Price discrimination in organic food markets: the case of ready-to-eat cereal*

Thiago Cacicedo†
University of Alicante

Abstract

In this paper, price discrimination with respect to the organic attribute in the ready-to-eat cereal industry is quantified. I estimate a random coefficient discrete choice demand model to obtain the price-elasticity of each product. Then, with the estimated elasticities and a supply model, I recover the marginal costs, which allows to disentangle whether the amount of price difference between organic and non-organic products is due to price discrimination or due to different production costs. I find that around 6% of the price difference is due to price discrimination with respect to this attribute. Counterfactual exercises show that: i) a tax on non-organic products is welfare detrimental and does not substantially reduce price discrimination; ii) price discrimination happens due to the existence of high income households.

Keywords: organic food, price discrimination, BLP, structural model

JEL classification: L13, D22, D43

*I thank Pierre Dubois, Joel Sandonís, Pedro Albarrán, participants at the Applied Micro Workshop at the University of Alicante. Financial support from the Spanish Ministerio de Economía, Industria y Competitividad and the European Social Fund under project ECO2015-65820-P is acknowledged. All estimates and analysis in this paper, based on data provided by IRI are by the authors and not by IRI. All errors and omissions are my responsibility.

†Email: thiago.cacicedo@ua.es

1 Introduction

The market for organic food has been growing rapidly in the past decades. In the United States, the sales of organic food were of $52.5 billions in 2018, a growth of 6.3% compared to the previous year[1], and in the United Kingdom it rose to £2.45bn in 2019, a growth of 4.5%.[2] There are several reasons to explain these numbers. One of them is that traditional retailers have recently started to offer this type of product, and thus consumers can buy them in their usual grocery shopping trip. Moreover, retailers are also offering their own brand of organic products (Jaenick and Carlson 2015). Another important reason is that organic products are seen as a healthier and environmentally friendly option. To be considered organic, products cannot use synthetic fertilizers, pesticides, and livestock feed additives during the production process. Zepeda and Deal (2009) report that the health issue is the main reason for consumers to start purchasing organic products.[3]

An important characteristic of organic food products is that they are usually more expensive than their non-organic counterpart. Indeed, Thompson and Kidwell (1998) and Jaenick and Carlson (2015) report a significant price premium for organic foods. The former authors use data collected from two retail stores in Arizona in 1994 and find that organic fresh products have a price premium in the range of 40-175%, whilst the latter, using data from a representative sample of more than 40.000 households between 2004 and 2010, find price premiums in the range of 45-70%. These premiums are so important that even governments are trying to benefit from this industry. For example, The Economist reported that Turkey is subsidising the production of organic tea hoping that wealthy foreigners will then pay more for Turkish tea, and consequently increase the revenue of Turkish firms.[4]

[1] Source: Organic Trade Organization. Their report can be obtained at https://ota.com/what-ota-does/market-analysis/organic-industry-survey/organic-industry-survey?oprtid=012G0000001BAsuIAG&caid=701G0000000yqzN.

[2] Source:https://www.theguardian.com/business/2020/feb/05/organic-food-and-drink-sales-rise-to-245bn.

[3] On the other hand, the medical literature has not reached a conclusion if organic food is healthier or not (for a survey on this topic see, for example, Smith-Spangler et al. (2012).

[4] Source: https://econ.st/2EwlxLh. The state tea company plans to convert entirely to organic in 2023.

Governments have been urging households to eat healthier (Lan and Dobson 2017), and price is a key determinant to achieve this goal. Thus, it seems important to understand the determinants of such huge price premiums. The first candidate to explain this price premium is the production costs of organic foods. Mayen et al. (2010) find that American organic dairy farms are 13% less productive than non-organic ones. However, consumers do not buy directly from producers but from food manufacturers. It might also be the case that manufacturers are using price discrimination in order to extract additional consumer surplus, taking advantage of consumers' higher willingness to pay for organic food. For example, Akgüngör et al. (2010) find that the consumers are willing to pay more for organic food in urban Turkey, whilst Huang and Liu (2017) find that consumers buy more bottled water compared to other beverages when they receive more information about the water health status, which is also documented by Batte et al. (2007) for the U.S.

Concerning the organic attribute, Bonanno (2016) finds that consumers positively appraise it in the yoghurt market. Therefore, it seems important to determine the relevance of price discrimination in organic food markets, which is the empirical aim of this paper. To the best of my knowledge, this is the first paper to specifically analyse price discrimination regarding the organic attribute.

I study this question using data from the U.S. ready-to-eat cereal industry. The most challenging issue in answering this question is that costs are not observable, and thus, a priori, I cannot know whether the price premium of organic food is due to price discrimination or simply due to cost differences. To circumvent this problem, I use a discrete choice approach such as Berry (1994) and Berry et al. (1995).

I analyse this industry due to its importance for consumers. Cereal is the main breakfast food for most households in the United States. In 2019, 86.25% of the population consumed it.[5] Additionally, it is a industry that accounted for sales of 8.66 billion dollars in 2017.[6]

[5] Source:https://www.statista.com/statistics/279999/us-households-consumption-of-breakfa

[6] Source: https://www.statista.com/statistics/407906/us-ready-to-eat-cereal-sales/

My model follows closely Cohen (2008). The demand side is modelled as a random coefficient model and I obtain the demand elasticities from its estimation. Combining these elasticities with a supply side model, I recover the marginal cost and the margins for each product. And using the estimated margins, I am able to identify the existence of price discrimination with respect to the organic characteristic. Finally, I perform a counterfactual experiment and find that 6% of the price difference between organic and non-organic products is due to price discrimination with respect to the organic attribute.

The study of price discrimination is well documented in the literature. For example, Cohen (2008) uses aggregate data to analyse price discrimination with respect to the size of the product in the paper towel market. He finds that quantity discounts are consistent with second degree price discrimination and that consumers are better off when it happens. The reason is that the use of this pricing strategy results into more competition in the multi-roll package size segment of the market.

Although not directly comparable, the results of the present paper are similar to those in Wallace (2018), which analyses price discrimination in the U.S. fluid milk sector. Although it is a good that consumers purchase in grocery stores, this market presents several other differences when compared to the cereal market. Milk is a more homogeneous good and the price of raw milk is regulated by the government, which implies that it is easier to obtain the costs for fluid milk. He found that retailers are only able to price discriminate on their own private label products, and that 7.2% of the markup can be explained by price discrimination, whereas for regional brands this number reduces to 0.34%.

One of the main issues regarding price discrimination has to do with its welfare consequences. Although firms use price discrimination to increase profits, its impact on consumers is, *a priori*, not clear. The reason is that price discrimination can induce consumers to switch, buying the product that gives them a higher utility, and this could generate an increase in consumer surplus. For instance, Leslie (2004) analyses a specific Broadway show, *Seven Guitars*, and finds that price discrimination allows firms to increase profits by 5%, and it has little or no effect on consumer welfare. On the other hand, Miller and Osborne (2014) find that price discrimination hurts consumers. They analyse the Portland cement industry in Arizona, California, and Nevada and estimate a structural model of spatial differentiation and price discrimination, finding that consumer surplus would increase by $30 million dollars per year if price discrimination was forbidden.

In order to be able to analyse the welfare effects of price discrimination, I perform a policy experiment introducing a 5% tax on non-organic products. Notice that this type of experiment is interesting even in the absence of price discrimination. If governments want to induce a healthier lifestyle, and people believe that organic food is healthier, or the government and the population have environmental concerns, then taxation is a way to induce the consumption of organic food. The results suggest that this tax policy would be detrimental to welfare. The reason is that the tax is not enough to make consumers switch and consume more organic cereal, which implies that consumers pay more and keep buying the regular ones. The reason is that the elasticities obtained from the demand estimation suggest that regular and organic cereals are not close substitutes. Hence, in the cereal industry, taxing regular products might not be the best way to induce the consumption of organic products.

