Demand steering through the smokescreen of stockouts: evidence from cigarette vending machines*

Pablo Casas † Asis Martinez-Jerez ‡ Helena Perrone §

March 2023

Preliminary and incomplete

Abstract

We show evidence of retailers using stockouts to steer demand towards products with higher retailer margins. We consider a unique setting where regulation limits the use of vertical agreements that could weaken steering incentives, prices are set by manufacturers, and product assortment is fixed in the short run. Using data on cigarette vending machines, we find empirical evidence consistent with retailers making strategic product re-stocking decisions. They exert less re-stocking effort for low-margin products, prompting consumers to shift purchases toward high-margin products. In a setting where prices vary infrequently, we exploit variation in product availability as a source of identification to recover preference parameters. Estimated diversion ratios are high across products within the same vending machine and low towards outside retailers. We also recover manufacturers' marginal costs. Counterfactual exercises based on our model parameter estimates measure the welfare effects of demand steering for consumers and manufacturers. On average, welfare losses are economically relevant; however, some manufacturers are better off under strategic stockouts.

*Perrone gratefully acknowledges financial support by the Deutsche Forschungsgemeinschaft (DFG) through the CRC TR 224 (project B04).

†Arizona State, casas.arce@gmail.com

‡Cornell University, fm447@cornell.edu

§University of Mannheim and MaCCI, helenaperrone@uni-mannheim.de

1 Introduction

Product stockouts in retailing are widespread and happen frequently. For example, Gruen et al. (2002) find that around 8% of products are out of stock each day in shops in developed countries, (see Hickman and Mortimer, 2016, for many other examples in different retail contexts). These stockouts may significantly decrease manufaturers' profits due to the lost sales resulting from out of stock products, and also consumer surplus, if consumers cannot find their favorite products. So why do these stockouts happen so frequently? Part of it is likely due to unpredicted shocks to demand and the costs of restocking and maintaining inventories. However, part of it could be non-accidental, resulting from profit-maximizing decisions of the retailers responsible for the stocking decisions. Stockouts can act as a non-pricing tool for retailer profit maximization by affecting product availability and steering consumers towards certain products.

More specifically, if different products within a category have different retailer margins, retailers may benefit from diverting demand from lower-margin products that are out of stock towards higher-margin products. We show empirical evidence consistent with strategic stockout, i.e., retailers putting more effort into restocking higher-margin products than lower-margin ones. This behavior may negatively affect manufacturers' profits that lose sales due to stockouts and consumer substitution towards brands produced by competing manufacturers.[1] It can also hurt consumers by leading them to purchase products with lower indirect utility. We quantify the costs of stockouts for manufacturers and consumers and show that these costs can be relevant even when the profit gains from demand steering for retailers seem small.

Smoking out the steering motive behind stockouts is relevant for at least two reasons. First, in many markets, demand steering can be an antitrust concern (for example, demand steering in digital retailing, such as the recent case involving the Amazon Buy Box) due to possible exclusionary effects. Second, suppose the retailers' objectives leading to stockouts conflict with the profit-maximizing objectives of the manufacturers. In that case, we have an instance of downstream moral hazard that can lead to market inefficiencies that eat up manufacturers and consumer welfare.

However, the use of stockouts to steer demand is challenging to detect in habitual contexts. There are several reasons for that. It is hard to separate stockouts due to unpredictable high-demand realizations (unintentional stockouts) and intentional stockouts to steer demand away from certain products (smokescreen to exclusionary practices), especially when the stocking effort of the retailer is unobserved. Also, typically retailers concomitantly use other profit-maximizing tools such as pricing, discounts, promotions, and changes in the offered product line. Furthermore, conventionally used vertical agreements between manufacturers and retailers to solve conflicts of interest soften retailers' demand steering incentives, making it hard to empirically identify the strategic stockout mechanism.

We elude the challenges mentioned above by considering a context where we observe sales and latent demand (demand for the stocked-out product) and stockouts and restocks, which are

[1] We use the words brand and products interchangeably throughout.

highly correlated with the effort level of the retailer but cannot be contracted upon. Furthermore, in our setting, manufacturers set retail prices nationally, the retailer product line is fixed in the medium run, and retailers cannot engage in promotional activities. Hence none of the conventional profit-maximizing tools are available to retailers. Additionally, due to strict industry regulations, manufacturers and retailers cannot enter vertical agreements to solve potential agency conflicts.

Our analysis centers on cigarette vending machines in a major European metropolitan area. Each machine sells different cigarette brands produced by various multiproduct manufacturers. This setting is especially fit for our study for several reasons. First, due to the design of our cigarette vending machines, consumers only learn that a product is out of stock after actively trying to purchase the product, that is, after pressing the product button in the machine. Importantly, the machine records that someone pressed the button. Therefore, we observe latent demand for a product when that product is out of stock, not only actual sales. This information is key in relating stockouts and retail effort, as discussed further below. Second, vending machines are a standard retail format with fixed capacities for a discrete number of unique products. Hence the retailer's decisions regarding assortment and restocking are discrete and relatively straightforward. Third, cigarette prices are decided by manufacturers at the national level and infrequently vary (less than once a year). Therefore, retailers' primary focus and only profit maximization instrument is stocking decisions (that may affect product assortment in the short run).[2] Fourth, retail markups are fixed, so we observe markups and can study how stocking decisions and stockouts respond to them. Last, due to heavy regulation, vertical agreements between manufacturers and retailers that could alleviate agency conflicts, such as vertical rebates, are prohibited.

Using these data, we show reduced-form evidence consistent with demand steering through stockouts. We show that stockouts decrease with margins, controlling for product total demand (sales plus lost sales) and unobserved machine characteristics. Remark that for this exercise to be meaningful, it is essential to control for total demand. Otherwise, a negative correlation between stockouts and margins could mean that higher margin products have higher demand and, therefore, stock out more frequently.[3] We also show that the number of times a product is recharged in a machine and month increases with its margins. Furthermore, we show that monthly machine revenue is higher when there is a stockout, controlling for total sales and machine fixed effects. Remark that in this case it is critical to control for total sales: we want to compare two machines with the same number of sales, one that had a stockout that could potentially divert demand to higher margin products, and the other that did not benefit from this possibility. Moreover, we show that the probability that a specific product gets recharged increases with margins when at least one product in the same machine gets recharged.[4]

[2] Also, prices do not respond to short-run market-brand specific unobserved demand shocks, making identification of demand parameters easier

[3] Our empirical evidence points to the opposite, that higher margin products have higher demand than lower margin products

[4] We observe in the data instances when some products in the machine are recharged but not all, including stocked-out products.

