Do preferences for private labels respond to supermarket loyalty programs?

Jorge Florez-Acosta
Facultad de Economía-Universidad del Rosario

Seminario en Matemáticas Aplicadas
Quantil
Bogotá ⋆ May 10, 2018

Universidad del Rosario | Facultad de Economía

Intorduction

1. Private labels (*PL*) are retailers' own-branded products.
2. Nowadays, quality-equivalent products to National brands (*NB*).
3. Increase customers choice set, intensify intra-brand competition and stimulate upstream competition.
4. Increasing market shares: 46% in the UK, 35% in Germany, 33% in Spain and 19% in the US.
5. Sales growing around 4% on average in Europe.
6. Prices are 20% lower on average as compared to *NB*.
7. In addition, some nonprice strategies are used to promote PLs.

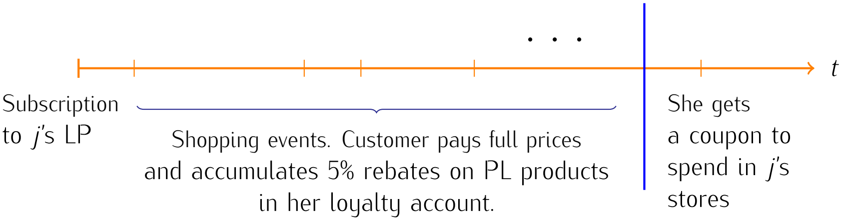
Supermarket loyalty programs

1. Retail chains give loyalty rewards on PLs and other selected products to “loyal” customers.

2. **Loyal:** who subscribe to a supermarket loyalty program (*LP*).

3. In France, rebates are mainly linked to private label purchases.

How do supermarket LPs work in France?



Why are LPs so widespread nowadays?

- **LPs (coupons) allow retailers to retain customers and induce repurchase:** Cremer (1984), Klemperer (1987a, 1987b), Caminal and Matutes (1990), Chen and Pearcy (2010).

- **LPs (coupons) are a way to exercise market power, in particular, they can be used as a discriminatory device:** Caillaud and De Nijs (2011), Fudenberg and Tirole (2000).

- **Coupons are more profitable than a second-period price reduction for loyal customers:** Caminal and Matutes (1990).

- **Empirical support for (1), (2), (3):** Nevo and Wolfram (2002), manufacturer-issued coupons.

Research question: Why do retailers give additional discounts on their lower-price PL? Why not on the whole range of products they carry?

Why are LPs so widespread nowadays?

- **LPs (coupons) allow retailers to retain customers and induce repurchase:** Cremer (1984), Klemperer (1987a, 1987b), Caminal and Matutes (1990), Chen and Pearcy (2010).

- **LPs (coupons) are a way to exercise market power, in particular, they can be used as a discriminatory device:** Caillaud and De Nijs (2011), Fudenberg and Tirole (2000).

- **Coupons are more profitable than a second-period price reduction for loyal customers:** Caminal and Matutes (1990).

- **Empirical support for (1), (2), (3):** Nevo and Wolfram (2002), manufacturer-issued coupons.

Research question: Why do retailers give additional discounts on their lower-price PL? Why not on the whole range of products they carry?

This paper

Objective

Empirically examine the link between loyalty programs and the demand for private labels.

Empirical strategy

1. Pick one non-durable product: **Yogurt.**
2. Specify a flexible structural demand model to capture causal effects.
3. Apply the demand estimation techniques developed recently: Berry (1994), BLP (1995), Nevo (2000, 2001), Knittel and Metaxoglou (2014).
4. Compute optimal instruments for identification.
5. Conduct counterfactual simulations.

Main results

- Marginal valuation of PL products increases with subscription to supermarket loyalty program.
- Multi subscription makes LPs effects weaker.
- Making subscription to LPs prohibitively costly harms consumers.

Previous research

Topic	Theory	Empirics
Introduction of PL	Raju et al. (1995)	Bonfrer & Chintagunta (2004)
	Chintagunta et al. (2002)	
PL demand determinants	Berges et al. (2009)	
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)	

Outline

Overview of the data

- **Kantar Worldpanel:** scanner data on consumer purchases of grocery products.
- 14,529 households in France during 2006.
- **I observe:**
  - Purchase occasions: date of purchase, product and description, total quantity, total value of purchase.
  - Multiple products: over 350 separate products sold by supermarkets in France.
  - Demographics and other household characteristics: household size, income category, education level, etc.
  - Membership: An indicator variable taking on 1 if a household is member to a specific supermarket loyalty program, and zero otherwise.
- **I do not observe:**
  - Data before subscription.
  - Coupons issuing/redemption rate.

