Consumer Welfare and Pricing Policies in a Supermarket Chain *

Cristobal Cerda Meneses†

March 2019

(version 22.03.a)

Abstract

Price discrimination can cause either positive or negative effects in the long run sustainability of a retail chain. Particularly in Supermarkets, where the profit margins are relatively low, this question arises as one of the most importants. I use the same approach that the authors of the original paper used, that is, a mixed logit demand function (a Random Coefficients model). Consumer welfare is calculated using the Hicksian Compensated Variation. Pricing policies are calculated following a profit maximization process. For each pricing configuration, the retailer can set a pricing policy that maximizes its profits, and in that way affecting the consumer welfare. Compared to a chain pricing policy, zone and store pricing generate benefits for the customers (i.e. there is a positive effect for going from chain to zone and from chain to store). While the results in the following report can be considered as preliminar (because particularly in the chain to store analysis the effects are far from those observed in the original paper), it is possible to appreciate that changes in pricing policies in the orange juice category are more sensitive than those of a zone category.

keywords: price discrimination, customer welfare, demand modeling.

---

*This paper is intended to be a partial replica of Chintagunta, Dube and Singh 2003, "Balancing Profitability and Customer Welfare in a Supermarket Chain", Quantitative Marketing and Economics 1, 111-147. I acknowledge the James M. Kilts Center, University of Chicago Booth School of Business, for providing the datasets used in this research.

†University of Chicago, The Kenneth C. Griffin Department of Economics, 1126 E. 59th Street, Chicago, IL 60637, USA, ccerda@uchicago.edu.

1 Introduction

Price discrimination can cause either positive or negative effects in the long run sustainability of a retail chain. Particularly in supermarkets, where the profit margins are relatively low, this question arises as one of the most important. I use the same approach that the authors of the original paper used, that is, a mixed logit demand function (a Random Coefficients model). Consumer welfare is calculated using the Hicksian Compensated Variation. Pricing policies are calculated following a profit maximization process.

I analyze the effects in customer welfare of the application of different pricing policies in Dominick's Finer Foods (DFF). I use data for the year 1992 for the categories of laundry detergent and refrigerated orange juice. In order to do that, first I estimated the consumer's demand starting from a log utility model. In particular I estimate a Random Coefficients model, following the work of McFadden and Train,2000. After estimating the demand parameters, I build the retailer's model of pricing, which is a profit maximization problem subject to a configuration of pricing, namely chain, zone or store. I compute the consumer welfare using the Hicksian compensated variation as developed by Small and Rosen,1981.

This research is also closely related to the paper of Berry, Levinson and Pakes, 1995. The authors develop a methodology for analyzing demand and supply in a differenciated product markets. They build the technique over a model of discrete choice with random coefficients.

Another related paper is Hoch et al,1995. In this research, the authors develop an analysis of 18 product categories in the same DFF chain, aiming to measure the price elasticity for the categories at store-level and the relation between demographic variables and price response, finding that a set of 11 demographic variables explain approximately two thirds of the variation in price response.

Li et al, 2018, develop an empirical analysis of national vs. local pricing in the market of digital cameras, focusing on competition dynamics and how this factor affects in the determination of the optimal pricing policy. Leveraging on the detailed availability of the data, the authors estimate for each local market an aggregate model of demand, with random coefficients and micro moments in order to capture the heterogeneity across geographic areas. For the supply side, the authors assume that firms compete in a two-stage game, first selecting a pricing policy and then setting period prices for each product. The authors found that for two of the three companies under analysis (national electronics retail chains), a national pricing policy increase profits, but for a national discount retailer a local pricing policy is better. Unlike the work of Chintagunta et al., the researchers does not deal with customer welfare as they are focused on the effects on competition.

