An empirical analysis of loyalty programs and private label demand

Jorge Florez-Acosta*

[Preliminary. Do not circulate without permission.]

February 14, 2014

Abstract

Loyalty programs (*LP*) are by now a predominant short-run nonprice strategy in retailing markets. Most work the same way: a member who purchases today gets a reward to be used next time she returns to the store (or after she crosses some threshold). Previous researchers have concluded that the purpose of LPs as a marketing strategy is customer retention. In the grocery retailing sector it seems to be as well the boost of private label (PL) demand. This paper empirically examines this conjecture. Using discrete-choice methods, I estimate brand-level demand taking into account household membership to loyalty programs. I find that although consumers give a lower value to private labels relative to quality equivalent national brands, loyalty programs have a positive and significant effect on PL choice, i.e. the marginal valuation for a PL is higher for LP members. Moreover, the more prone to subscribe to LPs a customer is, the larger her sensitivity to a price increase and the weaker the expected effects on PL demand.

JEL Codes: D12, L13, L22, L81.

Keywords: Grocery retailing, supermarket chains, loyalty programs, private labels, oligopolistic competition, discrete choice models, random coefficients.

*Toulouse School of Economics, jorge.florez@tse-fr.eu. I am grateful to my supervisors Bruno Jullien and Pierre Dubois for their guidance. I also thank Fabien Bergès, Yhingua He, Laura Lasio and the participants of the following academic meetings: EARIE 40th Annual Conference in Evora (2013), XXVIII Jornadas de Economía Industrial 2013 in Segovia, the *Seminario de Economía* at Banco de la República in Medellín (Colombia) and the Applied Microeconometrics and the Econometrics Lunch Workshops at the Toulouse School of Economics for helpful comments and suggestions. All errors are my own.

1 Introduction

Retailer own-branded products, also known as private labels (hereafter, *PL*),[1] have become of big importance for the agents involved in the grocery supply chain, on the one hand, and for researchers and policy makers, on the other hand. Evidence shows that PL market share has been increasing steadily in the last decade and have become a good alternative for consumers looking for better deals. For instance, in the United States PL's market share attained 19.1% on sales in 2010 (Turcic et al., 2013), whereas it was even larger for European countries such as the United Kingdom (46%), Germany (35%) and Spain (33%)(Steenkamp and Geyskens, 2013). On top of this, the European Commission reports that PL sales are growing around 4% per year on average (European Commission, 2011).

PL increase the rage of products available for consumers, intensifies the intra-brand competition and may stimulate upstream competition as well. Moreover, the fact that they are supplied at lower prices constitute a further advantage for consumers.[2] However, it has long been recognized that PL have some negative impact as well. In particular, retailers might be using them as a way to strengthen their position vis-à-vis manufacturers (buyer power) and obtain better deals at the expense of suppliers' profits.[3] From this perspective, one may think that retailers strategies to boost PL demand may help increasing their bargaining power.

One of this strategies seems to be loyalty programs (hereafter, *LP*), which are by now a predominant short-run practice in grocery retailing markets. The objective of this paper is to empirically examine the effects of loyalty programs on PL demand. Previous researchers have provided several explanations as to why retailers offer such costly programs. They can be summarized in two ways: *i*) consumer retention, as customers are more likely to return when there is a promised price reduction; and *ii*) the exercise market power, in particular, the subscription requirement acts as an explicit discriminatory device. In the case of grocery retailing, a third explanation seems to be the boost of PL demand. In France, for example, most supermarket chains link loyalty programs to the purchase of private labels(see Table 1).[4]

Table 1: France: some supermarket loyalty programs

Retailer	Loyalty program	Reward	Products included
Carrefour	Carte Fidelité	5%	Carrefour
Intermarche	Avantage Carte	5%/5 PL products min.	Several PL
Auchan	Carte Auchan	5%	Auchan and NB
Leclerc	Carte e.Leclerc	€/5 products	Marque Repère
Casino	Carte Casino	1 mile/3€	NB and PL

Source: Supermarkets web pages.

[1]They are retailers' own-branded products supplied exclusively in their stores and produced by a separate manufacturer. They are produced as quality-equivalent products respective to national brands (*NB*), which are the regular highly advertised and generally everywhere-available products. In general, they are more profitable for retailers than *NB*, even though they are of lower price.

[2]Evidence shows that PL are supplied at a 20% lower price on average relative to a quality-equivalent NB (Berges-Sennou, 2009).

[3]For further readings on this topic, see Scott-Morton and Zettelmeyer, 2004, Bergès-Sennou, 2006, and Meza and Sudhir, 2010.

[4]Most retailing chains give lagged rebates based on current purchases of selected PL products, conditional on the previous subscription to the program. Rebates are accumulated in customer's account and after a given time/money threshold is crossed, the acquired amount of money is given back to customers as a purchase coupons to be expended in any of the retailer's stores. Some programs, such as that of Casino's, work slightly different as they give "miles" to customers according to a predetermined exchange rate, and have a catalogue where members can pick a gift according to the accumulated number of miles.

I estimate brand-level demand taking into account household membership to loyalty programs. I use a three-dimensional panel of quantities and prices for up to 13 brands of plain yogurt, purchased from the 6 largest supermarket chains in up to 94 departments of France, weekly in 2006. I also observe demographic characteristics including household membership to supermarket LPs. In addition to the well documented challenges faced when estimating demand, such as the endogeneity of prices and the dimensionality problems implied by the large number of brands, I also deal with the correlation between LP membership indicator and the unobserved supermarket attributes. In a first stage of the estimation, I try some standard sets of instruments (BLP, 1995 and Nevo, 2001) to treat the endogeneity of prices and LP membership. Yet, due to a poor performance of those sets I overcome the endogeneity problems by computing the optimal instruments based on Chamberlain (1987).

The dimensionality problem related to a large substitution matrix resulting from the numerous products considered here is solved by using discrete-choice methods. In particular, I estimate a random coefficients Logit model following closely the standard literature, namely, Berry (1994), Berry, Levinsohn and Pakes (1995) and Nevo(2000a, 2001).

Results confirm that private labels are, on average, less valued relative to national brands. However, I find that the marginal valuation of PL products increases with subscription to the supermarket LP, which confirms the believe that LP serve as a way to boost store-brands demand. Moreover, when customers subscribe to several separate LPs of competing retailers, the expected effects are weaker, i.e., the marginal valuation of PL products decreases with the number of subscriptions and customers are more sensitive to price changes.

I use the estimates to conduct some demand-side counterfactuals playing with the subscription status of the customers. In a first scenario, I assume that everybody is member of at least one loyalty program. In a second scenario, I assume the contrary: no one is member to LPs. Results support the hypothesis of this paper that LP might be used as a way to boost PL demand as it increases by more than 75% when former 'non-loyals'[5] become members (first scenario). On the other hand, in the absence of LPs (second scenario) PLs seem to be less attractive products for both former 'loyals'[6] and 'non-loyals' as aggregate demand decreases by 26% in average. Welfare analysis shows that consumers are in general better off when they all join a LP and worse off when no one is member to LPs, as compared with the baseline scenario. However, the second scenario has a larger impact on consumer surplus: zero subscription to LP causes a reduction in consumer surplus that doubles the increase in consumer welfare when there is full subscription coverage.

There is a developed literature on topics related to either private labels, brand- and store-loyalty or loyalty programs. However, to my knowledge this paper is the first to provide empirical evidence on the link between supermarket loyalty programs and private label demand. Table 2 summarizes the most important contributions by topic.[7]

This article relates in particular to Lewis (2004), who makes an evaluation of the effects of loyalty programs based on the idea that they are addressed to "enhance [customer] retention". Bonfrer and Chintagunta (2004) study the effects of the introduction of PL on retailers' profits taking into account that consumers can be store- and brand-loyals. They propose measures for store- and brand-loyalty based on the number of trips to the same

[5]This is what I call those people who are not members of any LP.

[6]Similarly, this make reference to those who are linked to supermarket LP.

[7]For a complete survey about the theoretical and empirical literature on this theme, see Berges-Sennou et al. (2009).

Table 2: Contributions to the literature on Loyalty programs and PL

Topic	Theory	Empirics
Introduction of PL	Raju et al. (1995)	Bonfrer & Chintagunta (2004)
	Chintagunta et al. (2002)	
PL demand determinants	Berges et al. (2009)	
Competition & vertical relationships	Soberman & Parker (2006)	Bonnet & Dubois (2010)
	Bonnet & Dubois (2010)	
Store & brand loyalty	Berges (2006)	
Loyalty programs	Lal & Bell (2003)	Bolton et al. (2000)
		Lal & Bell (2003)
		Lewis (2004)
		Lederman (2007)
Coupons	Caminal & Matutes (1989)	Nevo & Wolfram (2002)
	Cremer(1989)	

store and the average "share of wallet" of a brand relative to the total expenditure on that category of goods. They find a significant negative correlation between both types of loyalty. Bolton, Kannan and Bramlett (2000) argue that loyalty programs members are more likely to do repeat buying, as they weigh less than others the best outside alternative. They conclude that LPs members are less sensitive to both quality changes and lower prices offered by competitors. Lal and Bell (2003) claim that there are two reasons explaining the "success" of loyalty programs: (*i*) reduced price competition and therefore higher profits due to switching costs, and (*ii*) reduced marketing expenses by focusing attention on retaining loyal (and known) customers. This is in line with the marketing idea that promotions should be addressed to customers that are more likely to stay. All these papers share a common basic question: What are the determinants of customer retention and repeat buying? This article asks a rather different question, as the interest is focused on the effects of loyalty programs on private label demand.

In particular, this paper closely relates to Nevo and Wolfram (2002). They provide empirical support on coupons issuing strategies by manufactures. Their objective is to describe manufacturers' motivations for issuing coupons. They evaluate some hypotheses such as price discrimination, dynamic demand effects and retailers' pricing strategies using data on breakfast cereal. The key difference with my article is that I focus on retailers' rather than manufactures' case with the particularity that the former give customers personalized "checks" that can be expended in any set of products in stock in the supermarket.

The remainder of this paper is structured as follows. Section 2 provides an overview of the theory predictions on loyalty programs. Section 3 outlines the data and a preliminary analysis. Section 4 sets out the empirical framework, the way I carried out the estimation process and presents the estimates. Section 5 displays the results of two simulated scenarios on the demand side. Finally, Section 6 concludes.

2 The theory of Loyalty Programs

We can summarize findings from theoretical research in two statements. First, LPs allow retailers to retain customers and induce repurchase, as long as it is a way to impose artificial switching costs on customers and, at the same time, they are more likely to come back when there is a promised price reduction. Second, LPs are a way to exercise market power, in particular, they can be used as an explicit discriminatory device as customers

must subscribe to the LP to be able to enjoy the benefits. In this Section, I present an overview of the main theoretical work supporting each of those conclusions. In Section A of the Appendix, I set out a simple two-period model which is in line with the previous literature and that gives useful insights on the LPs effects.

Cremer (1984) addresses the question of repeat buying induced by coupons that are valid only for next purchases of the same product, using a two-period model in which a monopolist produces a good that consumers have not yet tasted. He finds that, as consumers have elastic participation (they can take an outside option instead of buying the good), the monopolist optimal strategy is to charge a lower price to repeat buyers instead of precommitting to a second-period price. Klemperer (1987a, 1987b) lists repeat-purchase coupons and "frequent-flyer" programs (FPPs) as examples of artificial or contractual switching costs that make rational customers prone to display brand loyalty as demand becomes more inelastic and firm's market power increases. Caminal and Matutes (1990) use a similar framework as Klemperer (1987b) but endogenize switching costs to show that coupons valid for next period purchases perform better than price precommitment as they allow firms to get higher overall profits and reduce the intensity of competition.

In line with these results, Chen and Pearcy (2010) develop a dynamic pricing model to determine in which cases it is optimal to reward loyalty to retain current customers or to entice brand switching in order to attract new customers away from rival firms. They find that for those markets in which customers do not have a strong preference for a particular brand (because good substitutes are available or preferences are likely to change for future choices), firms tend to enroll them in loyalty programs to reward repeat purchases and discourage brand switching.

The conclusion of LPs serving as a discriminatory device follows from the exposition above. When firms reward repeat buying, they are at the same time setting a differentiated future price schedule where new customers will be charged with the full tariff whereas the repeat buyers will have a reduced one by redeeming the coupon they have got in a previous period.

A similar conclusion can be obtained from behavior-based price discrimination[8] literature (see for example Caillaud and De Nijs, 2011). Although most contributions on this topic conclude that firms should offer lower prices to new customers, Caillaud and De Nijs (2011) get the opposite result under the assumption that some firms cannot distinguish between old and new customers. Hence, they reward loyalty by offering lower tariffs to previous customers and charge full prices to the new ones.

Literature has been focused, hence, on the study of these two insights, based mainly on coupons issuing by manufacturers and FFPs as the motivating facts. However, I claim that other forms of rewards programs, promotions and coupons issuing made by retailers are not fully explained by the existing literature. Loyalty programs other than FFPs seem to be used for additional purposes other than customer retention or price discrimination. It is the case of supermarket LPs in France, where rewards are linked to store-brands suggesting that one of retailers' main motivations to enroll customers in LPs is the boost of their own-brands demand. An argument against this conjecture could be that when a supplier need to boost the demand for a good, it suffices to lower its price making suboptimal the use of indirect strategies. Yet, Caminal and Matutes (1990) provide the rational behind the advantages of loyalty discounts relative to price reductions. They find that a precommitment to a second-period discount is more profitable than a precommitment to a second-period price reduction for loyal customers.[9]

[8] It consists of offering different prices to different consumers depending on their past purchases.

[9] Another argument is given by the fact that, in a world in which store-brands are cheaper and perceived

3 Preliminary analysis: customer profile and the grocery retailing industry

This Section aims at giving a general view of the supermarket industry in France and the customer profiles, through an exploratory analysis of the data. As loyalty programs are a supermarket-level rather than a product-level marketing strategy, I use data on more than 350 different food products purchased by the households in the sample during 2006 to provide some descriptive evidence on customers behavior in the presence of loyalty programs offered by competing supermarkets. In Section 4 I will focus on one product to assess the effects of LPs on product choice.

