[17]

[16]It is worth noting that we do not perform a competition analysis between Newegg and Amazon UK in our exercise. Items observed on Amazon UK only acts as control group; in particular, the latter is comprised of identical products to the treated group in all aspects that could affect the outcome except for the treatment of interest, i.e. PMGs. The validity of the control group is supported in Section 5.2 by an in-depth analysis of parallel trends.

[17]In a DiD context, a classic model would be built like $Y = \alpha + \beta_1(Treated) + \beta_2(Post) + \beta_3(Treated * Post) + \varepsilon$. In all models we exclude $Treated$ and $Post$ terms, since they are multicollinear with time

$$\log Price_{i,l,t} = \alpha_{i,l} + \gamma\big(T_{i,l,t} * P_{i,t}\big) + X^{T}_{i,l,t}\beta + \mu_{i,l} + \tau_t + \epsilon_{i,l,t} \quad (1)$$

The dependent variable, $Price_{i,l,t}$ represents the price (natural logarithm) of good i, on platform l, at time t; $T_{i,l,t}$ denotes a binary variable equal to 1 for treated goods; $P_{i,t}$ is a binary variable that is equal to 1 for any day since policies are switched off and $\epsilon_{i,l,t}$ is an error term.

In order to rise the degree of comparability of the treatment and control groups, the model includes a full set of daily time dummies, τ_t, accounting for unobserved time-varying price determinants that are common to all goods.[18] Products fixed effects, $\mu_{i,l}$, control for any time invariant unobserved heterogeneity at the product and platform level that could be correlated with the included regressors and that could also drive prices. Equation (1) also contains a set of covariates, $X_{i,l,t}$, that accounts for product characteristics (products rating and popularity) derived by UGCs that might affect the outcome of PMGs.

The γ coefficient associated to the interaction term $(T_{i,l,t} * P_{i,t})$ represents the DiD estimate of the effect of PMGs switch off on treated products prices and it measures the average price differential between the treated and the control group.

We also explore the issue of heterogeneity in the effect of PMGs policies on prices. As discussed in the Introduction, we split the sample according to products features that might affect the outcome of PMGs policies and that can be recovered exclusively on online markets thanks to UGCs. In particular, platform data allow us to obtain information on products visibility that can provide evidence on product-specific search costs; moreover, we can recover data on products' popularity and quality and we argue that these characteristics may influence the impact of PMGs on prices. Indeed, we believe that products heterogeneity based on such features is associated to the presence of heterogeneous consumers in terms of information, willingness to search (and to invoke PMGs) and price elasticity, and such heterogeneity can have an impact on the effect of PMGs on prices.

After classifying products according to their visibility, as measured by the search index described above, we estimate Equation (1) on different sub-samples. In particular, we analyse separately products characterised by high (low) visibility in terms of the Google search rank, namely products whose search index is greater (lower) than 0,8. Moreover, given that products quality and visibility resulted to be highly correlated, we split the sample according to both characteristics jointly considered.

The heterogeneity issue is also investigated with a different approach by estimating a Triple Difference regression (DDD) on the full sample. In particular, we estimate the following models:

---

and product fixed effects.

[18] As noted in Section 3, the sample does not include important dates, like Thanksgiving or Christmas; moreover daily fixed effects allows to control for possible time effects associated to particular periods like "back to school" days.

$$\log Price_{i,l,t} = \alpha_{i,l} + \varphi\left(T_{i,l,t} * P_{i,t} * HV_{i,l}\right) + X^{T}_{i,l,t}\beta + \mu_{i,l} + \tau_t + \omega_{i,l,t} \quad (2)$$

$$\log Price_{i,l,t} = \alpha_{i,l} + \delta\left(T_{i,l,t} * P_{i,t} * HRHV_{i,l}\right) + X^{T}_{i,l,t}\beta + \mu_{i,l} + \tau_t + \omega_{i,l,t} \quad (3)$$

Equations (2) and (3) include additional components in the interaction term, namely $HV_{i,l}$ and $HRHV_{i,l}$, i.e. dummy variables assuming value 1 for high visible products and for high visible and high rating products respectively. The coefficients φ and δ of the triple interaction terms measure the average treatment effect of PMGs on prices for high visible and high visible-high rating products respectively.

All specifications are estimated by OLS with robust standard errors. Moreover, following Cameron and Miller (2015), we also compute bootstrapped standard errors allowing for a cluster structure (at product level). In addition, we perform an extensive robustness analysis and placebo tests.

5 Empirical Results

5.1 Main Results

In Table 2 are reported estimates of different models that do not include control variables, while Table 3 shows empirical results for more extended models.

Table 2: DiD Estimates of the Impact of PMGs on Prices (No Controls).

Products Prices (log)	(1) FULL SAMPLE	(2) L. VISIBLE	(3) H. VISIBLE	(4) LR-LV	(5) HR-HV
$T_{i,l,t} * P_{i,t}$	-0.0401***	0.0242***	-0.0543***	0.0331**	-0.0381***
	(0.00628)	(0.00769)	(0.00795)	(0.01370)	(0.00864)
Controls	NO	NO	NO	NO	NO
Product Dummies	YES	YES	YES	YES	YES
Time Dummies	YES	YES	YES	YES	YES
Observations	9,028	2,295	6,733	994	4,864
R-squared	0.986	0.990	0.984	0.983	0.983
F Test (p-value)	0.000	0.002	0.000	0.016	0.000

Notes: All specifications include time and product fixed effects. Controls include the absolute and the relative number of reviews as well as ratings. High (low) quality products have ratings higher (lower) than 4. High (low) visible products have a normalized search index higher (lower) than 0,8. LR-LV are low rating and low search index products, HR-HV are high rating and high search rank products. Robust Standard Errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Estimates of our baseline specification performed over the full sample (Column 1 of both Tables) provide evidence in favour of higher prices when PMGs policies are in place. In particular, estimated coefficients suggest that when PMGs policies are switched off, prices decrease of about 4% for treated products.[19] Rather comfortingly,

[19] It is worth noting that our identification strategy is based on a comparison of price levels before and after the policy implementation.

Table 3: DiD Estimates of the Impact of PMGs on Prices (With Controls).