2 Model

Consumer behavior

I will use the standard approach for modeling demand under a discrete choice model. It is well known that this model avoids the problem of dealing with an unfeasible number of variables, in other words, is a parsimonious model. Additionally, there are extensive literature that applies this model in the real world. According to McFadden (1974) the necessary elements for the study of choice behavior are: a set of alternatives, observable attributes and a model of individual behavior (including the distribution patterns of behavior).

I consider that the utility of consumer i for product j depends on the characteristics of the product and the consumer. Some of those characteristics are observable and others not by the econometrician.

For this purpose I will use the mixed logit model developed by McFadden (2000). A particular feature of this model is that it adds normally-distributed random coefficients to a standard Multinomial Logit Model. This model has shown to behave well in a discrete choice setup.

In particular, a mixed multinomial logit (MMNL) model is specified in the following way:

$$P_C(i|x,\theta) = \int L_C(i;x,\alpha) \cdot G(\alpha;\theta) d\alpha \tag{1}$$

with $L_C = e^{x_i\alpha} / \sum_{j\in C} e^{x_j\alpha}$

In this equation, C is the set of alternatives, x_i is the set (vector) of observed atributes for alternative i, part of the choice set; α is a vector of random parameters; L_C is the MNL model for the choice set, and θ is a vector of parameters for the distribution G.

The main difference between the standard multinomial model and the mixed multinomial model is that the later allows for random taste coefficients, solving the restriction of the substitutability pattern (IIA). Nevertheless the introduction of this feature in the model, the setup and the economic background of it (utility maximization) has remained immutable.

As is common in this kind of problems, it is necessary to solve the endogeneity of prices which arises due to the existence of some atributes that are unobserved by the researcher but are observed by the firm. The Instrumental Variables procedure allows to solve the endogeneity problem. Given the access to wholesale price, this measure is an interesting first approach of instrument in order to deal with this problem. First, wholesale prices are highly correlated with shelf prices, which is one of the conditions for solving endogeneity. Second, wholesale prices are exogenous for the retail company.

Let us consider that an individual i that in any given store and week (I assume that each combination of store and week is a market) can choose between purchase one product of the category or not purchasing (the outside option). The conditional utility of individual i of purchasing a product j in market (store-week) s is given by:

$$u_{ijs} = \alpha_{ijs} + x_{js}\beta_i + \theta_{is}p_{js} + \xi_{js} + \epsilon_{ijs} \tag{2}$$

where

$$\beta_i \sim N(\bar{\beta}, \lambda_x)$$

$$\theta_{is} \sim N(\bar{\theta} + D_s'\gamma, \lambda_p)$$

$$\alpha_{ijs} \sim N(\bar{\alpha_j} + D_s'\sigma, \Sigma)$$

Where D is a vector of demographic and competitive variables. Σ is the variance covariance matrix, which in this case is estimated in the following way:

$$\Sigma = L\omega\omega' L' \qquad \omega \sim N(0, I)$$

Where L can be interpreted as a vector of latent attributes (following Elrod, 1988). For this case, as it is in the original paper, I assume two attributes and arbitrarily set the outside option at the origin of the coordinates map (translational invariance). Additionally I set one of the attributes in the horizontal axis (rotational invariance).

2.1 Customer Welfare

Small and Rosen (1981) develop a theoretical model to measure changes in welfare in the environment of discrete choice models. This model is derived from the Hicksian compensated variation, and among other assumptions, it suppose that the marginal utility of income is independent of price and quality of the good; that discrete goods are sufficiently unimportant to the consumers so the income effects from quality changes are negligible (i.e. the compensated demand can appropriately be approximated by a Marshalian demand function), and that the marginal indirect utility w.r.t quality tends to zero as price tends to infinte. Therefore, the compensated variation, which is the amount of money that a consumer would have to receive in order to compensate a change in price can be estimated using the following relation:

$$\Delta E = \frac{-1}{\lambda} \int_{W_1^0}^{W_1^f} \pi_1(W_1, W_2) \mathrm{d}W_1 \tag{3}$$