We also estimate preference parameters in a demand model that allows for rich heterogeneity across machines. We take advantage of our panel's long time series and estimate machine-specific parameters combined with a nested logit model where the nests are tobacco type (black, "light", and regular). The demand model considers the observed variation in available products crucial to identification in a context where prices vary infrequently. Using our estimated preference parameters, we construct the counterfactual demands if there are no stockouts. In this way, we can measure the consumer welfare effects of stockouts. We then recover manufacturers' marginal costs, assuming they set prices following a multiproduct Bertrand model. We combine the marginal costs and estimated preference parameters to calculate manufacturers' profits if there are no stockouts.

Results demand

Counterfactuals results.

A natural question is why retailers resort to strategic stockouts to increase margins instead of adjusting their product assortment permanently, i.e., offering only high-margin products. First, the assortment decision is long-run as the retailer cannot change the machine's product line often. Hence, in the short run, the retailer can only change assortment through stockouts. Remark also that the motivation behind strategic stockouts is to divert demand from low-margin products towards high-margin products. Therefore the absence of a particular low-margin product has to be unexpected for the consumer. A consumer looking for a low-margin product would not enter a bar with a machine that never offers her preferred product. Hence, the retailer cannot benefit from demand diversion from consumers who, once at the bar, will prefer switching cigarette brands over traveling to another bar. But the retailer still needs to attract low-margin product consumers. (Also, the product line of the machine may be easier to contract upon –it is observable, so not subject to moral hazard.)

Even if the product line is fixed in the medium run, why would retailers sometimes stock up on low-margin products and sometimes not? The retailer wants to avoid building a reputation of never carrying low-margin products. Alternatively, we can assume that retailers must commit to stocking up occasionally. The justification for this assumption is that the retailer wants to attract low-margin product consumers to stores, who then should not be able to perfectly predict that low-margin products will be out of stock. We propose a model where consumers decide to visit a particular store depending on the probability of finding their preferred product. Retailers maximize on that probability, i.e., they decide how frequently they will offer the low margin product.

1.1 Literature

Our paper is close to Conlon and Mortimer (2021), who study the effect of vertical rebates on retail effort and assortment decisions. Vertical rebates are payments made by the manufacturer to the retailer conditional on some sales target. They are used to relieve downstream moral hazard by incentivizing retailers to exert more sales effort. However, it can also have anti-competitive effects if the retailer drops products from competing manufacturers to facilitate attaining the target. In their paper, they study a vertical rebate paid by Mars, the leading

US candy manufacturer, using data on vending machines selling snacks. They develop and estimate parameters of a model of consumer choice and retailers' dynamic restocking decisions. Identification exploits exogenous assortment variation created by a field experiment where they removed Mars products from vending machines. Empirical results indicate that the vertical rebate led to the foreclosure of competing products and higher profits for Mars but lower consumer welfare and aggregate producer surplus.

Our paper also relates to the literature on moral hazard in expert-client agency contexts. Iizuka (2007) studies the agency problem between doctors and patients in drug prescription using data from Japan, where some doctors prescribe and sell drugs to their patients. The paper examines whether prescription decisions are driven solely by concerns about the patient's welfare or are also influenced by drug markups. Results indicate that drug markups significantly affect prescription choices. In another paper, Iizuka (2012) focuses on Japanese doctors' choice between generics and brand-name pharmaceuticals. Results show that doctors do not internalize patients' costs, which explains the infrequent adoption of generics, and that doctors ' prescription decisions respond to markup differentials between generics and branded products. This evidence is consistent with Liu et al. (2009), who shows that financial incentives affect doctors' choice between generics and brand-name drugs in Taiwan, where, like Japan, doctors prescribe and sell drugs to their patients.

More generally, our work speaks to the current debate on the anti-competitive effects of demand steering and self-preferencing or own-content bias in platforms. Online platforms, for example, can divert demand by posting high-profit margin products (e.g., own products) more prominently or using their recommendation systems to steer consumers. The theoretical literature on demand steering and self-referencing has been active. See, for example, De Corniere and Taylor (2019); Hagiu and Jullien (2011); Hervas-Drane and Shelegia (2022), among others. There is also growing empirical evidence on steering in online platforms. Shih and Spinola (2007), for example, discusses how Netflix uses its recommendation system to drive customers towards movies that generate higher revenues for itself. See also Lee and Musolff (2021), Gutierrez (2021), Lam (2021), and Raval (2022).

The rest of this paper is organized as follows. In the next section, we describe the relevant institutional details of cigarette vending machines in the European metropolitan area that we study. section 3 develops a parsimonious model of strategic stockout that, although simple, unveils the main mechanism and trade-off we focus on. This section also derives testable implications of the model. Section 4 describes the data and shows basic statistics, whereas section 5 shows results form reduced-form tests derived from the model implications. In section 6, we develop our demand approach, and discuss identification and instruments, and show preference parameter estimation results. In section 7, we describe the counterfactual exercises that are used to measure the costs of stockouts for manufactures and consumers. Section 8 concludes.

2 Cigarette vending machines in an European market

We focus on the cigarette vending machine market in a preeminent metropolitan area of a European country. As with everywhere else in Europe, this country's cigarette market is highly regulated. Tobacco can only be sold in a primary official network of tobacconists who are specially licensed by the state to sell cigarettes. Tobacco can be sold in the secondary network only in vending machines, which are allowed in bars and nightclubs, restaurants, convenience stores, newsstands, and hotels. Tobacco products cannot be sold directly by manufacturers to consumers. In 2017, there were more than 250000 cigarette vending machines in the country. Around 40%-50% of total national tobacco sales go through vending machines, which are managed and stocked by tobacconists in the area. Retailers can only stock tobacco in the tobacco store or inside the machine (but never in the bar, restaurant, etc., outside of the machine, so the maximum stock kept in bars, restaurants, etc., where there is a cigarette vending machine equals the machine's capacity). There are different machine models, but on average, machines have around 14 channels (i.e., "windows" or buttons for the various brands) and 25-35 packs per channel.

The manufacturers fix cigarette prices at the national level. Price changes must be announced, approved, and published beforehand, and they infrequently happen (less than once a year, on average). Retailers' margins are fixed by legislation at 8.5%, and prices at the vending machine are the prices at the tobacco store plus 15 cents.