Summary statistics

Variable	Mean	Median	Sd	Min	Max
Demographics					
Household size	2.63	2	1.39	1	9
Income (€/month)	2337	2100	1175	150	7000
HH head's age	47.84	45	15.66	18	99
Car (=1 if yes)	0.92	1	0.27	0	1
Live in urban areas (=1 if yes)	0.75	1	0.43	0	1
Shopping behavior					
Private label (=1 if yes)	0.34	0	0.47	0	1
Total expenditure (€/week)	63.80	53.78	46.47	0.09	2249
PL share (% on total expenditure)	27.61	23.51	22.02	0	100
Number of visits to the same store in a week	1.54	1	0.88	1	7
Number of different stores visited in a week	1.72	2	0.88	1	9
Duration (days) between visits to stores	5.73	4	6.69	1	315
Loyalty-related information					
LP membership to at least one store (=1 if yes)	0.85	1	0.36	0	1
Number of LP memberships	2.21	2	1.65	0	12
Exp. share in stores if membership (% tot. exp.)	27.28	0.00	41.37	0	100
PL exp. share if membership (% tot. exp.)	29.91	24.43	25.02	0	100
Proportion of stores visited if membership	26.02	0	37.40	0	100

Preliminary descriptive results

	Price index	Number of visits
LP membership (=1 if yes)	-0.194	0.045
	(0.007)	(0.001)
Number of subscriptions	-0.012	-0.009
	(0.001)	(0.000)
Private label (=1 if yes)	-0.315	—
	(0.003)	
Private label×LP membership	0.038	—
	(0.007)	
Number of visits to a store	0.130	—
	(0.002)	
Number of stores visited	-0.214	0.023
	(0.002)	(0.000)
Hypermarket	0.180	-0.181
	(0.007)	(0.001)
Convenience	-1.101	0.170
	(0.008)	(0.002)
Constant	2.961	0.297
	(0.032)	(0.005)
R^2	0.297	0.065
Observations	9.53 million	9.53 million

Empirical strategy

1. Pick one non-durable product: **Yogurt.**
2. Aggregate data up to brand level by subgroup of population:
   $\Longrightarrow$ m=1: LP members and m=0: nonmembers.
3. Specify a flexible structural demand model to capture causal effects.
4. Exploit panel structure of the data to control for brand, supermarket and time fixed-effects.
5. Apply the demand estimation techniques developed recently:
   Berry (1994), BLP (1995), Nevo (2000, 2001), Knittel and Metaxoglou (2012).
6. Compute optimal instruments for identification:
   Chamberlain (1987), Newey (1990), BLP (1999), Reynaert and Verboven (2013).
7. Conduct counterfactual simulations.

The random coefficients Logit setup

Let the indirect utility be given by

$$u_{ijst} = \mathbf{x}_{js}\boldsymbol{\beta}_i - \alpha_i p_{jst} + \lambda_i M_{is} + \varphi_i M_{is} \times PL_{js} + \phi_t + \xi_{js} + \Delta\xi_{jst} + \varepsilon_{ijst} \qquad (1)$$

The parameters are functions of household observed and unobserved characteristics:

$$\begin{pmatrix} \alpha_i \\ \beta_i \\ \lambda_i \\ \varphi_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \\ \lambda \\ \varphi \end{pmatrix} + \boldsymbol{\Pi}\mathbf{D}_i + \boldsymbol{\Sigma}\mathbf{v}_i, \qquad \mathbf{v}_i \sim N(\mathbf{0}, \mathbf{I}_K) \qquad (2)$$

p_{jst}: price of product j at store s in period t.
$\mathbf{x}_{js}$: product and store observable attributes.
ξ_{js}: product and store fixed-effects.
ϕ_t: time fixed effects.
$\mathbf{D}_i$: observable household characteristics.
$\mathbf{v}_i$: unobserved household attributes.

