WORKING PAPERS

N° TSE-520

August 2014

"Analysis of Multiple Discrete-Continuous Choices: Empirical Evidence of Biased Price Elasticities under Standard Discrete Choice models on the Soft Drink Market"

Céline Bonnet and Olivier de Mouzon

Toulouse School of Economics

Analysis of Multiple Discrete-Continuous Choices: Empirical Evidence of Biased Price Elasticities under Standard Discrete Choice models on the Soft Drink Market

Céline Bonnet*and Olivier de Mouzon†

August 2014‡

Abstract

The New Empirical Industrial Organization (NEIO) literature allows to analyze competition, market power and welfare implications of public policies. This literature is based on the estimation of consumer substitution patterns, that is on the estimation of demand models. Standard discrete choice models are commonly used and assume that the consumer chooses only one unit of a single alternative. However, panelists are often shoppers making decisions for the entire household. A household is composed of several persons with various tastes. The current NEIO literature then often omits multiple brand choice and quantity choice of a product representing multiple tastes within the household. To tackle this point, we model multiple discrete/continuous choices of households following Bhat (2005a). Moreover, we deal with observed and unobserved household heterogeneity as well as omitted variable problem that occurs in this type of consumer choice model. Using French purchase data on the Soft Drink Market, we show that consumer substitution patterns could be biased if multiple brand and quantity choices are not taken into account. Indeed, our results suggest that own price elasticities could be significantly underestimated.

Key words: multiple purchases, discrete-continuous choices, soft drink market, price elasticities, household heterogeneity

*Toulouse School of Economics (GREMAQ, INRA), 21 Allée de Brienne, F-31000 Toulouse France, Tel: +33 (0)5 61 12 85 91, cbonnet@toulouse.inra.fr

†Toulouse School of Economics (GREMAQ, INRA), 21 Allée de Brienne, F-31000 Toulouse France, Tel: +33 (0)5 61 12 85 52, Olivier.deMouzon@toulouse.inra.fr

‡We want to thank the Scientific Committee of the Economics Department of INRA which has funded this project. We also thanks Fabrice Etile for his comments and participants of the EAAE-AAEA seminar in Freising, the Jornadas of Economia Industrial in Murcia, the Journées Jeunes Chercheurs in Montpellier for their comments. Any remaining errors are ours.

1 Introduction

The New Empirical Industrial Organization literature allows to analyze competition, market power and welfare implications of public policies. This literature is based on the estimation of consumer substitution patterns, that is on the estimation of demand models. Standard discrete choice models are commonly used and assume that the consumer chooses only one unit of a single alternative. These models are largely used since Guadani and Little (1983) and Train (1986). The first set of papers allowing to extend discrete choice models are the discrete-continuous models developed by Hanemann (1984), and Dubin and McFadden (1984) and used later in marketing by Krishnamurthi and Raj (1988), Chiang (1991), Chintagunta (1993), and Dillon and Gupta (1996). These papers deal not only with brand choice behavior but also with quantity choice for the chosen brand. However, panelists are often shoppers making decisions for the entire household. There may be different tastes within a household, which implies buying different products. For instance, children and parents may have very different tastes within a category for flavors, styles and colors or men and women for diet products. The variety of products bought during the shopping trip may also correspond to the shopping planed for different consumption occasions. Indeed the household may seek for variety from one consumption occasion to another.

A recent marketing literature deals with the multiple discrete choice models. Dubé (2004, 2005) and Hendel (1999) estimate a structural model that allows households to purchase a bundle of products and suppose the shopping purchase occasions correspond to several future consumption occasions. During each consumption occasion, they assume a standard discrete choice model where only one product is consumed. Therefore, due to varying tastes across individual consumption occasions, a household consumes a variety of goods at the current purchase occasion. They take into account this taste variation assuming a normal distribution in the specification of the model. The number of consumption occasions is assumed to follow a poisson distribution. Kim, Allenby and Rossi (2002) consider a Kuhn Tucker approach to model the multiple discreteness of demand of goods. They propose a translated additive utility structure which allows to obtain corner and interior solutions of the maximization utility problem. This approach is a Dubé/Hendel alternative approach but has the advantage to allow different satiation parameters or diminishing returns to differ across products. However, the model of Kim *et al.* assumes that the subutility of a good not purchased is different from zero. Hence, this model does not allow to take into account the weak complementarity assumption. On the other hand, they use a normal distribution for the error terms that does not allow to have a closed form expression for the probability. Their model is not very practicable. Bhat (2005b) and Bhat and Sen (2006) extend the paper of Kim *et al.* assuming a different assumption on the distribution of the error terms. An Independently and Identically Distributed (IID) Gumbel instead of IID Normal assumption of the error terms allows to have a simple closed form expression for the discrete-continuous probabilities. Bhat (2005a) extends the previous papers using a more easy-to-interpret and general utility form.

Another current literature of the multiple continuous/discrete choice models is on environmental economics (von Haefen and Phaneuf, 2005; von Haefen, 2003; Phaneuf *et al.*, 2000 Phaneuf and Herrigues, 2000; Herrigues *et al.*, 2004). They suppose a linear expenditure system form for the utility function contrary to previous papers which consider a constant elasticity of substitution (CES) form. The important problem of their model is that they assume a deterministic subutility for the benchmark product (the outside numeraire option) in order to simplify the expression of the probability. So their model depends on the benchmark product and this implies different expressions and values for the probability for the same consumption pattern according to the benchmark product chosen. Another problem about these models is that they employ a numerical gradient method to estimate the Jacobian whereas Bhat (2005a) finds an analytical expression for it. Their method implies a less precise and slower computation of the Jacobian.

An alternative approach, the Neoclassical Demand System approach, would allow to tackle quantity and discrete choices of consumers. It specifies a system of demand equations, one for each competing product in the market under investigation. This approach assumes that the demand for a product depends on its price and the price of competing products, and other variables as well. Several models are derived from this approach: the linear expenditure model (Stone, 1954), the Rotterdam model (Theil, 1965; Barten, 1966), the translog model (Christensen et al., 1975), the Almost Ideal Demand System

(Deaton and Malbauer, 1980), or the multilevel demand system (Hausmann, 1996). However they have two drawbacks. When the number of products becomes large, the number of parameters to be estimated dramatically increases. For example, dealing with a market composed of 100 differentiated products leads to estimate 10,000 parameters[1]. An additional problem is the heterogeneity in consumer tastes. This approach describes the preference for an average consumer and does not allow to take into account individual heterogeneity in purchasing behaviors. Hence, this kind of models does not reflect all the differentiation in a market and can thus affect policy conclusions.

The objective of this paper is to adapt the previous work of Bhat to an household choice behavior of food consumption. We want to guess the bias in the estimation of price elasticities by not taking account of multiple discrete-continuous choices of consumers. Moreover, we deal with omitted variable problem that occurs in this type of consumer choice models and with some observed and unobserved household heterogeneity as well. Our paper suggests that not taking into account quantity and multiple choices could significantly underestimate price elasticities and could then bias competition policy analyses. On the Soft Drink market, own price elasticities with standard discrete choice models could be more than 30% lower than with our model. We also find that dealing with omitted variables problem avoids to significantly underestimate own price elasticities.

The paper is organized as follows. Section 2 describes the Soft Drink market and the French available data on this sector. Section 3 presents the multiple continuous/discrete choice model and section 4 discusses the results. In section 5, we examine some robustness checks on demand estimates. Section 6 concludes.

2 The Soft Drink market and Data

In 2006, according to the National Association of Soft Drinks, turnover of this industry reaches more than 2 billion euros, that is 1.5% of the total turnover of the food industry. This sector is dynamic since its production rises each year (for instance, +4% between 2005 and 2006). Refreshing drinks mainly include

[1] This number could decrease imposing economic restrictions (symmetry and adding up restrictions implied by economic theory) but remains large.

colas, fruit drinks, ice tea, fruit juices and nectars. In France, the total consumption of soft drinks reaches in average 60 liters per year and per individual. This consumption is comparatively weak with respect to the mean European consumption of soft drinks by 94 liters and the mean US consumption by 160 liters.

We use data from a consumer panel data collected by Kantar World Panel. We have a French representative survey of 9,458 households over the year 2005. This survey provides information on purchases of all food products (quantity, price, date, store, characteristics of goods) and on characteristics of households (number of children, number of persons, weight, height, age and sex of each member of the household...). From the panel data, we select the 13 main national brands of the soft drink industry and an "aggregate" private label. These 14 products are differentiated according to the three main characteristics of products in this market: diet, pure juice and carbonated drink. We then analyze the consumer choice behavior through 34 differentiated products. We split the year 2005 into 13 periods of four weeks.

Our sample contains 253,253 observations over the 13 periods. Households buy 3.25 different products by period in average. This figure can vary from 1 to 14 and almost 65% of households buy more than one product in a period. This motivates that modeling only a unique consumer choice by period does not allow to account for the true consumption behavior. Table 7.1 in the appendix (7.1) gives mean prices and quantities for each product considered in this market. We see that prices are very heterogeneous according to the product and can vary from 0.3€/liter to 2.29€/liter. Mean and standard deviation for quantities suggest that some heterogeneity exists in the quantity choice. Indeed, the mean quantity vary across products (from 0.68 to 3.12 liters) and standard deviations are relatively large, meaning that the quantity choice for a given product is also very heterogeneous. Table 7.1 in the appendix (7.1) presents average prices and quantities according to main characteristics: regular and diet products, still and sparkling products, and pure juice products or not. Average price for diet products is lower than for regular ones, and still products are in average more expensive than sparkling ones. Pure juice products, both regular and still, are more expensive than soft drinks like colas or tonics. Furthermore, diet products are bought in larger quantities than regular ones, sparkling products are also bought in larger quantities than still ones, and pure juice products are bought in fewer quantities than other ones.

These data allow us to observe household and individual characteristics. Households are composed of 2.7 persons in average and this figure can vary from 1 to 9. We also know the socio-economic class of households. Our sample contains 13% of rich people, 31% of average upper, 40% of average lower, and 16% of poor people.

3 The model

We suppose that households are faced to J products during T time periods. We add a $(J+1)^{th}$ product: the outside good that represents either all the food and drink purchases or other brands in the soft drink sector purchased by households with very low market shares. We allow households to buy several products during the same period, that is households may select multiple products and multiple units of each contrary to traditional discrete choice models where a single alternative only is chosen among a set of available products. We then intend to model a multiple continuous-discrete choice behavior.