Tobacco advertising and promotion are generally prohibited, with limited exceptions at tobacco stores.

3 A parsimonious model of strategic stockout

In this section we develop a simple model with few parameters to illustrate the demand steering mechanism through stockouts. The model captures main trade-off of retailer: create a reputation of carrying low margin product to attract low valuation consumer to the store, and using stockouts to affect product assortment and steer consumers to high margin products. The consumer will decide to pay the transportation cost to go to a certain store depending on the probability of finding her preferred product. The probability of having a product in stock will depend on retail margin differences between the products, consumer valuation differences, and consumer' cost. The model is parsimonious but helps us explain how differences in consumer preferences (which determine substitution patterns), differences in product margins, and competition among stores (captured by transportation costs) can affect retailers' choice of how frequently to stock each product.

3.1 Details of the model

There are two product, a high margin product H and a low margin product L. Consider two consumers that differ in terms of their valuation θ for the high margin product (but coincide in their valuation for the low type product, such that, $\underline{\theta} < \bar{\theta}$, where the upper bar and lower bar distinguish the valuations of the high and the low valuation consumers. We normalize to 1 the valuation of consumers for the low margin product.

Let t_H, t_L =be price of the high and low margin products respectively, with $t_H < t_L$, q be the probability of finding the low margin (cheaper) cigarette in the machine, and s the consumer transportation cost to get to the machine.

The high margin product is always in stock, and the high valuation consumer always pays the transportation cot to go to the machine. The low valuation buyer walks s if:

$$q(1 - t_L) + (1 - q)(\underline{\theta} - t_H) > s \tag{1}$$

The retailer's profit is given by:

$$\pi = m(q)(qt_L + (1 - q)t_H) + t_H \tag{2}$$

where $m(q)$ is the probability that the low valuation consumer walks to the machine.

Then,

$$m(q) = P(s \leqslant q(1 - t_L) + (1 - q)(\underline{\theta} - t_H)) = \int_0^{g(q)} f(s)\, ds \tag{3}$$

We assume that s has some random distribution on the [0,1] interval, e.g., a uniform distribution.

Then,

$$P(s \leqslant q(1 - t_L) + (1 - q)(\underline{\theta} - t_H)) = F(s) \tag{4}$$

where $g(q) = q(1 - t_L) + (1 - q)(\underline{\theta} - t_H)$

The Uniform distribution assumption implies that

$$P(s \leqslant q(1 - t_L) + (1 - q)(\underline{\theta} - t_H)) = g(q) \tag{5}$$

The problem of the retailer is to choose q so as to maximize profits:

$$\max_q g(q)(qt_L + (1 - q)t_H) + t_H \tag{6}$$

The first order conditions is:

$$g'(q)[qt_L + (1 - q)t_H] - g(q)[t_H - t_L] = 0 \tag{7}$$

and

$$g'(q) = (1 - t_L) - (\underline{\theta} - t_H) \tag{8}$$

3.2 Testable implications

Three straightforward testable implications of the model are:

1. If $t_H > t_L$ (and the difference between the two is "large enough"), then the profit maximizing q is lower than 1.

2. The stocking out probability for high margin products is smaller than for low margin product.

3. The profit maximizing q increases with t_L.

4 Data and descriptive statistics

The data bring daily information on cigarette vending machines in a large metropolitan area in Europe during 2016, 2017, 2018, and 2019. They include brand, sales, prices, and recharge occasions per machine channel. The data also have information on lost sales, i.e., how many times a consumer tried to purchase a stocked-out product. An unusual feature of these cigarette vending machines and very convenient for our analysis is that consumers only find out a product is out of stock after trying to purchase it, i.e., after clicking the product's button. Hence the machine's information system can collect data on frustrated purchases, which we call lost sales. Thus, a unique feature of our data is that we observe what, in other studies, is the latent demand for out-of-stock products.

We also have data on the location of other cigarette vending machines and tobacco stores around each machine. For the empirical exercises in this paper, we work with a subsample of 261 machines with the most frequent sales. We consider 14 brands as "inside" products and bunch the other products in the outside good. The 14 inside products have a joint market share of close to 90%.

We observe that some machines have multiple channels for the same brand in the data. Hence, one channel may be out of stock (and the machine records lost sales for that channel), whereas the product is still available in another channel of the same machine. For that reason, we further restrict our subsample by eliminating those machine-day combinations when a product was out of stock in one specific channel but not in the machine as a whole. To check for robustness, we also performed the reduced-form tests using the larger subsample, and the results remain qualitatively the same.

Our final subsample has 5,604,370 observations; one observation is a channel-machine-day combination. There are 261 distinct machines and 266,923 machine-day combinations. We then aggregate the data at the machine-day-product level (that is, we sum sales and lost sales of different channels selling the same product). Table 1 shows summary statistics of the variables relevant to our study at the machine-day level. The mean number of distinct brands sold in a machine is between 13 and 14, and the mean total daily sales per machine is slightly above 16 packs, thus a bit more than one pack per brand. The mean daily revenue per machine is 77.5 euros, so the average price per pack sold is 4.8 euros. On any day, the probability that a machine has at least one product out of stock is around 50%. This probability is calculated as the average, across machines, days, and brands, of a variable equal to 1 if a product is out of stock and zero otherwise. The probability of recharge, calculated as the average across machines, days, and brands of a variable equal to one if the brand is recharged and zero otherwise, is around 18%, much lower than the probability of a stockout. Table 1 also shows summary statistics on the number of days between recharges (12 on average), lost sales at any given day (3.3 on average), number of brands that stockout on any given day (1.1 on average), and on the density of machines and of tobacco stores in a radius of 500 meters around each machine 27 machines and 3,5 tobacco stores on average).

Table 2 shows summary statistics of each of the 14 inside products we consider in the empirical analysis. Japan Tobacco produces 2 of the inside products (1 and 14), Imperial Tobacco produces 6 out of the 14 products (2, 3, 4, 5, 12, 13), Phillip Morris produces four products (6,7, 10, 11), and British Tobacco produces two products (8, 9). These five manufacturers are the leading tobacco producers in the world market.