This policy experiment is also directly related to a growing literature analysing the effects of taxation on unhealthy products ("sin tax"). For example, Zheng et al. (2013) analyse the theory and policy implications of sales and excise taxes on food and beverage consumption, when consumers may have imperfect tax knowledge. They find that the effects on demand are larger for excise taxes. However, in the long-run, after a period of learning about the taxes, the effects become more similar. Wang (2015) analyses the impact of taxing soda on consumer welfare, taking into consideration the storability of this product. Similar to the results obtained in the present paper, the author finds that it is unlikely that the taxation will impact the consumption of soda, although it increases tax revenues. Similarly, Flores and Rivas (2017) also find that taxes are relatively ineffective to reduce unhealthy food consumption. They compare this result to direct cash incentives, which despite being more effective on reducing the consumption, its higher cost can lead to significant monetary losses to the social security system. On the other hand, Dubois et al. (2020) analyse which segments of the population are more affected by the tax on sodas. They find that taxes are effective at targeting the sugar intake of younger individuals, but it is less effective on those with a high sugar diet.

Finally, the estimates of the demand model in the present paper suggest that organic cereal is preferred by high-income households. Therefore, price discrimination might arise due to the presence of these households in the market. In order to analyse if this is the case, I perform a counterfactual exercise where I set a uniform income for all households in the market. The result of this exercise shows that, when I eliminate the high-income households, price discrimination disappears. Therefore, the presence of this type of household can be seen as determinant of price discrimination.

The remainder of the paper is organised as follows. Section 2 describes the data and the cereal industry. Section 3 describes the empirical model, Section 4 discusses the estimation strategy, and Section 5 presents the results. Section 6 presents the policy experiment, and Section 7 concludes.

2 The Data and Industry

In this paper I use two different datasets to perform my analysis: the IRI Marketing Dataset[7] and the Consumer Population Survey (CPS) from the Bureau of Labour Statistics (BLS). From the IRI dataset I obtain information regarding prices, market shares, and products' characteristics, whilst from the CPS I obtain consumers' demographic characteristics.

The IRI dataset contains weekly store level data, for several Metropolitan Statistical Areas (MSAs) in the United States, and it provides prices, quantities, and some product characteristics. I aggregate this information to the quarterly level and define a market as the combination of MSA-quarter.

[7] see Bronnenberg et al. (2008) for a detailed description of this dataset.

I aggregate to the quarterly level for the following reasons. First, I need to observe products that were properly purchased during the period. Some products in the database are not purchased in a week or a month, but they all are purchased at least once in a quarter. This allows me to avoid the problem of zeroes in the market shares. Second, several costs components are determined for a period longer than a week or month. For example, marketing activities are planned at the quarterly level, and this is the same with some production decisions.[8]

Additionally, from all MSA available in the data, I focus on 19 of them. The IRI dataset does not contain information on household purchase data for all MSAs in the country, therefore I need to sample individuals from the CPS. In the CPS, some locations only have few individuals[9], therefore I choose the locations because that allow me to sample a greater number of individuals for my analysis.

I study the ready-to-eat cereal category. Cereal is the main breakfast food for most households in the United States. In 2019, 86.25% of the population consumed it.[10] It contains many vitamins and it is also a good source of fibres and iron, which makes it a great choice to start the day.[11] It comes in different flavours, and with options for both adults and kids.

An important characteristic of this industry is the degree of concentration and the number of brands (Nevo 2001). Although there are many different products, the industry is highly concentrated. The four main firms and grocery stores' own brands are responsible for almost the whole market share in the period under analysis (see table 1).

The raw dataset has information on 1937 different types of cereals defined by their Universal Product Code (UPC), which is the definition of product used in the present paper. I aggregate the weekly information to a quarterly level as follows: to obtain a quarterly price (measured in dollars per ounce), I average across all stores in the MSA

[8]See Nevo (2001) for the production process of ready-to-eat cereal.

[9]The smallest MSA at the IRI dataset only has 34 individuals in the CPS.

[10]Source:https://www.statista.com/statistics/279999/us-households-consumption-of-breakfa

[11]For a better description of the industry see Nevo (2001) and references therein. Notice that these benefits can be reduced depending on the quantity of sugar in the cereal.

Table 1: Market Share (volume) - main firms

Parent Company	Market Share (%)	Cum Market Share (%)
KELLOGG CO	34.50	34.50
GENERAL MILLS INC	30.18	64.68
PRIVATE LABEL	14.15	78.83
RALCORP HOLDINGS	11.48	90.30
PEPSICO INC	7 .63	97.93

Notes: Market share is based on the total volume, measured in ounces, sold by grocery stores. Private Label refers to stores' own brands.

and all weeks in the quarter.[12]

In my final analysis, I keep products that were available in at least one MSA for all quarters, and that had, at least, 0.001% market share in at least one market. In the end, I keep 1421 products, and they account for 35651 product-market observations, out of which 4200 are of cereal made with organic grains (11.78% of the total observations). Moreover, I only analyse cereals sold in boxes, because almost all observations belong to this category. From all the products, 76 correspond to cereal with organic grains; these are from nine national brands and from retail stores' own brands. Regarding the main brands, Kellogg's produces four organic cereals, General Mills produces 14, and the rest belong to smaller companies and retail brands.

An important definition in my analysis (see section 3 for more details) is the potential market of each product and I calculate it in a a way similar to Cohen (2008). I take the largest quarterly total sales from all products across the four quarters as the potential market size for that MSA. Then, the share of each product is calculated by dividing the total sales (in ounces) of each product by the potential market size.

Regarding the product's characteristics, besides price, my main interest is on a dummy variable indicating if a cereal contains organic grains or not. The other two main variables are related to advertising activities of each product. IRI provides weekly display and feature information at the UPC-store level. The variable Feature is defined by IRI as follows: no feature, small size ad, medium size ad, large size ad, and a retailer coupon or rebate. I classify this variable in a scale of 0-4. Display is defined as no display, minor, and major, and I transform it into a scale of 0-2.[13] Both variables are

[12] Prices are deflated by regional price deflators and measured in terms of 1982-84 dollars.

[13] This is the information provided in the dataset. The only extra information given about the definition of feature and display is that major display includes end of aisle and code lobby, and small ad is

aggregated to the quarterly level similarly to prices. Notice that, in the way I define these variables, they are representing the degree of in-store advertisement and describe the advertisement intensity (see Michel and Weiergraeber (2018) for the effects of advertisement intensity in prices).[14]

The summary statistics are presented on Table 2. They show some clear patterns. As expected, organic cereal is, on average, more expensive, and has a lower market share than regular cereal. Regarding the advertisement variables, regular products are much more advertised than organic ones.

Since organic products are usually more expensive, demographic characteristics might be correlated with its consumption. I do not observe individuals in my main dataset, therefore I use the CPS and randomly select 200 individuals in each MSA. The descriptive statistic of the (per capita) annual income are presented in Table 2.

Table 2: Descriptive Statistics

Statistic		Mean	Std.	Min	Max
Price ($/oz.)	Organic	0.138	0.039	0.050	0.300
	Regular	0.100	0.039	0.021	0.365
Size(oz.)	Organic	12.971	2.464	9.000	17.501
	Regular	15.188	4.184	5.000	64.000
Low Sugar	Organic	0.014	0.118	0.000	1.000
	Regular	0.089	0.285	0.000	1.000
Feature	Organic	0.153	0.272	0.000	2.000
	Regular	0.377	0.513	0.000	4.000
Display	Organic	0.065	0.150	0.000	1.576
	Regular	0.162	0.285	0.000	2.000
Share (%)	Organic	0.036	0.057	0.000	0.565
	Regular	0.208	0.320	0.000	3.449
Income ($)		11,470.90	11,871.52	15.25	221,992.22
Total Number of Observations			35,651		

usually one line of text.

[14]Although the original variables are categorical, I create a linear index to simply the computational burden in the estimation using this *ad hoc* scale. In future versions of this paper, I plan to reestimate the model using these variables in a qualitative approach.