Let u_{jht} be the subutility of household h at time t to consume product j. We have chosen an additive utility structure because we suppose that products are not jointly consumed and then we suppose that the utility gained by the consumption of one product is not affected by the consumption of the others. Hence, the utility of household h at period t is:

$$U_{ht}(q_{ht}) = \sum_{j=1}^{J+1} u_{jht}(q_{jht}),$$

where q_{ht} is the vector of quantities bought for each product $(j = 1, 2, ..., J+1)$.

Each household maximizes its utility subject to the budget constraint $\sum_{j=1}^{J+1} p_{jt}q_{jht} \leqslant y_{ht}$, where p_{jt} is the price of product j at period t, q_{jht} is the quantity of product j bought by household h at time t, and y_{ht} is the expenditure of the group of products under study for household h at period t.[2]

To solve for optimal demand, we form the Lagrangian and apply the Kuhn-Tucker first order conditions. The Lagrangian function for this problem is given by:

[2] We assume no variation in expenditures allocated to the market under investigation. We will compare in section 5 two markets: the soft drink market and the food and drink market. We assume for both that households choose soft drink products given a fixed budget per period. The household can only switch his consumption from a bundle of products to another one, outside option included.

$$\mathcal{L} = \sum_{j=1}^{J+1} u_{jht}(q_{jht}) - \lambda \left(\sum_{j=1}^{J+1} p_{jt} q_{jht} - y_{ht} \right),$$

where λ is the Lagrangian multiplier associated with the budget constraint.

The Kuhn-Tucker (KT) first order conditions are given by differentiating the Lagrangian:

$$u'_{jht}(q^*_{jht}) - \lambda p_{jt} \overset{q^*_{jht}}{\lhd} 0,$$

where q^*_{jht} is the optimal demand of household h for the product j at time t and $\overset{q^*_{jht}}{\lhd}$ is defined as:

$$\begin{array}{lll} \overset{q^*_{jht}}{\lhd} \text{ is} & = & \text{if } q^*_{jht} > 0, \text{ and} \\ \overset{q^*_{jht}}{\lhd} \text{ is} & \leqslant & \text{if } q^*_{jht} = 0 \end{array}$$

3.1 Random utility specification

As in Bhat (2005b), we specify a quasi-concave, increasing and continuously differentiable utility function with respect to the quantity bought $q_{ht} = (q_{1ht}, ..., q_{Jht}, q_{(J+1)ht})$ which belongs to the family of translated utility functions and we then assume the following form for the utility of household h at time t:

$$U_{ht}(q_{ht}) \quad = \quad \sum_{j=1}^{J+1} \frac{\gamma_j}{\alpha} \Psi(x_{jt}, \varepsilon_{jht}) \left\{ \left(\frac{q_{jht}}{\gamma_j} + 1 \right)^{\alpha} - 1 \right\},$$

where $\Psi(x_{jt}, \varepsilon_{jht})$ is the baseline utility for product j which captures the quality of product j through the characteristics x_{jt} of product j at time t and the idiosyncratic unobserved characteristics ε_{jht} of product j for household h at time t. We then suppose a random utility specification for the baseline utility Ψ. Besides, γ_j, associated to product j, and α are parameters to be estimated. Moreover, we normalize the subutility of the outside good (J+1) to zero and then assume that $\Psi(x_{(J+1)t}, \varepsilon_{(J+1)ht}) = e^{\varepsilon_{(J+1)ht}}$.

The parameter γ_j enables corner solutions for indifference curves[3] and governs the level of satiation of product j, i.e. the level of consumption for product j from which a consumer has had enough, that is the level from which they do not value so much an additional unit[4], and the parameter α is the global

[3] If γ_j is 0, the indifference curves are tangent to the axes, then there will be no corner solutions.

[4] We suppose that the level of satiation of product J+1 is equal to 1 in order to identify the other J+1 coefficients γ_j for $j = 1, ..., J$ and α

satiation parameter, which allows to decrease the marginal utility when the consumption increases (Bhat, 2005a). This specification supposes the assumption of weak complementarity, i.e. if household h does not consume product j at time t, the corresponding subutility will be zero: $u_{jht} = 0$. Then the household does not receive any utility from this product at this period.

If we suppose the consumer chooses only one unit of a single alternative only, $j = 1, ..., J+1$, i.e. $q_{jht} = 1$ and $\forall j' \neq j, q_{j'ht} = 0$, this specification is simply the expression of the multinomial logit model: $U_{jht} = \frac{\gamma_j}{\alpha}\Psi(x_{jt}, \varepsilon_{jht})\left\{\left(\frac{1}{\gamma_j}+1\right)^{\alpha} - 1\right\} \approx \Psi(x_{jt}, \varepsilon_{jht})$ when $\alpha \to 1$. Going back to the general case, note that $\alpha \to 1$ and high values of $\Psi(x_{jt}, \varepsilon_{jht})$ for product j only imply that we expect purchases of large quantities of this product only. On the other hand, small values of α imply multiple products purchased if the $\Psi(x_{jt}, \varepsilon_{jht})$s are not too different from one j to another.

This function is valid if $\Psi(x_{jt}, \varepsilon_{jht}) > 0$, $\gamma_j > 0$ and $0 \leqslant \alpha \leqslant 1$[5]. To impose the three conditions, we suppose that:

(i) $\Psi(x_{jt}, \varepsilon_{jht}) = e^{\beta' x_{jt} + \varepsilon_{jht}}$, where β is a vector of parameters to be estimated. The exponential form guarantees the positivity of the baseline utility.

(ii) $\gamma_j = e^{\mu_j}$, which ensures $\gamma_j > 0$, and μ_j is estimated.

(iii) $\alpha = \frac{1}{1+e^{\delta}}$ and we estimate δ to obtain $\alpha \in (0, 1)$.

3.2 Optimal demand

According to the previous specification of the utility function, the Lagrangian can be written as:

$$\mathcal{L} = \sum_{j=1}^{J+1} \frac{\gamma_j}{\alpha}\Psi(x_{jt}, \varepsilon_{jht})\left\{\left(\frac{q_{jht}}{\gamma_j}+1\right)^{\alpha} - 1\right\} - \lambda\left(\sum_{j=1}^{J+1} p_{jt} q_{jht} - y_{ht}\right)$$

and the KT first order conditions (for $j = 1, \ldots, (J+1)$) become:

$$\Psi(x_{jt}, \varepsilon_{jht})\left(\frac{q^*_{jht}}{\gamma_j}+1\right)^{\alpha-1} - \lambda p_{jt} \overset{q^*_{jht}}{\lhd} 0.$$

[5] $\alpha > 0$ comes from the division by α in the utility function, because we model a positive utility of consumption. $\alpha < 1$ because physiologically households should reach satiety after some quantities consumed: the marginal utility of consuming a larger quantity is positive, but decreasing in the quantity consumed.

Besides, the optimal demand q^*_{jht} satisfies the budget constraint. As we assumed $y_{ht} > 0$, at least one of the $J+1$ alternatives was bought. Let j^0_{ht} be such an alternative ($q^*_{j^0_{ht}ht} > 0$). Then, the previous equations (given for $j = 1, \ldots, (J+1)$) lead to:

$$\lambda = \frac{\Psi(x_{j^0_{ht}t}, \varepsilon_{j^0_{ht}ht}) \left(\frac{q^*_{j^0_{ht}ht}}{\gamma_{j^0_{ht}}} + 1 \right)^{\alpha-1}}{p_{j^0_{ht}t}}.$$

The previous equation enables to concentrate only on the J ($\forall j \neq j^0_{ht}$) remaining KT first order conditions and taking logarithms and replacing the perceived quality of the product considered by its expression ($\Psi(x_{jt}, \varepsilon_{jht}) = e^{\beta' x_{jt} + \varepsilon_{jht}}$), we obtain the following simplified expression, $\forall j = 1, \ldots, (J+1)$ and $j \neq j^0_{ht}$:

$$V_{jht} + \varepsilon_{jht} \overset{q^*_{jht}}{\lhd} V_{j^0_{ht}ht} + \varepsilon_{j^0_{ht}ht},$$

where

$$V_{jht} = \beta' x_{jt} + (\alpha - 1) \ln \left(\frac{q^*_{jht}}{\gamma_j} + 1 \right) - \ln p_{jt}.$$

As $q^*_{j^0_{ht}ht}$ is determined by using the budget constraint ($y_{ht} = \sum_{j=1}^{J+1} p_{jt} q^*_{jht}$), the optimal quantity for alternative j^0_{ht} depends on the vector of optimal quantities of the other alternatives. Therefore the above KT conditions can be rewritten using:

$$V_{j^0_{ht}ht} = \beta' x_{j^0_{ht}t} + (\alpha - 1) \ln \left(\frac{y_{ht} - \sum\limits_{j \neq j^0_{ht}} p_{jt} q^*_{jht}}{\gamma_{j^0_{ht}} p_{j^0_{ht}t}} + 1 \right) - \ln p_{j^0_{ht}t}.$$

Let $\mathcal{J}_{ht} = \left\{ j = 1, \ldots, (J+1) \mid q^*_{jht} > 0 \right\}$ be the set of alternatives which were bought by household h at time t, and $K_{ht} = |\mathcal{J}_{ht}|$ the number of alternatives which were bought by household h at time t. As we consider only cases where at least one alternative was bought ($K_{ht} \geqslant 1$) since the budget of households is fixed per period, $\exists j^0_{ht} \in \mathcal{J}_{ht}$. Then, as previous equations were given for any j^0_{ht} and are symmetric in all other $j \neq j^0_{ht}$, we can, without loss of generality, reorder the alternatives so that $\widehat{j^0_{ht}} = 1$ and $\widehat{\mathcal{J}_{ht}} = \{1, \ldots, K_{ht}\}$, where $\widehat{\cdot}$ denotes the reordering operator.