The random coefficients Logit setup (cont'd)

Assuming that:

1. An **outside option** exists. I normalize its utility to zero

$$u_{i0t} = 0,$$

2. The individual shock to utility ε_{ijst} is distributed iid Type I Extreme Value,

Then, product j's market share at supermarket s in t, by subgroup m (=1 if consumers are members to s's LP) is given by:

$$s_{mjst} = \int \frac{\exp\left(\mathbf{x}_{js}\boldsymbol{\beta}_i - \alpha_i p_{jst} + \lambda_i M_{is} + \varphi_i M_{is} \times PL_{js} + \phi_t + \xi_{js} + \Delta\xi_{jst}\right)}{1 + \sum_{k=1}^{S}\sum_{l=1}^{J} \exp\left(\mathbf{x}_{lk}\boldsymbol{\beta}_i - \alpha_i p_{lkt} + \lambda_i M_{ik} + \varphi_i M_{ik} \times PL_{lk} + \phi_t + \xi_{lk} + \Delta\xi_{lkt}\right)} dP(D)dP(v)dP(M)$$

The Yogurt data

1. Plain yogurt: 33% market share in yogurt category.
2. Non-storable and of “unit” demand.
3. I considered the leading 6 supermarket chains in France, with a LP.
4. Brands: 13 leading (6 PL and 7 NB).
5. Supermarket as a brand characteristic: 120 products (brand-supermarket combination).
6. **Final sample:** 31 differentiated products, based on their national market share in 2006 (66.5% overall).

Logit

Estimation

Let $\theta = (\theta_1, \theta_2)'$, where $\theta_1 = (\alpha, \beta, \lambda, \varphi, \phi_t, \xi_{js})'$ are the linear parameters, and $\theta_2 = (\text{vec}(\Pi), \text{vec}(\Sigma))'$ are the nonlinear parameters.

Estimation relies on the moment condition:

$$E[h(z)'\rho(x, \theta_o)] = 0, \tag{3}$$

$(z_1, \ldots, z_M)$: set of instruments.

A GMM estimator is

$$\widehat{\theta} = \arg\min_{\theta} \rho(\theta)' h(z) \hat{\Lambda}^{-1} h(z)' \rho(\theta), \tag{4}$$

with $\Lambda = Var(h(z)'\rho)$.

Given $\delta(\cdot)$, the error term writes as

$$\rho_{mjst} = \delta_{mjst}(\mathbf{x}, M, p_{.t}, S_{.t}; \boldsymbol{\theta}_2) - (\mathbf{x}_{js}\boldsymbol{\beta}_i - \alpha_i p_{jst} + \lambda_i M_{is} + \varphi_i M_{is} \times PL_{js} + \phi_t + \xi_{js}) \tag{5}$$

▸ More ▸ S_{mjst}

Two endogenous variables

For price: IV's used in the literature

- **Brand characteristics of other products and of rivals' products:** BLP (1995).
- **Regional average prices:** Hausman (1996), Nevo (2000, 2001).
- **Costs proxies and cost shifters:** Nevo (2000, 2001), Armstrong (2012).
- **Polynomials of prod. characteristics and costs shifters:** Dubé et al (2012).

For LP membership:

- I use self-collected data on characteristics of each store's LP.
- As BLP, I assume those characteristics do not respond to demand shocks.

Problem: The estimation resulted in huge s.e. and some convergence problems.

Mixed Logit: preliminary results

	Means	Std. Deviations	Interactions with Demographic variables		
Variable	(β's)	(σ's)	HH size	Income	# Subscriptions
Constant*	-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*	-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*	-0.035	2.114			
	(0.483)	(1.266)			
Sugar*	-0.223	0.562			
	(0.240)	(4.803)			
Wholemilk*	-2.998	0.670			
	(1.035)	(3.779)			

* Estimated by minimum distance.

Solución: Optimal instruments

Following BLP (1999) and Reynaert and Verboven (2014), optimal instruments for a single-equation problem are

$$h^*(z_{.t}) = D(z)'$$

More

Recall $\theta = (\theta_1, \theta_2)$, then:

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \theta_1'}\bigg| z_t\right] = E\left[\frac{\partial(\mathbf{x}_{js}\boldsymbol{\beta}_i - \alpha_i p_{jst} + \lambda_i M_{is} + \varphi_i M_{is} \times PL_{js} + \phi_t + \xi_{js})}{\partial(\alpha, \beta, \lambda, \varphi, \phi_t, \xi_{js})'}\bigg| z_t\right] \quad (6)$$

More

$$E\left[\frac{\partial \rho_{.t}(\theta)}{\partial \theta_2'}\bigg| z_t\right] = E\left[\frac{\partial \delta_{mjst}(s_{.t}, \theta_2)}{\partial \theta_2'}\bigg| z_t\right] \quad (7)$$