3.1 The data

This study uses the TNS Worldpanel data base provided by the TNS-Sofres Institute.[10] It is homescan data on grocery purchases made by a sample of 14,529 households in France during 2006. These data were collected by the household members themselves with the help of scanning devices. As the TNS Worldpanel is a continuous panel database, they keep most households originally sampled and renew the sample every year by changing a quarter of the households in the sample, removing those rarely reporting data or increasing the sample size.

The data set contains information on 350 different food products from around 90 retailers including hyper-, super- and mini-market stores, hard-discounters and specialized retailers. Every observation in the dataset corresponds to the purchase of a food product by a household in a given date. The entries are coded with the household identification number. This is to say, if the household made purchases of a bundle of products the same day, there will be an entry for every single product purchased with three levels of information:

- Household level: size, number of children, location, income, number of cars, etc.
- Individual level: characteristics of each household member (age, height, gender, schooling, etc.).
- Product level: price, quantity, retailer info (store's name, surface), brand, type of product (PL or NB), manufacturer, etc.

Furthermore, there is information on household membership to retailers' LPs. It is a dummy variable taking on a value of one if the household is member of the retailer's LP and zero otherwise. Unfortunately, detailed information on loyalty coupons issuing and redemption is not available.

3.2 Customer profile

Table 3 displays summary statistics on household demographics, purchases, and loyalty information. The survey includes people aged between 19 and 75 years old. In average, a household consists of two to three members and has an income of around 2300 € per

as lower quality goods compared to similar national brands, additional price cuts can be interpreted as a bad signal for consumers and cause the reverse effect.

[10] I am grateful to the *Institut National de la Recherche Agronomique*, INRA, for giving me access to the data base.

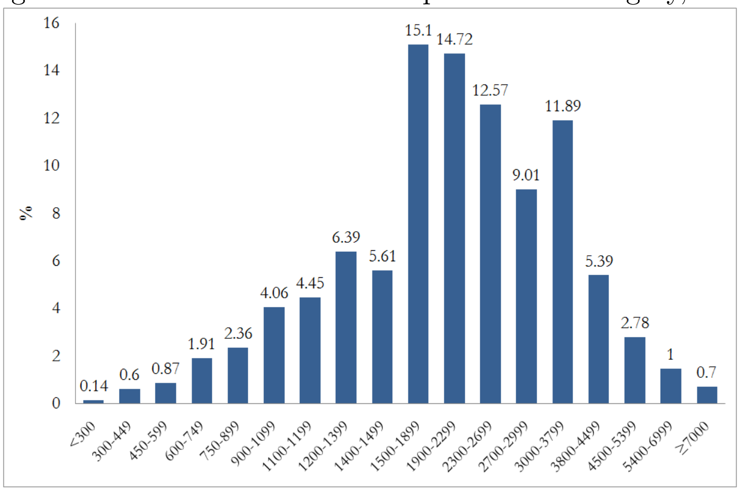
month.[11] Figure 1 displays the distribution of households across income categories. We can see that households are asymmetrically distributed being concentrated mainly in the upper-middle categories of income. A 75% of the households included in the dataset live in city areas of France.

A 34% of products purchased by a household are PL, which corresponds to 24.75% of its total expenditure. Moreover, the average French household members are one-stop shoppers, as they only visit one store within a week. Finally, a 85% of households are members to at least one supermarket loyalty program and, in average, they are subscribed to two separate programs.

Table 3: Summary statistics on household characteristics and purchases

Variable	Mean	Median	Sd	Min	Max
Demographics					
Size of household	2.63	2.00	1.39	1	9
Live in city	0.75	1	0.43	0	1
Income	2337	2100	1175	150	7000
# of cars	1.47	1.00	0.82	0.00	9.00
Purchases					
Private label purchases	0.34	0.32	0.17	0	1
Total expenditure (€/day)	39.10	28.56	35.40	0.01	2,221
PL share (% total exp.)	24.75	15.40	38.6	0	63.63
NB share (% total exp.)	75.24	72.49	83.17	0	100
Store-related information					
# different stores visited the same day	1.13	1.00	0.38	1.00	7.00
Duration (days) between visits	8.10	6.63	6.17	1	126
LP membership	0.85	1.00	0.36	0	1
# of different memberships	2.60	2	1.48	1	12

Figure 1: Households distribution per Income category, 2006



In a month households visit, in average, two different retailers and only 16.65% of the times they go to the store owning the LP to which the household is member. Using this information along with the average duration of 1.89 months a household takes to return

[11] I constructed the numeric variable *Income* from a categorical variable that originally displays income level of the household distributed in 18 income ranks. I took the average income per category and assigned the same value to all households in the same category. See Figure 1 for further information on the distribution of household income.

to the same retailer, we get seven weeks as the average duration a LP member takes to go back to its “patronized” supermarket, which is not so frequent taking into account that household members go shopping at least once a week in average (see Table 4).

Table 4: Summary statistics of monthly visits to stores

Variable	Mean	Median	Sd	Min	Max
# different stores visited	2.27	2.08	0.98	1	9.25
# visited if loyalty subscription	0.36	0	0.50	0	3
Loyalty ratio[a]	16.83	0	25.40	0	100
Duration (months) between visits	1.89	1.33	1.48	1	12

[a] Computed as the number of visits to stores where customer is subscribed over the total of stores visited.

“Loyal” vs. “non loyal” customers

A simple exploratory analysis of the data by subgroup of population gives no evidence to support the hypothesis of loyalty programs as a discriminatory device. In effect, some descriptive statistics (see Table 5) suggest no important differences between subscribers to a LP relative to non subscribers.

One can say that the fact of being a member of a LP does not say much about customer types as long as LPs are available to everybody and subscription costs might be lower than benefits for most customers.[12] Table 5 shows that these two groups are not only similar in demographic characteristics, but also in consumption patterns: the portion of PL products purchased is similar in average (34%), it is so the total expenditure per trip to the supermarket, the distribution of this expenditure between PL and NB, the fact that both are one-stop shoppers and that the average duration in days between trips to a supermarket is similar (around eight days).

Table 5: Loyals and non-loyals average characteristics

Variable	LP members	Non-members
% on total	85	15
Size of household	2.63	2.62
Income	2333	2361
# of cars	1.47	1.48
Private label purchases	0.34	0.35
Total expenditure (€/day)	39.09	39.19
PL share (% total exp.)	25.87	26.36
NB share (% total exp.)	74.13	73.64
# different stores visited the same day	1.12	1.12
Duration (in days) between visits	8.04	8.41

To go a little further in the exploration of the data, I regressed consumer weekly expenditure per retailer on the LP membership dummy and some demographic and supermarket characteristics. Table 6 displays the results. The coefficient for membership is positive and significant telling that a customer tends to expend more in those supermarkets where he is subscribed to a loyalty program than in those where he is not. Another interesting

[12] Although it is argued that subscription to a LP is free and so it is expected from a rational consumer to always subscribe, I believe that there are several sources of costs that may not be negligible depending on the consumer valuation of time, personal information, advertising spam, etc. It is actually an empirical fact that not all the consumers subscribe to LPs: 15% of French households who frequent supermarkets do not have any LP card.

estimates are those of number of subscriptions to different LPs and number of separate retailers visited within a week. As the descriptive statistics before suggested, people normally are members of two distinct LPs in average and some of them use to be multi-stop shoppers. The estimates are both significant and negative, showing that multi subscription affects negatively the expenditure in a given supermarket whereas, as expected, visiting several supermarkets within the same week decreases expenditure per retailer.

Table 6: Results for HH weekly expenditure per supermarket

	Log of total expenditure in s	
Variable	OLS	Fixed effects
LP member	0.372	0.499
	(0.003)	(0.004)
# of LP subscriptions	-0.0477	-0.0578
	(0.001)	(0.001)
# visits to	-0.123	-0.112
different retailers	(0.001)	(0.001)
Log of Income	0.202	0.205
	(0.003)	(0.003)
Log of age	0.109	0.104
	(0.004)	(0.004)
HH size	0.155	0.151
	(0.001)	(0.001)
Hypermarket	0.138	
	(0.003)	
Minimarket	-0.252	
	(0.006)	
Hard-discount	0.0211	
	(0.004)	
Constant	1.099	1.151
	(0.027)	(0.027)
adj. R^2	0.145	0.179

Notes: Regressions are based on 658,866 observations. The two regressions include time dummy variables. Asymptotically robust s.e. are reported in parenthesis.

3.3 The grocery retailing industry

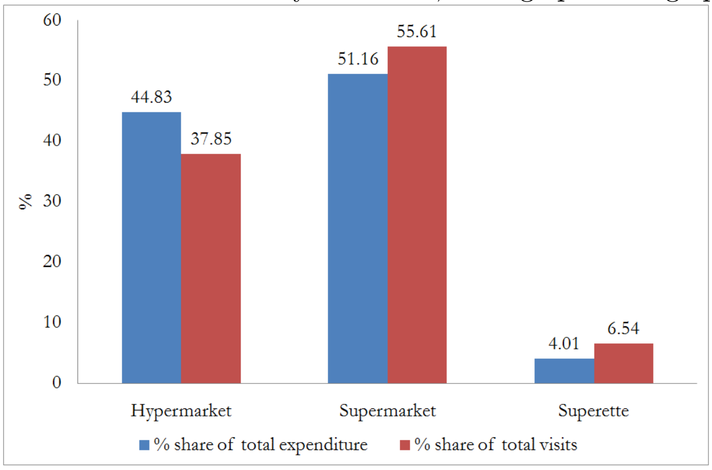
In France, grocery retailing is classified in three categories depending on the size of the store: Hypermarkets, Supermarkets and Minimarkets.[13] Among them, supermarkets are preferred by most consumers: the average market share of supermarkets in terms of the total consumer expenditure per day is of 51.16% against 44.83% of hypermarkets and 4% of minimarkets. In terms of the daily number of visitors, supermarkets also leads the industry: a 55.61% of the total number of customers went to supermarkets whereas a 37.85% went to hypermarkets and a 6.57% to minimarkets (see Figure 2).

With respect to regular stores, Hard-discount retailers have a share of 14% on total household expenditure in groceries in a day and 16% share on total visits per day.

In terms of LP subscribers, the largest market share is 21% suggesting that the market

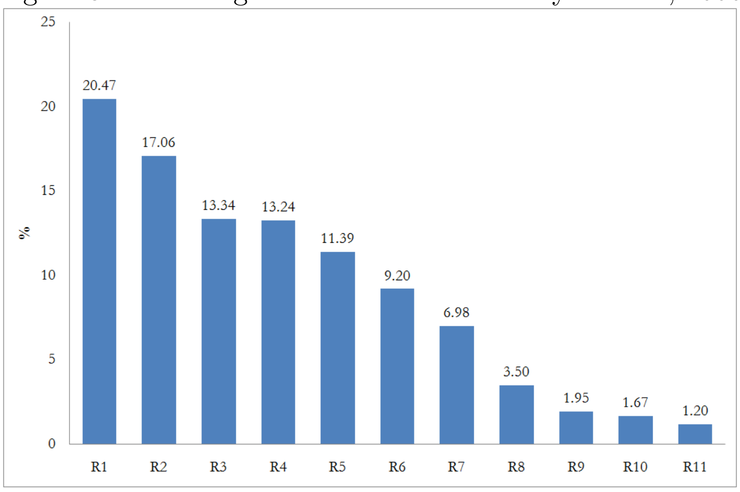
[13]According to the French law, a grocery retailing store is considered a Hypermarket when, among other characteristics, its surface is greater or equal to $2,500\text{m}^2$, a Supermarket if the surface is in the interval $[400, 2500)$ m^2 and a Minimarket if the surface is in $[120, 400)$ m^2. Hard-discount stores are also included in these three categories as they also have shops of all sizes.

Figure 2: Market share indicators by store size, average percentage per day, 2006



is not very concentrated.[14] Still, around 75% of LP subscribers is concentrated by the five leading retailers (see Figure 3). Moreover, around 62% of the households have multiple LPs subscriptions (in some cases 8 in total), reason why it is necessary to look for other indicators such as the percentage of members visiting the store in a day.

Figure 3: Percentage of subscribers to LP by retailer, 2006

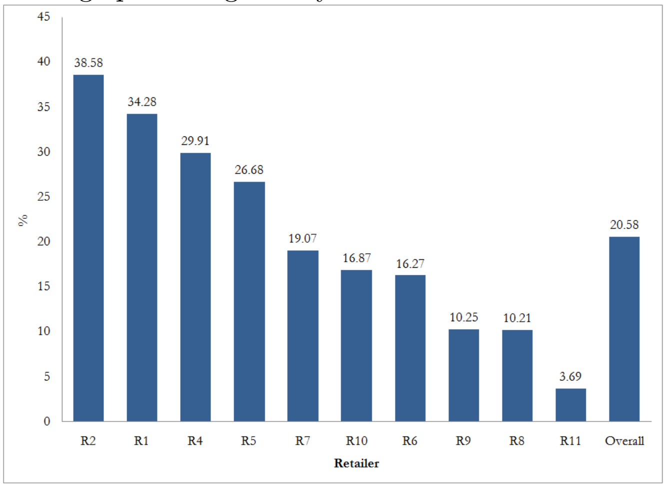


In fact, the ability of retailers to attract their "loyal" customers differs from the previous results. Figure 4 shows that retailer "R2" is visited by a higher percentage of its loyal customers than its rival "R1", whereas the latter has a higher portion of subscribers than the former. Overall, the average portion of LP members visiting the retailer owning the program is 20.6%.[15]

[14]There are cases of retailers that have several ways/cards that serve to the same end like, for example, Carrefour that has both "Carte fidelité" and "Carte Pass". In these cases, I aggregate up subscribers under the same store label, e.g., Carrefour loyalty program.

[15]The actual names of the retailers are hidden at the request of the TNS-Sofres, the institution providing the data.

Figure 4: Average percentage of loyal visitors on the total visitors per day



4 Empirical approach and results

As stated before, loyalty programs appear to be used by grocery retailers as a non-price strategy to boost private label demand. In the following I provide empirical support to this hypothesis.

Here I follow the discrete-choice literature and estimate two models: a multinomial Logit and a random-coefficients Logit (hereafter, I will use random-coefficients Logit, mixed Logit and the full model exchangeably). The reason for conducting these two estimations is that the Logit model is useful to take a first glance as it is easy to estimate and gives important preliminary information about the explanatory power of the variables of interest on consumer behavior. However, as it is well known, this model has some limitations due to its restrictive assumptions, in particular, it gives unrealistic information about substitution patterns. Then, the full model although computationally challenging, is useful to make closer-to-reality inference and counterfactual analysis.