Let $f(\varepsilon_{1ht}, \ldots, \varepsilon_{(J+1)ht})$ be the joint probability density function of ε_{jht} $(j = 1, \ldots, (J+1))$. The probability that household h purchases the first K_{ht} of the $(J+1)$ alternatives at time t is given by:

$$\Pr(\overset{\frown}{q^*_{1ht}}, \ldots, \overset{\frown}{q^*_{K_{ht}ht}}, \overset{\frown}{0}, \ldots, \overset{\frown}{0}) = \int\limits_{\varepsilon_{1ht}=-\infty}^{+\infty} \cdots \int\limits_{\varepsilon_{(J+1)ht}=-\infty}^{+\infty} f(\varepsilon_{1ht}, \ldots, \varepsilon_{(J+1)ht}) d\varepsilon_{1ht} \ldots d\varepsilon_{(J+1)ht}.$$

According to the KT conditions, $\forall j \in \overset{\frown}{\mathcal{J}_{ht}}$, $\varepsilon_{jht} = V_{1ht} + \varepsilon_{1ht} - V_{jht}$, leading to the following expression for the previous probability:

$$\begin{aligned}
&\Pr(\overset{\frown}{q^*_{1ht}}, \ldots, \overset{\frown}{q^*_{K_{ht}ht}}, \overset{\frown}{0}, \ldots, \overset{\frown}{0}) = |Jac_{ht}| \times \\
&\int\limits_{\varepsilon_{(K_{ht}+1)ht}=-\infty}^{V_{1ht}+\varepsilon_{1ht}-V_{(K_{ht}+1)ht}} \cdots \int\limits_{\varepsilon_{(J+1)ht}=-\infty}^{V_{1ht}+\varepsilon_{1ht}-V_{(J+1)ht}} \int\limits_{\varepsilon_{1ht}=-\infty}^{+\infty} f(\varepsilon_{1ht}, V_{1ht} - V_{2ht} + \varepsilon_{1ht}, ..., V_{1ht} - V_{K_{ht}ht} + \varepsilon_{1ht}, \varepsilon_{(K_{ht}+1)ht}, \ldots, \varepsilon_{(J+1)ht}) \\
&\qquad d\varepsilon_{(K_{ht}+1)ht} \ldots d\varepsilon_{(J+1)ht} d\varepsilon_{1ht}
\end{aligned}$$

where Jac_{ht} is the $(K_{ht}-1) \times (K_{ht}-1)$ Jacobian matrix which has for element (j, k):

$$Jac_{htjk} = \frac{\partial \left[V_{1ht} - V_{(j+1)ht} + \varepsilon_{1ht}\right]}{\partial q^*_{(k+1)ht}} = \frac{\partial \left[V_{1ht} - V_{(j+1)ht}\right]}{\partial q^*_{(k+1)ht}}$$

Then, assuming that $\forall j \in [1, \ldots, (J+1)]$ the ε_{jht}s are independently distributed across alternatives and have a centered (location parameter 0) Gumbel (also named type I extreme value) distribution of scale parameter σ and independent of the vector of variables x, price p and quantities q, the probability that household h purchases the first K_{ht} of the $(J+1)$ alternatives at time t takes this final expression (see Appendix for more details):

$$\begin{aligned}
&\Pr(\overset{\frown}{q^*_{1ht}}, \ldots, \overset{\frown}{q^*_{K_{ht}ht}}, \overset{\frown}{0}, \ldots, \overset{\frown}{0}) \\
&= \frac{(K_{ht}-1)!}{\sigma^{K_{ht}-1}} \left(\prod_{j \in \mathcal{J}_{ht}} \frac{(1-\alpha)}{q^*_{jht} + \gamma_j} \right) \left(\sum_{j \in \mathcal{J}_{ht}} \frac{q^*_{jht} + \gamma_j}{(1-\alpha)} \cdot \frac{p_{jt}}{p_{j^0_{ht}t}} \right) \frac{\prod_{j \in \mathcal{J}_{ht}} e^{\frac{V_{jht}}{\sigma}}}{\left(\sum_{j=1}^{J+1} e^{\frac{V_{jht}}{\sigma}} \right)^{K_{ht}}}.
\end{aligned}$$

The expression of the probability depends on all product prices and especially on product j^0_{ht} price. However, this price is constant in each individual likelihood function and then the estimation of parameters does not depend on this price, only the individual probability value will change[6]. The scale parameter σ

[6] The choice of the 'first' good consumed impacts on the purchase probability of consumers. This could be problematic if our results depended on probabilities estimated. However, not only parameter estimates do not change whatever the choice

should be positive because a price increase for product $j \in \widehat{\mathcal{J}_{ht}}$ should lead to a lower probability that consumer h purchases the first $\widehat{\mathcal{K}_{ht}}$ of the $(J+1)$ alternatives at period t. Indeed, the probability is an increasing function of $\frac{\widehat{V_{jht}}}{\sigma} = \frac{\beta'}{\sigma}\widehat{x_{jt}} - \frac{1}{\sigma}\ln\widehat{p_{jt}} + \frac{(\alpha-1)}{\sigma}\ln\left(\frac{\widehat{q^*_{jht}}}{\widehat{\gamma_j}}+1\right)$. This parameter becomes identifiable as the inverse of the estimated price parameter and could depend on household characteristics.

It has to be noted that we obtain a closed form expression from this probability which can be simplified to the standard multinomial logit model when $K_{ht} = 1$ (i.e. only one good is chosen: j^0_{ht}) and $\sigma = 1$ (using the standard Gumbel distribution): $P_{j^0_{ht}ht} = \frac{e^{V_{j^0_{ht}ht}}}{\sum_{j=1}^{J+1} e^{V_{jht}}}$, $\forall j^0_{ht} \in [1, \ldots, (J+1)]$.

3.3 Omitted variable problem

The previous expression of household h purchases probability is deduced from the assumption that x, p and q are independent of the error disturbances ε_{jht}. The individual error term could be split up in two components: $\varepsilon_{jht} = \xi_{jt} + e_{jht}$ where ξ_{jt} is product-specific characteristics varying in time and observed by both consumers and producers, but not included in the estimated specification and e_{jht} is a consumer specific idiosyncratic taste varying across products and time. Some omitted product characteristics included in ξ_{jt} could be correlated with prices. For instance, we don't know the amount of advertising that firms invest each month for their brand or the display on shelves in stores. This is then included in the error term because the publicity could be a determining factor in the choice process of households. As advertising is a non negligible part of the cost of soft drinks, it is obviously correlated with prices. To solve the endogeneity problem of prices, we use a two-stage residual inclusion approach as in Terza et al. (2008) or Petrin and Train (2010). We then regress prices on instrumental variables W_{jt}, that is input prices, as well as the exogenous variables of the baseline utility function, x_{jt}:

$$p_{jt} = W_{jt}\gamma + \beta_\eta x_{jt} + \eta_{jt}.$$

The estimated error term $\widehat{\eta}_{jt}$ of the first stage includes some omitted variables as advertising variations, promotions. Introducing this term in the indirect utility V_{jht} allows to capture unobserved characteristics.

of the first good but our elasticity results also do not depend on it. Indeed, we see in section 4.2 that elasticities depend only on demand parameters (through the maximization of the utility to recover quantities) and not on the probability estimates.

We then write

$$V_{jht} = \beta' x_{jt} + (\alpha - 1)\ln\left(\frac{q^*_{jht}}{\gamma_j} + 1\right) - \ln p_{jt} + \theta\widehat{\eta}_{jt},$$

where λ is the estimated parameter associated with the estimated error term of the first stage. Now, the new error term $\xi_{jt} + e_{jht} - \theta\widehat{\eta}_{jt}$ is not correlated with prices.

3.4 Household heterogeneity

We are able to account for household heterogeneity through different forms. Taste of product characteristics can vary across households assuming that $\varepsilon_{jht} = \zeta_{jht} + \beta'_h x_{jt}$, where $\beta_h = \Sigma\nu_h$, and where Σ is a matrix of parameters, and ν_h captures the unobserved household characteristics. We also are able to allow for the scale parameter to vary across households (Scarpa, Thiene and Train, 2008). Hence, we could introduce some heterogeneity in price sensitivity as follows: $\sigma_h = \sum_{c=1}^{C} \tau_c I_{ch}$ where I_{ch} is the dummy representing the income class c of household h and τ_c its related coefficients.

The probability that household h purchases the first K_{ht} of the $(J+1)$ alternatives at time t then becomes:

$$\Pr(q_{ht}) =$$

$$\int_{\nu_h}\int_{D_h} \frac{(K_{ht}-1)!}{\sigma_h^{K_{ht}-1}} \left(\prod_{j\in\mathcal{J}_{ht}} \frac{(1-\alpha)}{q^*_{jht}+\gamma_j}\right)\left(\sum_{j\in\mathcal{J}_{ht}} \frac{q^*_{jht}+\gamma_j}{(1-\alpha)} \cdot \frac{p_{jt}}{p_{j^0_{ht}t}}\right) \frac{\prod_{j\in\mathcal{J}_{ht}} e^{\frac{V_{jht}}{\sigma_h}}}{\left(\sum_{j=1}^{J+1} e^{\frac{V_{jht}}{\sigma_h}}\right)^{K_{ht}}} dF_\nu(\nu_h)dF_I(I_h),$$

where F_ν represents idd multivariable cumulative normal distributions for each variable in the vector ν_h and F_I represents idd parametric distributions for the income class of households.

To estimate parameters $(\alpha, \tau_1, ..., \tau_C, \beta, \Sigma, \theta, \gamma_1, ..., \gamma_J)$, we maximize the following log likelihood:

$$LL\left(\alpha, \tau_1, ..., \tau_C, \beta, \Sigma, \theta, \gamma_1, ..., \gamma_J\right) =$$

$$\sum_{\substack{(h,t)\in \\ H\times T \mid \\ K_{ht}\geqslant 1}} \ln\left(\int_{\nu_h}\int_{D_h} \frac{(K_{ht}-1)!}{\sigma_h^{K_{ht}-1}} \prod_{j\in\mathcal{J}_{ht}} \frac{(1-\alpha)}{q^*_{jht}+\gamma_j}\left(\sum_{j\in\mathcal{J}_{ht}} \frac{q^*_{jht}+\gamma_j}{(1-\alpha)} \cdot \frac{p_{jt}}{p_{j^0_{ht}t}}\right) \frac{\prod_{j\in\mathcal{J}_{ht}} e^{\frac{V_{jht}}{\sigma_h}}}{\left(\sum_{j=1}^{J+1} e^{\frac{V_{jht}}{\sigma_h}}\right)^{K_{ht}}} dF_\nu(\nu_h)dF_I(I_h)\right)$$

and we approximate it using simulation techniques (Train, 2002).