2.1 Price difference between organic and non-organic products

Table 2 shows that organic products are more expensive, on average, than regular ones. However, organic products differ from their non-organic counterparts not only in price but also in some other characteristics. In order to see if the price difference is due to observable characteristics, I run the following regression (in level and log):[15]

$$price_{jqm} = \lambda_1 organic_j + \lambda_2 low\ sugar_j + \lambda_3 size_j + \lambda_4 feature_{jt} + \lambda_5 display_{jt} + \eta_{b(j)} + \kappa_q + \nu_m + \epsilon, \tag{1}$$

where organic is defined as above, subindexes j stands for the product and q for the quarter, m represents the MSA, η is a brand fixed effect, κ is a quarter fixed effect, ν is a MSA fixed effect, and ϵ is an error term.[16] Price is defined as price per ounce, size is the volume, in ounces, of the box, and low sugar is a dummy indicating that the cereal has low or zero sugar.

The results are presented in Table 3. They suggest that, even controlling for other observable characteristics, there is still a difference between the price of organic and non-organic products. However, this price premium could stem either from different production costs or from price discrimination.. In the next Section, I proceed to disentangle these two factors.

Regarding the other variables, as expected, Table 3 shows that more advertised products are, per ounce, cheaper (for both measures of advertisement used). Additionally, it shows that the bigger the box of cereal, the lower is its unit price. There is also a price difference between low sugar and sugary cereals.

[15] I use the observable characteristics available in the dataset.

[16] The results remain the same if instead of using MSA and quarter fixed effects additively, I use them as a multiplicative market fixed effect.

Table 3: Conditional price difference between organic and regular products

	Dependent Variables	
	Price	log(Price)
	(1)	(2)
Organic	0.024***	0.221***
	(0.003)	(0.019)
Low Sugar	0.010***	0.078***
	(0.001)	(0.008)
Size	−0.005***	−0.044***
	(0.0002)	(0.002)
Feature	−0.012***	−0.076***
	(0.001)	(0.015)
Display	−0.016***	−0.155***
	(0.002)	(0.016)
Observations	35,651	35,651
Adjusted R^2	0.793	0.7855

Notes: * denotes significance at 10%, ** denotes 5%, and *** denotes 1%. All regressions include brand, quarter, and MSA fixed effects. Standard errors are clustered at the market level. The dependent variable is measured as price per ounce.

3 Model

I use a discrete choice approach to analyse the possible existence of price discrimination. The choice for a structural model follows the arguments presented by Berry (1994). Although a reduced form approach is feasible to estimate demand elasticities, if there is a market with N products, then there are N^2 elasticities to estimate. A structural model allows the parametrisation of the consumer utility function and all those cross-price elasticities can be identified estimating fewer parameters. Moreover, a structural model allows to perform counterfactuals. Finally, in this setting, it is easy to go from statements about aggregate demand to statements about consumer utility, which eases the welfare analysis. I use the well known random coefficient discrete choice approach - BLP - such as in Berry et al. (1995) and Nevo (2001). My approach follows closely Cohen (2008), and an advantage of this type of model is that it implies a very general substitution pattern among different products, and it does not suffer from the Independence of Irrelevant Alternatives (IIA) critique that other types of discrete choice models do.

3.1 Demand

I observe $t = 1, \dots, T$ markets, each with $i = 1, \dots, I_t$ individuals. A market is defined as a quarter-MSA and, in each market, consumers have to choose among J products. Following the same notation as in Nevo (2000) and Nevo (2001), consumer's i utility of consuming product j in market t is expressed as

$$u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_j + \Delta\xi_{jt} + \varepsilon_{ijt}, \tag{2}$$

where x_{jt} is a K-dimensional vector of observable characteristics of product j in market t, including the organic feature of this product. In the same way, p_{jt} is the price of product j in market t, ξ_j is the mean of the unobserved (to the econometrician) product characteristics, $\Delta\xi_{jt}$ is the market specific deviation from this mean, and ε_{ijt} is a mean zero stochastic term, which, as usual, is assumed to follow a type I extreme-value distribution.

The individual specific coefficients, α_i and β_i, are

$$\begin{pmatrix} \beta_i \\ \alpha_i \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix} + \Pi D_i + \Sigma \nu_i, \tag{3}$$

where α and β are the mean taste parameters, D_i is a $d \times 1$ vector of observable demographic variables in market t, Π is a $(K+1) \times d$ matrix of coefficients that measure how the consumer taste characteristics vary with demographics characteristics, ν_i is a $K \times 1$ vector representing the unobserved random part of consumer taste, which is assumed to be $\nu_i \sim N(0, I_{k+1})$, and Σ is a $K \times K$ scaling matrix.

The utility from the outside good is

$$u_{i0t} = \xi_0 + \pi_0 D_i + \sigma_0 \nu_{i0} + \varepsilon_{i0t}. \tag{4}$$

As usual, since the mean utility of the outside good is not identified, ξ_0 is normalised to zero.[17]

Let $\theta = (\theta_1, \theta_2)$ be the vector that contains all the parameters of the model. The vector $\theta_1 = (\alpha, \beta)$ contains the linear parameters, and the vector $\theta_2 = (\Pi, \Sigma, \pi_0, \sigma_0)$ contains the non-linear parameters.[18] I can express equation (2) as a sum of mean utility, δ_{jt}, and a mean zero heteroskedastic deviation from the mean that captures the effects of random coefficients, $\mu_{ijt} + \varepsilon_{ijt}$. It can be expressed as

$$u_{ijt} = \delta_{jt}(x_{jt}, p_{jt}, \xi_j, \Delta\xi_{jt}; \theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, \nu_i, D_i; \theta_2) + \varepsilon_{ijt}, \tag{5}$$

where $\delta_{jt} = x_{jt}\beta_i - \alpha p_{jt} + \xi_j + \Delta\xi_{jt}$, and $\mu_{ijt} = [p_{jt}, x_{jt}]' * (\Pi D_i + \Sigma \nu_i)$, where $[p_{jt}, x_{jt}]$ is a $(K+1) \times 1$ vector.

Consumers purchase only the good that gives them the highest utility. To obtain the

[17]Notice that π_0 and σ_0 are not identified separately from the coefficients of an individual-specific constant term in equation (2), so I also normalise them to zero.

[18]I can separate this way because the model generates a non-linear system of the market shares. Note that the non-observed products' characteristics enter the market share equation in a non-linear way.

market shares of each j-th product, I have to sum up over the mass of consumers in each market, and they are given by the following expression:

$$s_{jt}(\delta_t, \theta_2) = \int_{D_i, v_i} \frac{exp\{\delta_{jt} + \mu_{ijt}(D_i, v_i; \theta_2)\}}{\sum_{k=0}^{J} exp\{\delta_{kt} + \mu_{ikt}(D_i, v_i; \theta_2)\}} dF(D_i, \nu_i; \theta_2). \tag{6}$$

From these results, I obtain the own and cross-price elasticities, which are given by the following expression:

$$e_{jkt} = \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}}, \tag{7}$$

where s_{jt} is defined as above.

Due to the multidimensional integral in equation (6), an equation that is linear in the parameters cannot be obtained. Berry et al. (1995) proved that, given the parameter values and the observed market shares, this equation system is a contraction mapping in δ_t, and hence a unique vector can be obtained using iterations.

Notice that ξ_j, the non-observed product' characteristics, are related with prices, which generates the problem of simultaneity. Therefore, it is necessary to use instruments to estimate the demand system (in Section 4.3 I describe the instruments used.)

In Section 4 I provide more details on the estimation procedure and in the Annex I describe the algorithm used.

3.2 Supply

In order to recover the margins, I assume, as in Nevo (2001), that firms engage in differentiated products Bertrand-Nash competition. The profit maximisation problem for firm f in market t is as follows:

$$Max\ \Pi_{f_t} = \sum_{j \in F_{f_t}} (p_{jt} - mc_{jt}) M_t s_{jt}(\mathbf{p_t}), \tag{8}$$

where F_{f_t} is the subset of products that are produced by firm f in market t, which varies by firm and market, M_t is the potential size of market t, and $\mathbf{p_t}$ is a vector containing prices for all products in market t.