Where ΔE is the amount of income an individual must be given to make him indifferent to a change in price (denoted as the superscripts 0 and f in the integral), λ is the marginal utility of income, π_1 is the probability of choosing product one, which is a function of the "universal" utility functions for products 1 and 2, W_1 and

W_2 (see Small and Rosen). The form of the function W is identical across consumers and is a function of the price and characteristics of the product and the income and demographic characteristics of the individual. In the case of the multinomial logit model, the equation can be expressed in the following way:

$$\Delta E = \frac{log(\sum_{j=0}^{J} exp(W_1^f)) - log(\sum_{j=0}^{J} exp(W_1^0))}{\lambda} \tag{4}$$

And this relation can be aggregated as follows:

$$\Delta E_M = M \cdot \int \Delta E \phi(\upsilon) \partial \upsilon \tag{5}$$

Where $\phi(\cdot)$ is the pdf of a multivariate standard normal, and υ is a vector of normal deviates of the random coefficients.

Firm behavior

The approach for the company profit is the common fashion of profit maximization:

$$\max_{(p_j)_{j=1}^{J}} \Pi = \sum_{j=1}^{J} (p_{jt} - w_{jt}) Q_{jt} - F_t \tag{6}$$

Where w_{jt} is the wholesale price of product j in period t and F_t are the fixed costs of the period t. This analysis will be performed using the common first and second order conditions in order to find and verify that the solution correspond to an optimum for the company.

3 Data

I use data of DFF which is publicly available on The James M. Kilts Center for Marketing webpage[1]. Dominick's Finer Foods operated near of 100 stores in the Chicago area. Those stores were clustered in 16 zones according to a own definition

[1] https://www.chicagobooth.edu/research/kilts

of the chain. The data for the products consists of weekly information about volume sales, prices, profits, promotions and total store daily traffic, by store and zone. Additionally, there is availability of demographic information for each store such as average income (log), % of population over 60 year-old, ethnicity, etc.

The data for the products spans a period of 52 week of 1992, for the categories of laundry detergent and refrigerated orange juice. Several differences with the original paper are in this research. The first and most important is that the public dataset does not include information for the Dominick's brand of refrigerated orange juice, so I replaced that brand for the brand "HH" which appears in the dataset. Additionally, and due to timing limitations I aggregated products in a different way than the original paper. Instead of using a lower boundary in the pricing correlation, I just grouped the products by brand and similar size. This generate a different aggregated set of data which affects the results.

Table 1 shows the descriptive statistics for both categories. As in the original paper, actual market shares were computed assuming that each customer buys only one product per category. This is consistent with other research findings. With that, the market share used for estimation is equal to the total brand sales divided by the total weekly store traffic.

The demographic data consisted in an available dataset which source is the 1990 Census. the demographic data is used in order to capture heterogeneity across stores. One difference with the original paper is that one variable with information about distance (in miles) from one competitor is not available (JEWELDIST, see Chintagunta et al, 2003), so I use 6 instead of 7 demographic variables.

There is evidence for the same supermarket chain that at least two thirds of the elasticities are explained by demographic variables (see Hoch et al, 1995). Particularly important are variables related to education, family size and household value. In this case only one of those variables is included (Home Value).

Table 2 provides a summary of the statistics for the demographic and competitive variables. It is possible to appreciate that there is considerable variation in the