Optimal instruments for θ_2

1. Obtain an initial estimate for $\theta = (\theta_1, \theta_2)$. I use the sets of inefficient instruments.
2. Compute the predicted price $\hat{p}_{.t}$ and membership indicator $\hat{M}_{ms}$.
3. Retrieve the predicted mean utility as $\hat{\delta}_{mjst} \equiv \mathbf{x}_{js}\hat{\boldsymbol{\beta}}_i - \hat{\alpha}_i p_{jst} + \hat{\lambda}_i M_{is} + \hat{\varphi}_i M_{is} \times PL_{js} + \hat{\phi}_t + \hat{\xi}_{js}$ 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

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

Back

Mixed Logit results

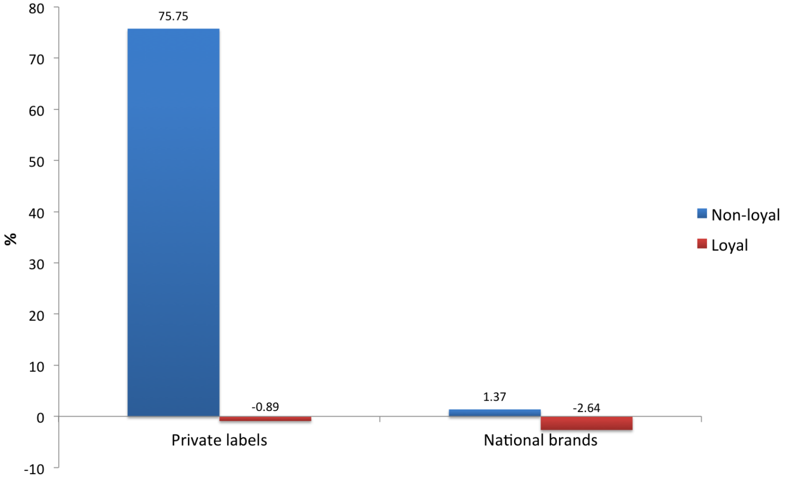
Variable	Means (β's)	Std. Deviations (σ's)	Interactions with Demographic variables: HH size	Income	# Subscriptions
Constant*	-10.067	0.478		0.131	-0.968
	(0.043)	(0.382)		(0.236)	(0.346)
Price	**-12.963**	1.057	0.505	0.900	**-0.535**
	(0.453)	(0.485)	(0.132)	(0.085)	(0.255)
LP member (M_{ms})	**-0.825**	0.685		0.182	
	(0.207)	(0.286)		(0.536)	
PL dummy*	**-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*	10.877	0.144			
	(0.060)	(0.406)			
Sugar*	6.840	0.171			
	(0.053)	(0.409)			
Wholemilk*	-1.513	0.466			
	(0.013)	(0.156)			

* Estimated by minimum distance.

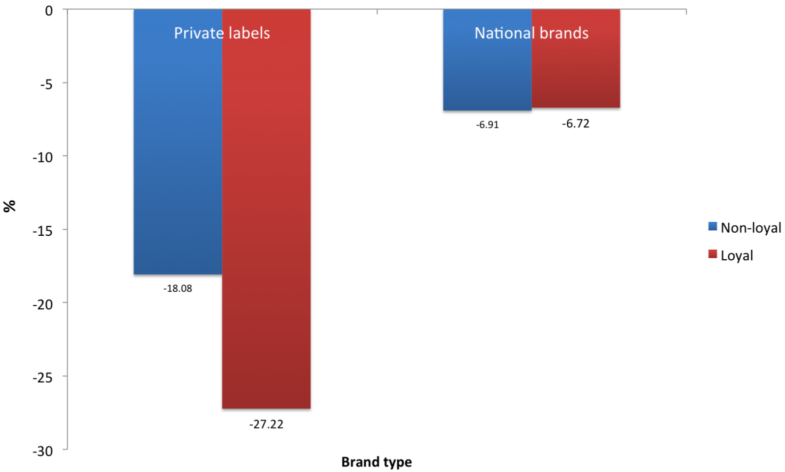
Counterfactual analysis

- A simulated experiment with two counterfactual scenarios.
- **Scenario 1:** All consumers are subscribers of at least one supermarket LP.
- **Scenario 2:** No one subscribes to any supermarket LP.
- It corresponds to a change in the variable M_{is} in the model.
- I compute expected changes in aggregate demand and Consumer surplus.