4 Results

We now describe estimation results of several multiple discrete-continuous choice models (MDCCM): with and without taking into account the endogeneity problem, and with and without taking into account household heterogeneity. Finally, we show price elasticities and compare them with standard discrete choice models.

We choose the following baseline utility specification:

$$\Psi(x_{jt}, \varepsilon_{jht}) = e^{\delta_j + \beta_h^D D_j + \beta_h^S S_j + \beta_h^{PJ} PJ_j + \varepsilon_{jht}},$$

where δ_j is a brand fixed effect that captures the unobserved variation across brands, D_j is a dummy that takes one if the product is diet and β_h^D is the parameter associated that represents the taste of household h for this product characteristic, S_j is a dummy for the sparkling characteristic and β_h^S is the parameter associated that can vary across households, and finally PJ_j is a dummy that takes one if the product is a pure juice product and β_h^{PJ} is the parameter associated for this characteristic.

4.1 Estimation results

Table 1 and Table 2 present demand results without controling for endogeneity problem and controling for it respectively. We see in Table 2 that λ is positive and significant, meaning that the unobserved part explaining prices is positively correlated with brand choice. This result is consistent whether this unobserved part is advertising or displays. We find that consumer price sensitivity is larger and then not taking into account endogeneity problem could decrease the consumer price sensitivity estimated. Indeed, the price sensitivity of consumers is measured by $\frac{1}{\sigma}$, which gives 1.71 in the uncorrected model and 1.99 in the two-stage residual inclusion approach. Tastes for the diet and sparkling characteristics are negative in average, meaning that consumers prefer non-diet and non-sparkling beverages. On the contrary, consumers prefer pure juice beverages (as their taste for the pure juice characteristic is positive in average).

	Mean (ste 10^{-3})		Mean (ste 10^{-3})		Mean (ste 10^{-3})
α	0.001 (0.015)	γ_1	5.178 (2.778)	γ_{20}	4.024 (1.835)
σ	0.585 (0.019)	γ_2	8.140 (4.494)	γ_{21}	2.095 (2.720)
β_{Diet}	-1.382 (0.062)	γ_3	5.738 (0.890)	γ_{22}	1.909 (1.551)
$\beta_{Sparkling}$	-1.260 (0.058)	γ_4	8.549 (2.351)	γ_{23}	11.105 (4.907)
$\beta_{purejuice}$	0.129 (0.042)	γ_5	3.483 (1.066)	γ_{24}	3.770 (1.764)
δ_1	-0.889 (0.164)	γ_6	3.105 (2.135)	γ_{25}	3.184 (3.841)
δ_2	0.834 (0.097)	γ_7	3.838 (1.222)	γ_{26}	3.037 (1.484)
δ_3	-1.454 (0.148)	γ_8	3.654 (1.767)	γ_{27}	1.408 (1.938)
δ_4	-0.193 (0.140)	γ_9	3.385 (0.096)	γ_{28}	4.175 (0.483)
δ_5	-0.047 (0.134)	γ_{10}	3.161 (1.472)	γ_{29}	3.968 (0.541)
δ_6	-1.673 (0.152)	γ_{11}	4.378 (1.366)	γ_{30}	2.823 (2.014)
δ_7	-0.585 (0.164)	γ_{12}	3.199 (3.339)	γ_{31}	1.841 (2.562)
δ_8	-1.344 (0.135)	γ_{13}	3.147 (1.128)	γ_{32}	5.059 (0.794)
δ_9	-1.722 (0.177)	γ_{14}	2.355 (1.781)	γ_{33}	1.129 (2.289)
δ_{10}	-2.375 (0.195)	γ_{15}	3.905 (1.311)	γ_{34}	5.022 (1.374)
δ_{11}	-0.952 (0.131)	γ_{16}	4.144 (1.590)		
δ_{12}	-1.911 (0.191)	γ_{17}	1.799 (3.356)		
δ_{13}	-1.387 (0.191)	γ_{18}	2.718 (2.262)		
δ_{14}	-0.480 (0.086)	γ_{19}	3.632 (1.672)		
Log Likelihood			-703,497		

Table 1: **Demand results for the MDCC model without endogeneity.**

Moreover, comparing Table 1 and Table 2, the estimates of the other variables affecting utility are robust to instrumentation except for the global satiation parameter. Parameter α is 0.14 in Table 2, meaning that consumers do not value so much an additional unit of beverage of the same product and prefer to get a unit of another product (if consumers approximatively value and saturate for both products in the same way). Indeed, an additional unit of the same product increases less the utility of consumers than a unit of another product. Without controlling for endogeneity problem, this global satiation parameter falls to 0.001, meaning that we underestimate the utility of an additional unit of the same product. Parameters γ_j vary between 1.13 and 10.81. Then, the threshold from which the consumer does not value so much an additional unit of product j is heterogeneous across products. This threshold is quite high for pure juice products and cola products, especially for diet colas.

We now present some results of demand estimations with unobserved and observed heterogeneity. Table 3 presents two models, one without endogeneity control and one with endogeneity control. We introduce some unobserved heterogeneity in the taste of product characteristics (diet, sparkling, and

	Mean (ste 10^{-3})		Mean (ste 10^{-3})		Mean (ste 10^{-3})
α	0.140 (0.140)	γ_1	5.194 (2.786)	γ_{20}	4.056 (1.853)
σ	0.502 (0.080)	γ_2	8.207 (1.968)	γ_{21}	2.179 (2.858)
β_{Diet}	-1.212 (0.160)	γ_3	5.751 (0.894)	γ_{22}	1.922 (1.564)
$\beta_{Sparkling}$	-1.131 (0.110)	γ_4	8.602 (2.370)	γ_{23}	10.81 (4.907)
$\beta_{purejuice}$	0.166 (0.052)	γ_5	3.499 (1.072)	γ_{24}	3.770 (1.764)
δ_1	-0.106 (0.809)	γ_6	3.131 (2.157)	γ_{25}	3.184 (3.841)
δ_2	1.414 (0.607)	γ_7	3.829 (1.218)	γ_{26}	3.037 (1.484)
δ_3	-0.591 (0.868)	γ_8	3.651 (1.763)	γ_{27}	1.408 (1.938)
δ_4	0.512 (0.764)	γ_9	3.392 (0.964)	γ_{28}	4.175 (0.483)
δ_5	0.638 (0.750)	γ_{10}	3.179 (1.482)	γ_{29}	3.968 (0.541)
δ_6	-0.865 (0.880)	γ_{11}	4.382 (1.367)	γ_{30}	2.823 (2.014)
δ_7	0.147 (0.810)	γ_{12}	3.130 (3.245)	γ_{31}	1.841 (2.562)
δ_8	-0.366 (0.854)	γ_{13}	3.154 (1.131)	γ_{32}	5.059 (0.794)
δ_9	-0.774 (0.939)	γ_{14}	2.369 (1.793)	γ_{33}	1.129 (2.289)
δ_{10}	-1.403 (0.965)	γ_{15}	3.984 (1.346)	γ_{34}	5.022 (1.374)
δ_{11}	-0.100 (0.844)	γ_{16}	4.094 (1.563)	λ	0.870 (0.668)
δ_{12}	-0.991 (0.939)	γ_{17}	1.543 (2.741)		
δ_{13}	-0.419 (0.937)	γ_{18}	2.681 (2.219)		
δ_{14}	0.186 (0.675)	γ_{19}	3.637 (1.674)		
Log likelihood			-702,761		

Table 2: **Demand results of the MDCC model with endogeneity control.**

pure juice) and observed heterogeneity across income class on the scale parameter. This will allow to get different price sensitivities across households according to income classes. Results suggest that some unobserved heterogeneity exits in the taste of product characteristics. Households do not like the diet characteristic on average and as the standard deviation is quite small relatively to the mean coefficient, a large part of households in our sample do not like it. The same conclusion could be done for the sparkling characteristic. On the contrary, as the mean coefficient for the pure juice characteristic is positive on average and the standard deviation of the normal distribution is larger, only 69% of households like pure juice products. The other coefficients are robust to the introduction of the unobserved heterogeneity. Results of model 4 suggest that the richer the household is, the more price sensitive he is. Even if this result is surprising, it is consistent with other works on this market as for instance Bonnet and Réquillart (2013). The other coefficients nearly remain the same.

MDCCM	Model 3	Model 4
	Mean (ste 10^{-3})	Mean (ste 10^{-3})
α	0.001 (0.015)	0.103 (0.172)
$\sigma\times$Well_Off	0.539 (0.026)	0.482 (0.091)
$\sigma\times$Average Upper	0.558 (0.022)	0.499 (0.094)
$\sigma\times$Average Lower	0.580 (0.022)	0.519 (0.098)
$\sigma\times$Modest	0.610 (0.030)	0.546 (0.104)
β_{Diet}	-1.562 (0.104)	-1.413 (0.236)
$\sigma_{\beta_{Diet}}$	0.590 (0.120)	0.527 (0.146)
$\beta_{Sparkling}$	-1.270 (0.060)	-1.166 (0.134)
$\sigma_{\beta_{Sparkling}}$	0.319 (0.072)	0.286 (0.083)
$\beta_{purejuice}$	0.115 (0.047)	0.145 (0.063)
$\sigma_{\beta_{purejuice}}$	0.350 (0.086)	0.301 (0.098)
θ		0.854 (0.229)
Brand fixed effects not shown		
Satiation parameters γ_j not shown		
Log likelihood	-701,867	-701,164

Table 3: **Demand results for the MDCC model with househol heterogeneity.**

4.2 Price elasticities

The objective of this paper is to compare consumer substitution patterns between the multiple discrete-continuous choice model presented above and the standard logit model representing discrete choices of households and hence to assess the bias in price elasticities and the policy conclusions.

The quantities consumed by household h at time t are based on the following problem:

$$\underset{q_{ht}}{Max}\ U_{ht}(q_{ht}) = \sum_{j=1}^{J+1} \frac{\gamma_j}{\alpha} \Psi(x_{jt}, \varepsilon_{jht}) \left\{ \left(\frac{q_{jht}}{\gamma_j} + 1 \right)^{\alpha} - 1 \right\},$$

subject to $y_{ht} = \sum_{j=1}^{J+1} p_{jt} q_{jht}$ and $\forall j = 1, ..., J+1, q_{jht} \geqslant 0$.

