Economics Bulletin

Volume 35, Issue 4

A simplified mixed logit demand model with an application to the simulation of entry

Sergio Aquino DeSouza
Universidade Federal do Ceara

Abstract

Given additional information on economic meaningful measures, such as price elasticity or margin, this paper proposes an alternative empirical approach to determine the parameters of a simplified (aggregate) Mixed Logit Model. This empirical method is particularly useful when valid instruments are difficult to find. The model is applied to uncover the demand parameters and simulate the competitive and welfare effects of the introduction of new products in the ready-to-eat cereal industry in the U.S.

I am thankful to CNPq and Universidade Federal do Ceará (CAEN and DTE) for funding this research.
Citation: Sergio Aquino DeSouza, (2015) "A simplified mixed logit demand model with an application to the simulation of entry", *Economics Bulletin*, Volume 35, Issue 4, pages 2333-2344
Contact: Sergio Aquino DeSouza - srgdesouz@gmail.com.
Submitted: March 17, 2015. **Published:** November 20, 2015.

1. Introduction

Demand estimation in product-differentiated industries from aggregate[1] data has been the central object in many studies in the industrial organization field. Indeed, after pinning down the preference parameters it is possible to analyze issues related to innovation, antitrust (mergers and divestitures), calculation of quality adjusted price-indices and prediction of the competitive effect of entry and exit of products. However, uncovering demand parameters from aggregate data on product-differentiated markets imposes several challenges: number of parameters, incorporation of consumer heterogeneity and price endogeneity.

There are basically two categories of aggregate demand models that are taken to data: representative consumer and discrete-choice demand models. Models in the former category are based on a representative consumer who has preference over a set of differentiated products and purchases simultaneously more than one variety. However, for markets characterized by the presence of many brands the representative consumer models may be too restrictive. Indeed, with many brands such models imply a demand system with many equations (the number of brands is equal to the number of demand equations), which results in a system with too many parameters. Furthermore, by construction, representative consumer models cannot properly deal with the presence of consumer heterogeneity. The second set of demand models is based on the theory of discrete-choice, in which it is assumed that the consumer chooses only one variety (i.e., simultaneous consumption of different varieties is not allowed in this setup). Further, the product choice is made indirectly as the consumer has preferences over attributes and picks the product that offers the best combination of such attributes. Using the literature jargon, the choice is made on the attribute space rather than on the product space as assumed in representative consumer models. This projection onto the attribute space makes the discrete-choice model a very attractive option of modeling product differentiation for empirical purposes. Indeed, the number of parameters depends on the number of attributes rather than the number of products. This can substantially reduce the size of the parameter set. In addition, consumer heterogeneity can be incorporated into the model in a natural way.

However, discrete-choice models do not avoid all the problems associated with the estimation of demand. As in representative consumer models, the endogeneity problem emerges as prices are expected to be correlated with unobserved determinants of demand (e.g., omitted attributes, unobserved quality). Then, as predicted by standard econometric theory, the researcher is likely to face inference problems regarding the estimation of the price coefficient.

The common solution to this problem is to find instruments that are correlated with the endogenous variable (prices) but not with the unobserved determinants of demand (regression error term). Berry, Levinsohn, and Pakes (1995) - BLP henceforth- propose a GMM method based on three sets of instruments. These instruments are based on the product attributes, which are assumed to be exogenous. The first set is formed by the attributes (excluding potentially endogenous ones). The second is composed the sum of the values of the attributes across own-firm products. Finally, the third set of instruments is calculated by the sum of the values of the same attribute across rival firm products. An alternative to the BLP instruments was introduced by Hausman et al. (1994) who exploit the panel structure of the data (geographically separated markets are observed through time) and the assumption that, given the cost structure and after controlling for demand shifters and fixed effects (observed and unobserved), the price of a brand j in market r is a valid instrument for the price of the same brand j in another market r'.

[1] Data is observed at the product-level, not at the consumer level. The two literatures are related, but the empirical methods are different. McFadden (1981) is classic reference for latter class of models. In turn, Berry (1994) e Berry, Levinsohn, and Pakes (1995) provide the seminal contributions for the former class of models.