The first order condition is given by:

$$s_{jt}(\mathbf{p}_t) + \sum_{r \in F_f} (p_{rt} - mc_{rt}) \frac{\partial s_{rt}(\mathbf{p_t})}{\partial p_{jt}} = 0, \; j = 1, \ldots, J_t. \tag{9}$$

Let us define $S_{jrt} = -\frac{\partial s_{rt}(\mathbf{p_t})}{\partial p_{jt}}$, $j, r = 1, \ldots, J_t$. Additionally, let $\omega^*_{jrt} = 1$ if j and r are produced by the same firm, and $\omega^*_{jrt} = 0$ otherwise. ω^*_t is a $J_t \times J_t$ matrix with $\omega_{jrt} = \omega^*_{jrt} * S_{jrt}$. Thus, the set of the J_t first order conditions can be written as the following vector:

$$s_t(\mathbf{p_t}) - \omega_{\mathbf{t}}(\mathbf{p_t} - \mathbf{mc_t}) = \mathbf{0}. \tag{10}$$

From this equation, I can calculate the margins, $(p_t - mc_t)$, and from them I can recover the marginal cost, which is

$$\mathbf{mc}_t = \mathbf{p}_t - \omega_t^{-1} * \mathbf{s}_t(\mathbf{p}_t). \tag{11}$$

3.3 Measuring price discrimination

With the margins obtained before, I can investigate if firms price discriminate with respect to the organic feature of the products. There are different ways to measure price discrimination.[19]

One possibility, similar to Cohen (2008), is to use the ratio between the difference in markups between organic and non-organic products and the difference in unit prices between them. However, a caveat of this approach is that it measures price discrimination with respect to all characteristics of each variety of cereal. For example, if firms price discriminate with respect to both size and organic characteristics of the cereal, I would not be separating the contribution of each attribute.

[19]See Clerides (2004).

Therefore, I proceed in a different way. First, I measure price discrimination as the difference between the average margin for organic and non-organic products across all markets. The following equation express this difference:

$$PD_{All} = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{ot}}\sum_{j=1}^{J_{ot}}(p_{jt} - mc_{jt}) - \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{not}}\sum_{j=1}^{J_{not}}(p_{jt} - mc_{jt}), \tag{12}$$

where J_{ot} and J_{not} are, respectively, the number of organic and non-organic products in market t. This expression measures price discrimination with respect to all characteristics of organic products.

In order to isolate price discrimination with respect to the organic attribute, I conduct the following counterfactual experiment. First, I remove the organic attribute of each product from all organic products in my data. The assumption behind this experiment is that the organic characteristic is independent of other characteristics, except the price.

However, as discussed in the Introduction, organic products usually have higher production costs. Hence, this characteristic is an important determinant of the marginal cost and, thus, I need to adjust these marginal costs. I estimate the following regression:

$$mc_{jt} = \lambda_0 + \lambda_1 o_j + \gamma_{jt} + \nu_{jt}, \tag{13}$$

where γ_{jt} contains other variables that are likely affecting the marginal cost (size, sugar content, MSA dummy, quarter dummy). The variable o_j is a dummy that has value one if a product is organic. The parameter of interest is λ_1, which gives an estimate for the part of the marginal cost which is due to the organic characteristic. Therefore, the adjusted marginal cost is $mc_{jt}^{ex} = mc_{jt} - \hat{\lambda}_1$.

Now, without the organic attribute and the new marginal costs, I can solve again the first order condition and find the new set of equilibrium prices ($\mathbf{p_t^{ex}}$). Then, I can re-calculate the difference between the average margins for organic and non-organic products. Equation (12) becomes:

$$PD_{Restricted} = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{ot}}\sum_{j=1}^{J_{ot}}(p_{jt}^{ex} - mc_{jt}^{ex}) - \frac{1}{T}\sum_{t=1}^{T}\frac{1}{J_{not}}\sum_{j=1}^{J_{not}}(p_{jt}^{ex} - mc_{jt}). \tag{14}$$

Notice that, since non-organic products have not lost an attribute, their marginal cost have not changed. This equation can be interpreted as the measure of price discrimination due to other characteristics, except the organic one, of each product. PD_{All} measures price discrimination taking all characteristics into consideration. Therefore, price discrimination with respect to the organic attribute is given by $PD = PD_{All} - PD_{Restricted}$.

An implicit assumption is that the difference marginal costs between organic and non-organic product is constant across products and markets. This is a plausible assumption due to the production process of cereal. The production of different brands, including competitors, usually occurs at the same plant. See Nevo (2000) for a complete description of the production process.

4 Estimation

4.1 Estimation Details

The key issue in the demand model is to estimate equation (6). This is a system of J_t equations and J_t unknowns. where J_t is the number of products in each market. The insight of Berry et al. (1995) is to perform a non-linear change of variables, i.e. $\delta_t \equiv D_t^{-1}(S_t, \theta_2)$, where S_t is the vector of observed market shares.

I use simulations to solve equation (6). In practice this equation becomes:

$$s_{jt} = \frac{1}{I_t}\sum_{i=1}^{I_t}\frac{exp\{\delta_{jt} + \mu_{ijt}\}}{\sum_{k=0}^{J_t} exp\{\delta_{kt} + \mu_{ikt}\}}, \tag{15}$$

where I_t is the number of simulations, and $\mu_{ijt} = [p_{jt}, x_{jt}]' * (\Pi D_i + \Sigma \nu_i)$, D_i is the i-th draw of the observed consumers characteristics from the CPS, at market t, and ν_{it} is the i-th random draw of the unobserved consumer characteristics variables, ν_i.

In order to perform this simulation, I assume that D_i and and ν_i are independent and I use Halton sequences to simulate the integral.

There are other issues that must be resolved to perform the estimation.

First, matrix Σ is restricted to be diagonal, which means that the unobserved consumer preference for different product characteristics are independent from one another.[20]

D_i is specified to include only one variable: income. I use only this demographic characteristic because it does not capture only heterogeneity in purchase power, but also other characteristics. It is likely that richer households are also more educated, and therefore, have more information and/or are environmental concerned. Therefore, they should be willing to pay more for organic food.

Regarding the observed products' characteristics, I use the following variables: price, Display, and Feature. Note that these variables vary per market. I also add quarter dummies.[21]

Finally, I do not estimate all parameters in the coefficient matrix for product and household characteristics interaction, Π.Specifically, Feature and Display, like the advertisement variable in Nevo (2000), do not present a random coefficient and are not interacted with the households' characteristic.

However, there are still non-observable characteristics of products that do not vary per market. Also, ξ_j in equation (2) is not observed in the data. Therefore, I include a set of product dummies to replace $x_j\beta + \xi_j$ in the following way:

$$Product_j = x_j\beta + \xi_j. \tag{16}$$

This strategy is similar to the minimum-distance procedure used by Nevo (2000). In

[20] I also estimated the model with covariances terms included. The estimated elasticities are similar to the ones obtained in the present paper and are available upon request.

[21] As Nevo (2000) I do not add MSA dummies. The market is well represented by the income distribution from the demographic variable.

order to retrieve the taste coefficients, β, I regress the estimated product effects, from the GMM estimation, on the observable characteristics.

4.2 Identification

I estimate the parameters of the demand system presented in Section 3 by exploiting a population moment condition and a structural error term to form a non-linear GMM estimator. Let $Z = [z_1, \ldots, z_M]$ be a set of instruments such that $E[Z'\dot{\phi}(\theta^*)] = 0$, where θ, a function of the model parameters, is an error term defined below and θ^* denotes the true value of these parameters. The GMM estimate is

$$\hat{\theta} = \text{arg min}_{\theta}\phi(\theta)ZA^{-1}Z'\phi(\theta), \tag{17}$$

where A is a consistent estimate of $E[Z'\phi\phi'Z]$.

4.3 Instruments

I need a set of exogenous instrumental variables to properly estimate the demand system. Since I have different products in different markets, I follow the literature (e.g. Berry et al. 1995; Bresnahan et al. 1997; Sudhir 2001; Petrin 2002) and use the well known BLP instruments. This set of instrumental variables contains variables capturing the average product characteristics of the UPCs in the same market, and the number of products, by firm ownership and product category, in the same market. The main idea behind this type of instrument is that firms make pricing decisions based on the number of competing products they face in each market, and how similar these products are to theirs.