Table 1: Summary Statistics at Machine x Day level

	mean	sd	min	max
Number of machines	261.000	0.000	261.000	261.000
Number of brands x machine	13.714	1.601	1.000	15.000
Daily sales x machine	16.308	15.974	0.000	249.000
Daily revenue x machine	77.510	75.628	0.000	1190.900
Daily lost sales x machine	3.384	7.154	0.000	92.000
Daily stockout x machine	1.120	1.490	0.000	12.000
At least one product stocked-out	0.543	0.498	0.000	1.000
Prob of recharge	0.177	0.381	0.000	1.000
Number of days between machine recharges	11.959	38.504	1.000	522.000
Density of machines in radius	27.559	12.548	1.000	78.000
Density of tobacco store in radius	3.510	2.014	1.000	20.000
N	119919			

One observation is one machine-day; Density of machines/tobacco stores in radius equals he number of other machines/tobacco stores in 500 m radius around each machine.

Table 2: "Inside" products' summary statistics

Products	Market Shares	Mean Price	Nb of price changes	Mean Sales	Prob of stockout	Mean lost sales	Prob of recharge	Between recharges nb of days	Between recharges nb of sales	Between recharges revenue
1	0.13	4.80	2.17	0.16	0.64	0.15	2.23	19.07	16.44	78.97
2	0.02	4.79	0.45	0.03	0.05	0.07	0.66	113.99	9.10	43.66
3	0.08	4.99	1.70	0.20	0.78	0.14	1.89	28.84	11.97	59.88
4	0.04	4.55	0.83	0.06	0.15	0.10	1.00	47.76	9.32	42.41
5	0.07	4.70	1.14	0.08	0.21	0.11	1.30	38.62	12.34	58.00
6	0.04	4.55	0.83	0.06	0.14	0.10	1.02	53.01	10.09	45.98
7	0.04	4.55	0.67	0.06	0.14	0.09	0.87	66.86	9.45	43.09
8	0.02	4.57	0.50	0.08	0.19	0.08	0.71	71.10	9.03	41.32
9	0.07	4.70	1.13	0.08	0.23	0.12	1.28	35.09	12.18	57.29
10	0.06	5.10	1.07	0.08	0.23	0.11	1.23	40.56	13.08	66.75
11	0.15	5.10	2.60	0.13	0.46	0.16	2.65	16.76	17.89	91.20
12	0.03	4.70	0.48	0.05	0.11	0.07	0.69	96.41	9.01	42.39
13	0.04	4.25	0.64	0.06	0.14	0.09	0.82	67.06	8.10	34.48
14	0.07	4.70	1.14	0.07	0.20	0.11	1.30	40.64	11.81	55.55
Total	0.83									

Notes: Means are per brand x machine x day; "prob of stockout " "prob of recharge" is the mean probability that the brand stocks-out and gets recharged, respectively, in a day x machine; "Nb of price changes" is the number of times we observe a price change for a certain product in our dataset (during the 4 years covered by the data); "Between recharges" counts the mean number of days ("nb of days"), men number of sales ("nb of sales"), and mean revenue between product recharge event for a certain brand x machine x day.

5 Reduced-form evidence consistent with moral hazard

In this section, we show empirical evidence consistent with strategic stockouts where retailers make a lower effort to avoid stockouts of products with lower retail margins relative to the effort made to avoid stockouts of higher product margins. In particular, we check whether: (i) stockout probabilities decrease with margins, controlling for total demand; (ii) products with higher margins recharged more frequently, controlling for total demand; (iii) machine revenue is higher when there are stockouts, controlling for total sales; and (iv) stockouts decrease with competition, controlling for total sales.

5.1 Stockout probabilities and margins

Suppose retailers indeed make a lower effort to restock lower-margin products. Then, lower-margin products should stock out more frequently, controlling for product total demand (product sales plus lost sales). It is crucial to control for total demand, not only sales. Otherwise, a negative correlation between margins and stockouts could be due to lower-margin products having higher demand, for example, even if restocking decisions are independent of retail margins. [5]

Table 3 shows results of linear regressions of product stockout on product margin (results from probit are qualitatively the same). The dependent variable in each regression is a discrete variable equal to 1 if product j is stocked out in day t and machine m, and zero otherwise. The main right-hand-side variable is product j's margin, equal to 8.5% of the product's price. The first column shows the results of a regression of the stockout variable on margins, without any other control. Columns 2, 3, and 4 include controls for the product's total demand in day t and machine m. Columns 3 and 4 also include controls for whether the day is a weekend and manufacturer and machine fixed effects (only column 4).

When we do not control for total demand, the estimated coefficient for the margin is positive and significant, indicating a positive correlation between margins and stockouts. However, once we control for total demand (columns 2, 3, and 4), the margins coefficient becomes negative and significant. This result implies that, if we compare two products with the same demand, the product with lower margins will stock out more frequently, consistent with retailers exerting higher effort to avoid or reduce stockouts of higher-margin products.

5.2 Probability of getting recharged and margins

In the previous exercise, we used stockout probabilities controlling for total demand to examine retail recharge effort indirectly. Now we look at recharge opportunities directly. There are two ways retailers could exert less effort to avoid low-margin product stockouts. First, retailers could refrain from recharging a stocked-out low-margin product even when recharging other products in the machine. Second, retailers could wait longer to visit a machine to restock a

[5] Remark that lower-margin products have lower prices than higher-margin products. Hence, a negatively sloped demand would explain higher stockout rates.

low-margin product. Hence, we start by checking whether the probability that a stocked-out product gets recharged increases with margins, conditional on the machine getting recharged. Table 4 shows results for a regression on the linear probability of a product getting recharged (the dependent variable is equal to 1 if product j is recharged at period t in machine m) and the product's margin, conditional on the machine being recharged. In the last two columns, we show the results of regressions that also control for whether the product is out of stock. Coefficient estimates indicate a positive and significant correlation between margins and the probability of a product getting recharged. Furthermore, the absolute value of the estimated margin coefficients increases as we include more controls: column 2 includes machine fixed effects, and column 3 additionally controls whether the product was out of stock.

Next, we study whether the time interval between two product recharges decreases with the product's margin, controlling for total demand. It is critical to control for total demand because: if recharges do not respond to margins, we expect products with a higher demand to be recharged more frequently because stocks decrease faster. Table 6 show estimated coefficients for a regression of the number of days between product j's recharges in machine m and its margins. Results indicate that the duration between recharges decreases with margins. However, this could be due, for example, to higher margin products having higher demand, requiring more frequent recharges. However, the negative correlation between the duration between recharges and margins remains once we control for total product demand and machine fixed effects.