4.1 The empirical framework

Here I will basically follow the standard literature but I will introduce some new notation, given the nature of my problem. For instance, I will assume that retailers are a source of product differentiation, i.e., the same brand sold by two separate retailers becomes two differentiated products.

Assume we observe $t = 1, 2, ..., T$ markets and $i = 1, 2, ..., I_t$ consumers per market. I define a market as a week-Department[16] combination where consumer purchases are observed. Every time a consumer goes shopping to a given supermarket $s = 1, 2, ..., S$, he faces a multiple-choice decision among J brands. The conditional indirect utility of consumer i from choosing product j at supermarket s in market t writes as

$$u_{ijst} = x_j\tilde{\beta}_i + r_s\lambda_i - \tilde{\alpha}_i p_{jst} + \tilde{\varphi}_i M_{is} + \tilde{\eta}_i M_{is} \times PL_{js} + \xi_j + \Delta\xi_{jt} + \xi_s + \Delta\xi_{st} + \epsilon_{ijst} \quad (1)$$

where x_j and r_s are a K- and R-dimensional (row) vectors of observable product j and

[16]In France, a Department (or *Departement* in French) makes part of the administrative division of the national territory being the second level of the government at the local area, after the Administrative Regions which are groups of departments.

supermarket s characteristics, respectively;[17] p_{jst} is the unit price of product j in supermarket s, M_{is} is a dummy indicating whether the individual i is a member of supermarket s loyalty program or not and PL_{js} is a dummy taking on 1 if j is a private label of retailer s. ξ_j and ξ_s are the mean (across individuals and markets) valuations of the unobserved (by the econometrician) product and supermarket characteristics and $\Delta\xi_{jt} = \xi_{jt} - \xi_j$ and $\Delta\xi_{st} = \xi_{st} - \xi_s$ are market deviations from the respective mean under the assumption that in each market people value differently those characteristics. Finally, individual heterogeneity enters the model through the set of individual-specific parameters $(\tilde{\alpha}_i, \tilde{\beta}_i, \tilde{\varphi}_i, \tilde{\eta}_i)$ and an additive separable mean-zero random shock ϵ_{ijst}.

As Nevo (2000a, 2001), I model consumer taste parameters as a function of observed and unobserved demographics and assume that the latter are normally distributed

$$\begin{pmatrix} \tilde{\alpha} \\ \tilde{\beta} \\ \tilde{\varphi} \\ \tilde{\eta} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \varphi \\ \eta \end{pmatrix} + \Pi D_i + \Sigma v_i, \qquad v_i \sim N(0, I_{K+3}) \tag{2}$$

where D_i is a $d \times 1$ vector of demographic variables, Π is a $(K+3) \times d$ matrix of coefficients measuring the change in tastes with demographics, Σ is a $(K+3) \times (K+3)$ of coefficients and v_i captures additional demographic characteristics that influence consumer choice but are generally not included in surveys.

In addition to the characterization of the choice among J products, it is necessary to introduce the "outside good" as consumers can decide to not purchase any of the available brands. Therefore, as its price is not set in response to the inside goods prices, when a general price increase of the J products takes place, the aggregate output may decrease as consumers can substitute to the outside alternative (see Berry, 1994, for a detailed discussion).

The indirect utility from the outside option is thus modelled in the following way

$$u_{i0t} = \xi_0 + \pi_0 D_i + \sigma_0 v_{i0} + \epsilon_{i0t} \tag{3}$$

where the mean outside good characteristics ξ_0, and the parameters π_0 and σ_0 are not identified. The standard solution to this problem consists of normalizing them to zero.

With all these elements along with the assumption that $\theta = (\theta_1, \theta_2)$ is a vector containing all the parameters of the model ($\theta_1 = (\alpha, \beta, \varphi, \eta, \lambda)$ contains the linear parameters and $\theta_2 = (\Pi, \Sigma)$ the nonlinear ones), we can rewrite the indirect utility of a consumer i from purchasing brand j at supermarket s in market t as the sum of two components: a mean utility common to all consumers of the same type (defined here by the subscript $m = \{0, 1\}$ of "loyals" and "nonloyals", respectively)

$$\delta_{mjst} = x_j\beta + r_s\lambda - \alpha p_{jst} + \varphi M_{ms} + \eta M_{ms} \times PL_{js} + \xi_j + \xi_s + \Delta\xi_{jt} + \Delta\xi_{st}, \text{ and} \tag{4}$$

and a mean-zero heteroscedastic deviation, $\mu_{ijst} + \epsilon_{ijst}$ with

$$\mu_{ijst} = [p_{jst}, x_j, r_s, M_{ms}, M_{ms} \times PL_{js}]' * (\Pi D_i + \Sigma v_i), \tag{5}$$

where vector of covariates is of order $K \times 1$. This yields

[17] Unlike product characteristics, I do not allow supermarket characteristics to interact with demographics but introduce them as controls instead, i.e. λ is fixed rather than a random coefficient.

$$u_{ijst} = \delta_{mjst}(\cdot;\theta_1) + \mu_{ijst}(\cdot;\theta_2) + \epsilon_{ijst} \tag{6}$$

This general framework is best known as the *random-coefficients Logit model*, as the taste parameters are allowed to vary over individuals and are determined by both observed and unobserved demographic characteristics, with some distributional assumptions on the latter.

A key assumption of this model is that consumers choose at most one unit of the brand that gives the highest utility. Let the following be the set of observed and unobserved variables determining the preference for brand j at store s

$$A_{mjst}(x, r, p_{.t}, \delta_{.t}; \theta_2) = \{(D_i, v_i, \epsilon_{ijst}) | u_{ijst} \geq u_{ilkt},\ \forall\, l = 0, 1, ..., J; k = 0, 1, ..., S\} \tag{7}$$

where x are the characteristics of all products, r are the characteristics of all supermarkets, and $p_{.t}$ and $\delta_{.t}$ are $J \times S$ matrices of prices and mean utilities of the J products available in the S supermarkets, respectively. Assuming ties occur with zero probability, the market share of the jth brand purchased from supermarket s in market t as a function of the mean utility levels of all the $J+1$ products, given the parameters, by group of population m is

$$s_{mjst}(x, r, p_{.t}, \delta_{.t}; \theta_2) = \int_{A_{mjst}} dP(D, v, \epsilon) = \int_{A_{msjt}} dP(\epsilon) dP(v) dP(D) \tag{8}$$

where $P(\cdot)$ denotes population distribution function. The last equality is a result of the independence assumption of D, v and ϵ. The market shares in (8) can be computed by making assumptions on the distribution of each of the individual variables $(D_i, v_i, \epsilon_{i.t})$.[18]

4.2 The yogurt data

Among all the products in the data described in Section 3, yogurt seems to be a good candidate to estimate the effects of LP subscription on PL demand. It matches pretty well the assumptions of the Logit setup in the sense that it can be considered a non-storable good as it should be consumed soon after purchased, and consumers demand only one unit of it at a time,[19] which is also a key assumption for the empirical framework I use.

The original database of yogurt had information on purchases of 174 varieties of yogurt sold by an average of 20 separate retailers in the 94 metropolitan departments of France. In addition, different flavors are branded under the same label, which increases the dimensionality of the space of products as long as for research purposes, such characteristics should be taken into account as consumer tastes might vary across flavors for the same branded product. In the French market there are around 144 different flavors available, being 5 the average number of flavors by brand. Here we have clearly a dimensionality

[18] The simplest distributional assumptions on this model lead to the (aggregate) Logit model, namely, that individual heterogeneity is only accounted for by the idiosyncratic error terms, ϵ_{ijst}, which are distributed i.i.d. Type I extreme-value. In spite of its tractability and easy estimation, the Logit implies important limitations as predicted substitution patterns are unrealistic. The random coefficients solve these problems (see BLP, 1995; Nevo, 2000a; 2001).

[19] It is true that people do not necessarily buy only one brand of yogurt at a time, instead, they can buy several varieties of the same product to have different choices at home (different flavors, fruit contents, thickness, etc.). However, following Nevo (2001), I claim that an individual normally consumes one yogurt (assuming 125gr per portion) at a time, so that the choice is discrete in this sense. Of course there could be cases in which some people consume more than one brand of yogurt at a time. Although I believe this is not the general case, the assumption can be seen as an approximation to the real demand problem.

problem coming from the large amount of yogurt brands and varieties by brand, which would result in a huge substitution matrix to be estimated.

To overcome this dimensionality problem, in the first place, I focused only on the plain yogurt variety and included the flavored yogurts in the outside good. Second, I kept only the leading 6 grocery retailing chains with a loyalty program. Then, I picked the 13 leading brands based on the national market shares on total sales in 2006. The aggregate market share of the selected brands is of 66.5%, six of them being PL, which guarantees the representativeness of the subsample and still keeps a reasonable variation in the data. Summary statistics on price and market shares of the selected brands are displayed in Table 7.

Table 7: Summary statistics for price and market shares of brands in sample

Variable	Mean	Median	sd	Min	Max
Total					
Price (€/125gr)	0.229	0.187	0.091	0.151	0.463
Market share (% on total sales)	5.113	4.092	4.386	1.816	17.960
Private labels					
Price (€/125gr)	0.174	0.170	0.019	0.151	0.203
Market share (% on total sales)	3.750	4.077	1.120	2.200	4.788
National brands					
Price (€/125gr)	0.277	0.246	0.103	0.165	0.463
Market share (% on total sales)	6.282	4.092	5.828	1.816	17.960

A particularity in this empirical work comes from the fact that loyalty programs are intended to be a supermarket-level program, i.e., they are a marketing strategy to make customers loyal to a supermarket and not necessarily to a particular brand. Therefore, the way to assess the effects of a LP on the demand for a specific product is to think of the supermarket selling the product as a differentiation source. Put other way, each variety should be defined here as a supermarket-brand combination.

I generated an ID variable for identifying each brand variety as the combination of a yogurt brand and the supermarket where it was purchased from. For instance, a Danone yogurt purchased from Casino appears in the database under the composite variety label "Danone-Casino", whereas the same brand purchased from Carrefour is identified under the composite variety label "Danone-Carrefour", meaning that equally branded products are now differentiated varieties due to the retailers. This exercise resulted in more than 120 varieties, but most had a very low market share. Based on this, I kept a final subsample with the leading 31 varieties of brands with market shares varying between 0.9% to 5.8% and an aggregate share of 72.9%, which were purchased from the biggest 6 supermarket chains in France in 2006.[20]

4.3 Variables description

The data set used for the estimation of the models previously described was aggregated to the brand level and contains information on total sales, unit price (per 125gr serving), product and store characteristics and the distribution of the household demographic characteristics. In particular, the following variables were constructed:

[20] The lack of randomness of the final sample considered in this paper does not lead to inconsistent estimates of the parameter as long as I include brand-supermarket dummy variables in the estimation. See Manski and Lerman (1977) and Bierlaire, Bolduc and McFadden (2008) for a detailed discussion on consistent estimation of choice probabilities from choice-based samples.

- *Brand market share* (S_{mjst}): It was computed per subgroup of population of LP members ($m = 1$) and non LP members ($m = 0$), as the percentage of 125gr servings sold in a market (in this paper a Department-week combination) on the total potential number of 125gr portions that could have been consumed in that market.[21] The serving was determined by converting volume sales originally in kg into a 125gr unit which is the size best sold in France. The potential volume sales per market was computed by multiplying the average national plain yogurt consumption of 1.14 125gr servings per person per week in 2006 obtained from the whole data of yogurt purchases by the total population in a department.[22]

- *Market share of the outside good* (S_{0t}): It was defined as the difference between one and the sum of the inside products market shares.

- *Price €/125gr* (p_{jst}): It was generated by dividing the total expenditure on yogurt products over the total servings purchased.

- *LP membership × PL dummy* ($M_{ms} \times PL_{js}$): It is an interaction variable between the membership to LPs indicator and a dummy variable taking on the value 1 if the brand variety is a private label and zero otherwise. It aims at capturing the marginal effect of LP membership on private label demand.

- *Other interactions of interest*: the regressions include other interactions between demographic variables and product characteristics such as: *#Subscriptions × PL dummy*, which combines information on the number of separate LP cards held by a household and a dummy for private label, and *Price × #Subscriptions* which will be useful to see the marginal effects of a price change on loyalty programs' members.

4.4 Estimation

The estimation of the models described in Subsection 4.1 was conducted following the standard methods originally proposed by Berry, Levinsohn and Pakes (1995) and improved some years later by Nevo (2000a, 2001).

As Nevo (2000a, 2001), I control for brand and supermarket fixed-effects by including dummy variables. This implies three main differences with respect to the original estimation algorithm by BLP: *i*) the dummies capture both brand and store unobserved characteristics, *ii*) brand characteristics are not anymore valid instruments, and *iii*) demand is identified without the need to characterize the supply side. I follow the standard literature to estimate my model. Hence, in this section I will briefly describe the generalities of the estimation process and the main particularities of my model. For a detailed discussion of the estimation algorithm and the differences *i*) and *iii*) with BLP procedure, see Nevo (2000a, 2001).

[21] I computed the market shares by subgroup of people in order to preserve the meaning of the membership indicator in a context of aggregate data. This also allows me to see the problem from the viewpoint of the treatment evaluation where the treated are those who subscribed to supermarket s LP and the non treated are those non-member customers.

[22] To be sure this would be a good approximation to the population average consumption of plain yogurt, I looked for an alternative source of information: the data on average consumption on food products from the National Accounts by *Institut National de la Statistique et des Etudes Economiques* (INSEE, Comptes nacionaux, base 2000). Here the reported average quantity consumption of 125gr servings of yogurt per person per week was 3.3217, and provided that around 34% of that number is plain yogurt, the average consumption of this variety per person per week is of 1.287 servings, which is similar to the one reported by the home scan data.

Estimation relies on the population moment conditions given by $E[h(z)'\rho(x,\theta_o)] = 0$, where $z_1, ..., z_M$ are a set of instrumental variables, i.e., a set of variables which are not included in the main regression equation and that are correlated with some (or all) included covariates but not with the error term; ρ is a function of the parameters of the model and θ_o is the true value of the parameters.