Products Prices (log)	(1) FULL SAMPLE	(2) L. VISIBLE	(3) H. VISIBLE	(4) LR-LV	(5) HR-HV
$T_{i,l,t} * P_{i,t}$	-0.0424***	0.0322***	-0.0577***	0.0532***	-0.0398***
	(0.00629)	(0.00786)	(0.00799)	(0.01310)	(0.00879)
Controls	YES	YES	YES	YES	YES
Product Dummies	YES	YES	YES	YES	YES
Time Dummies	YES	YES	YES	YES	YES
Observations	9,028	2,295	6,733	994	4,864
R-squared	0.986	0.992	0.985	0.989	0.983
F Test (p-value)	0.000	0.000	0.000	0.000	0.000

Notes: All specifications include time and product fixed effects. Controls include the absolute and the relative number of reviews as well as ratings. High (low) quality products have ratings higher (lower) than 4. High (low) visible products have a normalized search index higher (lower) than 0,8. LR-LV are low rating and low search index products, HR-HV are high rating and high search rank products. Robust Standard Errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

the inclusion of control variables into Model (1) does not significantly affect results.

These findings are consistent with those obtained by Zhuo (2017) on a large sample of products observed on Amazon in 2012. However, the author observes price changes (on the non-adopting platform) before and after the implementation of PMGs by competitors, while we focus on price changes observed on the adopting platform. Moreover, we innovatively build the control sample with platform price data for the same treated products but observed on another platform (Amazon UK).

In order to explore whether the impact of PMGs on prices is heterogeneous across products, we estimate Equation (1) on different sub-samples built according to products quality and visibility. We first distinguish products according to the Google search index as a proxy for products visibility.

Columns (2) and (3) in Tables 2 and 3 show results of this analysis and suggest that, when the PMGs are off, products characterised by a low search rank experience a price increase of roughly 2,5%, while for high visible products prices decrease of about 5%.[20] The latter result drives estimates obtained over the full sample, given that about 70% of observations are associated to high visible products.

Since quality and visibility are correlated in our sample, we estimate Equation (1) after splitting the sample according to both product characteristics jointly considered. Results shown in column (4) and (5) of Table 2 suggest that, when PMGs are switched off, prices for high quality and high visible products decrease of about 3,9%, while prices of low quality and low visible ones raise of about 3,3%.[21] These findings are confirmed when we include control variables (Table 3) and when we perform heterogeneity analysis with a DDD regression approach (Equations 2 and 3), as

[20] Similar results also arise from estimates that include controls.

[21] Again, the latter category represents just 10% of total observations.

Table 4: DDD Estimates of the Impact of PMGs on Prices.

Product Prices (log)	(1) DDD	(2) DDD	(3) DDD	(4) DDD
$T_{i,l,t} * P_{i,t} * HV_{i,l,t}$	-0.0522***	-0.0540***		
	(0.00779)	(0.00781)		
$T_{i,l,t} * P_{i,t} * HRHV_{i,l,t}$			-0.0537***	-0.0556***
			(0.00808)	(0.00810)
Controls	NO	YES	NO	YES
Product Dummies	YES	YES	YES	YES
Time Dummies	YES	YES	YES	YES
Observations	9,028	9,028	9,028	9,028
R-squared	0.986	0.986	0.986	0.986
F Test (p-value)	0.000	0.000	0.000	0.000

Notes: All specifications include time and product fixed effects. Controls include the absolute and the relative number of reviews as well as ratings. High (low) quality products have ratings higher (lower) than 4. High (low) visible products have a normalized search index higher (lower) than 0,8. $HV_{i,l,t}$ is a dummy equal to 1 for high visible products. $HRHV_{i,l,t}$ is a dummy equal to 1 for high quality and high visible products. Robust Standard Errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

reported in Table 4.

Finally, we investigate the possibility that the effects of the treatment may speed up, stabilize, or mean revert over time. In order to explore this issue, we estimate (over the full sample) a specification of Equation (1) that includes lags à la Autor (2003). The model takes on the following form:

$$\log Price_{i,l,t} = \alpha_{i,l} + \sum_{j=0}^{5+} \gamma_j \big(T_{i,l,t} * P_{i,t+j}\big) + X_{i,l,t}^T \beta + \mu_{i,l} + \tau_t + \epsilon_{i,l,t} \tag{4}$$

Specification (4), where $P_{i,t+j}$ assumes the value of 1 in day $t+j$, and 0 otherwise, allows the treatment to generate different effects over time. In order to lower the number of parameters of the model, we estimate such effect from the implementation day ($j = 0$) until five days later and onward.

According to results shown in Table 5, coefficients related to lagged variables are always negative and statistically significant. However, point estimates suggest that the impact of the treatment reaches its maximum after one day and starts decreasing afterwards. Figure 3 graphically shows parameter estimates patterns.

Overall, our empirical findings obtained on the full sample are in line with those reported by the previous applied literature for different markets (e.g. Cabral et al., 2018; Chilet, 2018; Byrne and De Roos, 2019) and online platforms (Zhuo, 2017). However, results obtained on different sub-samples provide novel evidence that cannot be compared with the previous empirical literature. We believe that the interpretation of such findings can be discussed by considering consumers heterogeneity that is reflected into UGCs. We argue that consumers are heterogeneous in terms of informa-

Table 5: DiD Estimates of the Impact of PMGs on Prices with lags à la Autor (2003).