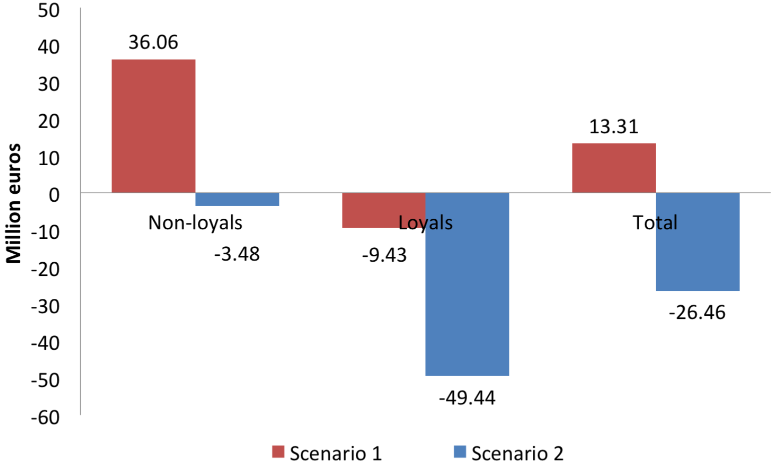
Changes in aggregate demand: Scenario 1



Changes in aggregate demand: Scenario 2



Changes in Consumer surplus



Summary and Conclusion

- Theory does not fully explain the case of supermarkets.
- This paper gives empirical support to the idea that LPs are used as a nonprice strategy to boost PL demand.
- In fact, results show that consumer valuation for PL increase with subscription to LP.
- The impact on consumer welfare of a no subscription scenario is twice as much as the gain when all consumers subscribe.

To explore in the future

1. **Loyalty from an objective measure**
   - Marketing measures: share of wallet and number of visits a month.
   - Are LP making subscribers loyal?
   - Are non subscribers more loyal than subscribers?

2. **With the ideal data set on LPs**
   - Model LP membership decision and supermarket choice
   - Structurally model the supply side, recover manufacturers margins.
   - Counterfactuals playing with the amount of the loyalty reward, immediate vs. lagged rebates.
   - Are large retailers using PL and LPs to increase their buyer power?

Questions? Comments? Suggestions?

jorge.florez@urosario.edu.co

Thank you!

Universidad del Rosario | Facultad de Economía

More details

Indirect utility components:

$$\delta_{mjst} = x_j\beta + r_s\lambda - \alpha p_{jst} + \varphi M_{ms} + \eta PL_{js} \times M_{ms} + \xi_j + \xi_s + \Delta\xi_{mjt} + \Delta\xi_{mst},$$

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

Outside good utility:

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

back

Descriptive regressions: Logit model

Dependent variable: $\ln S_{mjst} - \ln S_{0t}$

	OLS	IV	
Variable	(1)	(2)	(3)
Price(€/125gr)	-4.502	-11.18	-5.965
	(0.159)	(1.181)	(0.463)
LP membership	-0.252	-0.438	-0.281
	(0.015)	(0.032)	(0.018)
PL dummy	-0.504	-1.149	-0.637
	(0.062)	(0.131)	(0.074)
LP membership×PL dummy	0.210	0.366	0.235
	(0.022)	(0.034)	(0.023)
#Subscriptions×PL dummy	-0.0561		-0.0974
	(0.015)		(0.019)
Price×#Subscriptions	-0.696		-0.396
	(0.056)		(0.102)
Constant	-17.45	-5.888	-16.91
	(0.680)	(1.128)	(0.698)
Brand-supermarket fixed-effects	Yes	No	Yes
Brand characteristics	No	Yes	No
Demographics	Yes	No	Yes
Other controls	Yes	No	Yes
Instruments		Brand dummies	Prices
Fit/Test of over	0.297	5,337	131.4
Identification[b]		(1.145)	(10.851)
1st Stage R^2		0.749	0.909

▸ S_{mjst} ▸ Back

Brand-specific dummy variables

- Brand dummy variables are included to capture brand fixed-effects.
- The coefficients on the dummies capture (Nevo, 2000, 2001):
  1. The mean valuation of observed characteristics that do not vary by market: $x_j\beta + r_s\lambda$.
  2. The overall mean valuation of the unobserved characteristics: $\xi_j + \xi_s$.
- We can write this as

$$d = X\beta + \xi \tag{8}$$

- Let V_d be the covariance matrix of $\hat{d}$. Assuming $E[\xi|X] = 0$,

$$\hat{\beta} = (X'V_d^{-1}X)^{-1}X'V_d^{-1}\hat{d} \tag{9}$$

$$\hat{\xi} = \hat{d} - X\hat{\beta} \tag{10}$$

back

The estimation algorithm

1. Retrieve δ by solving the nonlinear system

$$s_{.t}(x, r, M, p_{.t}, \delta_{.t}; \theta_2) = S_{.t} \tag{11}$$

BLP (1995) proposed a contraction mapping such as

$$\delta_{.t}^{(k+1)} = \delta_{.t}^{(k)} + \ln S_{.t} - \ln s_{.t}(x, r, M, p_{.t}, \delta_{.t}^{(k)}; \theta_2), \tag{12}$$