In case we have more instruments than needed for identification (i.e. the matrix $h(z)$ has more columns than the matrix of covariates of the model) a generalized method of moments estimator is obtained by solving the problem

$$\min_{\theta} \rho(\theta)' h(z) \hat{\Lambda}^{-1} h(z)' \rho(\theta), \tag{9}$$

where $\hat{\Lambda}$ is a consistent estimator of $E[h(z)'\rho\rho' h(z)]$ and plays the role of the optimal weighting matrix in expression (9).

Now, according to the empirical framework explained before, once supermarket-brand dummy variables are included, the error term of the model is $\Delta\xi_{jt} + \Delta\xi_{st}$ which can be computed as a function of the mean utilities $\delta_{.t}$, the data and the parameters. Following Berry (1994), this computation requires solving first for the $\delta_{.t}$ from the system of equations resulting from the match of observed and predicted market shares

$$s_{.t}(x, r, M, p_{.t}, \delta_{.t}; \theta_2) = S_{.t} \tag{10}$$

where $s_{.t}(\cdot)$ is the predicted market share function defined in (8). As the system in (10) does not have a closed-form solution for the the mixed Logit case, it should be solved numerically.

After inverting (10) in order to express $\delta_{.t}$ as an explicit function of the observed market shares, the error term in (9) can be defined as

$$\rho_{jst} = \delta_{mjst}(x, r, M, p_{.t}, S_{.t}; \theta_2) - (x_j\beta + r_s\lambda - \alpha p_{jst} + \varphi M_{ms} + \eta PL_{js} \times M_{ms} + \xi_j + \xi_s)$$

The estimation of the parameters minimizing the expression in (9) is performed using a non-linear search. Nevo (2000a) was the first to publish a MATLAB computer code with the estimation algorithm, and nowadays several improvements have been proposed by authors who have tried to solve problems such as the accuracy of the estimates, convergence issuess or computational efficiency in the process. In particular, Knittel and Metaxoglou (2012) detected some problems in the optimization algorithm such as convergence at nonoptimal points, multiple local optima and convergence failure, which affected the parameter estimates with consequences for economic predictions. They propose an improved code with eleven alternative optimization routines, the possibility of running the nonlinear search process several times with different sets of starting values (they use 50) and provide several stopping rules playing with tolerance levels. Dubé et al. (2011) offer an alternative way to estimate the model, letting aside the Nested Fixed Point (NFP) algorithm used by previous works and proposing a mathematical program with equilibrium constraints (MPEC) which can be more accurate as it avoids some numerical computations and performs faster compared to the NFP.

In this paper I use Knittel and Metaxoglou's code. The algorithm works as follows: to solve for $\delta_{.t}$, BLP(1995) proposed a contraction mapping which uses starting values for δ and θ_2 and iterates up until it stops at some value of δ determined by a stopping rule provided by the econometrician

$$\delta_{.t}^{(k+1)} = \delta_{.t}^{(k)} + \ln S_{.t} - \ln s_{.t}(x, r, M, p_{.t}, \delta_{.t}^{(k)}; \theta_2), \tag{11}$$

where $\delta_{.t}^{(k)}$ denotes the kth iterate. The starting values for θ_2 can be random draws from a particular parametric distribution, and given those values, a starting value for δ can be obtained from a Logit regression as $\delta_{.t}^{(0)} = \ln S_{.t} - \ln S_{0t}$ which is obtained aftermatching the predicted and obserd market shares and linearizing by taking logs (see Berry, 1994).

With a value for δ and the estimates from the linear part (θ_1) of the model obtained from a 2SLS estimation, the error term appearing in the objective function can be computed and the optimization process can be performed to obtain consistent estimates of θ_2.

4.5 Optimal Instruments

As previously stated, the inclusion of brand-supermarket fixed-effects captures the unobserved brand-supermarket characteristics and then the error term of the model becomes $\rho_{jst}(\theta) = \Delta\xi_{jt} + \Delta\xi_{st}$, which are the group-market deviations from the mean valuation of product and supermarket unobserved characteristics, respectively. Under the assumption that both supply-side agents and customers observe those characteristics and, consequently, their decisions account for these local deviations, we have then two sources of correlation with some explanatory variables.

On the one hand, we have that prices are correlated with the department deviation of the mean valuation of product unobserved characteristics, $\Delta\xi_{jt}$. This problem is always present in this kind of demand estimation models, for which a set of solutions can be found in the standard literature. According to Nevo (2000a), there are two reasons explaining the endogeneity of prices: *i*) differentiated products pricing models assume that firms know the unobserved (by the econometrician) characteristics of the good and use them to set the prices of products, which are a function of marginal cost and a markup depending on demographics, and *ii*) the specification of the model in (1) assumes that product characteristics are the same for all individuals, including the price, which is not consistent with the fact that observed prices paid by consumers are different and thus, leads to measurement error bias. The Logit framework solves the last problem due to the mean utility structure, but the first source of correlation is still present.

On the other hand, $M_{.s}$ seems to be correlated with the local deviation from the mean valuation of supermarket unobserved characteristics, $\Delta\xi_{st}$. Given that this sort of programs are of the firm level[23] and under the assumption that the choice of a LP is driven, among other things, by the supermarket choice, i.e. it depends on the consumer's valuation of supermarket characteristics, there is some information about households subscription decisions contained in the error term.

The key identifying assumption, being the population moment condition described previously, requires to find some valid instruments to deal with those endogeneity problems. Here, I use an approximation to the optimal instruments following BLP (1999) and Reynaert and Verboven (2013).

Instrumental variables estimation for problems with conditional moment restrictions and i.i.d observations were proposed, among others, by Kalejian (1971), Amemiya (1974, 1977), and Jorgenson and Laffont (1974). Amemiya (1977) proposed the computation of the optimal instruments in a by now standard way (see equation (15) below). This developments assumed parametric forms for the error terms. It was Chamberlain (1987) who studied the asymptotic properties of the IV estimator for nonparametric models, where all that we know is that the distribution function of the data satisfies the equality

[23] A customer joins the whole supermarket chain program and not only a given store or a given product category sell by the supermarket.

of the expected value of the residual to zero when multiplied by appropriate functions of the exogenous variables, and that efficiency bounds are attained when these functions are replaced by the optimal instruments. Finally, Newey (1990) proposed nonparametric estimation methods of the optimal instruments for nonlinear simultaneous equations models.

However, inefficient sets of instruments rather than optimal ones have been commonly used by previous research in empirical IO literature to treat the endogeneity of prices. This is the case of BLP (1995), Nevo (2000a, 2001) and Dubé, Fox and Su (2012). BLP (1999) were the first to use a sort of optimal instruments computed in an approximate way, and Reynaert and Verboven (2013) propose a method to get a more accurate version of Chamberlain's instruments for random coefficient models.

Following Newey (1990), consider an econometric model with the following moment restriction[24]

$$E[\rho(x_i, \theta_o)|z_i] = 0 \tag{12}$$

where $\rho(x, \theta)$ is a $Q \times 1$ residual vector, z are instrumental variables and θ is a $p \times 1$ vector of parameters where θ_o stands for the true value of this set of parameters, and $x_1, ..., x_n$ are i.i.d. observations on the data vector x_i where z_i makes part of its components. Assuming homoscedasticity, the variance-covariance matrix conditional on the instrumental variables is

$$E[\rho(x, \theta_o)\rho(x, \theta_o)'|z_i] = \Omega \tag{13}$$

where Ω is constant. Estimation of the parameters of the model rely in principle on the conditional moment restriction in (12). However, it implies that we have an infinite number of moments one of each corresponding to a given set of z_i. This problem can be solved by using functions of the data and the parameters, which reduces the set of equations to a finite set of ordinary moment restrictions. Let $h(z_i)$ be this function, then estimation relies now on the unconditional moment

$$E[h(z_i)\rho(x_i, \theta_o)] = 0 \tag{14}$$

The optimal instruments are a particular function $h(\cdot)$ that allows us to consistently estimate the parameters of the model and attain the efficiency bound of the asymptotic variance-covariance matrix. This is[25]

$$h(z_i) = D(z)'\Omega^{-1} \tag{15}$$

where

$$D(z) = E\left[\left.\frac{\partial \rho_{\cdot t}(x, \theta_o)}{\partial \theta}\right| z_i\right] \tag{16}$$

According to Newey (1990), for the particular case of a single equation system where the residual is a scalar (as is the case of the demand model described in subsection 4.4

[24]For the sake of exposition in the general formulation of optimal instruments I replace panel subscripts by a single subscript i indicating a particular observation of the data. I will go back to the usual notation when I derive the particular optimal IVs for the model in this paper.

[25]For a complete discussion about IV estimation methods of nonlinear models, optimal instruments and efficiency bounds for nonlinear models and other models of interest, see Newey (1990).

of this paper) the optimal instruments become $h(z) = D(z)'$ as in (16).[26] The number of instruments is equal to the number of parameters to be estimated in the model $\theta = (\theta_1, \theta_2)$, where θ_1 is the vector of parameters in the linear part of the model and θ_2 is the vector of nonlinear parameters, that is

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \beta'}\bigg| z_t\right] = E[x_j|z_t] = x_j \tag{17}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \alpha}\bigg| z_t\right] = E[p_{jst}|z_t] = x_j\gamma_1 + r_s\gamma_2 + w_{jst}\gamma_3 \tag{18}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \varphi}\bigg| z_t\right] = E[M_{ms}|z_t] = r_s\tau_1 + l_s\tau_2 \tag{19}$$

$$\begin{aligned} E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \eta}\bigg| z_t\right] &= E[M_{ms} \times PL_{js}|z_t] \\ &= E[M_{ms}|z_t] \times PL_{js} = (r_s\tau_1 + l_s\tau_2) \times PL_{js} \end{aligned} \tag{20}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \theta_2'}\bigg| z_t\right] = E\left[\frac{\partial \delta_{mjst}(s_{.t},\theta_2)}{\partial \theta_2'}\bigg| z_t\right] \tag{21}$$

The instruments resulting for the identification of β are just product observed characteristics which are assumed to be exogenously set by producers and supermarkets, i.e., those attributes do not vary in response to department-specific demand shocks. As for the parameters of the endogenous variables α, φ and η, the instruments are the corresponding predicted values from first-stage OLS regressions for $p_{.t}$ and $M_{.s}$.

To compute, first, the predicted price in (18), I follow Reynaert and Verboven (2013) and assume for simplicity that marginal costs are linear and depend on product and store characteristics and cost shifters, $w_{.t}$, and that markets are competitive so that firms set prices at marginal cost.[27] However, as I do not observe any cost shifters in my homescan data set, I use average regional prices of the same product in all the 21 French administrative regions (excluding the department to be instrumented from the average price of the region it is located in) as proxies for marginal costs information. Following Nevo (2001), I claim that after controlling for brand-specific means, regional-specific valuations are independent of product valuations of people from other regions. This implies that in case a demand shock happens in one region, only the local price will be affected and the others will remain equal. This guarantees the exogeneity condition of prices. Now, prices of two departments in a country are linked by common marginal costs as long as they are produced (supplied) by the same manufacturer (retailer) or under a standardized process, reason why I believe prices of the same brand in other departments contain useful information on common marginal costs.[28] Table B.1 in the Appendix displays first-stage results of price regressed on average regional prices.

[26]The error term denoted by $\rho(x, \theta_o)$ corresponds to a general formulation of a model with simultaneous equations. As the model to be estimated here consists of a single equation, the $\rho(\cdot)$ function equals the error term of the demand model $\Delta\xi_{jt} + \Delta\xi_{st}$. For BLP models, for instance, where demand and supply equations are estimated simultaneously, the residual is a vector containing both the demand- and the supply-side error terms, $\rho = (\xi, \omega)'$.

[27]Reynaert and Verboven (2013) examine both perfect and imperfect competition cases and obtain similar results.

[28]Although the independence assumption seems reasonable, there may be cases were it cannot hold as, for example, a national demand shock as pointed out by Nevo (2001).

Second, the predicted value for the subscription decision is assumed to be a linear function of supermarket characteristics and the characteristics of the loyalty program.[29] Here, however, the solution cannot be as elaborated as for the problem with prices due to the lack of information, namely, I do not observe in the data set neither household consumption patterns before subscription nor any other additional information on the intensity of purchases motivated by loyalty rewards, the effective amount of rewards obtained by households and the rate of coupon redemption.

Given this, I collected myself some information on the characteristics of each loyalty program and use them to have a very rough approximation to the predicted subscription decision indicator $\hat{M}_{.s}$. Table B.2 in Appendix displays estimation results for the regression of M_{is} as dependent variable on some covariates describing the LP characteristics set by supermarkets, i.e. exogenous for the consumers, and controls for demographics. Although the information collected is not enough to fully account for the variation in M_{is}, the results are quite appealing showing that the exogenous structure of the programs have some explanatory power. Under the argument that supermarket LPs are the same over the whole country, most LP characteristics are not set in response to market specific deviations of the mean valuation of the program, hence, they can be used as exogenous instruments for a first-stage estimation of the model.

Finally, the instruments for the nonlinear parameters in (21) are more difficult to compute as they are nonlinear functions of product characteristics and the expectation in (21) is a function of the true parameters of the demand function, namely, both linear θ_1 and nonlinear θ_2. This means that the instruments for the nonlinear parameters of the model cannot be directly computed from the data and require instead a first-stage estimation of the model. Here I follow Reyneart and Verboven (2013) notation and compute the "optimal" instruments à la BLP (1999), i.e., the population moment in (21) is replaced directly by an approximation of the Jacobian for the delta function and not by the empirical analogue of the moment restriction[30]

1. Obtain an initial estimate for $\theta = (\theta_1, \theta_2)$. I used the sets of explanatory variables previously described as instruments.

2. Compute the predicted price $\hat{p}_{.t}$ and membership indicator $\hat{M}_{is}$ as in equations (18) and (19) respectively.

3. Retrieve the predicted mean utility as $\hat{\delta}_{mjst} \equiv x_j\hat{\beta}+r_s\hat{\lambda}-\hat{\alpha}p_{jst}+\hat{\varphi}M_{ms}+\hat{\eta}PL_{js}\times M_{ms}$ and use it to recover the predicted market shares $\hat{s}_{mjst} = s_{mjst}(\hat{\delta}_{.t}, \hat{\theta}_2)$.