Table 1: Categories descriptive statistics

Category	Brand	Size	Share	Price	Cost
Refrigerated Orange Juice	Florida	64 OZ	2.92%	2.31	1.57
	Minute Maid	64 OZ	23.21%	1.80	1.53
	Minute Maid	96 OZ	1.62%	3.45	2.54
	HH	64 OZ	30.58%	1.34	0.86
	HH (128)	96 OZ	6.45%	3.17	2.03
	Tropicana Premium	64 OZ	14.46%	2.39	1.73
	Tropicana Premium	96 OZ	4.0%	4.16	3.07
	Tropicana SB	64 OZ	16.7%	1.64	1.39
Laundry Detergent	Tide	64 OZ	21.09%	3.24	2.98
	Tide	128 OZ	28.22%	7.53	7.26
	Cheer	64 OZ	7.74%	3.49	3.05
	Cheer	128 OZ	4.01%	7.22	6.52
	Wisk	64 OZ	14.95%	4.32	4.16
	Wisk	128 OZ	4.71%	7.43	7.20
	All	64 OZ	8.11%	3.32	3.19
	All	128 OZ	5.22%	5.62	5.15
	Surf	64 OZ	5.95%	3.80	3.83

demographic across stores. For example, looking at the “ETHNIC” variable, some stores are in zones predominantly white while others are located in almost pure non white neighborhoods. In a similar way income goes from around $19,300 to $76,100 a year. The competitive variable (EDLPDIST) also shows an interesting degree of dispersion, going from as near as 0,13 miles to as far as 17.85 miles.

AGE60 represent the % of the population over 60 year old in each store neighborhood. HVAL corresponds to the mean household value. the SHOPINDX variable measures the ability of a household to shop (it represent the % of the population with car and single family house)

Table 2: Demographic and competitive variables

Variable	Mean	Std. Dev	Min.	Max.
INCOME (log)	10.61	0.29	9.87	11.24
AGE60	17.5%	6.3%	5.8%	30.7%
ETHNIC	16.0%	19.1%	2.5%	99.6%
HVAL (th)	144.5	44.42	64.35	267.40
SHOPINDX	74.5%	23.1%	0.0%	98.6%
EDLPDIST (mi)	5.26	3.46	0.13	17.86

Table 3: List of Parameters to be estimated

Parameters	Description	Mean	Std Deviation
α_j	Brand Intercept	$\bar{\alpha_j} + D_s'\sigma$	L$\cdot\omega$
β_{pr}	Promotion Dummy	$\bar{\beta_{pr}}$	λ_{pr}
β_{size}	Size Dummy	$\bar{\beta_{size}}$	λ_{size}
θ_s	Marginal Ut. of Income	$\bar{\theta}+ D_s'\gamma$	λ_{price}

4 Estimation

4.1 Market Demand

The estimation procedure for the demand system starts with calculated the instrumental price. I use the standard Instrumental Variables technique, with 2 steps least squares, to build a "synthetic price". This is neccesary to ensure that the price used in the estimation is not correlated with the unobservables, ξ. For this first step I use as instruments the observable variables (promotion and size), the brand intercepts and the wholesale price which can be obtained from the profit information provided in the dataset.

The mean utility δ, and the deviations from the mean μ, are calculated separately and according to equation [...]. Given that the calculation of the utility involves an integral, this is computed by simulation, using 30 draws for each observation (1 observation per brand and size per week per store). I used 4 series of standard normal draws for the simulation. The integral is approximated in the following way:

$$s_{jt} = \frac{1}{ns} \sum_{i=1}^{ns} \frac{exp[\delta_{jt} + \mu_i jt]}{1 + \sum_{k=1}^{J} exp[\delta_{kt} + \mu_{ikt}]} \tag{7}$$

Where j={1,...,J} is a product of the category, t={1,...,T} is the market, and i={1,...,ns} is an element of the simulation set.

I used the Inner and Outer loop resolution methodology explained by Berry, Levinston and Pakes (1995), so I refer to the interested reader to that paper and to Nevo (2000b) in order to know details about the methodology. However, just to give a flavor of the main challenges, I ended up using around 30,000 data-points for each variable. The numbers of parameters to be estimated is 32, so I used a pre-estimation approach in order to be efficient in the use of computational time. First, I estimated a first approximation to the mean utility parameters (20 parameters)in order to have an "educated guess" for the BLP method. This is the same that solving the model without heterogeneity. With those parameters fixed, I solve for the deviation parameters (12 parameters) using the BLP method in order to have a first approximation of those variables. Finally, I solve all the parameters together. This simple procedure allows me to reduce the computational process time in more than 50%.