which uses random numbers as starting values for θ_2, and the δ from

$$\delta_{.t}^{(0)} = \ln S_{.t} - \ln S_{0t} \tag{13}$$

2. Estimate the linear parameters $\theta_1 = (\alpha, \beta, \varphi, \eta, \lambda)$ by 2SLS.
3. Compute the error term $\rho(\theta)$.
4. Perform the optimization process to estimate θ_2.

back

The observed brand market shares

- Computed by subgroup of population of LP members ($m = 1$) and non LP members ($m = 0$).

$$S_{mjst} = \frac{\text{No. 125gr servings sold in } t}{\text{Total 125gr servings potentially consumed in } t} \tag{14}$$

- The potential volume sales per market was computed as the av. national plain yogurt consumption of 1.185 125gr servings per person per week in 2006.

- The outside option market share was computed as

$$S_{0t} = 1 - \sum_{j,s,m} S_{mjst} \tag{15}$$

back

First-stage regressions for price

	Est.	s.e.
ivp1	–0.632	0.132
ivp2	–2.902	0.194
ivp3	–1.218	0.227
ivp4	–1.546	0.148
ivp5	–1.591	0.100
ivp6	–2.717	0.194
ivp7	–2.820	0.288
ivp8	–1.915	0.216
ivp9	–1.940	0.170
ivp10	–0.895	0.140
ivp11	–6.149	0.276
ivp12	–2.236	0.191
ivp13	–0.863	0.164
ivp14	–2.346	0.167
ivp15	–3.253	0.212
ivp16	–1.008	0.135
ivp17	–1.660	0.117
ivp18	–1.233	0.139
ivp19	–1.310	0.174
ivp20	–2.881	0.283
ivp21	–6.385	0.338
adj. R^2	0.840	

Controls: brand and time dummy variables.

back

A Logit for LP membership

Variable	OLS	Logit
Log total Sup. LP members excluding location department	0.141	0.892
	(0.028)	(0.185)
NB included	0.0499	0.322
	(0.020)	(0.137)
Online access to inscription form	0.0410	0.261
	(0.012)	(0.074)
Log of average reward per 100 €spent	0.0110	0.0692
	(0.005)	(0.033)
Controls for Demographics	Yes	Yes
adj. R^2	0.034	

Robust standard errors in parentheses.
All displayed coefficients are significant at 5% level.

back1

back2

Optimal instruments: theory

Following Chamberlain (1987) and Newey (1990), assume

$$E[\rho(x_i, \theta_o)|z_i] = 0 \tag{16}$$

Assuming homoscedasticity, the conditional var-cov matrix is

$$E[\rho(x, \theta_o)\rho(x, \theta_o)'|z_i] = \Omega \tag{17}$$

Estimation relies on

$$E[h(z_i)\rho(x_i, \theta_o)] = 0 \tag{18}$$

The optimal choice of $h(\cdot)$ is

$$h^*(z_i) = D(z)'\Omega^{-1} \tag{19}$$

where

$$D(z) = E\left[\frac{\partial \rho_{.t}(x, \theta_o)}{\partial \theta}\middle| z_i\right] \tag{20}$$

Optimal instruments for θ_1

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \beta'}\middle| z_t\right] = E[x_j|z_t] = x_j \tag{21}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \lambda'}\middle| z_t\right] = E[r_s|z_t] = r_s \tag{22}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \alpha}\middle| z_t\right] = E[p_{jst}|z_t] = x_j\gamma_1 + r_s\gamma_2 + w_{jst}\gamma_3 \tag{23}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \varphi}\middle| z_t\right] = E[M_{ms}|z_t] = r_s\tau_1 + l_s\tau_2 \tag{24}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \eta}\middle| z_t\right] = E[M_{ms}|z_t] \times PL_{js} \tag{25}$$

$$E\left[\frac{\partial \rho_{.t}(x,\theta)}{\partial \theta_2'}\middle| z_t\right] = E\left[\frac{\partial \delta_{mjst}(s_{.t},\theta_2)}{\partial \theta_2'}\middle| z_t\right] \tag{26}$$

back