4. Compute the Jacobian of the inverted market share system $\delta_{mjst}(\hat{s}_{.t}, \theta_2)$ as

$$\left.\frac{\partial \delta_{mjst}(\hat{s}_{.t}, \theta_2)}{\partial \theta_2'}\right|_{\theta_2=\hat{\theta}_2} \tag{22}$$

As $\delta_{.t}$ is an implicit function of the system of equations defined in (10), the Jacobian is computed as the product of two matrices containing the derivatives of the predicted

[29] Demographics are definitely key factors determining the decision to join a given LP. However, in order to have valid instruments I exclude any variable correlated with the error term of the model.

[30] Reynaert and Verboven (2013) provide an algorithm to compute the optimal instruments more accurately by replacing the population moment by its empirical analogue and compare its performance with those of BLP (1999) using montecarlo simulations. They conclude that the gains in efficiency are small when using their proposed method whereas the computational burden increases, reason why I stick to the BLP (1999) approximate version of the optimal instruments.

market shares with respect to the mean valuation $\delta_{.t}$ for each market and the derivatives of the market shares with respect to the nonlinear parameters. Nevo (2000a) provides an Appendix with all the details and a MATLAB code for the computation of the Jacobian, which is originally used in the GMM optimization process and the computation of the variance-covariance matrix of the estimates.

4.6 Logit results

Let S_{mjst} be the observed market share of brand j in supermarket s by subgroup of population m in market t. The market shares were computed, as indicated previously, using the potential consumption of 125gr servings of plain yogurt by the whole population in every market t as the potential market.

Table 8 displays the results for the estimation of a Logit demand model for yogurt by regressing $ln(S_{mjst}) - ln(S_{0t})$ on the variables described in subsection 4.3 as main covariates. Additionally, depending on the estimation method I control for product characteristics (columns (1) and (4)), brand fixed effects (all columns but (1) and (4)), and demographics (columns (3), (6) and (7)). Columns (4)-(7) in Table 8 display the results of 2SLS regressions using instrumental variables to treat the endogeneity of prices. As the purpose of this subsection is purely descriptive, I deal just with the endogeneity of prices using sets of inefficient instruments: in column (4) I use brand dummy variables, which play a similar role as BLP (1995) instruments, and in columns (5) through (7) I use the set of average regional prices as in Hausman (1996) and Nevo (2000a, 2001). Column (4) includes brand characteristics, whereas columns (5) through (7) include brand fixed-effects, and that is why it is no longer possible to use brand dummies as instruments. Even though I do not instrument for the endogeneity of LP membership for now, I use characteristics of each supermarket LP as controls in all regressions.

For every IV regression I conducted a Hausman test of over identifying restrictions, which always rejects the null hypothesis of the exogeneity of instruments, even in column (7) where demographics were replaced by department dummies which are supposed to capture local demand shocks in a better way. An explanation could be that given that the test is distributed as a chi-square, the large number of observations will cause any model to be rejected (Nevo, 2001). However, the IVs are individually and jointly significant at 1% level and the high first-stage R-squareds and F-statistics suggest that they have some power (Nevo, 2001).

A result of special interest is the estimate for the interaction $LPmember \times PLdummy$. In all regressions, the coefficient is positive and significant, and the estimate does not vary importantly when adding demographic controls, further interaction variables or when the prices are instrumented. The coefficient means that the marginal valuation of PL products increases with the subscription to the supermarket owning the brand. In other words, LP subscription has a positive impact on the demand for PLs, supporting the hypothesis of supermarket chains benefiting from this kind of programs as a way to boost PL demand.

Additionally, regressions (3) and (6) include the interactions $Price \times \#Subscriptions$ and $\#Subscriptions \times PLdummy$. On the one hand, the former has a negative and significant coefficient but it will be positive in the full model with optimal instruments, which suggests that there is a downward bias in the estimate maybe coming from the use of inefficient instruments for price. On the other hand, the latter interaction has a negative coefficient indicating that the marginal valuation of PL products decreases with the number of LP subscriptions. This means that multi subscription weakens the effects of LPs on private label demand.

Table 8: Results from Logit for PL demand[a]

Variable	OLS (1)	(2)	(3)	IV (4)	(5)	(6)	(7)
Price(€/125gr)	-4.345	-6.216	-4.502	-11.18	-7.335	-5.965	-6.386
	(0.094)	(0.114)	(0.159)	(1.181)	(0.280)	(0.463)	(0.196)
LP membership	-0.298	-0.282	-0.252	-0.438	-0.300	-0.281	-0.268
	(0.018)	(0.016)	(0.015)	(0.032)	(0.017)	(0.018)	(0.011)
PL dummy	-0.405	-0.140	-0.504	-1.149	-0.460	-0.637	-2.049
	(0.025)	(0.052)	(0.062)	(0.131)	(0.088)	(0.074)	(0.066)
LP membership×PL dummy	0.255	0.236	0.210	0.366	0.249	0.235	0.151
	(0.025)	(0.023)	(0.022)	(0.034)	(0.023)	(0.023)	(0.015)
Plastic	-0.451			-2.024			
	(0.034)			(0.273)			
Sugar	-0.125			-0.402			
	(0.020)			(0.051)			
Wholemilk	0.133			0.155			
	(0.011)			(0.012)			
#Subscriptions×PL dummy			-0.0561			-0.0974	
			(0.015)			(0.019)	
Price×#Subscriptions			-0.696			-0.396	
			(0.056)			(0.102)	
Average HH size			0.181			0.167	
			(0.009)			(0.010)	
Log income			-1.055			-1.050	
			(0.023)			(0.023)	
Car			1.306			1.287	
			(0.064)			(0.064)	
# stores visited the			0.110			0.105	
same week			(0.038)			(0.038)	
# of trips to the same			0.0357			0.0412	
store within a month			(0.010)			(0.010)	
Constant	-12.28	-35.37	-17.45	-5.888	-34.78	-16.91	10.54
	(0.284)	(0.560)	(0.680)	(1.128)	(0.576)	(0.698)	(0.801)
Fit/Test of over	0.142	0.235	0.297	5,337	138.2	131.4	79.9
Identification[b]				(1.145)	(10.851)	(10.851)	(10.851)
1st Stage R^2				0.749	0.841	0.909	0.843
1st Stage F-test				1,676	1,896	3,351	1,019
Instruments				Brand dummies	Prices	Prices	Prices

[a]Dependent variable $ln(S_{mjst}) - ln(S_{0t})$. Based on 37,662 observations. All parameters are significant at 5% level. All regressions include week dummy variables and with the exception of columns (1) and (4) all regressions include brand dummy variables. The regression in (7) includes department dummy variables. Asymptotically robust s.e. are reported in parentheses. All regressions include characteristics of each supermarket's loyalty program as controls.
[b] Adjusted R^2 for the OLS regressions, and a Hausman test of over identification for the IV regressions with the 0.95 critical values in parentheses.

4.7 Results from the Mixed Logit model

Table 9 displays the results of the full model with "optimal" instruments. A first-stage version of the random coefficients model with the same specification was estimated with inefficient instruments for price (average regional prices) and loyalty membership (LP characteristics) in order to be able to compute the set of 20 "optimal" instruments which is the same number of parameters to be estimated (see Table B.3 in Appendix).

The first column contains the estimated means of the distributions of the individual marginal utilities. They are all significant and most preserve the same sign as in the descriptive Logit model. The interaction between LP membership and the PL dummy is positive, supporting the previous result of a positive impact of loyalty programs on private label demand. Provided that promotions are mainly addressed to those store-branded products, customers are more willing to consume them as members of a LP compared to non members. On the contrary, the coefficients for the variables in levels, namely PL dummy and LP membership dummy, are both negative, which reflect the fact that people in general people value less store branded products with respect to NBs and that being member of a loyalty program implies some costs, respectively.

The second column displays the estimated standard deviations of the mean coefficients referred above. Put other way, these estimates are the coefficients of the interactions of the right-hand side variables of the model with the unobserved demographics. All but two are significant (those for *Plastic* and *Sugar*), meaning that the unobserved demographics v has some explanatory power for the heterogeneity in consumer tastes. As for *Plastic* and *Sugar*, the non significance of the estimates mean that included demographics are enough to explain the variation in consumer tastes.

Most included demographics have significant estimates. One interesting result is the negative coefficient of the interaction between PL dummy and income, confirming the intuition that higher income households value less PL products relative to lower income ones. The interactions with the total number of loyalty cards held by households (*#Subscriptions*) are also interesting. They are all negative, meaning in the first case (interaction with the constant) that the more memberships a household has, the higher the costs they face, this is consistent with the estimate of the indicator for LP membership.[31] In the second case, the interaction with price show that multi-subscribers are more price sensitive as a marginal increase in price would have a larger impact for those holding more cards. The intuition behind this result is that the retention effect of a LP is weaken by the fact that a consumer holding several cards can substitute now among supermarkets without the cost of losing the points for a future rebate provided that in other places she might as well get them. Last, but at least, is the coefficient of the interaction of the number of subscriptions and PL dummy: a negative estimate indicates that the more memberships the consumer has, the lower the valuation for private label products, mitigating the positive impact that LP have in general on the demand for PLs.

Table 9: Results from the Mixed Logit model[a]

	Means	Std. Deviations	Interactions with Demographic variables		
Variable	(β's)	(σ's)	HH size	Income	# Subscriptions
Constant[b]	-10.067	0.478		0.131	-0.968
	(0.043)	(0.382)		(0.236)	(0.346)
Price	-12.963	1.057	0.505	0.900	-0.535
	(0.453)	(0.485)	(0.132)	(0.085)	(0.255)
LP member (M_{ms})	-0.825	0.685		0.182	
	(0.207)	(0.286)		(0.536)	
PL dummy[b]	-11.063	0.916	-0.862	-0.446	-1.243
	(0.025)	(0.192)	(0.357)	(0.229)	(0.255)
LP member×PL dummy	0.584	0.765			
	(0.211)	(0.245)			
Plastic[b]	10.877	0.144			
	(0.060)	(0.406)			
Sugar[b]	6.840	0.171			
	(0.053)	(0.409)			
Wholemilk[b]	-1.513	0.466			
	(0.013)	(0.156)			
GMM Objective	1.69E-06				
MD χ^2	9707686251				

[a] Based on 37,662 observations. Except where noted, parameters were estimated using GMM. All regressions include brand and week dummies. Asymptotically robust s.e. are given in parentheses.
[b] Estimated using a minimum-distance procedure.

[31] I claim that enrolling in a loyalty program is costly (the time expended in getting information about how the loyalty program works, filling up forms, etc., the fact that you need to give your personal information to a supermarket with the advertising and e-mail spamming consequences, taking the card in the wallet in the like. Although for most people these costs might be very low with respect to the expected benefits, I believe they are not negligible and this is what the estimates are reflecting.

5 Demand response to changes in LP membership

In this Section I show the results of two counterfactual experiments based on the previous estimates. Under the assumptions that both prices and unobserved product characteristics do not respond to the membership decisions of the consumers (at least in the short run), and that the utility from the outside good remains the same, I consider two scenarios.[32] First, all the households in the sample are members to at least one loyalty program, i.e. I set M_{ms} in equation (6) to one for all $m = 0, 1$ and replace the demographics variable *#Subscriptions*, defined in subsection 4.3, by one for those who were non LP members. Intuitively, this situation may arise when either the reward system of a LP is good enough so that those customers not having strong preferences for such programs are better off joinning or when real and perceived costs of subscription are zero (or even negative).

Second, I assume no household is member of a loyalty program whatsoever, i.e. I set M_{ms} in equation (6) to zero for all $m = 0, 1$. Such a case may arise when consumers perceive either subscription as prohibitively costly (because supermarkets set too high a membership fee[33]) or rewards are not worthwhile or just unattainable.[34]

Table 10 displays the results of the change in aggregate demand for each counterfactual scenario with respect to the baseline predicted demand. Interestingly, the model predicts that when everybody is member of at least one loyalty program ($M_{ms} = 1 \, \forall \, m = \{0, 1\}$) aggregate demand for private labels increases in 75.8% for people who were non members before the change.[35] The demand for national brands increases also by 1.4%, which may be interpreted as an indirect effect of loyalty programs. This result supports the initial hypothesis that LPs may be used as a marketing strategy to boost PL demand, apart from other objectives.

As for 'Loyals', i.e. those who were already LP members, demand seems to not be affected as it decreases by less than 1% with respect to the predicted demand in the baseline scenario. This is in line with what is expected as 'loyal' members may not have additional benefits but worse conditions instead as supermarket may reduce the benefits of LP once everybody has joined.[36]

In the second scenario ($M_{ms} = 0 \, \forall \, m = \{0, 1\}$) overall aggregate demand decreases with respect to that in the baseline scenario. Both demand for PL and NB decrease for

---

[32]Due to the data limitations referred previously, this is the best I can do to exploit my model and the data. The results of the demand side might as well be exploited to recover retailers' price-cost margins according to some assumptions on the conduct model of the industry, as in previous literature. However, the nature of the grocery retailing sector where the supermarkets are not just distributors but also rivals of upstream suppliers, and the fact that the presence of PL implies both vertical differentiation between PL and NB and horizontal differentiation across PLs, adds new dimensions to the problem that cannot be treated in the standard way. A structural model addressing all these features is needed. This is out of the scope of this paper and is left for future research.

[33]In some retailing sectors in France such as clothing or department stores, loyalty programs giving permanent and special rebates, exclusive offers, and other benefits are offered to customers who should pay a subscription fee either once or yearly that can go up to 30 € in some cases.

[34]Some loyalty programs require customers to pay a fraction in cash of the full price of the reward, such as FPP. Airlines ask FPP members to do a large number of "qualifying" (generally international) flights or to accumulate a given number of miles (which use to be high) in a short period of time so as to reach better status and enjoy extra benefits. Some set short deadlines to expend the accumulated miles also.

[35]To obtain the annual aggregate demand I proceed the following way: first, I compute the per brand-market (week-Department) aggregate demand, q_{jst}, as the predicted market share of each brand in a market, s_{jst}, times the size of that local market (total number of consumers in a Department in a week), $\mathcal{M}_t$. Then, I sum up local per-brand demands by brand across markets to get the annual aggregate demand per brand, q_{js}. Finally, I aggregate across brands to obtain the total annual demand of yogurt.

[36]Recall that I am assuming that prices and supermarket-brand unobserved characteristics remain unchanged in the counterfactual scenarios.

the two groups. However, the impact is lower than that of the fisrt scenario. Loyals' aggregate demand for PL decreases by 27.2% as well as non-loyals' which decreases by 18.1%. The intuition behind this is that the existence of the loyalty programs make PL more interesting for both groups members and non-mebers, even if the latter do not get any rewards from them, so that when they desappear everybody substitute PL for NB and the outside good. Similarly, NB demand decreases for both groups indicating that there might be an effect of loyalty programs over the whole the product category: the idea that some yogurt brands purchases may give you rebates and other rewards, leads customers to increase their demand in general for the product category.