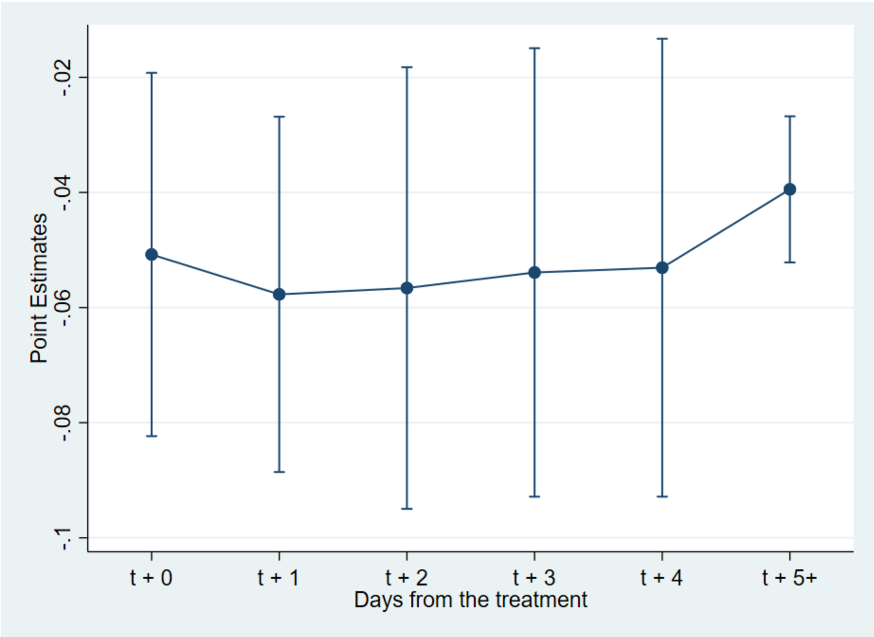
Products Prices (log)	(1) DiD	(2) DiD
$T_{i,l,t} * P_{i,t+0}$	-0.0500***	-0.0508**
	(0.01610)	(0.01610)
$T_{i,l,t} * P_{i,t+1}$	-0.0566***	-0.0577**
	(0.01580)	(0.01570)
$T_{i,l,t} * P_{i,t+2}$	-0.0558***	-0.0566**
	(0.01960)	(0.01960)
$T_{i,l,t} * P_{i,t+3}$	-0.0530***	-0.0539**
	(0.01990)	(0.01990)
$T_{i,l,t} * P_{i,t+4}$	-0.0529***	-0.0531**
	(0.02030)	(0.02030)
$T_{i,l,t} * P_{i,t+5+}$	-0.0368***	-0.0395**
	(0.00646)	(0.00648)
Controls	NO	YES
Product Dummies	YES	YES
Time Dummies	YES	YES
Observations	9,028	9,028
R-squared	0.986	0.986
F Test (p-value)	0.000	0.000

Notes: All specifications include time and product fixed effects. Controls include the absolute and the relative number of reviews as well as ratings. $P_{i,t+j}$ assumes the value of 1 in day $t+j$, and 0 otherwise. Robust Standard Errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

tion, price elasticity and availability to search. In particular, consumers characterized by high search (hassle) costs are most likely to focus on products that are high ranked by the Google search index (high visible products) and, according to our results, will pay higher prices during the policy implementation period. Moreover, in our sample, such products are likely to have high quality rating. On the other side, customers that put more search effort focus on products that do not have high Google search rank and are more likely to invoke the price matching clause; such customers are like to have a lower willingness to pay and a higher price sensitiveness so that they manage to pay lower prices.

Results obtained for sub-samples based on products visibility are broadly in line with the hypothesis of price matching motivated by price discrimination, so that for high (low) visible products prices are higher (lower) during the policy implementation days. Indeed, according to such hypothesis, "firms offering price-matching guarantees provide discounts selectively to customers who are aware of lower prices in the market while charging a high list price for non searchers" (Moorthy and Winter, 2006). Moreover, the price discrimination hypothesis requires that a significant percentage of customers invoke PMGs rights. Moorthy and Winter (2006) observe redemption

Figure 3: DiD Estimates of the Impact of PMGs on Prices with lags à la Autor (2003).



Notes: Point estimates and confidence intervals of lagged variables. The specification includes time and product fixed effects. Controls include product popularity, number of reviews and rating. See column (2) of Table 5 (Full Sample, with Controls).

rates ranging between 1% and 25%, on a sample of 46 Canadian retailers and suggest that percentages above 10% are compatible with the aforementioned hypothesis. We believe that it is reasonable to expect high redemption rates on online markets, given that "hassle" cost should be lower for e-commerce.

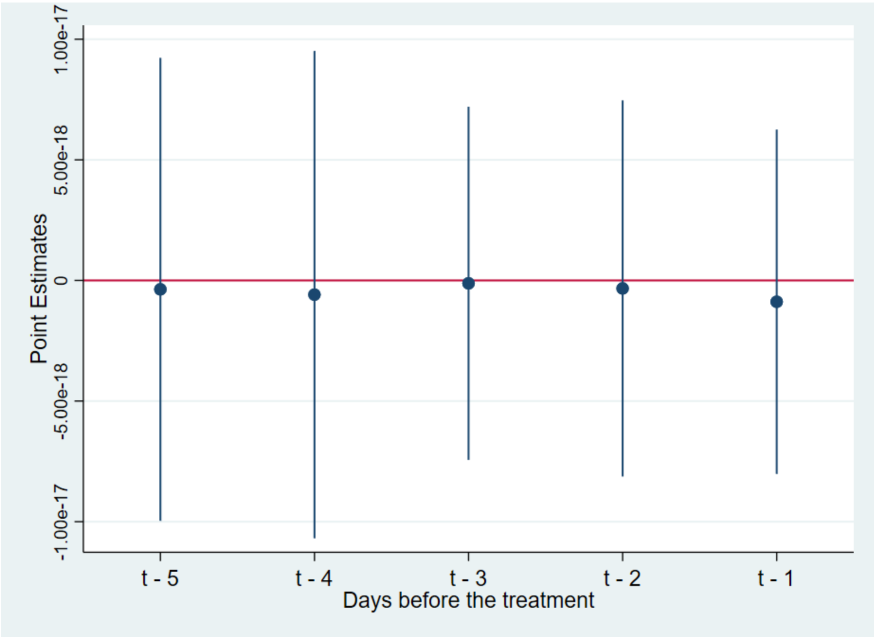
Despite some of the assumptions of theoretical models that interpret PMGs as discriminating tools seem to be satisfied in the analyzed context, platform markets differ from brick and mortar ones along different dimensions. Among others, the process of quality evaluation of search goods is cheaper and faster thanks to the availability of UGCs, so that commercial policies, like PMGs, are designed according to such features. Indeed, the characteristics of PMGs observed in our sample significantly differ from those described in theoretical models that are more likely to fit brick and mortar markets and it would be interesting to modify such models in order to account for the peculiarities of online markets.

5.2 Robustness Analysis

In this Section, we discuss empirical results obtained by conducting an in-depth robustness analysis of our results.

The first concern we tackle is an important issue in a DiD research design, i.e. the presence of pre-treatment common trends for treated and control units. This assumption is indeed fundamental for the validity of the counterfactual policy evaluation analysis. In order to explore this issue, in Figure 4 we show point estimates and

Figure 4: Price Differentials Between Treated and Control Groups Before PMGs Switch Off.