We use the algorithm of Pinjari and Bhat (2011) to forecast the quantities for Multiple Discrete-Continuous Choice model (see Appendix 7.4 for details about the algorithm). Once we have estimated optimal quantities, we recover price elasticities of the aggregate demand estimated by evaluating a centered numerical derivative of quantities estimated. The elasticities reported in Table 4 are mean across periods and products sold under each of the fourteen brands. Models 1, 2, and 3 are standard multinomial logit (ML) models, the second and third ones correct the omitted variable problem and the third one introduces some consumer heterogeneity. Estimation results of the three models are presented in Table 7.3 in the appendix. Models 4, 5 and 6 are Multiple Discrete-Continuous Choice (MDCC) models. Model

5 and 6 solve the endogeneity problem and Model 6 introduces some household heterogeneity.

We see that omitting the multiple discrete-continuous choice would lead to underestimate own price elasticities when we compare Model 1 to Model 4, Model 2 to Model 5, and Model 3 to Model 6. Indeed, own price elasticities of Model 4 are almost 30% larger than the ones of Model 2. Moreover, we can see that the MDCC models allow to highlight the heterogeneity of own price elasticities across brands. Indeed, they are different in the MDCC models whereas they are quite similar in the ML models.

We also show that not taking into account omitted variable problem would also lead to underestimate price elasticities by comparing Models 1 and 2 in the case of MLM and Models 4 and 5 in the cas of MDCCM. Our model allows to introduce some heterogeneity in consumer substitution patterns across products without introducing some household heterogeneity and this is due to the modelisation of multiple choices and quantities that is very heterogenous across households. We see that own price elasticities can vary across brands and standard deviations in brackets show some heterogeneity within the brand, that is across products which are from the same brand but with different characteristics as diet, sparkling and pure juice characteristics. Regarding cross price elasticities, we get significant differences between MLM and MDCCM but we can not conclude about the sence of the bias.

Using observed heterogeneity to describe choice behavior on soft drink market allows to compute price elasticities according to the observed consumer characteristics introduced in the model. In this paper, we permit different effects of the income class on the scale parameter of the Gumbel distribution of the random part of the baseline utility. We then allow for different deviations of the mean quality of products according to the income class. This leads to significant own price elasticities. Table 5 presents the average own price elasticities for the four income class. Given that richer households are more sensitive to price variations, their price elasticities are higher. The differences reach up to 10% between rich and poor households. A price policy will then affect less poor individuals in terms of variation of consumption.

5 Robustness checks

	MLM			MDCCM		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Brand1	-1.57 (0.01)	-1.93 (0.00)	-1.93 (0.00)	-2.09 (0.06)	-2.38 (0.06)	-2.33 (0.06)
Brand2	-1.49 (0.07)	-1.82 (0.09)	-1.82 (0.09)	-2.17 (0.08)	-2.45 (0.06)	-2.39 (0.06)
Brand3	-1.57 (0.01)	-1.92 (0.01)	-1.92 (0.01)	-1.99 (0.06)	-2.26 (0.07)	-2.21 (0.06)
Brand4	-1.57 (0.02)	-1.92 (0.01)	-1.92 (0.01)	-2.08 (0.05)	-2.36 (0.03)	-2.31 (0.04)
Brand5	-1.57 (0.01)	-1.92 (0.01)	-1.92 (0.01)	-2.06 (0.04)	-2.34 (0.03)	-2.29 (0.04)
Brand6	-1.57 (0.01)	-1.92 (0.01)	-1.92 (0.01)	-1.95 (0.04)	-2.19 (0.07)	-2.14 (0.06)
Brand7	-1.57 (0.01)	-1.93 (0.00)	-1.93 (0.00)	-2.01 (0.04)	-2.29 (0.03)	-2.25 (0.03)
Brand8	-1.57 (0.01)	-1.92 (0.01)	-1.92 (0.01)	-2.04 (0.04)	-2.32 (0.02)	-2.28 (0.02)
Brand9	-1.58 (0.00)	-1.93 (0.00)	-1.93 (0.00)	-2.05 (0.02)	-2.33 (0.03)	-2.29 (0.02)
Brand10	-1.58 (0.00)	-1.93 (0.00)	-1.93 (0.00)	-1.88 (0.03)	-2.14 (0.02)	-2.09 (0.02)
Brand11	-1.56 (0.00)	-1.91 (0.00)	-1.91 (0.00)	-2.15 (0.15)	-2.42 (0.13)	-2.38 (0.12)
Brand12	-1.58 (0.00)	-1.93 (0.00)	-1.93 (0.00)	-2.00 (0.05)	-2.27 (0.06)	-2.23 (0.06)
Brand13	-1.58 (0.00)	-1.93 (0.00)	-1.93 (0.00)	-2.05 (0.03)	-2.35 (0.03)	-2.30 (0.04)
Brand14	-1.51 (0.06)	-1.84 (0.08)	-1.84 (0.08)	-1.89 (0.04)	-2.15 (0.06)	-2.09 (0.05)
Endogeneity	No	Yes	Yes	No	Yes	Yes
Heterogeneiy	No	No	Yes	No	No	Yes

Mean values across periods and products sold under brands considered.

Standard deviations are in parenthesis.

Table 4: **Own Price elasticities in multinomial logit models (MLM) and multiple discrete-continuous choice models (MDCCM).**

	Rich	Average upper	Average lower	Poor
Brand1	-2.40 (0.07)	-2.37 (0.06)	-2.30 (0.07)	-2.25 (0.08)
Brand2	-2.49 (0.12)	-2.43 (0.09)	-2.37 (0.06)	-2.29 (0.10)
Brand3	-2.28 (0.07)	-2.24 (0.07)	-2.18 (0.06)	-2.13 (0.07)
Brand4	-2.40 (0.11)	-2.36 (0.07)	-2.29 (0.06)	-2.20 (0.08)
Brand5	-2.38 (0.10)	-2.34 (0.04)	-2.26 (0.07)	-2.21 (0.08)
Brand6	-2.22 (0.09)	-2.17 (0.08)	-2.12 (0.06)	-2.06 (0.05)
Brand7	-2.36 (0.11)	-2.29 (0.06)	-2.22 (0.04)	-2.17 (0.08)
Brand8	-2.38 (0.03)	-2.31 (0.02)	-2.25 (0.02)	-2.17 (0.04)
Brand9	-2.37 (0.03)	-2.32 (0.02)	-2.26 (0.02)	-2.19 (0.04)
Brand10	-2.16 (0.03)	-2.11 (0.03)	-2.07 (0.03)	-2.02 (0.03)
Brand11	-2.47 (0.13)	-2.41 (0.13)	-2.36 (0.12)	-2.29 (0.12)
Brand12	-2.31 (0.08)	-2.25 (0.06)	-2.20 (0.06)	-2.13 (0.05)
Brand13	-2.39 (0.06)	-2.32 (0.06)	-2.28 (0.04)	-2.20 (0.07)
Brand14	-2.16 (0.07)	-2.12 (0.06)	-2.06 (0.05)	-2.01 (0.04)

Mean values across periods and products sold under brands considered.

Standard deviations are in parenthesis.

Table 5: **Own Price elasticities for the multiple discrete-continuous choice model according to the income class of consumers.**

5.1 Outside option

ML and MDCC models assumed that the consumer can choose an additional alternative to the choice set, that is the outside good. The introduction of this outside option allows to consider a variable size of the market. Indeed, if the price of the J alternatives increases, the consumer can choose another product, supposed to be substitute. The J alternatives and the outside option define the total size of the market and we assume this total size does not change per period. Foncel and Ivaldi (2005) show that the total size of the market affects the level of utility. Hence, the larger the size is, the lower the utility is and the larger the elasticities are. We then test different sizes of the market. Results presented in Table 4 assume that the total size of the market is the total expense for food and drink products. We have then supposed that the consumer can substitute a soft drink product to some other drink or food products. We have estimated the same MDCC models restricting the size of the market to the other soft drink products only and we found that the estimates of the demand model and then the elasticities do not change so much (Table 6). On average, the difference of products' own price elasticities is 1.6% and the maximum reaches to 4.1%. 85% of the differences are negative and strengthen the result of Foncel and Ivaldi (2005). However, this difference is quite small and could be explained by the fact that French consumers do not substitute soft drinks with other food and drink products. Our results do not thus depend on the size of the chosen market.

5.2 Random draws

The baseline utility $\Psi(x_{jt}, \varepsilon_{jht}) = e^{\beta' x_{jt} + \varepsilon_{jht}}$ of the product j at period t for household h that represents the quality associated to the product is composed of a deterministic part that depends on product characteristics x_{jt} and of a random part ε_{jht} that is the deviation for household h of the mean quality of product j at period t $e^{\beta' x_{jt}}$. We assume that ε_{jht} follows an independant Gumbel distribution of scale parameter σ. We then need to randomly evaluate ε_{jht} in order to compute forecasted quantities via the algorithm as explained in Appendix 7.4. The elasticities in Table 4 are computed from 50 random numbers for the Gumbel distribution. We present the estimated elasticities in Table 6 computed from

Size of the market	Soft drink products	Food and drink products	
Random Numbers	100	100	50
Brand1	-2.03 (0.05)	-2.09 (0.05)	-2.09 (0.06)
Brand2	-2.10 (0.05)	-2.18 (0.05)	-2.17 (0.08)
Brand3	-1.96 (0.04)	-1.99 (0.06)	-1.99 (0.06)
Brand4	-2.01 (0.02)	-2.06 (0.02)	-2.08 (0.05)
Brand5	-2.00 (0.03)	-2.06 (0.03)	-2.06 (0.04)
Brand6	-1.92 (0.02)	-1.94 (0.04)	-1.95 (0.04)
Brand7	-1.96 (0.03)	-2.01 (0.02)	-2.01 (0.04)
Brand8	-2.00 (0.04)	-2.03 (0.04)	-2.04 (0.04)
Brand9	-2.01 (0.01)	-2.05 (0.01)	-2.05 (0.02)
Brand10	-1.89 (0.02)	-1.88 (0.03)	-1.88 (0.03)
Brand11	-2.11 (0.14)	-2.14 (0.15)	-2.15 (0.15)
Brand12	-1.96 (0.03)	-2.00 (0.05)	-2.00 (0.05)
Brand13	-2.02 (0.03)	-2.06 (0.03)	-2.05 (0.03)
Brand14	-1.88 (0.02)	-1.89 (0.04)	-1.89 (0.04)

Mean values across periods and products sold under brands considered.