Table 10: Change in aggregate demand as a result of changes in LP market coverage (in % per year)

Baseline	Total		Private label		National brand	
group	$M_{ms} = 1$	$M_{ms} = 0$	$M_{ms} = 1$	$M_{ms} = 0$	$M_{ms} = 1$	$M_{ms} = 0$
Non-loyals	29.70	-11.17	75.75	-18.08	1.37	-6.91
Loyals	-1.74	-17.30	-0.89	-27.22	-2.64	-6.72
Total	1.50	-16.67	5.09	-26.51	-2.13	-6.74

Column headers indicate counterfactual scenarios: everybody is member of LP, $M_{ms} = 1 \, \forall \, m = \{0, 1\}$, and nobody is, $M_{ms} = 0 \, \forall \, m = \{0, 1\}$. Row labels stand for the two original population subsamples according to their membership status.

I also compute the change in total consumer surplus (CS) for each scenario. Following Train (2009), the expected change in consumer surplus for individual i, provided that the price coefficient, $\tilde{\alpha_i}$, do not depend on income, i.e. the coefficient does not change when either income or price changes,[37] can be easily calculated as:

$$\Delta\mathbb{E}(CS_i) = \frac{1}{\tilde{\alpha}_i}\left[\ln\left(\sum_{s=1}^{S}\sum_{j=1}^{J}\exp(\delta^1_{mjst})\right) - \ln\left(\sum_{s=1}^{S}\sum_{j=1}^{J}\exp(\delta^0_{mjst})\right)\right], \tag{23}$$

where δ_{msj} is defined by (4) and the superscripts 0 and 1 make reference to before and after the change in loyalty membership, respectively. The total mean change in consumer surplus is obtained by taking the average (11) over the whole sample and multiplying by the size of the national market which is the total population in France in 2006 (according to the official statistics by the INSEE, 61.1 million people) times 52 weeks (Nevo, 2000b).

Table 11 displays the results of the change in total consumer surplus for each of the two simulated cases. In the first counterfactual scenario (first column of Table 11), where LP market is fully covered, former non-loyals' CS increases by about 1.5% (36 million euros a year), whereas that of former loyals slightly decreases (0.6%). A similar trend is shown by the results of the second counterfactual scenario. The former non-members are not affected by the fact that now no one is willing to subscribe to a LP. However, former loyals are worse off with a reduction in total CS by 3.1% (49.4 million euros a year). Overall, customers are better off when everybody subscribes to at least one loyalty program, with a predicted increase of 13 million euros in CS, and worse off if nobody subscribes, with a predicted decrease of 26 million euros in total CS.

[37] For a complete discussion on this, see Train (2009). For the computation of consumer surplus when the marginal utility of income varies with income, see McFadden (1999).

Table 11: Change in Consumer Surplus as a result of changes LP market coverage (in millions of € and percentage per year)

	$M_{ms} = 1$		$M_{ms} = 0$	
Group	millions of euros	%	millions of euros	%
Non-loyals	36.06	1.54	-3.48	-0.15
Loyals	-9.43	-0.59	-49.44	-3.11
Total	13.31	0.68	-26.46	-1.34

Column headers indicate counterfactual scenarios: everybody is member of LP, $M_{ms} = 1 \, \forall \, m = \{0, 1\}$, and nobody is, $M_{ms} = 0 \, \forall \, m = \{0, 1\}$. Row labels stand for the two original population subsamples according to their membership status.

6 Conclusions and further research

This paper studies the effects of supermarket loyalty programs on private label demand. Prior research has concluded, among other things, that loyalty programs are generally used to retain customers as it induces repeat buying, and as a discriminatory device. Evidence also suggests that retailers may be using LPs as a way to boost PL demand. An empirical fact is that PL products are in average 20% of lower price relative to quality equivalent national brands. Yet, a common feature of supermarket loyalty programs is to reward private label purchases. This motivates the question I hoped to address in this paper: Why do profit-maximizing retailers give additional rebates for their lower-price own-brands?

This article adds to the literature of both private label and loyalty programs topics by using advanced methods of demand estimation, and provides empirical support to a question that, to the best of my knowledge, have not yet been addressed. Using a random coefficients Logit model to estimate a brand-level demand system for yogurt in France, I find that loyalty program members have a higher valuation for PL as compared with non members, despite that my results confirm the general believe that store brands are perceived as of lower quality relative to similar national brands. When multi subscription is present, the effects of LPs are weaker: marginal valuation of private labels decreases with the number of subscriptions to different supermarket LPs and customers are more price sensitive.

By conducting some counterfactuals on the demand side, assuming exogenous changes in LP subscription, I find evidence that unambiguously supports the hypothesis of this paper that LP have a direct impact on PL demand. In fact, aggregate demand for private labels of those people who formerly were not members of any LP increases by more than 75% when they become members. On the other hand, in the absence of LPs (second scenario) private labels become less attractive products as aggregate demand decreases by 26%. Welfare analysis shows that consumers are in general better off when they all join at least one LP and worse off when no one is member to LPs, as compared with the baseline scenario. Results suggests that making subscription to LPs prohibitively costly may harm all consumers, even those who are non members in the baseline scenario.

Due to very limited data on loyalty programs, this paper cannot go further on the evaluation of LP effects on customer purchasing behavior. However, the results can be used as a motivation for pursuing further research on these topics in case richer data sets are available. In particular, with the help of descriptive results this paper sheds some light on questions that are worth asking. For instance, with the appropriate data modelling customer subscription decisions to supermarket LPs would help to understand why, if subscription is free and economic intuition indicates that each consumer able to

subscribe should do so, not everybody subscribes or, among those who subscribe, not everybody does it to as many programs as available in the market. In fact, in France 15% of customers are not members of any supermarket LP whatsoever and, among the members, 20% are affiliated to only one supermarket. Another interesting question to be explored is what are the effects of LPs on supermarket choice. This paper provides some descriptive results suggesting that members of a given supermarket loyalty program expend in average a higher portion of their income in that supermarket compared to non member customers. I believe that in many cases supermarket choice comes first when people need to go shopping, and brand choice comes next. In such cases LPs may play a key role in consumer's decision, as they are supermarket-level rather than brand-level marketing devices.

Further research imply going deeper in the meaning of making customers loyal to a store or a particular brand and understanding the role played by loyalty programs in this context. By adopting a measure of loyalty, such as the share of wallet or the shopping intensity, it can be explored whether LPs are making customers loyal. A question that remains to be answered is thus: are non members more loyal than members?

Finally, from a policy perspective this paper may be seen as a first step for assessing more complex and interesting questions: Are loyalty programs being used by retailers as a strategy to strengthen their buyer power? Moreover, taking into account the theory result that loyalty programs allow retailers to increase PL prices, it is worth asking: Can loyalty programs be harmful for consumers at some point? Should policy authorities be cautious about the implementation of such marketing strategies? In order to answer such questions a structural model of the supply side is needed, which should include the horizontal (competition among retailers with their own brands) as well as the vertical dimension (competition among retailers and manufacturers, i.e, PL against NB). A counterfactual experiment on the supply side would allow to assess whether retailers are trying to exert some buyer power or not, under the premise that retailers might use their bargaining power to squeeze manufacturers' margins. I hope to think further about this in the future.

Appendix

A A theoretical background

This is a two-period model of oligopolistic competition with horizontal differentiation.[38] Consider two symmetric retailers ($j = 1, 2$) differentiated à la Hotelling and located at the two extremes of a line of length one. They supply the good at a constant marginal cost $c > 0$ and compete for a mass of customers indexed by $i \in [0, 1]$ uniformly distributed along the Hotelling line. We assume unit demands, i.e., each consumer buys at most one unit of the good in each period. Moreover, we assume the good is not storable.[39]

Consumer location x_t is independently and identically distributed over time. In particular, consumers cannot anticipate their second period location in period 1.[40]

At the beginning of the first period, retailers offer a loyalty program (LP) to consumers consisting of a coupon B of a given amount of money determined by period 1 purchases and redeemable in period 2 in the same retailer's stores. This rebate is given to a consumer conditional on his subscription to the retailer's LP in the first period. We assume, for simplicity, that consumers can only subscribe to one loyalty program and that the market is fully covered, i.e., the utility of joining a LP is large enough so that it is better to subscribe than the outside option.[41]

Finally, we assume non-strategic forward-looking consumers with rational expectations. The structure of the model defines a two-stage game as follows:

- *First stage:* Retailers determine simultaneously first-period prices and loyalty rebates to maximize their expected two-period profit function.

- *Second stage:* Prices, loyalty rebates and first-period market shares are realized. Each firm then decides her second-period price to maximize her second-period profit taking into account her subscribers and the eventual "non-loyal" customers (switchers) she could have in every period.

[38] In their 1990 paper R. Caminal and C. Matutes, who analyse the effects of alternative pricing policies on duopoly competition with endogenous switching costs, consider the commitment to a future rebate through a coupon as one of the pricing policies. The results coincide with those I obtain in this Section with the difference that they derive the main results computationally whereas I solve for the equilibrium tariffs analytically. I hope this will provide a richer insight. In any case, this is not the main contribution of this paper and is intended to help mainly to derive objective economic intuition to frame the empirical analysis.

[39] Given the structure of the food retailing industry in which there are commonly more than two firms, one can argue that a circular city framework might better model this problem. However, for the purpose of this paper, where we try to explain consumer choice under the existence of loyalty rebates instead of a full entry-location game, a linear city model with two differentiated retailers seems to be enough. Moreover, the empirical evidence shows that in average most people visit at most two different stores when they go shopping.

[40] Although this assumption is admittedly made for simplicity in order to introduce uncertainty in location and to avoid consumers anticipating second-period results from the first period, there are intuitive interpretations of it. We can think, for example, of a person who remembers running out of toothpaste in the morning when going out of her workplace, and go to buy to the closest supermarket or on the way home; another example is that of a person having a craving for pizza after work and stops by to buy one on the way home. We can also think of this assumption as uncertainty of tastes depending on the day or the weather: maybe in a cold-rainy day I can feel like eating lasagna but in a sunny day I would prefer eating salad or gazpacho instead.

[41] The subscription to a given LP does not imply that consumer has committed herself to repeat purchases from the same retailer in the second period. As a consequence, consumers have always the choice to go to the retailer with the lower price given their second period location and the size of B, which becomes a switching cost for those who decide to go to a different retailer.

We are looking for subgame-perfect equilibria.

Definition 1. *The set of choices* $\{p_{11}^*, B_1^*, p_{21}^*\}$ *for firm 1 and* $\{p_{12}^*, B_2^*, p_{22}^*\}$ *for firm 2 is said to be a subgame-perfect equilibrium of the game if*

- *Second stage: For any given first-period prices and loyalty rebates* $\{p_{11}, B_1; p_{12}, B_2\}$ *and for any given first-period market shares* $\{\alpha_1(p_{11}, p_{12}), \alpha_2(p_{11}, p_{12})\}$*, the prices* p_{21}^* *and* p_{22}^* *constitute a Nash equilibrium.*
- *First stage: Given the second-period optimal prices* $\{p_{21}^*, p_{22}^*\}$ *and market shares* $\{D_1, D_2\}$*, the quadruple* $\{p_{11}^*, B_1^*; p_{12}^*, B_2^*\}$ *is a Nash equilibrium.*

We are interested in the set of interior equilibria, therefore we impose an upper bound on B_j, which must be small enough so that the consumers' demands remain interior:

$$0 \leq B_j \leq \bar{B}$$

A.1 Second period

Let us assume $\{p_{1j}, B_j\}_{j=1,2}$ as given. Moreover, let us assume that a mass $\alpha_j \in (0,1)$ of consumers subscribed to firm j's loyalty program in the first period. The remaining $(1-\alpha_j) \equiv \alpha_k$ subscribed to rival's LP. α_j and α_k denote, hence, firm j, k's first-period market share, respectively.

The utility of a consumer i located at x_2 who has joined firm j's LP is:

$$\begin{array}{ll} u - (p_{2j} - B_j) - \tau x_2 & \text{if he continues buying from retailer } j \\ u - p_{2k} - \tau(1 - x_2) & \text{if he switches to retailer } k,\ \forall k = 1,2 \end{array}$$

where u is consumer i's valuation of the good and τ is a transportation cost parameter which is constant over time and symmetric across consumers.

Assuming that the consumer who is indifferent between the two retailers is located at $x_2 = D_j$, firm j's market share is defined by:

$$p_{2j} - B_j + \tau D_j = p_{2k} + \tau(1 - D_j)$$

Solving for D_j yields firm j's second-period market share:

$$D_j(p_{2j}, p_{2k}, B_j) = \frac{1}{2} + \sigma\left[p_{2k} - (p_{2j} - B_j)\right],\ \forall j, k = 1,2 \tag{24}$$

where

$$\sigma \equiv \frac{1}{2\tau} \tag{25}$$

is a parameter indicating the degree of substitutability between the two retailers.

Firm j's profit maximization problem writes as:

$$\max_{p_{2j}} \quad \pi_{2j} = (p_{2j} - c)(\alpha_j D_j + (1 - \alpha_j) D_k) - \alpha_j D_j B_j$$

FOC:

$$\alpha_j D_j + (1 - \alpha_j) D_k + (p_{2j} - c)\left[\alpha_j \frac{\partial D_j}{\partial p_{2j}} + (1 - \alpha_j)\frac{\partial D_k}{\partial p_{2j}}\right] - \alpha_j \frac{\partial D_j}{\partial p_{2j}} B_j = 0 \tag{26}$$

Plugging $\frac{\partial D_j}{\partial p_{2j}} = \frac{\partial D_k}{\partial p_{2j}} = -\sigma$ into (26) and solving for p_{2j} yields:

$$p_{2j} = \frac{c}{2} + \frac{\tau}{2} + \alpha_j B_j + \frac{p_{2k} - (1-\alpha_j)B_k}{2} \tag{27}$$

By the symmetry of the model, p_{2k} has a similar expression to the previous equation. Plugging it to (27) and solving for p_{2j} we have:

$$p_{2j}^*(\alpha_j) = c + \tau + \alpha_j B_j,\ \forall j = 1,2 \tag{28}$$

or

$$p_{21}^*(\alpha_1) = c + \tau + \alpha_1 B_1 \quad \text{and} \quad p_{22}^*(\alpha_2) = c + \tau + \alpha_2 B_2 \tag{29}$$

Note that if we set $B_j = 0$, we get $p = c + \tau$ which is the standard Hotelling price with unit demands.