Notes: The figure shows point estimates values and relative confidence intervals of the difference in the level of prices between treated and control products from five days before the treatment to the day when the policy is switched off. In order to obtain these values, we estimate a panel model where we regress average daily price differences between the two samples on lead terms for five days before the treatment. We control for product fixed effects and daily fixed effects.

confidence intervals of price differentials between treated and control products, from five days before the treatment to the switching day.[22] Plotted values suggest that, before the treatment, price levels for the treated sample do not seem to be significantly different from control sample prices. This result provides evidence in favor of the validity of parallel trends assumption for our samples.

In order to further analyse this issue, we follow Autor (2003) and we estimate Equation (4) after including some leads of the treatment interaction variable:

$$\log Price_{i,l,t} = \alpha_{i,l} + \sum_{j=-1}^{-5} \gamma_j \left(T_{i,l,t} * P_{i,t+j}\right) + X_{i,l,t}^{T}\beta + \mu_{i,l} + \tau_t + \epsilon_{i,l,t} \tag{5}$$

Indeed, if leads coefficients turn out to be statistically significant, there may be anticipatory effects of the policy and a failure in the parallel trend assumption. According to Table 6 and Figure 5, estimated coefficients of the anticipatory effects are not statistically significant, thus providing further evidence in favor of the existence of a parallel trend between treatment and control sample.

In order to further validate our research design we test for potential treatment endogeneity by leveraging on a regression where the probability of being treated is a

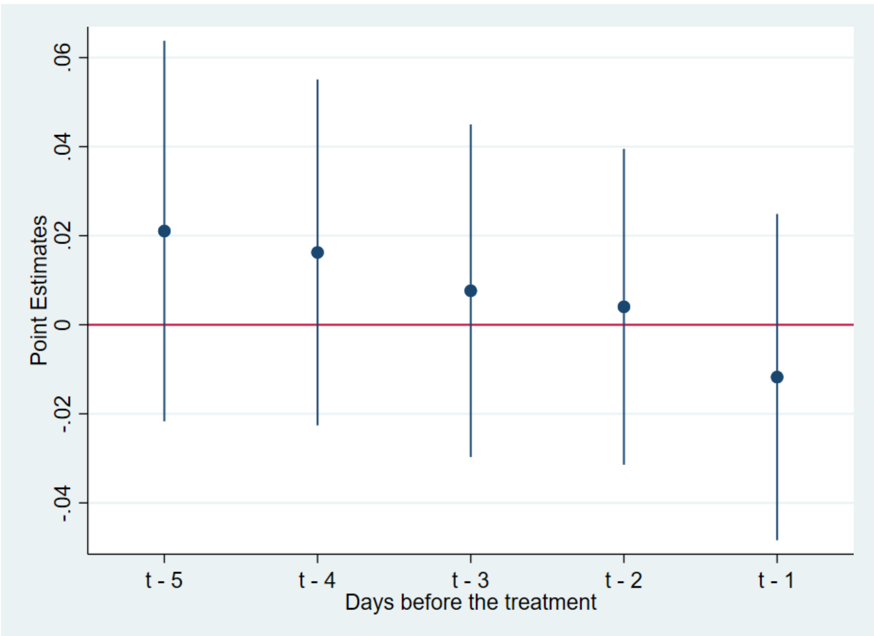
[22] In particular, in order to obtain these values we estimate a panel model where we regress average daily price differences between treated and control samples on lead terms for five days before the treatment. We control for product fixed effects and daily fixed effects. Results are available upon request.

Table 6: DiD Estimates of the Impact of PMGs on Prices with leads à la Autor (2003).

Dep. Var.: Prices (log)	DiD (1)	DiD (2)
$T_{i,l,t} * P_{i,t-1}$	-0.0121	-0.0117
	(0.0187)	(0.0187)
$T_{i,l,t} * P_{i,t-2}$	0.0030	0.0040
	(0.0181)	(0.0181)
$T_{i,l,t} * P_{i,t-3}$	0.0067	0.0076
	(0.0191)	(0.0191)
$T_{i,l,t} * P_{i,t-4}$	0.0155	0.0162
	(0.0199)	(0.0198)
$T_{i,l,t} * P_{i,t-5}$	0.0205	0.0211
	(0.0219)	(0.0218)
Controls	NO	YES
Product Dummies	YES	YES
Time Dummies	YES	YES
Observations	9,028	9,028
R-squared	0.986	0.986
F Test (p-value)	0.842	0.000

Notes: All specifications include time and product fixed effects. Controls include the absolute and the relative number of reviews as well as ratings. $P_{i,t+j}$ assumes the value of 1 in day $t+j$, and 0 otherwise. Robust Standard Errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