BLP is solved using GMM, and the weighted matrix is built using the instruments. In other words, the estimated parameters are those that solve the following concept:

$$\min \xi' Z(Z'Z)^{-1} Z' \xi \tag{8}$$

Where Z is the matrix of instruments.

4.2 Category Pricing and Profit Maximization

After estimating the parameters of the demand function it is possible to estimate the price level for a given pricing policy by calculating the prices that maximize the profits on each market (store-week).

In order to estimate the pricing under certain configuration (i.e store, zone and chain level) it is necessary to compute the own and cross price elasticities for each product and each market. to do that I measure the effect in the purchasing unconditional probabilities (i.e the market share) for each brand and each market upon an infinitesimal (1e-5) change in price. For zone and chain configuration I use weighted average prices for every zone and for the whole chain per week. This procedure allows me to obtain the Ω^{-1} matrices just by simply inverting the elasticities matrices for each market under each pricing configuration.

With the foregoing I am in a position to calculate the optimal pricing for each product on each market under the three different criteria of configuration.

4.3 Consumer Welfare

The estimation of the consumer welfare is quite straight forward after having the optimal set of prices and the parameters of the demand system. The expected utility for each market and each pricing configuration is calculated using the draws previously obtained. In the same way that I calculate the market shares by simulation it is possible to integrate the customer welfare across markets by simulation. In other words, the aggregate costumer welfare can be approximated as follows:

$$\Delta W_s = M_s \cdot \frac{1}{ns} \sum_{i=1}^{ns} \frac{log(\sum_{j=0}^{J} exp[V_{jis}^{conf_x}]) - log(\sum_{j=0}^{J} exp[V_{jis}^{conf_y}])}{\theta_{is}} \tag{9}$$

Where $V_{jis}^{conf_x}$ is the expected utility for the market (store-week) s, product j and draw i given a 'x' pricing configuration.

5 Results

I present in this version the main results and some preliminary results for the consumer welfare.

Table 4: GMM - main parameters estimated

Parameter	Laundry Det.	Ref. Juice
$\bar{\beta}_{promo}$	0.7279	-0.7175
$\bar{\beta}_{size}$	0.4371	2.0022
σ_{INC}	-0.6170	-0.0482
σ_{AGE}	-1.2000	2.9029
σ_{ETH}	-0.4308	0.1446
σ_{HVAL}	0.0009	-0.0022
σ_{SHOP}	0.1295	-0.6129
σ_{EDLP}	-0.0566	-0.0094
$\bar{\theta}$	-5.423	-4.7870
γ_{INC}	0.4679	0.5275
γ_{AGE}	0.5537	-2.7953
γ_{ETH}	0.0594	0.0173
γ_{HVAL}	0.0049	0.0041
γ_{SHOP}	0.1271	0.4229
γ_{EDLP}	0.0355	-0.0087
$\bar{\alpha}_{brand1}$	0.680	-8.9642
$\bar{\alpha}_{brand2}$	-0.5967	-7.3520
$\bar{\alpha}_{brand3}$	-0.6804	-6.3738
$\bar{\alpha}_{brand4}$	-0.9257	-7.2199
$\bar{\alpha}_{brand5}$	-1.2107	-7.6181
$\bar{\lambda}_{promo}$	2.18	0.6546
$\bar{\lambda}_{size}$	2.59	0.6546
$\bar{\lambda}_{price}$	6.96	0.6545

5.1 Demand System Results

As I commented before, and as it is shown in Table A-1 of the appendix, first I run a model with no heterogeneity in order to get a first approximation and an "educated guess" to the mean utility parameters. Table 4 shows the parameters so obtained for the two models.