A.2 First period

A consumer located at x_1 will have the following instantaneous utility:

$$\begin{array}{ll} u - p_{1j} - \tau x_1 & \text{if he goes to retailer } j \\ u - p_{1k} - \tau(1-x_1) & \text{if he goes to retailer } k \end{array}$$

Conditional on buying from retailer j in the first period, he will buy again in the second period to this retailer if x_2 satisfies:

$$x_2 \leq D_j \equiv \frac{\tau + p_{2k} - p_{2j} + B_j}{2\tau}$$

The consumer i's expected second-period surplus is then:

$$\delta\left[\int_{x=0}^{D_j} (u - p_{2j} + B_j - \tau x)dx + \int_{D_j}^{x=1} (u - p_{2k}(\alpha_k) - \tau(1-x))dx\right] \tag{30}$$

performing the integrals and rearranging, yields:

$$\delta\left\{(p_{2k} - p_{2j} + B_j + \tau)D_j - \tau D_j^2 + u - p_{2k} - \frac{\tau}{2}\right\}$$

Hence, consumer i's lifetime utility is:

$$u - p_{1j} - \tau x_1 + \delta\left\{(p_{2k} - p_{2j} + B_j + \tau)D_j - \tau D_j^2 + u - p_{2k} - \frac{\tau}{2}\right\} \tag{31}$$

Similarly, conditional on buying from retailer k in the first period, the consumer continues buying from her if:

$$x_2 \geq D_k \equiv \frac{\tau + p_{2k} - p_{2j} - B_k}{2\tau}$$

Then, consumer i's expected total surplus conditional on buying from k is:

$$u - p_{1k} - \tau x_1 + \delta\left\{(p_{2k} - p_{2j} - B_k + \tau)D_k - \tau D_k^2 + u - p_{2k} + B_k - \frac{\tau}{2}\right\} \tag{32}$$

A consumer located at $x_1 = \alpha_j$ who is indifferent between retailers makes (32) = (31), which results in firm j's market share:

$$\alpha_j = \frac{1}{2} + \tilde{\sigma}\left[(2p_{1k} - \delta B_k) - (2p_{1j} - \delta B_j)\right] \tag{33}$$

where

$$\tilde{\sigma} \equiv \frac{\tau}{4\tau^2 + \delta(B_j + B_k)^2}$$

is a 'modified' substitutability parameter that takes into account the dynamic effects of LPs. Note that if we set $B_j = B_k = 0$, we obtain the standard market share in the static Hotelling model.

Equation (33) tells us that even though the consumer does not receive an immediate rebate, he perceives current prices lower as if the loyalty discount were immediate. Also, that a delayed rebate does not have the same impact on the demand as an immediate one as long as the discount factor reduces the impact of this loyalty rebate, stimulating only a fraction of demand increase per unit of price reduction.

Assuming that firms discount the future at the same rate δ as consumers, firm j's overall profit function is given by:

$$\Pi_j = \pi_{1j}(p_{1j}, p_{1k}) + \delta\pi_{2j}\left(\alpha_j(p_{1j}, p_{1k}, B_j, B_k)\right) \tag{34}$$

where:

$$\begin{aligned}\pi_{1j} &= (p_{1j} - c)\alpha_j(p_{1j}, p_{1k}, B_j, B_k)\\ &= (p_{1j} - c)\left(\frac{1}{2} + \tilde{\sigma}\left[(2p_{1k} - \delta B_k) - (2p_{1j} - \delta B_j)\right]\right)\end{aligned} \tag{35}$$

and

$$\begin{aligned}\pi_{2j} &= (p_{2j} - c)(\alpha_j D_j + (1 - \alpha_j)D_k) - \alpha_j D_j B_j\\ &= \frac{\tau^2 - \alpha_j(1 - \alpha_j)(B_j + B_k)B_j}{2\tau}\end{aligned} \tag{36}$$

Firm j's problem is then:

$$\max_{\{p_{1j}, B_j\}} \Pi_j$$

FOCs:

$$\begin{aligned}\frac{\partial \Pi_j}{\partial p_{1j}} = \frac{1}{4\tau^2 + \delta(B_j + B_k)^2}&\left[2\tau\left(p_{1k} - 2p_{1j} + \tau + \delta\frac{(B_j - B_k)}{2} + c\right) + \delta\frac{(B_j + B_k)^2}{2}\right]\\ &+ \frac{\delta}{2\tau}\left[(B_j + B_k)B_j\frac{\partial \alpha_j}{\partial p_{1j}}\left(-(1 - \alpha_j) + \alpha_j\right)\right] = 0\end{aligned}$$

In a symmetric equilibrium, $p_{1j} = p_{1k} = p_1^*, \alpha_j = 1/2, B_j = B_k = B$:

$$p_1^* = c + \tau + \frac{\delta B^2}{\tau}, \tag{37}$$

$$p_2^* = c + \tau + \frac{B}{2} \tag{38}$$

And the respective profits are:

$$\pi_1^* = \frac{1}{4\sigma} + \delta\sigma B^2 \qquad\qquad \pi_2^* = \frac{1}{4\sigma} - \frac{\sigma B^2}{2} \tag{39}$$

$$\Pi^* = (1+\delta)\frac{1}{4\sigma} + \delta\frac{\sigma B^2}{2} \tag{40}$$

Where σ is the standard substitutability parameter as defined in (25). Note that if we set $B = 0$ we get $\pi_1 = \pi_2 = \frac{1}{4\sigma}$, which is the Hotelling profit in a model with unit demands.

As shown in the previous equations, retailers' instantaneous profits can be some periods higher and some others lower than those they would obtain in absence of LPs. However, overall profits are higher thanks to the rise in prices induced by LPs. According to equation (40), overall profits increase in the loyalty rebate. Given the transportation costs, the higher the B the better for the retailer.

B First-stage results

In this Section I present the regressions conducted as a first step to obtain predicted price, membership and the deltas necessary to compute the optimal instruments.

B.1 Modelling prices

As pointed out previously, I assume that the price of a particular brand in a department is a linear function of brand and supermarkets characteristics and the average prices of the same brand in all the regions of the country (excluding the department from the average price of the region the market is located in). These prices are used as proxies of marginal costs under the assumption that a supermarket chain has some common marginal costs in all its stores across the country, and then retail prices contain information on that common costs (see Section 4 for more details).

Table B.1 displays the results of the linear regression of the price of the brand variety j in supermarket s at t on average prices of the same brand in the 21 administrative regions of France, and brand dummy variables and week dummy variables as controls. The fact that all the estimates are significant at 1% level and the high R-squared suggest that prices have important explanatory power.

B.2 Modelling the membership decision

As explained before, the information I have about loyalty programs is very limited: a dummy variable indicating whether a given household was member of the loyalty program of the supermarket it purchased from in 2006 is all I have. We do not know anything about how long ago the household is member of it, nor if the household gets the promised coupons or even if it redeems them.

Due to these data limitations, I cannot fully explain the determinants of the individual decision to subscribe to a loyalty program. Ideally, a proper way to identify the parameters for the subscription decision modelling is having observations on demographics, the purchasing behaviour of the households, and loyalty programs' and supermarkets' characteristics including the reward system, before and after the subscription decision is made.

Table B.1: Results from a linear regression for price

	Est.	s.e.
pR1	-0.632	0.132
pR2	-2.902	0.194
pR3	-1.218	0.227
pR4	-1.546	0.148
pR5	-1.591	0.100
pR6	-2.717	0.194
pR7	-2.820	0.288
pR8	-1.915	0.216
pR9	-1.940	0.170
pR10	-0.895	0.140
pR11	-6.149	0.276
pR12	-2.236	0.191
pR13	-0.863	0.164
pR14	-2.346	0.167
pR15	-3.253	0.212
pR16	-1.008	0.135
pR17	-1.660	0.117
pR18	-1.233	0.139
pR19	-1.310	0.174
pR20	-2.881	0.283
pR21	-6.385	0.338
adj. R^2	0.840	

Based on 37,662 observations.
Robust s.e. in parentheses.
Regressions include brand
and time dummy variables.

What I can do, nevertheless, is exploiting the data I have to give a clue of how loyalty program characteristics and demographics are correlated with the membership status of a household to a supermarket LP. In addition, this exercise allows me to find instruments to treat the endogeneity of the loyalty membership indicator caused by its correlation with department-week specific deviation from the overall mean valuation of the supermarket characteristics ($\Delta\xi_{st}$).

I collected some data on the characteristics of each loyalty program that do not vary by market, namely, a dummy variable equal to one if the program includes both NB and PL in the reward system and zero otherwise, a dummy variable equal to one if the customer can either subscribe to the program by Internet or at least download the inscription form and zero otherwise, the average rebate in euros per 100 € expended in the supermarket, and dummies for whether the program has an upper bound for accumulating rebates (Cap), a threshold to be able to redeem certain amount of points or money accumulated ($Threshold$) and a variable equal to the number of months for the card to expire in case of lack of use ($Expire\_date$).[42]

From the data set described in section 3.1, I used demographic variables as household size, income, a dummy equal to one if the household has at least one car, age of the head

[42] To obtain this information, I reviewed carefully each supermarket program's "Terms and Conditions" document and identify some common categories describing the general structure of the programs. These documents were obtained from each supermarket website in 2011 and 2012. Although the information is used to instrument a decision that could have been taken even before 2006 (the year of the homescan data base used in this article), all the programs where launched before 2006 and I believe the structure has not changed a lot since then.

of household and a dummy equal to one if the family lives in urban areas. In addition, I constructed variables such as the number of different stores visited by the household members within a week, the number of shopping trips to a same retailer in a month and the number of members to the LP of a given supermarket in the rest of the country (excluding the department where the household was located).

Table B.2 displays the results of OLS and Logit regressions of the dummy variable indicating household membership to a supermarket loyalty program on the variables previously described. After a first specification with all the covariates, the LP characteristics *Cap*, *Threshold* and *Expire_date* were either non significant at 5% level or had a different sign than expected or both. Thus, I excluded them form the final specification displayed in the table.

Table B.2: Results from OLS and Logit regressions for LP membership

Variable	OLS	Logit
LPs characteristics		
Log total Sup. LP members	0.141**	0.892**
excluding location depart.	(0.028)	(0.185)
NB included	0.0499**	0.322**
	(0.020)	(0.137)
Online access to inscription	0.0410**	0.261**
form	(0.012)	(0.074)
Log of average reward per	0.0110**	0.0692**
100 €spent	(0.005)	(0.033)
Demographics		
# of different stores	−0.272**	−1.643**
visited a week	(0.044)	(0.235)
# trips to the same	0.156**	1.127**
retailer per month	(0.011)	(0.093)
Household size	0.0288**	0.190**
	(0.003)	(0.023)
Log income	-0.0132	−0.0928*
	(0.009)	(0.057)
Log age	0.0365**	0.221**
	(0.015)	(0.088)
Car	0.0683**	0.369**
	(0.022)	(0.114)
Lives in city	0.00869	0.0615
	(0.009)	(0.058)
Constant	−0.632**	−7.515**
	(0.254)	(1.660)
N	9,976	9,976
R^2	0.035	
adj. R^2	0.034	

Robust standard errors in parentheses.
** Significant at 5% and * at 10% level.

Although we cannot say that this variables explain the whole variation of the dependent variable, and this is confirmed by the low adjusted R^2, the results are quite appealing. On the one hand, in the Logit regression (column 2) all the parameters are significant, and on the other hand, the signs of the variables are according to the intuition.

In the first place, the probability for a customer to join supermarket j loyalty program increases with the number of members in the rest of the country, which means that the more popular the program is, the more attractive is for a customer to join most probably

because this is a signal of the program quality. Moreover, this probability also increases with household characteristics such as the number of trips the household members do to the retailer j, being an indicator of shopping intensity, the more often a customer goes to the same store, the more informed about prices and promotions she is and, consequently, the more likely to benefit from rebates; household size, age, car ownership and living in urban areas which facilitates the access to the retailer.

Second, individual characteristics such as the income and the multi-store shopping behavior make consumers to be less interested in joining LPs because either they are less sensitive to small price cuts or they are less likely to get rewards. In fact, one can expect that as income increases supermarket loyalty programs become less attractive because wealthier people may be less sensitive to little price cuts and also because loyalty rebates are generally addressed to middle quality goods rather than to high quality branded products. Regarding the multi-store shopping behavior, it is more difficult to accumulate points or money in a loyalty account when purchases are not concentrated in one retailer, which makes individuals with this shopping profile either less interested to join any loyalty program or willing to join all of them, but not likely to use them as rewards will never be attainable.

As for the the LP characteristics included in the regression: Logit results show that the probability of joining a LP increases with the fact that both NB and PL purchases give the customers points to get a future rebate compared to those programs including only PL, also with the possibility of joining the program online (or at least having access to the inscription form) and the average reward in euros, i.e, the better rewarding programs are the more likely to attract customers to their loyalty program.

B.3 First-stage estimation of the full model

The results of the first-stage estimation of the full model using the set of inefficient instruments for the endogenous variables are basically of the expected sign. However, most estimates are statistically non significant due to the very large robust standard errors obtained. In addition, some convergence problems were experienced as the size of the estimates and in some cases the sign were quite sensitive to the provided set of starting values. All this suggests that the instruments used to identify the parameters of the full model are not performing well, even though they are similar to the instruments used in some standard papers. This may be due to the fact that they are just helping to identify the parameters in the linear part of the model (θ_1) and not those of the whole model (θ_1, θ_2). This fact empirically supports the need to use optimal instruments as the best way to overcome endogeneity problems in this kind of models.

B.4 Own- and cross-price elasticities

Table B.4 presents the estimated own- and cross- price elasticities for the leading 12 brand varieties based on their national market shares. The columns in the table indicate de prices and the rows the market shares used to compute the variation, so that for instance, the entry located in row 2 column one gives the elasticity of brand 2 with respect to a change in the price of brand 1.