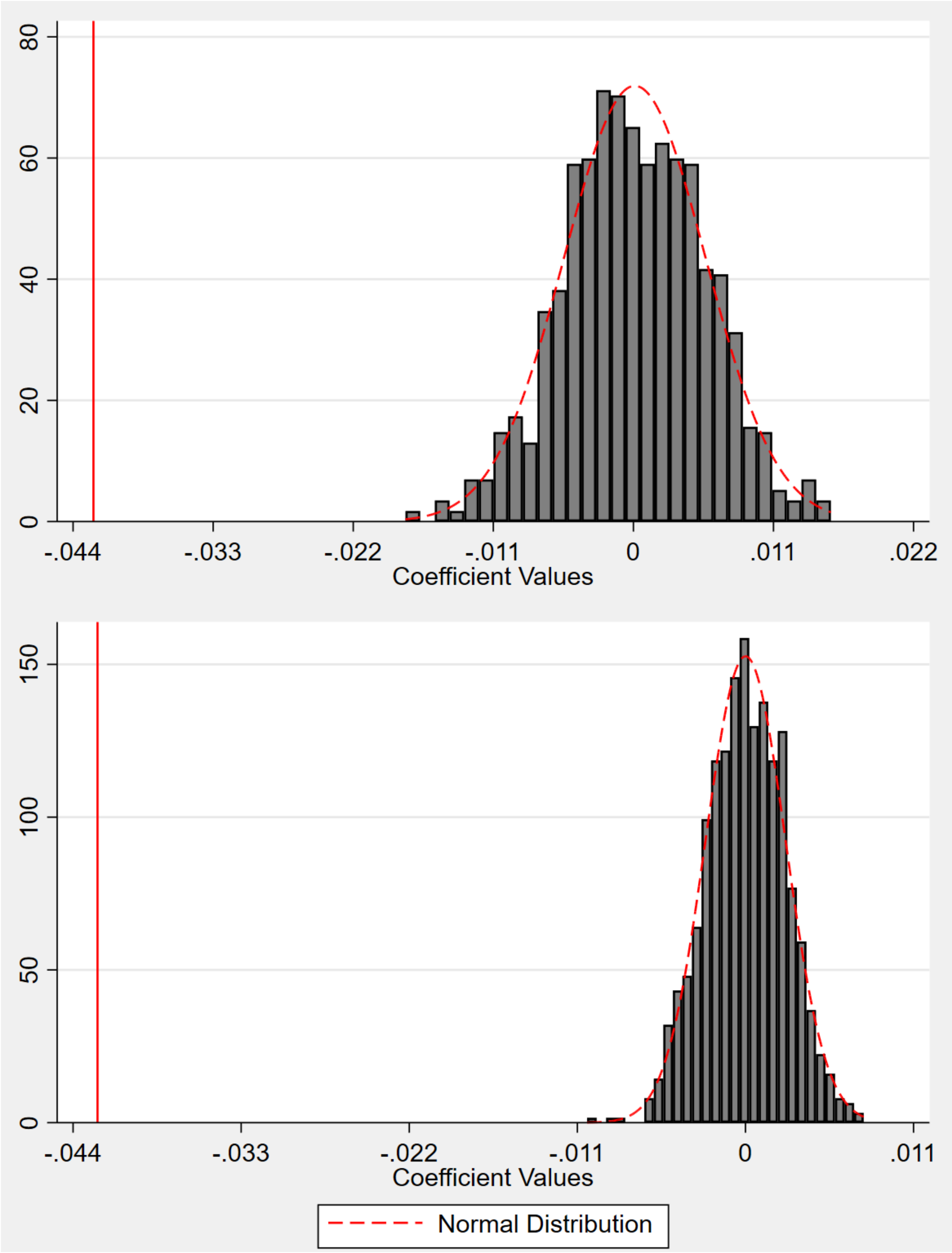
Figure 5: DiD Estimates of the Impact of PMGs on Prices with leads à la Autor (2003).



Notes: The figure shows point estimates values and confidence intervals from the estimation of Equation (5). See column (2) of Table 6 (Full Sample, with Controls).

function of products prices and visibility. Rather comfortingly, results suggest that

Figure 6: Placebo Plot Test.



Notes: In the top panel we randomly assign treatments timing by generating simulated values for the dummy $Post_{i,t}$. In particular, we build a *fake* $Post_{i,t}$ equal to 1 if random numbers drawn from a uniform distribution [0,1] are greater than the sample treatment probability. We then estimate Model (1) after including the new interaction term, $Treated_{i,l,t} * Post_{i,t}^{fake}$ and we iterate such procedure 1000 times in order to obtain a distribution of placebo γ coefficients to compare with the estimated value shown in Table 3. The bottom panel reports coefficient estimates distribution obtained with the same iterative method, where both treatments timing and treated products have been randomly assigned. Dark bars represent the distribution of estimated placebo interaction term coefficients. Vertical red lines represent the estimated γ coefficient of $Treated_{i,l,t} * Post_{i,t}$ ($\gamma = -0.0424$) shown in Table 3.

treatment probability is independent from such variables.[23]

[23]Results are available upon request.

Treatment exogeneity is also investigated by means of a complete set of placebo tests.

First, we randomly assign treatments timing by generating simulated values for the dummy $Post_{i,t}$. In particular, we build a placebo variable $Post_{i,t}^{fake}$ equal to 1 if random numbers drawn from a uniform distribution [0,1] are greater than the sample treatment probability. We therefore estimate Model (1) after including the new interaction term, $Treated_{i,l,t} * Post_{i,t}^{fake}$ and we iterate such procedure 1000 times in order to obtain a distribution of placebo γ coefficients to compare with the estimated value shown in Table 3. Indeed, a statistically significant treatment effect should be significantly different with respect to placebo estimates. In the top panel of Figure 6 dark bars represent the distribution of estimated placebo coefficients obtained with such iterative method, while the vertical red line shows the γ coefficient value estimated in the baseline specification. It is worth noting that the latter lies outside the density distribution of placebo effects, thus being statistically significantly larger, in absolute value, than those obtained using randomly assigned treatments. Moreover, placebo γ coefficients are normally distributed, with zero mean, thus suggesting the robustness of our main findings.

Second, we follow the same procedure to build up another placebo test; in particular, we build *fake* treatments timing as well as *fake* treated products in order to obtain a different placebo interaction term, i.e. $Treated_{i,l,t}^{fake} * Post_{i,t}^{fake}$ (again, placebo dummies are built according to draws from uniform distributions). The bottom panel of Figure 6 shows density estimates for such analysis. Once again, the coefficient inferred from our main specification is significantly different from values obtained from the placebo study, thus confirming the robustness of our previous results.[24]

It is worth noting that our results are confirmed when we estimate our baseline and DDD specifications after introducing *fake* products and *fake* treatments randomly drawn from Bernoulli distributions and when we estimate our models after substituting the dependent variable with a *fake* outcome, where product prices are drawn from product specific random distributions resembling sample ones (same mean and variance).

In order to analyze if our main findings are robust to the exclusion of a particular product, we estimate the baseline Equation (1) after dropping one product at a time and all previous results are confirmed; same conclusions arise when we estimate Equation (1) after balancing the panel and when we compute bootstrapped standard errors allowing for a cluster structure (at product level).[25]

Finally, we replicate our main analysis relying on an alternative control sample. In particular, we built such control sample following the same approach explained

[24]Moreover, the distribution of coefficient values from placebo studies in Figure 6 resemble a Normal one with zero mean, thus highlighting no treatment effects under the hypothesis of products and/or treatments *fake* assignments.

[25]Results are reported in Appendix A.

in Section 3 but relying on Amazon US prices. Although prices observed on Amazon US might not be completely independent from the policy under scrutiny (because of possible price-tracking practices within the same country), all results are confirmed.