The second and definite step is to run the full model (inner and outer loop) and minimize the GMM function. Table 4 present the results of the minimization. It is worth to note that this sequential procedure allows me to greatly reduce the value of the initial iteration of the GMM function, compared to a naive alternative in which each initial guess is just set to zero.

5.2 Customer Welfare

The results for the Customer Welfare are now presented. Starting from a chain pricing setting I compare the variation in customer welfare of both, a zone and a store pricing policy. The model shows just a modest increase in the case of Laundry detergent (almost $500) for a zone pricing, while in the case of the store configuration the increase in customer welfare is more than $ 1 million. This last result is far from the one obtained in the original paper, where the authors estimated approximately $40,000 for the store pricing.

In the case of the orange juice, both results are by far inconsistents to those of the original paper. For a zone pricing, the model estimated an increase in customer welfare of $40,000 approx (versus a loss of $20,000 approx). In the case of store configuration the result makes no sense.

While I still investigated the causes for the divergence in results, particularly in the store configuration, I present those results that I consider still in a preliminary stage.

6 Conclusion

The sequential estimation of the parameters dramatically reduces the value of the initial GMM function and is also a more efficient method for calculating the parameters of the model.

For each pricing configuration, the retailer can set a pricing policy that maximizes its profits, and in that way affecting the consumer welfare. Compared to a chain pricing policy, zone and store pricing generate benefits for the customers (i.e. there is a positive effect for going from chain to zone and from chain to store). While the results in the following report can be considered as preliminar (because particularly in the chain to store analysis the effects are far from those observed in the original paper), it is possible to appreciate that changes in pricing policies in the orange juice category are more sensitive than thos of a zone category.

References

Daniel McFadden and Kenneth Train, Mixed MNL Models for Discrete Response, Journal of Applied Economics, 2000,(15, 5), 447-470

Kenneth A. Small and Harvey S. Rosen, Applied Welfare Economics with Discrete Choice Models, Econometrica, 1981, (19, 1), 105-130

Steven Berry, James Levinsohn, Ariel Pakes, Automobile Prices in Market Equilibrium, Econometrica, 1995, (63, 4), 841-890

Stephen J. Hoch, Byung-Do Kim, Alan L. Montgomery, Peter E. Rossi, Determinants of Store-Level Price Elasticity, Journal of Marketing Research, 1995, (32, 1), 17-29

Yang Li, Brett R. Gordon, Oded Netzer, An Empirical Study of National vs. Local Pricing by Chain Stores Under Competition, Marketing Science, 2018, (37, 5), 812-837

APPENDIX

A-1 Appendix

Table 5: First approximation to mean utility parameters

Parmeter	Laundry Det.	Ref. Juice
$\bar{\beta}_{promo}$	0.7338	0.7175
$\bar{\beta}_{size}$	0.4730	2.0022
σ_{INC}	-0.5265	-0.0480
σ_{AGE}	-1.2057	2.9029
σ_{ETH}	-0.4378	0.1446
σ_{HVAL}	0.0004	-0.0022
σ_{SHOP}	0.1349	-0.6129
σ_{EDLP}	-0.0126	-0.0094
$\bar{\theta}$	-5.330	-4.7870
γ_{INC}	0.4136	0.5275
γ_{AGE}	0.5516	-2.7953
γ_{ETH}	0.0596	0.0173
γ_{HVAL}	0.0005	0.0041
γ_{SHOP}	0.1153	0.4229
γ_{EDLP}	0.0088	-0.0087
$\bar{\alpha}_{brand1}$	0.6922	-8.9642
$\bar{\alpha}_{brand2}$	-0.5926	-7.3520
$\bar{\alpha}_{brand3}$	-0.6875	-6.3738
$\bar{\alpha}_{brand4}$	-0.9222	-7.2199
$\bar{\alpha}_{brand5}$	-1.2102	-7.6181