Another set of instruments is constructed following Hausman (1994) and Nevo (2001). It is the average price of the same UPC in other markets. These are correlated to the price variable, but not to the error term under the assumption that the unobserved UPC-market specific demand shifting factors are independent across markets. If this assumption does not hold then this instrument is not valid. For instance, national advertising campaigns can shift demand in different markets at the same time. To lessen this concern, I control for advertising at the UPC-market level.

Whether or not a good faces close substitutes is an important determinant of its demand. Thus, a third set of instruments is based on how similar a product is to its competitors. These instruments were proposed by Gandhi and Houde (2019) and they are based on exogenous measures of product differentiation. These instruments exploit products' relative isolation in the characteristic space and they perform well to identify heterogeneous consumer preferences.

As a last set of instruments, I also follow Nevo (2001) and use supply shifters. I use regional dummies and population density to control for the land cost of retailing, and also hourly wage in the supermarket sector to control for the labour cost of retailing. Since these are supply shifters, they are not likely to be correlated with demand shifters.

Finally, I use optimal instruments, as proposed by Reynaert and Verboven (2014). They showed that the use of Chamberlain (1987) optimal instruments increases the estimator's efficiency and stability in a BLP framework, especially allowing for a more precise estimation of the standard deviations of the random coefficients.

5 Results

First, notice that, unlike in subsection 2.1, I do not use the variables size and low sugar in the main model. They are characteristics that matter for the production cost, but I have no evidence of its relevance for consumers. This happens for two reasons. First, almost all products have a box size of 16oz, with some products being outliers with respect to this characteristic. These outliers represent less than 2% of total sales, thus there are not many options available, and consequently I do not assume that consumers, on average, have heterogenous valuation about the size of the box when deciding to purchase. Second, the coefficient of the variable "low sugar" was not statistically different from zero, and including it or not does not have a significant impact on the estimated price elasticity.

Table 4 presents the results from the GMM estimation. First, notice that all mean parameters are statistically significant and have the expected sign. The price coefficient is negative and both variables regarding advertisement activity (display and feature) have a positive sign, suggesting a positive effect of advertisement on demand. The coefficient of the organic characteristic is negative, as expected, since organic cereal has a lower share when compared to regular products. Notice, however, that this result does not mean that households do not like the organic attribute. The total mean effect of this characteristic must be combined with the interaction of this characteristic and income. After this operation it does have a positive sign, meaning that above a certain income threshold consumers start to value this characteristic.

Regarding the standard deviation, matrix Σ, the coefficients are statistically significant. This suggests that there is a non-observed heterogeneity regarding preferences for organic cereal.

Finally, the coefficients of the interaction between income and product characteristic are statistically significant. Consumers with above-average income are less price sensitive than below-average income ones. Moreover, higher income individuals prefer cereal with organic grains. This was an expected result since richer households usually care more about healthier and environmentally friendlier options.

Using the estimates on the utility parameters, I then calculated the own-price elasticity for each UPC. I plot the mean of these elasticities, in each market, in Figure 1a. Overall, the results are in line with the elasticities estimated in Nevo (2001) and Michel and Weiergraeber (2018), for example, and they imply the individuals are quite price sensitive.[22] In Figure 1b, I present the aggregate elasticity of demand in each market. Aggregate elasticities reflect the change of total sales under a proportional sales tax. The results imply that demand is rather inelastic to a a proportional sales tax, and this will be clear with the simulation performed in section 6.

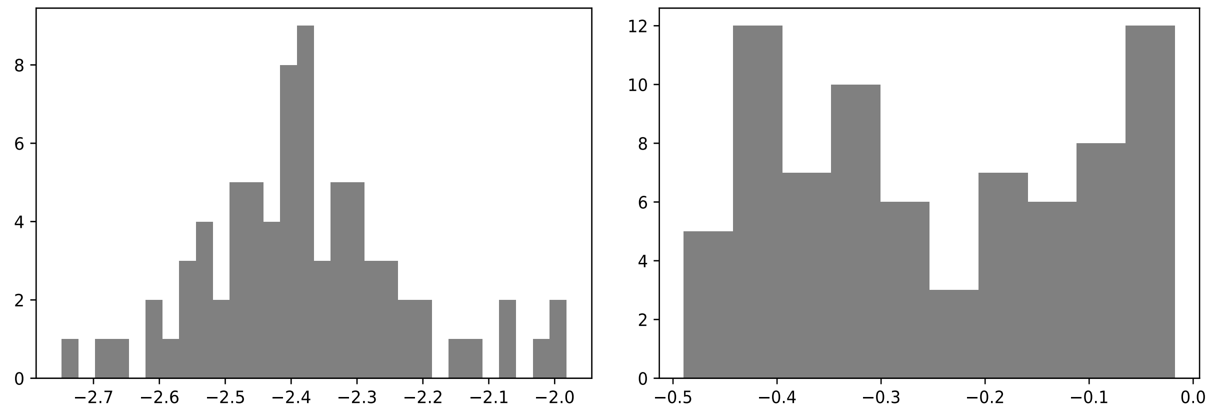
An important point to highlight is the cross-price elasticity between organic and regular products. These values are smaller than the cross-price elasticities within each

[22]Note that Michel and Weiergraeber (2018) estimate a nested logit model.

Table 4: Random Coefficients Logit Estimation Results

Variable	Mean	Std.Dev	Income
Price	-30.476^{***}	7.728^{***}	2.886^{***}
	(1.578)	(1.927)	(0.359)
Organic	-0.837^{a***}	5.737^{***}	1.053^{**}
	(0.016)	(0.524)	(0.358)
Feature	0.231^{***}	–	–
	(0.026)	–	–
Display	0.887^{***}	–	–
	(0.048)	–	–
Constant	0.649^{a***}	0.275^{*}	0.746^{**}
	(0.126)	(0.163)	(0.334)

Notes: 1) * denotes significance at 10%, ** denotes 5%, and *** denotes 1%.
2) Quarter and MSA fixed effects were included in the estimation.
3) "a" denotes estimates from the minimum distance procedure using the estimated UPC dummy coefficients from the GMM estimation.



(a) Mean own-price elasticities **(b)** Aggregate Demand Elasticities

Figure 1: Demand Elasticity

category. Almost all cross-price elasticities between a regular product and an organic product are close to zero. This implies that, if the price of a regular product increases, consumers prefer to switch to another regular product instead of to an organic product. The same is true for consumers of organic products.

I also calculated the diversion ratio of each product with respect to the outside good. A diversion ratio gives us how much the consumer is willing to switch to the outside good when the price of a product increases. This ratio is calculated in the following way:

$$D_{jj} = -\frac{\frac{\partial s_{0t}}{\partial x_{jt}}}{\frac{\partial s_{jt}}{\partial x_{jt}}}. \tag{18}$$

The ratios computed in the present paper were close to zero. This implies that consumers hardly switch to the outside good. This will be clear in section 6, where I simulate the effects of a tax on regular products. The matrix of diversion ratios and of all cross-price elasticities are available upon request.

Finally, I use the elasticities to calculate the marginal costs and the margins for all UPCs in each market. These results are summarised in table 5. First, as expected, the average marginal cost and average margin are higher for organic products. Also, notice that around 1% of the estimated marginal costs are negative. This situation usually occurs when the product had a high level of promotion activities during the quarter. However, these values can mislead the estimates of counterfactual prices, thus products with negative estimated marginal costs were excluded from the experiment I conduct to measure price discrimination.