5.3 Machine revenues and stockouts

We argue that stockouts result from profit-maximizing strategies of retailers who allocate less effort in avoiding stockouts of low-margin products. This strategy can be profitable to retailers when demand for the out-of-stock product is diverted towards higher-margin products. Suppose this form of strategic stockout is indeed happening in our data. Compare two identical machines with the same sales in a certain period, but one has stockouts, and the other does not. Then, although the number of sales is the same between the two machines, the composition of the sales differs because, on average, the machine using strategic stockout sells a higher share of high-margin products. Hence, machines that experienced stockouts should have higher revenues than machines that did not.

To test this possibility, we run regressions of machine per period revenue on whether there was a stockout in the period (stockout==1). Critically, we control for sales, machine unobservable characteristics (machine fixed effects), and common unobservable factors (year and week fixed effects).[6] Table **??** shows the results from these regressions. The first column shows estimated coefficients from regressing monthly machine revenues on whether there was a stockout

[6]Remark that when we are looking at the machine level, we should look at total sales, not sales plus lost sales because as there is substitution from lost sales to other brands within the machine, the sum of lost and actual sales will overestimate actual demand at the machine level. When the analysis is at the product level, the relevant measure of demand is sale plus lost sales because this is the variable that measures the total number of consumers that wished to purchase that product in that machine that day (and some could not purchase it when the product was out of stock).

that month without further controls. In column 2, the regression includes machine fixed effects. Column 3 also control for monthly machine sales, and column 4 adds week and year fixed effects.

The coefficient for stockout in column 1 is positive and significant. The coefficient sign in column 1 is as expected because stockouts increase with sales. Hence unobserved positive demand shocks raise revenues and stockouts. Thus the importance of controlling for total sales and including fixed effects to control for remaining unobserved shocks. When we include further controls and the coefficient estimates are identified from comparable machines and occasions, the sign of the stockout coefficient becomes negative and significant. These results imply that all else equal, machines that have stockouts have higher revenue. The magnitude of the estimated stockout coefficients decreases as we add controls but remain positive and significant even when we include week and year fixed effects.

Note that the estimated coefficient for stockout in the last column is small. But we should be careful when interpreting it. It does not mean that using strategic stockouts increases revenues by the amount of that coefficient compared with not using them. Why? We are comparing machines that have no stockout in the month with machines that have one or more stockouts, so the coefficient does not measure the marginal value of having a stockout. Also, remember that stocking up a machine has a cost. So the profit increase due to a strategic stockout is necessarily larger than the revenue increase.

5.4 Competition and stockouts

Retailers hope to increase profits by steering demand from lower to higher-margin products if they are monopolists or if the diversion ratios from out-of-stock products to other retailers are low. This insight implies that we should observe more stockouts in machines with lower competition, all else equal. More geographically isolated machines (away from other machines or tobacco stores) can profit more from stockouts because the transportation costs for consumers to look for alternative retailers are higher than when there are many machines around. We look, therefore, at how stockout probabilities relate to the density of competitors at a radius of 500 meters around the machine. Table 8 show the results of regressions where the dependent variable is equal to the number of stockout events (day x product) in machine m in a specific month. The main right-hand-side variables are the number of other machines in a radius of 500m around machine m and the number of tobacco stores in a radius of 500m around machine m.

Remark that it is essential to control total sales because a high number of machines and tobacco stores in an area can be due to high demand for cigarettes in that area. Therefore, if we do not control for total sales, there could be a positive correlation between the number of competitors and stockouts because high-demand areas have a higher number of cigarette machines and stores. The density of tobacco stores negatively correlates with the monthly stockout frequency. This result is consistent with our demand steering theory. Nevertheless, proximity to tobacco stores could also decrease stockouts because it may decrease the cost of recharging the machine (because the retailer responsible for recharging it is closer by, for example). Hence it is crucial also to control for recharge frequency. Otherwise, a negative correlation between stockouts and the number of retailers could be solely due to retailers recharging nearby machines more frequently. Once we control for machine recharge frequency (column 3), the estimated coefficient for tobacco stores' density remains negative and significant, indicating that lower recharge costs cannot explain the estimated negative correlations.

However, the density of other cigarette machines has no significant effect on the machine-level frequency of stockouts. In interpreting this result, we should consider that the same retailer can own other machines in the area. Hence, in this case, the other machines are not actual competitors and therefore do not impact strategic stockout decisions, which depend on expected diversion towards other retailers.

6 Demand

Our demand model takes advantage of a high T in our panel to estimate machine-specific preference parameters. Our approach allows for full unobserved consumer heterogeneity across machines combined with a 1-level-nested logit, where groups are defined by tobacco type (regular, "light', and black). Compared to the standard random coefficient approach, we model consumer preferences as individual level parameters that we estimate, rather than treating then as random effects drawn from a known distribution.[7] We also allow for consumption sets to vary per machine and period, which represents an essential source of data variation for parameter identification in a setting where prices vary infrequently.

6.1 Consumer choice

Indirect utility and individual purchase probability

$$u_{ijt} = \alpha_i p_{jt} + \xi_{ijt'} + \gamma_{it} + \tilde{\epsilon}_{ijt} \tag{9}$$

where p_{jt} is the price of product j at period t, α_i is consumer i's marginal utility of income, $\xi_{ijt'}$ is consumer i's taste for unobserved product characteristics of product j that may vary over time, γ_{it} captures common factors , and $\tilde{\epsilon_{ijt}}$ are consumer- and product-specific per-period unobserved shocks.

We make the nested logit model distributional assumption on $\tilde{\epsilon}_{ijt}$, which allow consumer product valuations to be correlated among products in a same group (See Verboven, 1996, for a detailed discussion of the nestel logit model)We consider three groups depending on the type of tobacco, plus the outside good. We assume that

$$\tilde{\epsilon}_{ijt} = \zeta_{igt} + (1 - \sigma_i)\epsilon_{ijt} \tag{10}$$

where ϵ_{ijt} is iid extreme value and ζ_{igt}, where g idexes the group, has the distribution such that $\tilde{\epsilon}_{ijt}$ is extreme value. The parameter σ_i measures consumer i's taste correlation across products

[7] Grigolon and Verboven (2014) brings a thorough comparison between the nested logit and random coefficients models

in the same group and its value should lie between zero and 1 (remark that the standard nested logit nesting parameter σ does not vary across consumers. But in our demand model allows for a rich pattern of consumer heterogeneity).