Standard deviations are in parenthesis

Table 6: **Own price elasticities from the MDCEV model for different outside goods and different number of random draws.**

100 random numbers and we can see that the results stay the same.

6 Conclusion

Standard discrete choice models do not allow to characterize the full choice of consumers as they frequently choose a set of alternatives and different quantities for each of them. Most analyses of consumer substitution patterns use them as they allow to get analytic expressions of price elasticities. However, assuming that households choose one unit of a unique alternative could bias substitution patterns. The objective of this paper is to assess what the bias could be. We then compare estimates of consumer substitution patterns between multiple discrete-continuous choice models and multinomial logit demand models, which are a particular case of the first ones. Therefore, we model multiple brand and quantity choices of households in each purchase occasion. We then allow for taste variety within the household, that is different tastes in a household composed of several persons. Using French purchase data on the Soft Drink Market, we show that consumers seek variety. Indeed, they prefer to buy several products rather than a large quantity of a single product. Our results also suggest that consumer substitution patterns could be biased if multiple brand and quantity choices are not taken into account. We find that consumers have a more elastic behavior allowing for a multiple discrete-continuous choice model since own price elasticities are larger. This result has important implications when analyzing competition policies between products and consumer welfare effects of regulations, price policies. For example, the estimation of market power in industries strongly depends on estimates of prices elasticities. The more price sensitive the consumers are, the lower the market power is. As our results suggest that own price elasticities could be underestimated by more than 20% with standard discrete choice models, the analysis of market power would be overestimated.

Note that this model could be extended to allow satiation parameters to depend on observed and/or unobserved characteristics as the consumer behaviour of satiation and the taste for variety could vary across men and women, or parents and children, for example.

7 Appendix

7.1 Descriptive Statistics

Table 7.1 shows the descriptive statistics on prices and quantities for each product. We compute the price dividing total expenditures by the total quantities for each product at each time period. The standard deviations of prices show the distribution across time periods. For each product, time period and household, we compute the aggregated bought quantity in liters. Conditionning households have a positive consumption, the mean value represents the mean quantity (in liters) bought by households during one time period and the standard deviation represents the heterogeneity across households and time periods. Table 7.1 shows the same results, but aggregated by product characteristics (diet or not, sparkling or not, pure juice or not) and the average for all products.

Products	Brand	Pur Juice	Diet	Carbonated	Price (std)	Quantity (std)
OG other soft drinks		yes/no	yes/no	yes/no		6.01 (8.04)
OG other food and drink products						134.66 (37.14)
1	1			yes	0.75 (0.07)	5.47 (6.20)
2	1		yes	yes	0.74 (0.01)	7.16 (7.91)
3	2			yes	0.98 (0.01)	5.88 (6.22)
4	2		yes	yes	0.99 (0.01)	5.78 (6.65)
5	3				1.14 (0.02)	3.36 (3.02)
6	3		yes		1.11 (0.02)	3.09 (2.72)
7	4			yes	1.09 (0.02)	3.69 (3.31)
8	4		yes	yes	1.00 (0.02)	3.61 (2.96)
9	5			yes	1.14 (0.01)	3.05 (2.53)
10	5		yes	yes	1.03 (0.02)	3.02 (2.54)
11	6				1.00 (0.05)	4.11 (3.67)
12	6		yes		0.77 (0.01)	2.72 (1.39)
13	7			yes	1.02 (0.05)	3.28 (2.81)
14	7		yes	yes	0.87 (0.03)	2.61 (2.11)
15	8				1.26 (0.01)	3.16 (2.54)
16	8	yes			1.76 (0.01)	2.94 (2.59)
17	8		yes		1.79 (0.02)	1.72 (1.24)
18	9				1.80 (0.04)	2.33 (2.04)
19	9	yes			1.82 (0.02)	2.65 (2.29)
20	10				1.01 (0.05)	3.69 (3.45)
21	10	yes			1.26 (0.11)	2.59 (2.20)
22	11				1.80 (0.05)	1.95 (1.61)
23	11	yes			2.29 (0.02)	3.31 (2.76)
24	12				1.18 (0.01)	3.13 (2.41)
25	12		yes		1.10 (0.09)	2.56 (1.88)
26	13				1.99 (0.02)	2.52 (2.32)
27	13		yes		2.06 (0.02)	1.45 (0.92)
28	PL				0.80 (0.01)	4.79 (5.11)
29	PL	yes			1.17 (0.01)	3.50 (3.37)
30	PL		yes		0.55 (0.03)	3.09 (2.30)
31	PL	yes	yes		0.93 (0.07)	2.12 (2.15)
32	PL			yes	0.46 (0.01)	5.11 (5.60)
33	PL	yes		yes	0.91 (0.10)	5.49 (6.81)
34	PL		yes	yes	0.31 (0.00)	4.68 (5.62)

Descriptive Statistics on Prices and Quantities.

	Price	Quantity
Regular	1.07 (0.38)	4.60 (5.46)
Diet	0.86 (0.29)	5.36 (6.70)
Still	1.15 (0.40)	4.19 (4.94)
Sparkling	0.88 (0.24)	5.48 (6.49)
Not pure juice	0.93 (0.25)	5.04 (6.04)
Pure Juice	1.42 (0.45)	3.59 (3.82)
Total	1.05 (0.37)	4.68 (5.62)

Descriptive Statistics on Prices and Quantities.

7.2 Computation details about the model

Computation of the determinant of the Jacobian matrix As we have seen,

$$V_{jht} = \beta' x_{jt} + (\alpha - 1)\ln\left(\frac{q^*_{jht}}{\gamma_j} + 1\right) - \ln p_{jt}$$

is a general expression and $q_{jht} > 0$ holds only for the K_{ht} first products (after re-ordering). And $q^*_{1ht} = \left(y_{ht} - \sum_{l=2}^{K_{ht}} p_{lt} q^*_{lht}\right)/p_{1t}$.

Hence, $\forall j = 2, \ldots, K_{ht}, V_{jht} + \varepsilon_{jht} = V_{1ht} + \varepsilon_{1ht}$, leading to:

$$\begin{aligned} V_{1ht} - V_{jht} &= \beta' x_{1t} + (\alpha-1)\ln\left(\frac{q^*_{1ht}}{\gamma_1} + 1\right) - \ln p_{1t} - \left(\beta' x_{jt} + (\alpha-1)\ln\left(\frac{q^*_{jht}}{\gamma_j} + 1\right) - \ln p_{jt}\right) \\ &= \beta' x_{1t} + (\alpha-1)\ln\left(\frac{y_{ht} - \sum_{l=2}^{K_{ht}} p_{lt} q^*_{lht}}{\gamma_1 * p_{1t}} + 1\right) - \ln p_{1t} - \left(\beta' x_{jt} + (\alpha-1)\ln\left(\frac{q^*_{jht}}{\gamma_j} + 1\right) - \ln p_{jt}\right) \end{aligned}$$

The Jacobian matrix has for element $(j-1, k-1), \forall j, k = 2, \ldots, K_{ht}$:

$$Jac_{htj-1,k-1} = \frac{\partial\left[V_{1ht} - V_{jht}\right]}{\partial q^*_{kht}} = \begin{cases} (1-\alpha)\frac{p_{kt}}{p_{1t}}\frac{1}{q^*_{1ht}+\gamma_1} & \text{if } j \neq k \\ (1-\alpha)\frac{p_{jt}}{p_{1t}}\left(\frac{1}{q^*_{1ht}+\gamma_1} + \frac{p_{1t}/p_{jt}}{q^*_{jht}+\gamma_j}\right) & \text{if } j = k \end{cases}$$

So, the Jacobian matrix can be rewritten in the following way:

$$Jac_{ht} = \begin{pmatrix} (c_{1ht} + c_{2ht})p_{2t} & c_{1ht}p_{3t} & . & . & c_{1ht}p_{K_{ht}t} \\ c_{1ht}p_{1t} & . & & & . \\ . & & . & & . \\ . & & & . & c_{1ht}p_{Kt} \\ c_{1ht}p_{1t} & . & & . \; . & (c_{1ht} + c_{K_{ht}ht})p_{K_{ht}t} \end{pmatrix},$$

where $c_{lht} = (1-\alpha)\frac{1}{p_{1t}}\frac{p_{1t}/p_{lt}}{q^*_{lht}+\gamma_l}$.

The determinant of this Jacobian matrix takes this simple expression:

$$\begin{aligned} |Jac_{ht}| &= \left(\prod_{l=2}^{K_{ht}} p_{lt}\right)\left(\prod_{l=1}^{K_{ht}} c_{lht}\right)\left(\sum_{l=1}^{K_{ht}} \frac{1}{c_{lht}}\right) \\ &= \left(\prod_{l=2}^{K_{ht}} p_{lt}\right)\left(\prod_{l=1}^{K_{ht}} \frac{(1-\alpha)}{p_{lt}q^*_{lht} + \gamma_l p_{lt}}\right)\left(\sum_{l=1}^{K_{ht}} \frac{p_{lt}q^*_{lht} + \gamma_l p_{lt}}{(1-\alpha)}\right) \\ &= \frac{1}{p_{1t}}\left(\prod_{l=1}^{K_{ht}} \frac{(1-\alpha)}{q^*_{lht} + \gamma_l}\right)\left(\sum_{l=1}^{K_{ht}} \frac{p_{lt}q^*_{lht} + \gamma_l p_{lt}}{(1-\alpha)}\right). \end{aligned}$$

Computation of the multiple integral In the expression of the probability that household h purchases at time t the first K_{ht} of the $J+1$ alternatives, we have to solve this multiple integral:

$$\int\limits_{\varepsilon_{(K_{ht}+1)ht}=-\infty}^{V_{1ht}+\varepsilon_{1ht}-V_{(K_{ht}+1)ht}} \cdots \int\limits_{\varepsilon_{(J+1)ht}=-\infty}^{V_{1ht}+\varepsilon_{1ht}-V_{(J+1)ht}} \int\limits_{\varepsilon_{1ht}=-\infty}^{+\infty} f(\varepsilon_{1ht}, V_{1ht}-V_{2ht}+\varepsilon_{1ht}, ..., V_{1ht}-V_{K_{ht}ht}+\varepsilon_{1ht}, \varepsilon_{(K_{ht}+1)ht}, ..., \varepsilon_{(J+1)ht}) d\varepsilon_{(K_{ht}+1)ht}...d\varepsilon_{(J+1)ht}d\varepsilon_{1ht}$$