Table 5: Marginal Cost and Margin Estimates

	Marginal cost ($)		Margin ($)	
	Organic	Regular	Organic	Regular
Mean	0.076	0.051	0.056	0.047
Std.Dev	0.033	0.031	0.006	0.006
Min	−0.006	−0.032	0.039	0.038
Max	0.204	0.276	0.078	0.062

Note: 1.15% of the estimated marginal costs were negative

Following the procedures described in section 3.3, I measured price discrimination and the results are presented in table 6. The price difference between organic and non-organic products is 0.0374 dollars. Out of this price difference, 0.0071 dollars (18.99%) comes from price discrimination, whilst the rest comes from difference in their production costs. When I isolate the price discrimination that comes exclusively from the organic attribute, I obtain that it corresponds to 0.0023 dollars, which is 6.14% of the price difference.

Table 6: Price discrimination with respect to the organic attribute

Average Price Difference ($)	0.0374
$PD_{All}(\$)$	0.0071
$PD_{Restricted}(\$)$	0.0048
$PD(\$)$	0.0023

Note: Products with negative marginal cost were not included in the experiment.

6 Policy Experiment and Robustness

6.1 Tax Policy

Organic products are environmentally friendlier. Hence, discouraging price discrimination may help to increase the consumption of organic cereal. Nevertheless, notice that even without price discrimination the government has incentives to increase if it is concerned about the environment.

My policy experiment consists of imposing a tax on regular products. I simulate the counterfactual for a 5% tax per ounce, on all regular products.[23]

If the government imposes a tax on non-organic products, firm f's profit function in market t becomes:

$$\Pi_{f_t} = \sum_{j \in F_{f_t}} \left[(p_{jt}^{cf} - mc_{jt}) \times M_t \times s_{jt}(\mathbf{p}_t^{cf} + \psi) \right], \tag{19}$$

where ψ is the tax imposed on product j if it is a non-organic product, and 0 otherwise. This tax creates a wedge between the price consumers pay and the one firms receive. In this exercise, I present the results for Charlotte, North Carolina, in the first quarter of 2011. The reason is twofold. First, by focusing on a selected market, I can discuss the results in a clearer way, which is not the case when I stack the results for all markets. Second, all marginal costs calculated for this market were positive, thus there

[23] Results for a tax of 1% or 15% were also simulated and are available upon request. They are qualitatively similar to the one presented in this paper.

is no bias, due to negative marginal costs, in the estimated conterfactual. I calculated the same results for other markets and a summary for all markets are available upon request.

Results are collected in table 7. The price of organic products remain, on average, unchanged. On the other hand, the prices of the regular products increase even more than the values of the tax. A possible explanation is that the aggregate elasticities calculated earlier showed that total sales are quite price-inelastic with a change in taxes, therefore firms take advantage that all firms have to increase prices and try to exploit this scenario to increase their revenue.

However, the shares of organic and regular products remain, on average, the same. This result implies two conclusions. First, consumers do not switch to the outside option, which was expected given the diversion ratio calculated in the previous section. The second conclusion is that, even though organic products become relatively less expensive due to the tax levied on regular products, it is not enough to induce consumers to switch from regular to organic cereal. According to the survey data[24], consumers are loyal to the brands they usually buy. Therefore, it would be necessary a huge change in the relative prices to induce a change in the consumers' purchase pattern.

It is important to discuss some implications of these results. First, it is clear that, on average, consumers keep buying the same product category. However, this does not mean that consumers are not changing the cereal brand they are buying. As discussed in section 5, consumers are price elastic and when they change their consumption, it happens from a regular cereal to a cheaper version within the regular cereal category.

Second, from a public policy point of view, the results suggest that taxes might not be the best instrument to induce the consumption of cereal with organic grains. This result coincides with most of the literature on this type of taxation, such as Wang (2015): taxes might not be the most effective way to change consumers' eating habits.

[24]The NDP Group/National Eating Trends, February 2018.

Next, I calculate the welfare change due to the tax. The expected change in consumer surplus, for a representative consumer, in market t, is given by the following expression[25]:

$$E[E(CS_t^{cf}) - E(CS_t)] = E\left(\frac{1}{\alpha_{itp}} \left[\ln \left(\sum_{j=1}^{J_t} exp(\delta_{jt}(\mathbf{p}_t^{cf}) + \mu_{ijt}(\mathbf{p}_t^{cf})) \right) - \ln \left(\sum_{j=1}^{J_t} exp(\delta_{jt}(\mathbf{p}_t) + \mu_{ijt}(\mathbf{p}_t)) \right) \right] \right), \tag{20}$$

where the superscript cf gives us the values of the counterfactual exercise.

The results are presented in Table 8. The first row shows that the average consumer loss, measured in dollars per ounce, is $-7.816425 \times 10^{-05}$. Since the potential market size for this specific market is 14,553,190.3 ounces, the total loss is around 1137.54 dollars. This result is easy to understand. Consumers are paying more for cereal, buying on average the same quantity, and not switching to other types of cereal. Hence, there is a decrease in the consumer surplus.

On the other hand, firms increase their profit. The reason is that consumers are substituting within the same product category. Therefore, even if they switch to a cheaper cereal, this product is also more expensive than in the scenario without the tax. In this specific market, the difference in profits is of 430.63 dollars. Therefore, the total social welfare decreases 706.91 dollars.

Table 7: Summary statistics of counterfactual for experiment: tax of 5% on regular products in market 4

	Organic				Regular			
	p	p^{cf}	share	$share^{cf}$	p	p^{cf}	share	$share^{cf}$
Mean	0.15107	0.15107	0.02732	0.02732	0.10042	0.11883	0.25516	0.25516
Std.	0.04222	0.04222	0.05299	0.05299	0.03686	0.26871	0.35520	0.35526
Min	0.08334	0.08334	0.00014	0.00014	0.04134	0.04134	0.00013	0.00001
Max	0.26962	0.26962	0.27177	0.27177	0.26913	0.79794	2.17919	2.17938

Notes: There are 427 UPC-market observations in this market. Share, and its counterfactual, are measured in percent. The prices variables are measured in dollars per ounce.

For completeness, I compute the price discrimination measure using these conterfactual prices. As expected, there is a marginal reduction in price discrimination regard-

[25] The expectations can be computed using simulations.

Table 8: Summary Statistics of Welfare for experiment: tax of 5% on regular products

	Mean
Δ CS ($ per person per ounce)	$-7.816425 \times 10^{-05}$
Δ CS ($)	-1137.54
Δ Profit ($)	430.63
Δ Welfare ($)	-706.91

ing the organic feature. This occurs because the price of organic products remains on average the same, whilst there is a slightly increase in the price of non-organic products. The reduction is much smaller than the value of the tax, and this result, associated with the loss on social welfare, shows that taxation is not an effective measure to fight price discrimination.

Table 9: Price discrimination with respect to the organic attribute with tax on regular products in market 4

Original Price Discrimination (%)	3.752
Counterfactual Price Discrimination (%)	2.911

Finally, it is also important to discuss if the tax revenues collected with this policy could be used to subsidise organic cereal. However, since the tax does not substantially reduce price discrimination, it is likely that it would result in a regressive policy. This is the case because the estimates presented in table 4 suggest that organic cereal is preferred by higher income households. Therefore, if the government uses this tax revenues to subsidise organic cereal, then it would be *the factor* taxing poorer households to subsidise richer ones.

6.2 Consumer heterogeneity as a source of price discrimination

After finding that price discrimination occurs in the cereal market, a natural question to address would be to analyse who is more affected by this practice. The results obtained in section 5 suggest that higher income households prefer organic products. Therefore, this source of consumer heterogeneity is a possible source of price discrimination. The introduction of organic products can even happen to satisfy these households' preferences. Hence, in this subsection, I analyse to what extent the existence of price discrimination occurs due to consumer heterogeneity.

I perform a counterfactual experiment setting the income of all households to the mean income in each market, i.e uniform income in all markets. The price discrimination result for this experiment is presented on Table 10. The result suggests that price discrimination with respect to the organic feature is -0.000196 cents, that is, without high-income households, the difference in average margins between the organic and non-organic cereal actually decreases -0.000196 cents when the organic feature is present. This would imply a price discrimination of 0.5% with respect to the non-organic cereal, which is basically equivalent to say that there is no price discrimination.