At period t, consumer i chooses to purchase the product that maximizes her indirect utility. Hence at every period t, the probability that i chooses j is

$$P(y_{ijt} = 1|\alpha_i, \gamma_{it}, \boldsymbol{\xi_{it}}, \boldsymbol{p_t}) = P(u_{ijt} \geq u_{ikt}, \forall k \in A_{it}|\alpha_i, \gamma_{it}, \boldsymbol{\xi_{it}}, \boldsymbol{p_t}) \tag{11}$$

where $y_{ijt} = 1$ if consumer i purchases j at period t and zero otherwise, and $\boldsymbol{\xi_{it}}$ and $\boldsymbol{p_t}$ are vectors stacking the unobserved taste parameters and product prices, respectively. Given the above distributional assumptions and normalizing the mean utility of the outside good to zero, this probability can be written as (Ivaldi and Verboven, 2005):

$$s_{ijt} = \frac{\exp(\delta_{ji})/(1-\sigma_i)}{D_{ig}} \frac{\exp D_{ig}^{(1-\sigma_i)}}{1 + \sum_{g=1}^{G} \exp D_{ig}^{(1-\sigma_i)}} \tag{12}$$

where D_{ig} is defined by:

$$D_{ig} = \sum_{k \in G_g} [\exp \delta_{ji}/(1-\sigma_i)] \tag{13}$$

We assume ϵ_{ijt} in equation (9) has the one-level "nested logit" distribution. Aggregating individual probabilities at machine-level and doing the standard inversion (Berry, 1994), we get the estimable equation:

$$\ln(s_{mjt}/s_{0mt}) = \alpha_m p_{jt} + \sigma_m \ln(s_{jmt|g_m}) + \xi_{mjt'} + \gamma_{mt} + \epsilon_{mjt} \tag{14}$$

where m indexes machine, $s_{jmt|g_m}$ is the market share of product j within its group in machine m and period t, and σ_m captures machine-specific unobserved taste correlations between products within the same group.

6.2 Empirical implementation

6.3 Identification and Instruments

In general, prices are endogenous in a demand equation. Firms' pricing decisions respond to unobserved product characteristics and per period unobserved shock to the willingness to pay. Therefore they are typically correlated with the error term in the demand equation. In (14), we explicitly control for products' machine-specific unobserved characteristics and allow them to vary over time ($\xi_{mj\tilde{t}}$). Conditional on this rich set of controls for consumers' product taste heterogeneity, prices can be considered exogenous in our setting. That is because, as mentioned earlier, cigarette prices in the country we study are set by the manufacturers at the national level and vary infrequently. Therefore, they do not respond to short-term machine-specific shocks to consumers' willingness to pay for the product. Thus, they are uncorrelated with ϵ_{ijt} conditional on including product fixed effects that capture unobserved product characteristics.

However, we should consider the endogeneity of the product j's market share within its group, $s_{jmt|g_m}$. This share can be correlated to unobserved socks affecting the probability of choosing product j at a certain period t and machine m. Therefore, unbiased and consistent estimates of model parameters require using instrumental variables correlated with the group shares but uncorrelated with the willingness to pay for product j. We follow the standard strategy in the literature (Ivaldi and Verboven, 2005) and exploit group-level variation in available sets.

In our application, we should also discuss another potential endogeneity problem. We argue that product stockouts are partly due to strategic profit-maximizing retailers' decisions. Hence, available products could be correlated with unobserved per-period and per-product demand shocks in some periods and machines. Assume retailers' stocking decisions are affected by a per-period and per-machine overall unobserved shock, not a brand-specific shock. An example of such a shock would be the arrival of consumers with high transportation costs unwilling to switch machines if their favorite product is out of stock. Then, including day fixed effects that control for unobserved common factors at the machine-day level solves the issue.

Remark that our strategic stockout story says that retailers stocking decisions are affected by products' retail margins and price. Hence there is a correlation between available products and prices in some periods. However, as we explicitly control for prices, this correlation does not challenge identification (it would be different if prices responded to unobserved period-and brand-specific shocks to demand. In this case, some unobservable shock to demand could affect both the set of available products and the prices, creating an endogeneity problem.)

Another potential threat to identification is the following. Suppose a retailer offers a product set Ω when she expects a selected group of consumers to show up at a particular day t in her machine (or the same consumers to behave differently in t). Hence the group of consumers on day t could be different in unobservable ways to consumers that show up when Ω' is offered. In this case, the preference parameters in t differ from those in t'. Our identification assumption is that conditional on machine-specific and date-specific unobserved shocks, consumers' preferences are comparable across periods in the same machine.

6.4 Product characteristics and Instruments

We estimate a demand equation for each machine, and hence estimate machine-specific taste parameters that should reflect taste differences between customers that shop in different machines. We include prices for each product, and brand fixed effects to capture unobserved product characteristics, as well as brand-year and brand-month fixed effects, allowing for brand perception to evolve over time and seasonally. We also include day-specific fixed effects.

The instruments we use for group market shares are the number of available products within each group, machine, and day.

6.5 Preference parameter estimates

We first estimate a pooled nested logit model, where the price and group market share coefficients are constrained to be the same, and consumers taste heterogeneity is allowed only through machine fixed effects. Table 9 shows the estimated preference parameter coefficients for such a model. The first column shows demand estimation results for a specification with no instrumental variables and no machine FE; the second column, shows results for a specification that includes instrumental variables; and the third column shows results for the full model that includes instrumental variables, machine fixed effects, and period fixed effects.

Table 10 shows means and standard deviation values for the estimated preference parameters α_m, the price coefficients, and σ_m, the group market share coefficient. Both means are within the expected intervals for the parameters to be consistent with utility maximization, i.e., $\alpha_m < 0$ and $\sigma_m \in [0, 1]$.

7 The costs of stockouts

In this session, we use our preference parameter estimates to construct the market shares under the counterfactual scenario where there were no stockouts. In that way, we can calculate the consumer surplus in this counterfactual scenario and compare it with the factual consumer surplus, measuring the costs of stockouts to consumers. Furthermore, we combine the counterfactual market shares with assumptions on manufactures conduct to measure how much surplus manufacturers loose due to stockouts.

7.1 Cost to consumers

Although we observe in the data the number of frustrated purchases of an out of stock product (intentions to purchase an out of stock product), we do not observe where these frustrated purchases went to (which other product, if any, in the machine). Therefore, we do not directly observe what per product sales would have been in the absence of stockouts. For that we need to use the estimated consumer preference parameters from our demand model to simulate consumers' choices when there are no stockouts (conditional on machine product line). We can then calculate consumer surplus when there are stockouts (observed choices) and when there are not stockouts (simulated choices using estimated preference parameters) to get a measure of the consumer welfare costs of stockouts.