6 Conclusions

In this work we empirically investigate the effects of Price Matching Guarantees (PMGs) commercial policies on daily prices of a representative sample of consumer electronics products observed on the US NewEgg platform, between May and October 2018. We apply a Difference-in-Differences (DiD) research design where the control sample is built by recovering price data for the same products affected by PMGs but sold on another platform, namely Amazon, that never offers such policies over the sample period. In particular, we choose to build the control sample with prices from Amazon UK, instead of Amazon US, in order to avoid possible confounding effects arising from platforms price tracking practices, that are more likely to take place among platforms belonging to the same country. However, overall findings are confirmed when we use the Amazon US control sample and when we perform an extensive robustness analysis.

Estimates provide evidence in favor of an average price reduction of about 4% after the PMGs validity period and such result is confirmed after controlling for products features, retrievable from User Generated Contents (UGCs), that might affect PMGs outcomes, like products quality and popularity. Moreover, in order to have a more detailed picture of the issue, we conduct an heterogeneity analysis by distinguishing products according to their platform visibility (and quality), as proxied by the Google search rank (and rating). Estimates conducted on specific sub-samples show that, when PMGs are switched off, low search rank (low rating) products experience a price increase of about 2.5%, while for high visible (high rating) products a price reduction of about 5% is observed.

These findings can be considered broadly consistent with the hypothesis of PMGs acting as a tool for price discrimination. In particular, the presence of consumers that differ in terms of information, price elasticity, hassle costs and willingness to search is reflected by UGCs shared on platforms. We argue that low (high) visible products are associated to consumers with a high (low) willingness to search, that in turn can be associated to different levels of price elasticity, and these assumptions are consistent with different prices observed for certain classes of products during the policy implementation period.

However, previous models that consider PMGs as price discriminating tools (Corts, 1996; Nalca et al., 2010, among others) have been proposed for brick and mortar retailers and are based on assumptions that do not fully reflect the context of on-line platform markets. Indeed, the development of a theoretical model, specifically designed to deal with online platform commercial policies, would fill a gap in the literature that has mostly focused on traditional markets.

References

Arbatskaya, M., Hviid, M. and Shaffer, G. (2000), 'Promises to match or beat the competition: Evidence from retail tire prices', advances in applied microeconomics (advances in applied microeconomics, volume 8)'.

Arbatskaya, M., Hviid, M. and Shaffer, G. (2006), 'On the use of low-price guarantees to discourage price cutting', *International Journal of Industrial Organization* **24**(6), 1139–1156.

Autor, D. H. (2003), 'Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing', *Journal of labor economics* **21**(1), 1–42.

Baye, M. R. and Kovenock, D. (1994), 'How to sell a pickup truck:'beat-or-pay'advertisements as facilitating devices', *International Journal of Industrial Organization* **12**(1), 21–33.

Belton, T. M. (1987), 'A model of duopoly and meeting or beating competition', *International Journal of Industrial Organization* **5**(4), 399–417.

Byrne, D. P. and De Roos, N. (2019), 'Learning to coordinate: A study in retail gasoline', *American Economic Review* **109**(2), 591–619.

Cabral, L., Dürr, N., Schober, D. and Woll, O. (2018), Price matching guarantees and collusion: Theory and evidence from germany, Technical report, Working Paper, New York University.

Cameron, A. C. and Miller, D. L. (2015), 'A practitioner's guide to cluster-robust inference', *Journal of human resources* **50**(2), 317–372.

Chilet, J. A. (2018), 'Gradually rebuilding a relationship: The emergence of collusion in retail pharmacies in chile'.

Chung, H. S., Kim, M. et al. (2016), 'Low-price guarantees and pricing behavior: evidence from hypermarkets in korea', *Economics Bulletin* **36**(2), 1223–1229.

Coad, A. (2009), 'On the distribution of product price and quality', *Journal of Evolutionary Economics* **19**(4), 589–604.

Constantinou, E. and Bernhardt, D. (2018), 'The price-matching dilemma', *International Journal of Industrial Organization* **59**, 97–113.

Corts, K. S. (1996), 'On the competitive effects of price-matching policies', *International Journal of Industrial Organization* **15**(3), 283–299.

Edlin, A. S. (1997), 'Do guaranteed-low-price policies guarantee high prices, and can antitrust rise to the challenge?', *Harvard Law Review* pp. 528–575.

Hay, G. A. (1981), 'Oligopoly shared monopoly and antitrust law', *Cornell L. Rev.* **67**, 439.

Hess, J. D. and Gerstner, E. (1991), 'Price-matching policies: An empirical case', *Managerial and Decision Economics* **12**(4), 305–315.

Hviid, M. and Shaffer, G. (1999), 'Hassle costs: the achilles' heel of price-matching guarantees', *Journal of Economics & Management Strategy* **8**(4), 489–521.

Hviid, M. and Shaffer, G. (2010), 'Matching own prices, rivals'prices or both?', *The Journal of Industrial Economics* **58**(3), 479–506.

Jain, S. and Srivastava, J. (2000), 'An experimental and theoretical analysis of price-matching refund policies', *Journal of Marketing Research* **37**(3), 351–362.

Logan, J. W. and Lutter, R. W. (1989), 'Guaranteed lowest prices: do they facilitate collusion?', *Economics Letters* **31**(2), 189–192.

Lu, Y. and Wright, J. (2010), 'Tacit collusion with price-matching punishments', *International Journal of Industrial Organization* **28**(3), 298–306.

Moorthy, S. and Winter, R. A. (2006), 'Price-matching guarantees', *The RAND Journal of Economics* **37**(2), 449–465.

Moorthy, S. and Zhang, X. (2006), 'Price matching by vertically differentiated retailers: Theory and evidence', *Journal of Marketing Research* **43**(2), 156–167.

Nalca, A., Boyaci, T. and Ray, S. (2010), 'Competitive price-matching guarantees under imperfect store availability', *Quantitative Marketing and Economics* **8**(3), 275–300.

OECD (2019), *An Introduction to Online Platforms and Their Role in the Digital Transformation*, OECD Publishing.
URL: *https://www.oecd-ilibrary.org/content/publication/53e5f593-en*

Png, I. P. and Hirshleifer, D. (1987), 'Price discrimination through offers to match price', *Journal of Business* pp. 365–383.