Let $\lambda(x) = \frac{1}{\sigma}e^{-\frac{x}{\sigma}}e^{-e^{-\frac{x}{\sigma}}}$ be the density function of the standard extreme value distribution and $\Lambda(x) = e^{-e^{-\frac{x}{\sigma}}}$ be the standard extreme value cumulative distribution function with σ a scale parameter. Assuming that the ε_js are independently distributed across alternatives and have a Gumbel or type I extreme value distribution and independent of the vector of exogenous variables of the baseline utility function, the multiple integral becomes:

$$\int\limits_{\varepsilon_{(K_{ht}+1)ht}=-\infty}^{V_{1ht}+\varepsilon_{1ht}-V_{(K_{ht}+1)ht}} \cdots \int\limits_{\varepsilon_{(J+1)ht}=-\infty}^{V_{1ht}+\varepsilon_{1ht}-V_{(J+1)ht}} \int\limits_{\varepsilon_{1ht}=-\infty}^{+\infty} \lambda(\frac{\varepsilon_{1ht}}{\sigma})\lambda\left(\frac{V_{1ht}-V_{2ht}+\varepsilon_{1ht}}{\sigma}\right)...\lambda\left(\frac{V_{1ht}-V_{(K_{ht}+1)ht}+\varepsilon_{1ht}}{\sigma}\right) \lambda\left(\frac{\varepsilon_{(K_{ht}+1)ht}}{\sigma}\right), ..., \lambda\left(\frac{\varepsilon_{(J+1)ht}}{\sigma}\right) d\varepsilon_{(K_{ht}+1)ht}...d\varepsilon_{(J+1)ht}d\varepsilon_{1ht}$$

$$\begin{aligned}
&= \int\limits_{\varepsilon_{1ht}=-\infty}^{+\infty} \prod_{i=2}^{K_{ht}}\lambda\left(\frac{V_{1ht}-V_{iht}+\varepsilon_{1ht}}{\sigma}\right) \prod_{i=K_{ht}+1}^{J+1}\Lambda(\frac{V_{1ht}-V_{iht}+\varepsilon_{1ht}}{\sigma})d\varepsilon_{1ht} \\
&= \int\limits_{\varepsilon_{1ht}=-\infty}^{+\infty} \left(\prod_{i=1}^{K_{ht}}\frac{1}{\sigma}e^{-\frac{V_{1ht}-V_{iht}+\varepsilon_{1ht}}{\sigma}}\right)\left(e^{-\sum\limits_{i=1}^{J}e^{-\frac{V_{1ht}-V_{iht}+\varepsilon_{1ht}}{\sigma}}}\right)d\varepsilon_{1ht} \\
&= \frac{1}{\sigma^{K_{ht}+1}}\prod_{i=1}^{K_{ht}}e^{-\frac{V_{1ht}-V_{iht}}{\sigma}} \int\limits_{\varepsilon_{1ht}=-\infty}^{+\infty} \left(e^{-\frac{\varepsilon_{1ht}}{\sigma}}\right)^{K_{ht}+1} e^{-\sum\limits_{i=1}^{J+1}e^{-\frac{V_{1ht}-V_{iht}+\varepsilon_{1ht}}{\sigma}}} d\varepsilon_{1ht}
\end{aligned}$$

Solving by full integration-by-parts the previous integral, we obtain

$$= \frac{(K_{ht}-1)!}{\sigma^{K_{ht}-1}} \frac{\prod\limits_{i=1}^{K_{ht}}e^{-\frac{V_{1ht}-V_{iht}}{\sigma}}}{\left(\sum\limits_{i=1}^{J+1}e^{-\frac{V_{1ht}-V_{iht}}{\sigma}}\right)^{K_{ht}}} = \frac{(K_{ht}-1)!}{\sigma^{K_{ht}-1}} \frac{\prod\limits_{i=1}^{K_{ht}}e^{\frac{V_{iht}}{\sigma}}}{\left(\sum\limits_{i=1}^{J+1}e^{\frac{V_{iht}}{\sigma}}\right)^{K_{ht}}}$$

Therefore, the probability that household h purchases the first K_{ht} of the $J+1$ alternatives at time t takes this final expression:

$$
\begin{aligned}
&\Pr(q^*_{0ht}, ..., q^*_{K_{ht}ht}, 0, ..., 0) \\
= \quad & |Jac| \int\limits_{\varepsilon_{(K_{ht}+1)ht}=-\infty}^{V_{0ht}+\varepsilon_{0ht}-V_{(K_{ht}+1)ht}} \cdots \int\limits_{\varepsilon_{Jht}=-\infty}^{V_{0ht}+\varepsilon_{0ht}-V_{Jht}} \int\limits_{\varepsilon_{0ht}=-\infty}^{+\infty} f(\varepsilon_{0ht}, V_{0ht}-V_{1ht}+\varepsilon_{0ht}, ..., V_{0ht}-V_{K_{ht}ht}+\varepsilon_{0ht}, \\
& \qquad \varepsilon_{(K_{ht}+1)ht}, ..., \varepsilon_{Jht}) d\varepsilon_{(K_{ht}+1)ht} ... d\varepsilon_{Jht} d\varepsilon_{0ht} \\
= \quad & \frac{(K_{ht}-1)!}{\sigma^{K_{ht}-1}} \left(\prod_{j \in \mathcal{J}_{ht}} \frac{(1-\alpha)}{q^*_{jht}+\gamma_j} \right) \left(\sum_{j \in \mathcal{J}_{ht}} \frac{q^*_{jht}+\gamma_j}{(1-\alpha)} \cdot \frac{p_{jt}}{p_{j^0_{ht}t}} \right) \frac{\prod\limits_{j \in \mathcal{J}_{ht}} e^{\frac{V_{jht}}{\sigma}}}{\left(\sum\limits_{j=1}^{J+1} e^{\frac{V_{jht}}{\sigma}} \right)^{K_{ht}}}.
\end{aligned}
$$

7.3 Estimation Results

Table 7.3 gives the estimation results of the standard discrete choice model in the cases where households are modeled without heterogeneity (Model 1 without endogeneity and Model 2 with endogeneity) and in the case where households are modeled with heterogeneity (model without endogeneity and Model 3 with endogeneity).

	Without household heterogeneity		With Household heterogeneity	
	Without endogeneity	With Endogeneity	Without Endogeneity	With Endogeneity
	Mean (std 10^{-3})	Mean (std 10^{-3})	Mean (std 10^{-3})	Mean (std 10^{-3})
σ	0.629 (0.112)	0.514 (0.078)		
$\sigma\times$Well_Off			0.627 (0.114)	0.512 (0.080)
$\sigma\times$Average Upper			0.628 (0.113)	0.514 (0.078)
$\sigma\times$Average Lower			0.631 (0.113)	0.515 (0.079)
$\sigma\times$Modest			0.629 (0.115)	0.514 (0.081)
β_{Diet}	-1.350 (0.205)	-1.138 (0.143)	-1.443 (0.234)	-1.214 (0.166)
$\sigma_{\beta_{Diet}}$			0.376 (0.164)	0.307 (0.131)
$\beta_{Sparkling}$	-1.211 (0.135)	-1.062 (0.095)	-1.211 (0.135)	-1.062 (0.095)
$\sigma_{\beta_{Sparkling}}$			0.376 (0.164)	0.693 (0.062)
$\beta_{purejuice}$	0.161 (0.060)	0.206 (0.046)	0.161 (0.060)	0.307 (0.131)
$\sigma_{\beta_{purejuice}}$			0.021 (0.077)	0.016 (0.046)
δ_1	3.331 (0.307)	3.586 (0.223)	3.330 (0.306)	3.586 (0.223)
δ_2	5.061 (0.074)	5.053 (0.060)	5.061 (0.074)	5.053 (0.060)
δ_3	2.823 (0.361)	3.167 (0.255)	2.822 (0.360)	3.167 (0.255)
δ_4	4.037 (0.242)	4.204 (0.178)	4.036 (0.241)	4.204 (0.178)
δ_5	4.187 (0.222)	4.330 (0.164)	4.187 (0.222)	4.330 (0.164)
δ_6	2.595 (0.373)	2.896 (0.267)	2.594 (0.373)	2.895 (0.267)
δ_7	3.623 (0.306)	3.836 (0.225)	3.624 (0.306)	3.836 (0.225)
δ_8	2.927 (0.350)	3.369 (0.242)	2.927 (0.350)	3.369 (0.242)
δ_9	2.514 (0.460)	2.971 (0.326)	2.515 (0.459)	2.971 (0.336)
δ_{10}	1.851 (0.501)	2.339 (0.355)	1.851 (0.500)	2.338 (0.355)
δ_{11}	3.334 (0.329)	3.653 (0.229)	3.334 (0.239)	3.653 (0.229)
δ_{12}	2.340 (0.464)	2.767 (0.330)	2.338 (0.464)	2.766 (0.330)
δ_{13}	2.864 (0.465)	3.334 (0.329)	2.864 (0.464)	3.334 (0.329)
δ_{14}	3.744 (0.112)	3.842 (0.081)	3.744 (0.112)	3.842 (0.081)
θ		0.797 (0.223)		0.797 (0.223)
Log likelihood	-550,980	-550,375	-550,872	-550,267

Estimation results of the standard discrete choice model.

7.4 Computation details about the forecasting algorithm

As shown at the beginning of section 3.2, if K specific products are bought, getting their optimal quantities is straightforward as it consists of solving a linear problem with K equations and K unknowns.

So all the problem is to know which products are bought. Theoretically, we should try the $2^{35}-1$ combinations (at least one product is bought and at most 35, including the outside good). For some bundles, at least one optimal quantity will be negative, which discards the bundle. For the remaining bundles, we should keep the one that maximizes the utility.