One of the advantages of a random-coefficients model, as it was said previously, is that it gives more realistic substitution patterns allowing for different brands to have different values for the elasticity due to the change in the price of a given price, as opposed to the simple Logit where all the elasticities in the same column would have the same cross-price elasticities. As a consequence, estimated elasticities vary by brand and by market.

Table B.3: First-stage results from the Mixed Logit model with inefficient instruments[a]

	Means	Std. Deviations	Interactions with Demographic variables		
Variable	(β's)	(σ's)	HH size	Income	# Subscriptions
Constant[b]	-3.631	1.353		-0.967	0.922
	(1.388)	(1.865)		(13.454)	(4.787)
Price	-6.399	2.301	-1.748	0.764	1.331
	(5.253)	(6.090)	(12.456)	(25.647)	(14.523)
LP member (M_{ms})	-8.338	6.223		-0.967	
	(6.896)	(2.246)		(9.011)	
PL dummy[b]	-2.634	2.755	-1.331	0.764	-7.803
	(2.270)	(3.569)	(4.305)	(4.120)	(5.237)
LP member×PL dummy	7.935	4.347			
	(3.381)	(3.344)			
Plastic[b]	-0.035	2.114			
	(0.483)	(1.266)			
Sugar[b]	-0.223	0.562			
	(0.240)	(4.803)			
Wholemilk[b]	-2.998	0.670			
	(1.035)	(3.779)			
GMM Objective	1,045				
MD χ^2	4139681				

[a] Based on 37,662 observations. Except where noted, parameters were estimated using GMM. All regressions include brand and week dummies. Asymptotically robust s.e. are given in parentheses.
[b] Estimated using a minimum-distance procedure.

In order to give an idea of the estimated substitution patterns, Table B.4 presents the medians of the own- and cross-price elasticities over the 7,704 in the sample. The median elasticities of the outside good with respect to column brand prices are also displayed.

Table B.4: Median own- and cross-price elasticities[a]

Brand	b1	b3	b7	b8	b9	b11	b12	b15	b16	b18	b21	b23	b24	b28	b29
PL															
1	-2.612	0.017	0.006	0.004	0.005	0.006	0.010	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.000
2	0.033	0.262	0.159	0.119	0.121	0.162	0.180	0.004	0.031	0.002	0.002	0.005	0.027	0.013	0.003
3	0.009	-15.995	0.045	0.029	0.022	0.032	0.035	0.001	0.007	0.000	0.000	0.002	0.008	0.003	0.001
4	0.030	0.269	0.125	0.081	0.081	0.119	0.163	0.002	0.027	0.001	0.001	0.006	0.025	0.009	0.002
5	0.041	0.387	0.271	0.144	0.148	0.239	0.244	0.006	0.043	0.003	0.003	0.009	0.050	0.016	0.005
6	0.009	0.157	0.053	0.035	0.036	0.061	0.049	0.001	0.011	0.001	0.000	0.002	0.009	0.003	0.001
7	0.005	0.052	-8.893	0.018	0.017	0.022	0.019	0.001	0.004	0.000	0.000	0.001	0.004	0.002	0.000
8	0.004	0.040	0.019	-7.310	0.012	0.017	0.021	0.000	0.003	0.000	0.000	0.001	0.004	0.002	0.000
9	0.003	0.023	0.016	0.007	-4.777	0.012	0.011	0.000	0.002	0.000	0.000	0.001	0.003	0.001	0.000
10	0.015	0.091	0.053	0.037	0.037	0.040	0.044	0.002	0.010	0.001	0.000	0.002	0.010	0.004	0.001
11	0.005	0.040	0.022	0.015	0.014	-10.365	0.027	0.001	0.004	0.000	0.000	0.001	0.004	0.002	0.000
12	0.009	0.049	0.031	0.017	0.020	0.031	-12.624	0.001	0.008	0.000	0.000	0.002	0.009	0.003	0.001
NB															
13	0.008	0.100	0.038	0.027	0.027	0.039	0.047	0.003	0.028	0.002	0.001	0.005	0.030	0.011	0.003
14	0.005	0.047	0.037	0.015	0.016	0.024	0.047	0.003	0.024	0.001	0.001	0.005	0.026	0.008	0.002
15	0.000	0.002	0.001	0.001	0.001	0.001	0.002	-1.834	0.001	0.000	0.000	0.000	0.001	0.000	0.000
16	0.001	0.012	0.007	0.004	0.004	0.005	0.009	0.001	-11.919	0.000	0.000	0.001	0.004	0.002	0.000
17	0.002	0.020	0.009	0.007	0.008	0.012	0.016	0.001	0.009	0.001	0.000	0.001	0.008	0.005	0.001
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.837	0.000	0.000	0.000	0.000	0.000
19	0.010	0.072	0.048	0.025	0.031	0.038	0.089	0.004	0.028	0.003	0.001	0.007	0.025	0.016	0.003
20	0.009	0.142	0.057	0.028	0.020	0.050	0.044	0.005	0.034	0.004	0.002	0.009	0.033	0.011	0.003
21	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.491	0.000	0.000	0.000	0.000
22	0.001	0.016	0.005	0.004	0.005	0.005	0.002	0.000	0.003	0.000	0.000	0.001	0.003	0.001	0.000
23	0.000	0.002	0.001	0.001	0.001	0.001	0.001	0.000	0.001	0.000	0.000	-2.694	0.001	0.000	0.000
24	0.001	0.007	0.003	0.003	0.003	0.003	0.006	0.000	0.002	0.000	0.000	0.001	-10.906	0.001	0.000
25	0.006	0.056	0.024	0.021	0.017	0.030	0.037	0.003	0.018	0.001	0.001	0.005	0.017	0.009	0.002
26	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
27	0.022	0.169	0.091	0.061	0.070	0.080	0.150	0.009	0.070	0.003	0.003	0.015	0.048	0.031	0.006
28	0.000	0.002	0.001	0.001	0.001	0.001	0.002	0.000	0.001	0.000	0.000	0.000	0.001	-4.371	0.000
29	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.852
30	0.004	0.047	0.026	0.015	0.014	0.022	0.025	0.002	0.019	0.001	0.001	0.003	0.018	0.007	0.002
31	0.003	0.025	0.012	0.010	0.008	0.014	0.016	0.002	0.009	0.001	0.000	0.002	0.009	0.004	0.001

[a] Each of the entries correspond to the median of the elasticities from the 7,704 markets. They give the percent change in market share of a given brand with a one percent variation in either own-price (for those in the main diagonal) or the price of another brand (for those outside the main diagonal). In the horizontal axis the brands with the largest market shares are displayed. The full matrix is available on demand.

[b] Corresponds to the outside good.

References

1. Amemiya, Takeshi (1974), "The nonlinear Two-Stage Least Squares Estimator", *Journal of Econometrics*, No. 2, pp. 105-110.

2. ________ (1977), "The maximum likelihood and the nonlinear three-stage least squares estimator in the general nonlinear simultaneous equation model", *Econometrica*, Vol. 45, No. 4, pp. 955-968.

3. Bierlaire, Michel, Bolduc, Denis and McFadden, Daniel (2008), "The estimation of Generalized Extreme Value models from choice-based samples", *Transportation Research Part B: Methodological*, vol. 42, No.. 4, pp. 381-394.

4. Berges-Sennou Fabian, Bontems Phillip and Requillart V. (2009), "L'impact économique du développement des marques de distributeurs", *Working paper*, Toulouse School of Economics, 30 p.

5. Berges-Sennou, Fabian (2006), "Store loyalty, bargaining power and the private label production issue", *European Review of Agricultural Economics*, Vol. 33, No. 3, pp. 315-335.

6. Berry, Steven (1994), "Estimating discrete-choice models of product differentiation", *The RAND Journal of Economics*, Vol. 25, No. 2, pp. 242-262.

7. Berry, Steven, Levinsohn, James and Pakes, Ariel (1995), "Automobile prices in market equilibrium", *Econometrica*, Vol. 63, No. 4, pp. 841-890.

8. ________ (1999), "Voluntary export restraints on automobiles: evaluating a trade policy", *The American Economic Review*, Vol. 89, No. 3, pp. 400-430.

9. Bolton, R., Kannan, P. and Bramlett, M. (2000), "Implications of loyalty program membership and service experiences for customer retention and value", *Journal of the Academy of Marketing Science*, Vol. 28, No. 1, pp. 95-108.

10. Bonnet, Céline and Dubois Pierre (2010), "Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance", *The RAND Journal of Economics*, Vol. 41, No. 1, pp. 139-164.

11. Caillaud, Bernard and De Nijs, Romain (2011), "Strategic loyalty reward in dynamic price discrimination", *Paris School of Economics Working Papers*, No. 2011-29, 38 p.

12. Caminal, Ramon and Matutes, Carmen (1990), "Endogenous switching costs in a duopoly model", *International Journal of Industrial Organization*, Vol. 8, No. 3, pp. 353-373.

13. Chamberlain, Gary (1987), "Asymptotic efficiency in estimation with conditional moment restrictions", *Journal of Econometrics*, No. 34, pp. 305-334.

14. Chen, Yongmin and Pearcy, Jason (2010), "Dynamic pricing: when to entice brand switching and when to reward consumer loyalty", *The RAND Journal of Economics*, Vol. 41, No. 4, pp. 674-685.

15. Chintagunta, Pradeep, Bonfrer, André and Song, Inseong (2002), "Investigating the effects of store-brand introduction on retailer demand and pricing behavior", *Management Science*, Vol. 48, No. 10, pp. 1242-1267.

16. Cremer, Jacques (1984), "On the economics of repeat buying", *The RAND Journal of Economics*, Vol. 15, No. 3, pp. 396-403.

17. European Commission (2011), "The impact of private labels on the competitiveness of the European food supply chain", available at:
$http://ec.europa.eu/enterprise/sectors/food/documents/studies/comments_privlab_en.htm$

18. Dubé, Jean-Paul, Fox, Jeremy and Su, Che-Lin (2012), "Improving the numerical performance of BLP static and dynamic aggregate discrete choice random coefficients demand estimation", *Econometrica*, Vol. 80, No. 5, pp. 2231-2267.

19. Fudenberg, Drew and Tirole, Jean (2000), "Customer poaching and brand switching", *The RAND Journal of Economics*, Vol. 31, No. 4, pp. 634-657.

20. Hausman, Jerry (1996), "Valuation of new goods under perfect and imperfect competitio", *The Economics of New Goods*, Studies in Income and Wealth, Vol. 58, ed. by T. Bresnahan and R. Gordon, National Bueau of Economic Research, Chicago.

21. Jorgenson, Dale and Laffont, Jean-Jacques (1974), "Efficient estimation of nonlinear simultaneous equations with additive disturbances", *Annals of Economic and Social Measurement*, 3/4, pp. 615-640.

22. Kelejian, Harry (1971), "Two-stage least squares and econometric systems linear in parameters but nonlinear in endogenous variables", *Journal of the American Economic Association*, Vol. 66, No. 334, pp. 373-374.

23. Klemperer Paul (1987a), "Markets with consumer switching costs", *The Quarterly Journal of Economics*, Vol. 102, No. 2, pp. 375-394.

24. ________ (1987b), "The competitiveness of markets with switching costs", *The RAND Journal of Economics*, Vol. 18, No. 1, pp. 138-150.

25. Knittel, Christopher and Metaxoglou, Konstantinos (2012), "Estimation of random coefficient demand models: two empiricists' perspective", non-published manuscript.

26. Lewis, Michel (2004), "The influence of loyalty programs and short-term promotions on customer retention", *Journal of Marketing Research*, Vol. 41, pp. 281-292.

27. Lal, Rajiv and Bell, David (2003), "The impact on frequent shopper programs in grocery retailing", *Quantitative Marketing and Economics*, No. 1, pp. 179-202.

28. Lederman, Mara (2007), "Do enhancements to loyalty programs affect demand? The impact of international frequent flyer partnerships on domestic airline demand", *RAND Journal of Economics*, Vol. 38, No. 4, pp. 1134-1158.

29. Manski, Charles and Lerman, Steven (1977), "The estimation of choice probabilities from choice based samples", *Econometrica*, Vol. 45, No. 8, pp. 1977-1988.

30. McFadden, Daniel (1999), "Computing willingness-to-pay in random utility models", in J. Moore, R. Riezman, and J. Melvin, eds., *Trade, Theory and Econometrics: Essays in Honour of John S. Chipman*, Routledge, London, pp. 253?274.

31. Nasser, Sherif, Turcic, Danko and Narisimhan, Chakravarthi (2013), "National Brand's Response to Store Brands: Throw In the Towel or Fight Back?", *Marketing Science*, Vol. 32, No. 4, pp. 591-608.

32. Nevo, Aviv (2000a), "A practitioner's guide to estimation of random-coefficients Logit models of demand", *Journal of Economics and Management Strategy*, Vol. 9, No. 4, pp. 513-548.

33. ________ (2000b), "Mergers with differentiated products: the case of the ready-to-eat cereal industry", *RAND Journal of Economics*, Vol. 31, No. 3, pp. 395-421.

34. ________ (2001), "Measuring market power in the ready-to-eat cereal industry", *Econometrica*, Vol. 69, No. 2, pp. 307-342.

35. Nevo, Aviv and Wolfram, Catherine (2002), "Why do manufacturers issue coupons? An empirical analysis of breakfast cereals", *RAND Journal of Economics*, Vol. 33, No. 2, pp. 319-339.

36. Newey, Whitney (1990), "Efficient instrumental variables estimation of nonlinear models", *Econometrica*, Vol. 58, No. 4, pp. 809-837.

37. Raju, Jagmohan, Sethuraman, Raj and Dhar, Sanjay (1995), "The introduction and performance of store brands", *Management Science*, Vol. 41, No. 6, pp. 957-978.

38. Reynaert, Mathias and Verboven, Frank (2013), "Improving the performance of the random coefficients demand models: the role of optimal instruments", *unpublished Working paper.*

39. Steenkamp, Jan-Benedict and Geyskens, Inge (2013), "Manufacturer and Retailer Strategies to Impact Store Brand Share: Global Integration, Local Adaptation, and Worldwide Learning", *Marketing Science*, forthcoming.

40. Soberman, David and Parker, Phillip (2006), "The economics of quality-equivalent store brands", *International Journal of Research in Marketing*, No. 23, pp. 125-139.

41. Train, Kenneth (2009), *Discrete Choice Methods with Simulation*, Second edition, Cambridge University Press.

42. Wooldrige, Jeffrey (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.