Given our demand model assumption, the net consumer surplus, CS, is measured as:

$$CS = \frac{1}{\alpha} ln \left(1 + \sum_{g=1}^{G} D_g^{1-\sigma} \right) \tag{15}$$

We find that in the absence of stockouts, consumer surplus would have been XXX % higher. In euro amounts, this represents, for a regular smoker (one pack a day) extra YYY euros per year.

7.2 Cost to manufacturers

We also calculate the difference in manufacture surplus, defined as the difference in variable profits, between the counterfactual and factual scenarios. We assume manufacturers play a static multiproduct Bertrand game and that marginal costs are constant. We need to recover marginal costs to calculate the manufacturer surplus. But, as a first approximation, we assume price cost margins for each product produced by a certain manufacturer is the same, that is, $(p_j - c_j) = (p_k - c_k)$ for all $j, k \in F_f$, where F_f is the set of product produced by manufacturer f. Under this assumption, the difference in counterfactual and factual surplus of a multi-product Bertrand manufacturer f with constant marginal costs is just the difference in market shares summed over the products in f's production set. The variable profit of manufacturer f is:

$$\Pi_f = \sum_{j \in F_f} (p_j - c_j) \sum_m s_j(p)^m M_m \tag{16}$$

where M_m is the market size of machine m and p is the vector of prices of all brands. Therefore, the percentage difference between counterfactual and factual variable profits for manufacturer is

$$\%\Delta MS_f = \frac{\sum_{j \in F_f} (p_j - c_j) \sum_m \tilde{s_j(p)}^m M_m - \sum_{j \in F_f} (p_j - c_j) \sum_m s_j(p)^m M_m}{\sum_{j \in F_f} (p_j - c_j) \sum_m s_j(p)^m M_m} \tag{17}$$

where $\tilde{s_j}^m$ is the counterfactual market share of product j in machine m. If we assume that all products produced by manufacturer f have the same price cost margins, that is, $(p_j - c_j) = (p_k - c_k) \forall j, k \in F_f$ the the above expression simplifies to

$$\%\Delta MS_f = \frac{\sum_{j \in F_f} \sum_m M_m (\tilde{s_j}^m - s_j^m)}{\sum_{j \in F_f} \sum_m s_j^m M_m} \tag{18}$$

To recover the manufacturers's marginal costs again we assume that manufacturers play a static multiproduct Bertrand game. Assume also that the cigarette demand at the machines is representative of the national cigarette demand in general (because prices are set nationally).Then, each manufacturer f set prices of each of the products $j \in F_f$ following first order condition:

$$\sum_m M_m s_j^m(p) + \sum_{k \in F_f} (p_k - c_k) \sum_m M_m \frac{\partial s_k^m(p)}{\partial p_j} = 0 \tag{19}$$

The set of J first order conditions implies price-cost margins for each product. the marginal costs can be solved for explicitly by defining $S_{jk} = -\sum_m M_m \frac{\partial s_k^m(p)}{\partial p_j}, j, k = 1, ..., J$,

$$\Gamma^\star_{jk} = \begin{cases} 1, & \text{if } \exists f : (k, j) \subset F_f, \\ 0, & \text{otherwise} \end{cases}$$

and Γ_{jk} is a JXJ matrix with $\Gamma_{jk} = \Gamma^\star_{jk} * S_{jk}$. In vector notation, the marginal costs are:

$$c = p + \Gamma_{jk}^{-1} s(p) \tag{20}$$

where $s(p)$ is a $JX1$ vector such that $s(p)_j = \sum_m M_m s_j^m(p)$, and p and c are $JX1$ vectors of prices and marginal costs, respectively.

Remark that here we make an implicit assumption that the price effects that we estimate using cigarette sales in vending machines identifies preference parameters for cigarettes in general in the country we study. That is, there is an implicit assumption that consumers respond to prices and availability the same way when they are buying in the machines and when they are buying in a tobacco store. This is not necessarily true. Consumers could be less price sensitive or less brand loyal when buying in machines, for example. However, we cannot identify substitution patterns usacross cigarette brands using conventional sales data from tobacco stores because these data do not have enough variation for parameter identification; that is, we do not observe frequent stockouts that observably vary the consumption set. Furthermore, as mentioned earlier, cigarette machine sales represent around 50% of total cigarette sales in the country we study.

Results on differences in manufacturer surplus

8 Conclusion

Using data form cigarette sales in vending machines, we show empirical evidence consistent with demand steering through strategic stockout. Retailers may benefit from stockouts as a way of changing product assortment in the short run when they can divert demand from products that are out of stock to products with higher retail margins. The cigarette market is highly regulated preventing the use of vertical agreements to solve agency problems. Therefore it constitutes a unique setting to study the extent and the costs of downstream moral hazard for manufacturers and consumers. Our setting is also unique in that we observe frequent variation in consumption sets due to frequent stockouts. This enables identification of preference parameter in an industry whit no price variation. Our counterfactual exercises indicate that the costs of stockouts to manufacturers and consumers are sizable, even in a situation where the retailer profit advantage of using demand steering is apparently small.

Our paper and results have relevant public policy and managerial implications. In terms of public policy, competition practitioners should consider that stockouts could be a smokescreen for demand steering with potentially exclusionary consequences. For management, our work implies that manufacturers should consider retailer strategic stockouts when setting prices. It also sheds light on a alternative non-price profit maximization tool for retailers, especially when the product line is fixed in the short run.

References

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, 242–262.

Conlon, C. T. and J. H. Mortimer (2021). Efficiency and foreclosure effects of vertical rebates: Empirical evidence. *Journal of Political Economy 129*(12), 3357–3404.

De Corniere, A. and G. Taylor (2019). A model of biased intermediation. *Rand 50*(4), 854–882.

Grigolon, L. and F. Verboven (2014). Nested logit or random coefficients logit? a comparison of alternative discrete choice models of product differentiation. *Review of Economics and Statistics 96*(5), 916–935.

Gruen, T., D. Corsten, and S. Bharadwaj (2002). Retail out of stocks: A worldwide examination of causes, rates, and consumer responses': Grocery manufacturers of america. *Washington DC, USA*.

Gutierrez, G. (2021). The welfare consequences of regulating amazon. Technical report, mimeo, New York University.