Price discrimination disappears in this scenario because because with no high income consumers, the sales of organic cereal decrease, which leads the firms not to engage in price discrimination regarding the organic feature. Income heterogeneity, as a result, is a significant source of price discrimination with respect to the organic attribute.

Finally, the effect of this counterfactual on prices and shares are presented on Table 11. On average, prices for the organic products decrease, whilst prices for regular products remain the same. This is the case because there are no longer high-income people. Also, the share of regular products increase, and this increase is similar, in magnitude, to the decrease in the share of organic products. This implies that those households, that previously purchased organic cereal, switch to the regular variety of cereal instead of switching to the outside good.

Table 10: Price discrimination with respect to the organic attribute

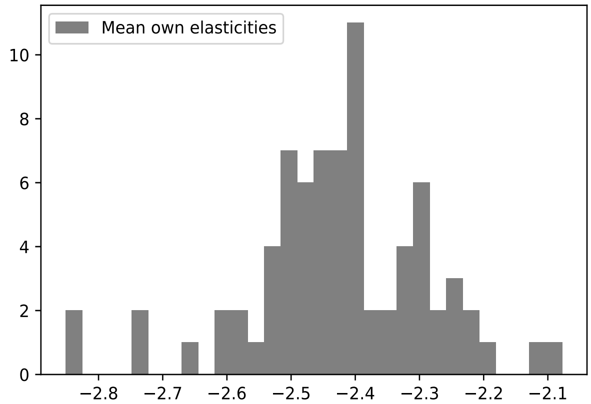
$PD_{All}(\$)$	-0.0051
$PD_{Restricted}(\$)$	-0.0049
$PD(\$)$	-0.000196

Note: Products with negative marginal cost were not included in the experiment.

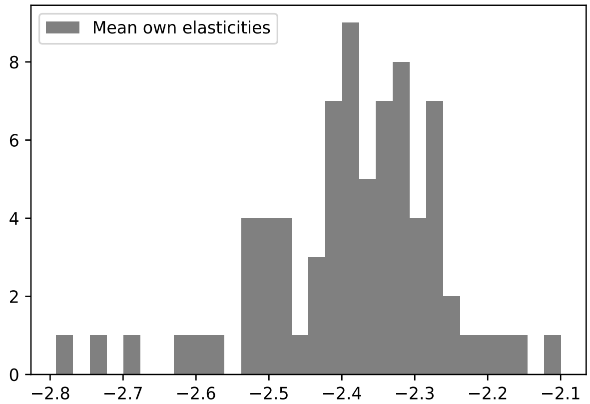
Table 11: Summary statistics of counterfactual for experiment: uniform income

	Organic				Regular			
	p	p^{cf}	share (%)	$share^{cf}$ (%)	p	p^{cf}	share	$share^{cf}$
Mean	0.138	0.0861	0.036	0.011	0.100	0.1001	0.208	0.2423
Std.	0.039	0.030	0.057	0.031	0.039	0.039	0.320	0.341

Notes: Products with negative marginal cost were not included in the experiment. The prices variables are measured in dollars per ounce.



(a) Mean own-price elasticities with market size changed by a factor of 2



(b) Mean own-price elasticities with market size changed by a factor of 3

Figure 2: Demand elasticities with different measures of market size

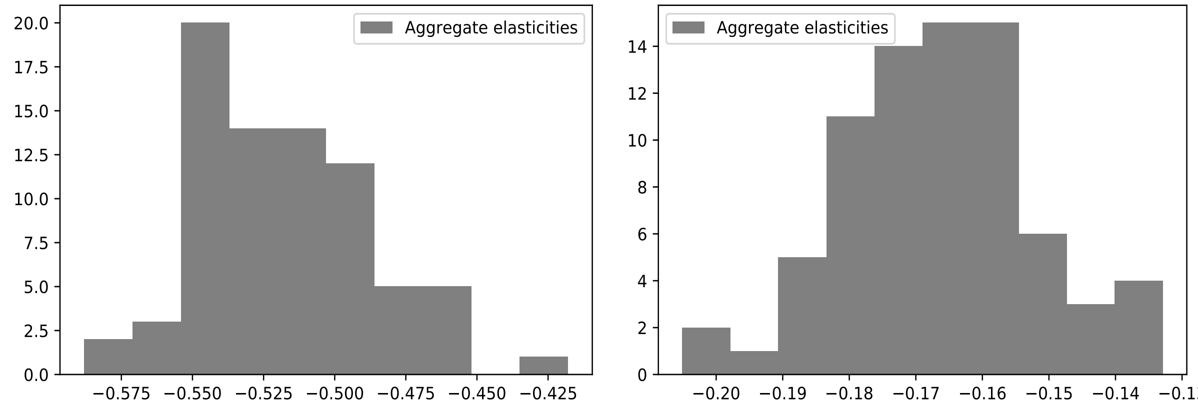
6.3 Robustness

An important definition in the present paper is how to measure the market size. This measure determines the shares and, consequently, the elasticities, which are the values that drive the calculation of the marginal costs and margins. In this subsection, I briefly present evidence that changing the market size does not generate substantial changes in the elasticities. The new market sizes are calculated by changing the original ones by factors of 2 and 3.

Regarding the mean own-price elasticities, the new definitions give very similar results, as shown in Figure 2. The aggregated elasticities (Figure 3) also point to the same direction: although there is a small change in the magnitude of these elasticities, they are still below -1 and suggest that consumers are inelastic with respect to taxation.

7 Conclusion

In this paper, I quantify price discrimination regarding the organic attribute in the ready-to-eat cereal market. In order to do it, I build a random coefficient discrete choice model of demand to obtain the price elasticity of each cereal product in my data. With the estimated elasticities, and a supply model, I recover the margins and the marginal costs. Then, I calculate how much of the price difference between organic



(a) Aggregate demand elasticities with market size changed by a factor of 2 **(b)** Aggregate demand elasticities with market size changed by a factor of 3

Figure 3: Aggregated demand elasticities with different measures of market size

and non-organic products is due to cost differences or due to price discrimination. I find that around 6% of the price difference is due to price discrimination with respect to the organic attribute.

I find that the marginal cost of organic cereal is, even in the absence of price discrimination, higher than for regular cereal. Even though price discrimination is an issue in this market, government policies might not have an effect in changing consumers' behaviour toward the organic options. A counterfactual exercise indeed shows that this is the case in the cereal industry and confirms the findings of the previous literature.

I analyse the effect of a 5% tax on non-organic products. There is virtually no change in consumption from regular options towards organic ones, although consumers substitute products within the regular category. Furthermore, consumers do not switch to the outside option, and thus they keep consuming cereal. Therefore, consumer surplus reduces since consumers are, on average, paying more for the most consumed cereal. Hence, firms increase revenues and profits. Nevertheless, social welfare decreases due to the decrease in consumer surplus.

Even imposing a tax on non-organic cereal, the price difference between organic and regular products is still too wide. Furthermore, the diversion ratio and the cross-price elasticity between organic and regular products are small. Therefore, taxes are not enough to change the consumers' purchase pattern. Price is not the variable that policy makers should focus on when analysing the incentives to change from regular to organic products in the cereal market. Other non-monetary policies are necessary to change the consumers' purchase pattern.

I also analyse a counterfactual where all households have the same income. In this scenario, price discrimination disappears, which suggests that one of the sources of price discrimination is the existence of high-income households in the market.

The results of this paper are subject to some caveats. First, no product level cost data are available to evaluate the accuracy of the recovered margins. Also, my analysis assumed a passive role of grocery stores in the pricing decision. This implicitly assumes that retailers' cost and margins are constant. Although a common used assumption, it could be the case that the bargaining power is on the part of retailers, and not the manufacturers (Villas-Boas 2007, 2009; Bonnet and Dubois 2010). In order to analyse this issue, one needs more data on the supply side. This might be a potential avenue for future research

References

Akgüngör, S., Miran, B. and Abay, C. (2010), 'Consumer willingness to pay for organic food in urban Turkey', *Journal of International Food and Agribusiness Marketing* **22**(3), 299–313.