Yet, we would have to do this several times: N (where $N = 1$ for Model 4 and N is at least $2,500$[7] in

[7] In order to capture both Gumbel distribution of σ (50 random numbers) and heterogeneity of households preferences (50 triplets of random numbers) regarding diet, pure juice and carbonated drinks.

the case of Models 5 or 6) for each household (9, 458), time period (13) and price scenarios (69: one with the current prices of soft drinks and the others were only one product price deviates from the current prices, either by +1% or by -1%, except for the outside numeraire option which keeps to the unit price).

So, this (resolving some 10^{21} linear problems, in the case of Models 5 or 6) is not computationally tractable[8]: It would still take around 8 years on the current (as for June 2014) world fastest computer (Tianhe-2, China, almost 34 petaFLOPS, see http://www.top500.org/lists/2014/06/). And anyway, we do not have access to such a computer. It would take almost 43 years on the fastest European computer (Piz Daint, Switzerland, 6 petaFLOPS, 6^{th} fastest computer in the world) and more than 1054 years on the fastest computer of the university of Toulouse (Eos, 12 teraFLOPS, 183^{rd} fastest computer in the world). On this last computer, Model 4 would still need more than 5 months for a complete resolution.

Hence, we use a greedy heuristic to forecast the bundle composition that will be as close as possible to the optimal one. This greedy algorithm can be found in (Pinjari and Bhat, 2011). Its adaptation to our model is the following: For a given household h, time period t, product j, position r in the random sets for observed household heterogeneity (R^{σ}_{htj}), position s in the random sets for unobserved household heterogeneity (R^D, R^S and R^{PJ}), the price normalized baseline utility is computed for each product ($j = 1, ..., J{+}1, J{+}1 = 35$) as $PNBU_{htjrs} = \frac{e^{\delta_j + (\beta_h^D + \sigma_{\beta_h^D} . R_s^D) D_j + (\beta_h^S + \sigma_{\beta_h^S} . R_s^S) S_j + (\beta_h^{PJ} + \sigma_{\beta_h^{PJ}} . R_s^{PJ}) PJ_j + \theta \widehat{\eta}_{jt} - \sigma_h \ln(\ln(R^{\sigma}_{htjr}))}}{\widetilde{p}_{jt}}$, where $\widetilde{p}_{jt}$ is the price considered for product j at time t. Then, price normalized baseline utilities are ranked for all products. Initialization ends with $M = 1$: number of products that are bought (at least one: the outside good is always bought).

Step "Number of Products bought":

Compute $\mu_{htrs} = \left(\frac{\sum_{m=1}^{M} \gamma_{j_m} . \widetilde{p}_{jmt} (PNBU_{htj_m rs})^{\frac{1}{1-\alpha}}}{y_{ht} + \sum_{m=1}^{M} \gamma_{j_m} . \widetilde{p}_{jmt}} \right)^{1-\alpha}$. If $M = J{+}1$ or else $M * \mu_{htrs} > PNBU_{htj_{m+1}rs}$, then go to Step "Quantity Estimation", else increment $M = M + 1$ and go back to Step "Number of Products bought".

Step "Quantity Estimation":

[8] This results in some 10^{25} floating-point operations: for each of the 10^{21} problems, its size n (number of unknowns and equations) leads to $n(1 + 2n(n-1))$ operations to compute the inverse matrix and $n(2n-1)$ operations to compute the solution. Problems sizes range from 1 to 35. There are C_{35}^{n} problems of size n to consider.

$\forall m = 1..M, \widehat{Q_{htj_mrs}}(\widetilde{p}) = \gamma_{j_m}\left((PNBU_{htj_mrs})^{\frac{1}{1-\alpha}} \frac{y_{ht}+\sum_{m=1}^{M}\gamma_{j_m}.\widetilde{p}_{jmt}}{\sum_{m=1}^{M}\gamma_{j_m}.\widetilde{p}_{jmt}(PNBU_{htj_mrs})^{\frac{1}{1-\alpha}}} - 1\right)$ and if $M < J + 1, \forall m = M+1..J+1, \widehat{Q_{htj_mrs}}(\widetilde{p}_t) = 0$. Finally, $\forall j = 1, ..., J+1, \widehat{Q_{htj}}(\widetilde{p}_t) = \underset{r,s}{\text{mean}}\, \widehat{Q_{htjrs}}(\widetilde{p}_t)$.

Once we have the estimated quantities, the elasticities are finally estimated using: $\epsilon_{i,j}^{t,cr} = \frac{Q_{tj}^{cr}(1.01*p_{it})-Q_{tj}^{cr}(0.99*p_{it})}{Q_{tj}^{cr}(p_{it})} * \frac{p_{it}}{1.01*p_{it}-0.99*p_{it}} = \frac{Q_{tj}^{cr}(1.01*p_{it})-Q_{tj}^{cr}(0.99*p_{it})}{0.02*Q_{tj}^{cr}(p_{it})}$, where $Q_{tj}^{cr}(x*p_{it})$ denotes $\sum_{CR(h)=cr} Q_{htj(p_t')}$, with $\widetilde{p}_{it} = x * p_{it}$ and $\forall k = 1, ..., J+1$, if $k \neq i$, then $\widetilde{p}_{kt} = p_{kt}$, $CR(h)$ giving the class of revenue of household h and cr denoting one of the four classes of revenue. Then, $\epsilon_{i,j}^{cr} = \frac{1}{T}\sum_{t=1}^{T}\epsilon_{i,j}^{t,cr}$, $\epsilon_{i,j}^{cr} = \frac{1}{CR}\sum_{cr=1}^{CR}\epsilon_{i,j}^{t,cr}$ and $\epsilon_{i,j} = \frac{1}{CR}\frac{1}{T}\sum_{t=1}^{T}\sum_{cr=1}^{CR}\epsilon_{i,j}^{t,cr}$.

8 References

Barten, A.P., 1966, “Theorie en Empirie van een Volledig Stelsel van Vraagvergelijkingen,” Doctoral Dissertation, Rotterdam: University of Rotterdam.

Berry S., J. Levinsohn et A. Pakes, 1995, Automobile Prices in Market Equilibrium, Econometrica, Vol 63, No. 4, pp. 841-890

Bhat C. R., 2005a, The Multiple Discrete-Continuous Extreme Value (MDCEV) Model: Role of Utility Function Parameters, Identification Considerations, and Model Extensions, forthcoming in Transportation Research Part B

Bhat, C.R., 2005b, "A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions," Transportation Research Part B, Vol. 39, No. 8, pp. 679-707

Bhat, C.R., and S. Sen, 2006, Household Vehicle Type Holdings and Usage: An Application of the Multiple Discrete-Continuous Extreme Value (MDCEV) Model, Transportation Research Part B, Vol. 40, No. 1, pp. 35-53

Bonnet C. and V. Réquillart (2013), Tax incidence with strategic firms on the soft drink market, Journal of Public Economics, 106, pp 77-88

Chiang J., 1991, The Simultaneous Approach to the Whether, What, and How much to Buy Questions, Marketing Science, Vol. 10, No. 4, pp.297-315

Chintagunta P., 1993, Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households, Marketing Science, Vol. 12, No. 1, pp. 184-208

Christensen, L.R., D.W. Jorgenson, and L.J. Lau, 1975, "Transcendental Logarithmic Utility Functions," American Economic Review, 65, 367–383.

Deaton, A., and J. Muellbauer, 1980, "An Almost Ideal Demand Sustem," American Economic Review, 70, 312–326.

Dillon, W.R. and S. Gupta, 1996, A Segment-Level Model of Category Volume and Brand Choice, Marketing Science, Vol 15, No. 1, pp. 38-59

Dubé J-P, 2004, Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks, Marketing Science, Vol. 23, No. 1, Winter 2004, pp. 66-81

Dubé J-P, 2005, Product Differentiation and Mergers in the Carbonated Soft Drink Industry, Journal of Economics & Management Strategy, 2005, 14(4), pp. 879-904.

Dubin, J. A. and D. L. McFadden, 1984, An Econometric Analysis of Residential electric Appliance Holdings and Consumption, Econometrica, Vol 52, No. 2, pp. 345-362

Guadani P.M. and J.D.C. Little, 1983, A Logit Model of Brand choice, Marketing Science, Vol 2, No. 3, pp. 203-238

Hausman, J., 1996, "Valuation of New Goods under Perfect and Imperfect Competition," in T. Bresnahan and R. Gordon, eds., The Economics of New Goods, Studies in Income and Wealth, Vol. 58, Chicago: National Bureau of Economic Research.

Hanemann, W. M., 1984, Discrete/Continuous Models of Consumer Demand, Econometrica, Vol 52, No. 3, pp. 541-561

Hendel I, 1999, Estimating Multiple-discrete Choice Models: An Application to computerization Returns, The Review of Economic Studies, Vol. 66, No. 2, April 1999, pp. 423-446

Kim J., G. M. Allenby and P. E. Rossi, 2002, Modeling consumer Demand for Variety, Marketing Science, Vol. 21, No. 3, Summer 2002, pp. 229-250

Krishnamurthi, L. and S.P.Raj, 1988, A model of Brand Choice and Purchase Quantity Price Sensitivities, Marketing Science, Vol 7, No. 1, pp. 1-20

Nevo A., 2000, 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

Petrin A. and K. Train, 2010, A Control Function Approach to Endogeneity in Consumer Choice Models, Journal of Marketing Research, Vol. 47, No. 1, pp. 3-13

Pinjari, A.R., and C.R. Bhat, "Computationally Efficient Forecasting Procedures for Kuhn-Tucker Consumer Demand Model Systems: Application to Residential Energy Consumption Analysis," Technical paper, Department of Civil & Environmental Engineering, University of South Florida, July 2009, 1st revision June 2010, 2nd revision July 2011

Scarpa R., M. Thiene and K. Train (2008), Utility in WTP space: a tool to address confounding random scale effects in destination choice to the Alps. American Journal of Agricultural Economics, 90, pp 994-1010

Stone, J., 1954, "Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand," Economic Journal, 64, 511–527.

Terza, J., A. Basu, and P. Rathouz. (2008),Two-Stage Residual Inclusion Estimation Addressing Endogeneity in Health Econometric Modeling. Journal of Health Economics 27 (3): 531–543.

Theil, H., 1965, "The Information Approach to Demand Analysis," Econometrica, 6, 375–380.

Train K., 1986, Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand, MIT Press, Cambridge, Massachusetts, USA

Train K., 2002, Discrete Choice Method and Simulations, MIT Press, Cambridge, Massachusetts, USA