Pollak, A. et al. (2017), Do price-matching guarantees with markups facilitate tacit collusion? theory and experiment, Technical report, Working Paper Series in Economics, University of Cologne.

Salop, S. C. (1986), Practices that (credibly) facilitate oligopoly co-ordination, *in* 'New developments in the analysis of market structure', Springer, pp. 265–294.

Wilhelm, S. (2016), Price-matching strategies in the german gasoline retail market, Technical report, Working Paper, Goethe Universität Frankfurt, erhältlich unter: http://ssrn . . . .

Zhuo, R. (2017), 'Do low-price guarantees guarantee low prices? evidence from competition between amazon and big-box stores', *The Journal of Industrial Economics* **65**(4), 719–738.

Appendix A

Additional Robustness Analysis

In order to further verify our results, we perform a full set of additional robustness tests.

We first estimate our baseline and DDD specifications by introducing artificially treatments timing and artificially treated products obtained with an alternative distribution. *Fake* assignments are drawn from two Bernoulli distributions with parameters p (probability of success) derived from the sample distributions of $Treated_{i,l,t}$ and $Post_{i,t}$ respectively. Within this setting, we should not observe any significant effect of PMGs switch off on prices. Comfortingly, results reported in columns (1) to (6) of Table A.1 confirm this prediction.

Next, we conduct another falsification test by estimating our models after substituting the dependent variable with a placebo outcome that should not be affected by PMGs switch off. In particular, we generate *fake* product prices drawn by random distributions resembling sample ones (same mean and variance). Results shown in columns from (7) to (12) of Table A.1 confirm the absence of any impact of PMGs on the fake outcome.

Third, in Figure A.1 we replicate the placebo test described for the bottom panel of Figure 6, in the Section 5.2 of the main text, but relying on Bernoulli distributions instead of uniform ones. Rather comfortingly, results in Figure A.1 confirm the robustness of our main findings.

Moreover, we re-estimate the baseline Model (1) after balancing the panel dataset. Precisely, we drop first 34 days in which we observe only some products and all results are confirmed. Lastly, it is worth noting that results do not change if we compute bootstrapped standard errors allowing for a cluster structure at product level. Table A.2 shows respective results.

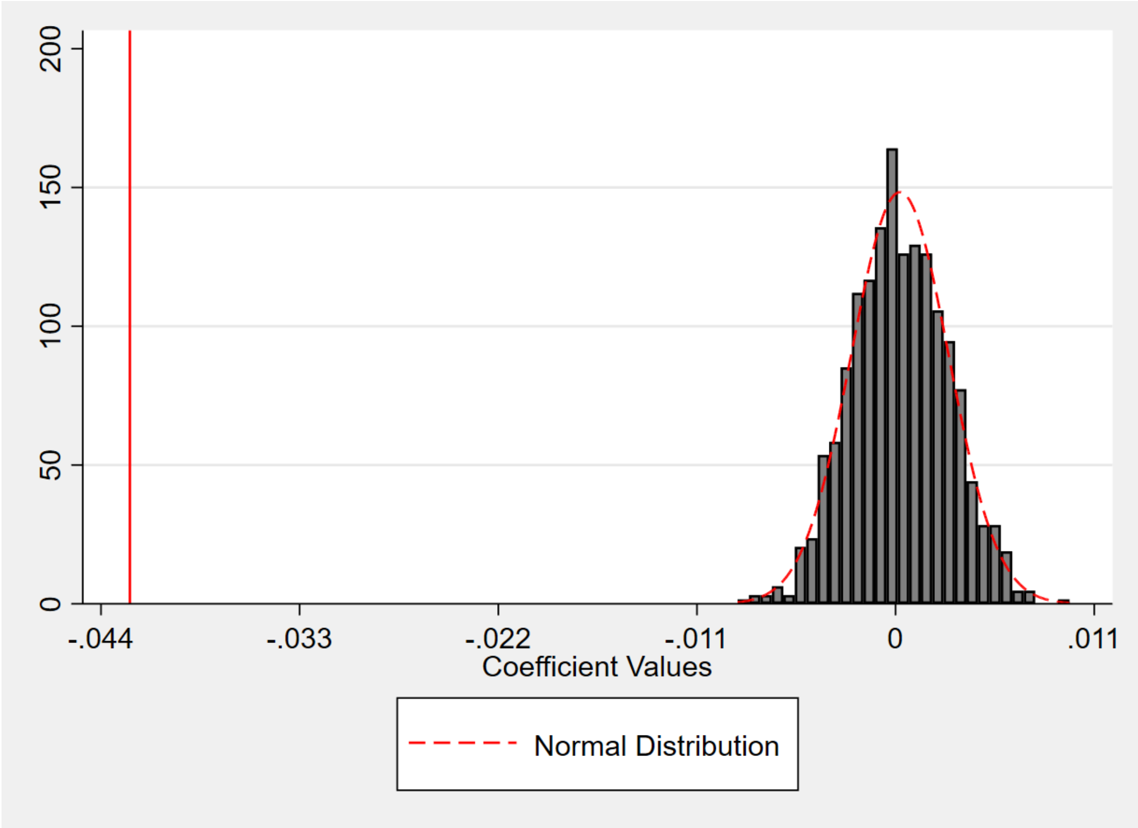
Finally, although we are aware that prices observed on Amazon US might not be completely independent from the policy under scrutiny, because of price-tracking practices frequently adopted by platforms within the same country, we replicate our main analysis also relying on such control sample. Comfortingly, all results are confirmed (results are available on request).