Hagiu, A. and B. Jullien (2011). Why do intermediaries divert search? *Rand 42*(2), 337–362.

Hervas-Drane, A. and S. Shelegia (2022). Retailer-led marketplaces.

Hickman, W. and J. H. Mortimer (2016). Demand estimation with availability variation. In *Handbook on the Economics of Retailing and Distribution*, pp. 306–340. Edward Elgar Publishing.

Iizuka, T. (2007). Experts' agency problems: evidence from the prescription drug market in japan. *The Rand journal of economics 38*(3), 844–862.

Iizuka, T. (2012). Physician agency and adoption of generic pharmaceuticals. *American Economic Review 102*(6), 2826–58.

Ivaldi, M. and F. Verboven (2005). Quantifying the effects from horizontal mergers in european competition policy. *International Journal of Industrial Organization 23*(9-10), 669–691.

Lam, H. T. (2021). Platform search design and market power. Technical report, Discussion paper, Northwestern University.

Lee, K. H. and L. Musolff (2021). Entry into two-sided markets shaped by platform-guided search. Technical report, mimeo, Princeton University.

Liu, Y.-M., Y.-H. K. Yang, and C.-R. Hsieh (2009). Financial incentives and physicians' prescription decisions on the choice between brand-name and generic drugs: Evidence from taiwan. *Journal of health economics 28*(2), 341–349.

Raval, D. (2022). Steering in one click: Platform self-preferencing in the amazon buy box. *Harvard Law School*.

Shih, Willy C., S. P. K. and D. Spinola (2007). Netflix. *Harvard Business School Case 607-138*.

Verboven, F. (1996). International price discrimination in the european car market. *The RAND Journal of Economics*, 240–268.

Table 3: Margins and the probability of a product being stocked-out in a machine x day

	Stockout of a product in a machine X day			
	(1)	(2)	(3)	(4)
margin	1.07***	-0.31***	-0.17***	-0.29***
	(0.01)	(0.01)	(0.01)	(0.01)
Total Demand		0.06***	0.06***	0.06***
		(0.00)	(0.00)	(0.00)
weekend			-0.01***	-0.01***
			(0.00)	(0.00)
N	1823789	1823789	1823789	1823789
r2	.0056548	.2175226	.2183712	.2254326
Manufacturer FE	no	no	yes	yes
Machine FE	no	no	no	yes

Notes: (i) Margin is the 8.5% of the price of the product; (ii) OLS regressions where the left-hand size variable is equal to 1 if the product j was out of stock in machine m and day t; (iii) Total demand is sales plus lost sales due to stock out and Weekend indicates whether the day of the week is either a Saturday or a Sunday.

Table 4: Product margins and the number of per product recharge events per machine and month

	Sum of recharge events per product x machine x month				
	(1)	(2)	(3)	(4)	(5)
margin	8.51***	-3.96***	7.06***	-0.60**	9.70***
	(0.41)	(0.35)	(0.41)	(0.30)	(0.34)
Total Demand		0.03***		0.03***	
		(0.00)		(0.00)	
stockout			1.18***		1.15***
			(0.02)		(0.02)
N	88816	88816	88816	88816	88816
r2	.0048087	.3119469	.0299524	.2984605	.0429053
Machine FE	no	no	no	yes	yes

Notes: blabla bla.

Table 5: Product margin and the probability a stocked out product gets re-stocked when the machine gets recharged

	Product got recharged = 1			
	(1)	(2)	(3)	(4)
price	0.26***	0.67***	0.26***	0.77***
	(0.02)	(0.06)	(0.01)	(0.03)
N	8163	8163	30935	30935
r2	.0152935		.0183481	
Machine FE	no	no	yes	

Table 6: Product margin and duration (in days) between product's recharges

	Duration between recharges		
	(1)	(2)	(3)
price	-37.54***	-22.88***	-35.32***
	(0.34)	(0.34)	(0.28)
Total Demand		-8.75***	-4.01***
		(0.03)	(0.03)
N	1963678	1963678	1963678
r2	.0061105	.0449482	.0217244
Machine FE	no	no	yes

Notes: (i) One observation is one product x machine x day; (ii) dependent variable is number of says between two recharges at the product level.

Table 7: Machine stockouts and machine monthly revenues

	Total machine revenue per month			
	(1)	(2)	(3)	(4)
stockout	849.57***	970.34***	2.38***	
	(33.09)	(29.14)	(0.87)	
sales			4.74***	
			(0.00)	
N	8197	8197	8197	
r2	.0744283	.4399572	.9995666	
Machine FE	no	yes	yes	yes
Month and Year FE	no	no	no	yes

Notes:

Table 8: Monthly frequency of stockouts in machine and number of other machines and other tobacco stores in radius around the machine

	(1)	(2)	(3)
Density of other machines	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Density of tobacco stores	-0.05	-0.08**	-0.09**
	(0.03)	(0.03)	(0.03)
Total Demand		0.00***	0.00***
		(0.00)	(0.00)
Machine recharge frequency			-0.00**
			(0.00)
N	6684	6684	6684
r2			

Notes: (i) One observation is one x machine x month x year; (ii) probit regressions and clustered errors at the machine level

Table 9: Demand model estimates with homogenous consumers

	OLS	Machine x Day FE	FE + IV
Price (α_i)	-0.02***	-0.04***	-0.06***
	(0.00)	(0.00)	(0.00)
Group market share (σ_i)	0.96***	0.78***	0.25***
	(0.00)	(0.00)	(0.01)
Machine and day FE	no	yes	yes
Product x Year FE	yes	yes	yes

Notes:

Table 10: Demand model estimates with consumer heterogeneity

		mean	sd
Price (α_i)	Estimate	-0.260	0.324
	Standard Error	0.089	0.077
Group market share (σ_i)	Estimate	0.530	0.345
	Standard Error	0.178	0.081
N	1728045		

One observation is one product-machine-day

Table 11: Difference between counterfactual and factual Consumer Surplus

consumer heterogeneity	N	% difference	std dev	difference in euros	
				day	year
yes	19,785	3.18%	0.022	0.54	197.47

Table 12: Difference between counterfactual and factual Manufacturer Surplus – model without consumer heterogeneity

Manufacturer	N	% Diff Manufacturer surplus	Std. Dev.
British American Tobacco	76	6.57%	0.113
Imperial Tobacco	152	4.21%	0.042
Japan Tobacco	65	8.34%	0.197
Phillip Morris	120	2.29%	0.053