Batte, M. T., Hooker, N. H., Haab, T. C. and Beaverson, J. (2007), 'Putting their money where their mouths are: Consumer willingness to pay for multi-ingredient, processed organic food products', *Food Policy* **32**(2), 145–159.

Berry, S. (1994), 'Estimating Discrete-Choice Models of Product Differentiation', *The RAND Journal of Economics* **25**(2), 242–262.

Berry, S., Levinsohn, J. and Pakes, A. (1995), 'Automobile Prices in Market Equilibrium', *Econometrica* **63**(4), 841–890.

Bonanno, A. (2016), 'A Hedonic Valuation of Health and Nonhealth Attributes in the U.S. Yogurt Market', *Agribusiness* **32**(3), 299–313.

Bonnet, Ć. and Dubois, P. (2010), 'Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance', *RAND Journal of Economics* **41**(1), 139–164.

Bresnahan, T. F., Stern, S. and Trajtenberg, M. (1997), 'Market Segmentation and the Sources of Rents from Innovation : Personal Computers in the Late 1980', *RAND Journal of Economics* **28**(0), 17–44.

Bronnenberg, B. J., Kruger, M. W. and Mela, C. F. (2008), 'The IRI marketing data set', *Marketing Science* **27**(4), 745–748.

Chamberlain, G. (1987), 'Asymptotic efficiency in estimation with conditional moment restrictions', *Journal of Econometrics* **34**(3), 305–334.

Clerides, S. K. (2004), 'Price discrimination with differentiated products: Definition and identification', *Economic Inquiry* **42**(3), 402–412.

Cohen, A. (2008), 'Package size and price discrimination in the paper towel market', *International Journal of Industrial Organization* **26**(2), 502–516.

Dubois, P., Griffith, R. and O'Connell, M. (2020), 'How well targeted are soda taxes?', *American Economic Review* (Forthcoming).

Flores, M. and Rivas, J. (2017), 'Cash Incentives and Unhealthy Food Consumption', *Bulletin of Economic Research* **69**(1), 42–56.

Gandhi, A. and Houde, J.-F. (2019), 'Measuring Substitution Patterns in Differentiated Products Industries', *Working Paper 26375. National Bureau of Economic Research* .

Hausman, J. (1994), 'Valuation of new goods under perfect and imperfect competition', *Working Paper 4970. National Bureau of Economic Research* .

Huang, L. and Liu, Y. (2017), 'Health information and consumer learning in the bottled water market', *International Journal of Industrial Organization* **55**, 1–24.

Jaenick, E. and Carlson, A. C. (2015), 'Estimating and Investigating Organic Premiums for Retail-Level Food Products Edward', *Agribusiness* **31**(4), 453–471.

Lan, H. and Dobson, P. W. (2017), 'Healthy Competition to Support Healthy Eating? An Investigation of Fruit and Vegetable Pricing in UK Supermarkets', *Journal of Agricultural Economics* **68**(3), 881–900.

Leslie, P. (2004), 'Price Discrimination in Broadway Theater', *The RAND Journal of Economics* **35**(3), 520.

Mayen, C. D., Balagtas, J. V. and Alexander, C. E. (2010), 'Technology adoption and technical efficiency: Organic and conventional dairy farms in the United States', *American Journal of Agricultural Economics* **92**(1), 181–195.

Michel, C. and Weiergraeber, S. (2018), 'Estimating Industry Conduct in Differentiated Products Markets: The Evolution of Pricing Behaviour in the RTE Cereal Industry', *Working Paper* pp. 1–77.

Miller, N. H. and Osborne, M. (2014), 'Spatial differentiation and price discrimination in the cement industry : evidence from a structural model', *RAND Journal of Economics* **45**(2), 221–247.

Nevo, A. (2000), 'A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand', *Journal of Economics & Management Strategy* **9**(4), 513–548.

Nevo, A. (2001), 'Measuring market power in the ready-to-eat cereal industry', *Econometrica* **69**(2), 307–342.

Petrin, A. (2002), 'Quantifying the benefits of new products: The case of the minivan', *Journal of Political Economy* **110**(4), 705–729.

Reynaert, M. and Verboven, F. (2014), 'Improving the performance of random coefficients demand models: The role of optimal instruments', *Journal of Econometrics* **179**(1), 83–98.

Smith-Spangler, C., Brandeau, M., Hunter, G. E., Bavinger, C., Pearson, M., Eschbach, P. J., Sundaram, V., Liu, H., Schirmer, P., Stave, C., Olkin, I. and Bravata, D. (2012), 'Are Organic Foods Safer or Healthier Than Conventional Alternatives?', *Annals of Internal Medicine* **157**(5), 348–366.

Sudhir, K. (2001), 'Competitive pricing behavior in the auto market: A structural analysis', *Marketing Science* **20**(1), 42–60.

Thompson, G. D. and Kidwell, J. (1998), 'Explaining the Choice of Organic Produce: Cosmetic Defects, Prices, and Consumer Preferences', *American Journal of Agricultural Economics* **80**(2), 277–287.

Villas-Boas, S. B. (2007), 'Vertical relationships between manufacturers and retailers: Inference with limited data', *Review of Economic Studies* **74**(2), 625–632.

Villas-Boas, S. B. (2009), 'An empirical investigation of the welfare effects of banning wholesale price discrimination', *RAND Journal of Economics* **40**(1), 20–46.

Wallace, B. (2018), 'Nonlinear Pricing: Evidence of Price Discrimination in the Fluid Milk Market.', *Mimeo. Available at SSRN Electronic Journal* pp. 1–39.

Wang, E. Y. (2015), 'The impact of soda taxes on consumer welfare: Implications of storability and taste heterogeneity', *RAND Journal of Economics* **46**(2), 409–441.

Zepeda, L. and Deal, D. (2009), 'Organic and local food consumer behaviour: Alphabet theory', *International Journal of Consumer Studies* **33**(6), 697–705.

Zheng, Y., McLaughlin, E. W. and Kaiser, H. M. (2013), 'Taxing food and beverages: Theory, evidence, and policy', *American Journal of Agricultural Economics* **95**(3), 705–723.

A Appendix

In this appendix, I describe the algorithm to estimate the demand system presented in this paper. This algorithm is a Nest Fixed Point Algorithm that consists of an outer loops that guess different values of the parameters (Π, Σ) and an inner loop the searches for the vector δ that equalises the predicted and observed vectors of market shares.

1. Obtain draws from the distribution of D_i, ν_i.
2. Initialise the values of the mean utility, δ_j.[26]
3. Guess θ_2
4. For the given θ_2 and the initial draws of demographic characteristics I compute the household deviation from the mean utility $\mu_{ijt}(x_{jt}, p_{jt}, \nu i, D_i; \theta_2)$
5. For the given mean utility and θ_2, I compute the predicted shares simulating equation (6). I use Halton draws to approximate the integral.

[26] I used the homogeneous logit to obtain the initial values.

6. Given θ_2, I look for δ_t that makes the predicted shares obtained in the previous step equal to the observed shares, $s_{jt} = s_{jt}(\delta_t, \theta_2)$. I use the contraction mapping procedure proposed by Berry et al. (1995) to solve this non-linear system of equations. The operator is the following: $\delta_t^{h+1} = \delta_t^h + lns_{jt} - lns_{jt}(\delta_t^h, \theta_2)$. I iterate until $||\delta_t^{h+1} - \delta_t^h||$ is below a chosen tolerance level.[27]

7. Now I can use δ_t to estimate the linear parameters θ_1 using the fact that $\delta_{jt}(s_{jt}, \theta_2) - (x_{jt}\beta - \alpha p_{jt}) = \xi_{jt}$. The IV moment conditions are $E[Z'\xi] = 0$

8. I then compute the GMM objective function and minimise it over θ_2 with steps 4-7 nested for every trial of θ_2.

9. Finally, after obtaining an estimated θ, I follow Reynaert and Verboven (2014) and compute the optimal instruments.

10. I use the optimal instruments to reestimate the model.

[27] I tried several tolerance levels and results are robust to changes on the level.