Table A.1: DiD and DDD Estimates of the Impact on Prices of *Fake* Treatments and *Fake* Treated Products. DiD and DDD Estimates of the Impact of PMGs on *Fake* Prices.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variables	Prices (log)	Prices (log)	Prices (log)	Prices (log)	Prices (log)	Prices (log)	*Fake* Prices (log)	*Fake* Prices (log)	*Fake* Prices (log)	*Fake* Prices (log)	*Fake* Prices (log)	*Fake* Prices (log)
$T_{i,l,t} * P_{i,t}(Fake)$	0.0018	0.0019										
	(0.00262)	(0.00261)										
$T_{i,l,t} * P_{i,t} * HV_{i,l,t}(Fake)$			0.0013	0.0014								
			(0.00320)	(0.00319)								
$T_{i,l,t} * P_{i,t} * HRHV_{i,l,t}(Fake)$					0.0033	0.0033						
					(0.00403)	(0.00403)						
$T_{i,l,t} * P_{i,t}$							-0.0011	-0.0011				
							(0.00154)	(0.00154)				
$T_{i,l,t} * P_{i,t} * HV_{i,l,t}$									-0.0012	-0.0011		
									(0.00196)	(0.00197)		
$T_{i,l,t} * P_{i,t} * HRHV_{i,l,t}$											-0.0013	-0.0013
											(0.00203)	(0.00204)
Controls	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES	NO	YES
Product Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time Dummies	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	9,028	9,028	9,028	9,028	9,028	9,028	9,028	9,028	9,028	9,028	9,028	9,028
R-squared	0.986	0.986	0.9859	0.9859	0.986	0.986	0.999	0.999	0.9989	0.9989	0.999	0.999
F Test (p-value)	0.502	0.000	0.687	0.000	0.418	0.000	0.466	0.722	0.5520	0.7564	0.516	0.746

Notes: All specifications include time and product fixed effects. Controls include the absolute and the relative number of reviews as well as ratings. High (low) quality products have ratings higher (lower) than 4. High (low) visible products have a normalized search index higher (lower) than 0,8. $HV_{i,l,t}$ is a dummy equal to 1 for high visible products. $HRHV_{i,l,t}$ is a dummy equal to 1 for high quality and high visible products. Columns from (1) to (6) show results of a placebo test with artificially timed treatments and artificially treated subjects. Columns from (7) to (12) provide estimates of another placebo test with fake product prices drawn by random distributions resembling sample ones (same mean and variance). Robust Standard Errors in in parentheses *** p<0.01, ** p<0.05, * p<0.1

Figure A.1: Placebo Plot Test.



Notes: The figure shows the result from a iterative placebo test with artificially timed treatments and artificially treated subjects. Model (1) is estimated relying on 1000 simulated datasets in which *fake* assignments are randomly reshuffled in each iteration by drawing from Bernoulli distributions with parameters p (probability of success) derived from the sample distributions of $Treated_{i,l,t}$ and $Post_{i,t}$ respectively. The vertical red line represents the effective coefficient of $Treated_{i,l,t} * Post_{i,t}$ ($\gamma = -0.0424$) in Table 3. Dark bars show the distribution of γ coefficient values from placebo tests.

Table A.2: DiD Estimates of the Impact of PMGs on Prices. Additional Robustness.

	(1)	(2)	(3)	(4)
Product Prices (log)	DiD	DiD	DiD	DiD
$T_{i,l,t} * P_{i,t}$	-0.0533***	-0.0555***	-0.0401*	-0.0424*
	(0.00611)	(0.00614)	(0.0226)	(0.0228)
Controls	NO	YES	NO	YES
Product Dummies	YES	YES	YES	YES
Time Dummies	YES	YES	YES	YES
Observations	8236	8236	9,028	9,028
R-squared	0.988	0.988	0.986	0.986
Standard Errors	ROBUST	ROBUST	BOOTSTRAP	BOOTSTRAP

Notes: Columns (1) and (2) show DiD estimates of Equation (1) on a balanced panel dataset (we drop first 34 days in which we observe only some products). Columns (3) and (4) provide results computing bootstrapped standard errors allowing for a cluster structure at product level on the full sample. All specifications include time and product fixed effects. Controls include the absolute and the relative number of reviews as well as ratings. Robust (Bootstrapped) Standard Errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Appendix B

Table B.1: Sub-Categories List.

Sub - Categories	# products
CPU Processor	6
Computer Case	4
Mobile Phone	2
Scanner	4
Speaker	4
Motherboard	2
Monitor	6
Headset	2
USB Flash	2
CPU Cooler	2
Speaker for Domotic	2
Tablet	2
Desktop PC	2
Laptop PC	2
Power Supply	2
Printer	4
Memory Card	4
Hard Disk	2
Smart Thing Domotic	4

Table B.2: Treated Products List.

Treated Products Titles
AMD Ryzen 5 1500X Processor
Corsair Crystal Series 570X RGB - Tempered Glass; Premium ATX Mid-Tower Case
BlackBerry PRIV (32GB) Verizon Factory Unlocked Phone
Fujitsu fi-7160 Color Duplex Document Scanner
Fujitsu ScanSnap S1300i Instant PDF Multi Sheet-Fed Scanner
Philips BT50B/37 Wireless Portable Bluetooth Speaker
Asus ROG MAXIMUS VIII FORMULA DDR4 ATX Motherboards
ASUS VS247H-P 23.6 Full HD 1920x1080 2ms HDMI DVI VGA Monitor
Samsung Hmd Odyssey Windows Mixed Reality Headset
Samsung 128GB BAR (METAL) USB 3.0 Flash Drive
Corsair CW-9060025-WW Hydro Series Liquid CPU Cooler
Echo Dot (2nd Generation) - Smart speaker with Alexa - Black
ASUS VivoMini Mini PC
Dell XF9PJ Latitude 7490 Notebook
Intel Core i7-8700 Desktop Processor 6 Cores
AMD Ryzen 7 2700X Processor Wraith Prism LED Cooler
Corsair RMx Series RM850 x 80 PLUS Gold Fully Modular ATX Power Supply
ASUS 24-inch Full HD FreeSync Gaming Monitor
Brother Monochrome Laser Printer; Compact All-in One Printer
Team 64GB microSDXC UHS-I/U1 Class 10 Memory Card with Adapter
LG Electronics 21.5 Screen LED-Lit Monitor
HP LaserJet Pro M227fdw All-in-One Wireless Laser Printer
Logitech Z313 Speaker System + Logitech Bluetooth Audio Adapter Bundle
PNY CS900 960GB 2.5 Sata III Internal Solid State Drive (SSD)
Samsung SmartThings ADT Wireless Home Security Starter Kit
Samsung SmartThings Smart Home Hub
Rosewill 2U Server Chassis Server Case (RSV-2600)
Corsair Apple Certified 16GB (2 x 8GB) DDR3 1333 MHz (PC3 10600) Laptop Memory
Acer Iconia One 10 NT.LDPAA.003 10.1-Inch Tablet