The types of instruments proposed by BLP and Hausman et al. (1994) are far from being a consensus among researchers in the IO field. Indeed, there are instances in which those instruments may fail. For instance, (Nevo, 1998) reports that in the ready-to-eat cereal industry the BLP instruments do not work, as they show little variation through space, time, or cross-section. In turn, the instruments proposed by Hausman et al. (1994) demand rich data sets, as they are require the observation of prices of the same brand in other geographical markets. Further, even if such detailed data set is available, the validity of prices in other markets as instruments may be questioned since there is always some common demand effect across markets that is not captured by usual controls (Bresnahan, 1996). Therefore, as described above, there are situations in which the researcher does not have applicable instruments[2]. Thus, either she abandons the research or proceed with typically downward biased estimates (in absolute value) of the price coefficient, which usually leads to implausible inelastic demands (see BLP) or, possibly, unreasonable positive own-price elasticities.

In this paper, I propose a new methodology to uncover the demand parameters. It offers an alternative to this uncomfortable dichotomous decision the researcher may face in some cases (abandon the research or proceed with biased estimates). By bringing additional information to the model, I demonstrate that one can retrieve the demand parameters of a particular class of mixed logit demand models without resorting to instrumental variables. The strategy can be summarized as follows. First, use this additional information to deterministically uncover (calibrate) the coefficient on the endogenous variable (price). Then project the residual (part of market shares that are not explained by prices) in the space of non-price attributes to econometrically estimate the remaining parameters.

This paper is organized as follows. In Section 2, the theoretical mixed logit demand model is presented. Section 3 presents the proposed methodology to uncover the parameters of this model. In section 4, the methodology is applied to uncover the Mixed Logit demand parameters and simulation of new entry is performed, using data on the U.S ready-to-eat cereal industry. And, finally, Section 5 presents the conclusion.

2. Model

In this section, I shall describe a mixed logit demand model with one random coefficient – henceforth MLOGIT[3]. Consumers rank products according to their characteristics and prices. There are $N+1$ choices in the market, N inside goods and one reference good (or outside good).

Consumer i chooses brand j, given price p_j, a K-dimensional row vector of observed characteristics (x_j), an unobserved characteristic (denoted by the scalar ξ_j), and unobserved

[2]Cost measures can also serve as instruments, one example is Villas-Boas (2007), but researchers rarely have access to such information for differentiated-product industries. Another approach relies on typical panel data methods (Arellano and Bover,1995 and Blundell and Bond, 1998). Di Giacomo (2008) applies this methodology to the Italian yougurt industry. However, a large data set is usually required and instruments may prove to be weak in many applications.

[3] It will be made clear why the restriction on the number of random coefficients is necessary in the methodology developed in this paper. The limitations arising from using a mixed logit model with only one random coefficient rather than its more general version with more than one random coefficient deserves further attention. However, it is important to stress that this restricted mixed logit model is superior to logit and nested logit models, which impose severe restrictions on price elasticities (see Nevo, 2000). Song (2007) uses a mixed logit with one random coefficient as a basis of comparison with pure characteristics models.

idiosyncratic preferences ε_{ij}, according to the following utility function:

(1) $$u_{ij} = g(\alpha, v_i, p_j) + x_j \beta + \xi_j + \varepsilon_{ij}$$

where $g(\alpha, v_i, p_j)$ is the utility term that defines how prices affect consumers´ preferences. In addition to price, this function depends on the parameter α and an unobserved (by the researcher) consumer-specific term v_i. The K-dimensional column vector β, whose typical element β_k represents the marginal utility of characteristic k, assumed invariant across consumers.

Alternatively, equation (1) can be rewritten as

(2) $$u_{ij} = g(\alpha, v_i, p_j) p_j + \delta_j + \varepsilon_{ij}$$

where $\delta_j = x_j \beta + \xi_j$ represents the mean utility of product j derived from characteristics other than prices. The utility derived from the consumption of the outside good can be normalized to zero u_{i0}=0. Assuming that ε_{ij} has a Type I extreme value distribution, the probability of individual i choosing good j (s_{ij}) takes the familiar logit form

(3) $$s_{ij}(\alpha, p, \delta(\beta, X, \xi), v_i) = \frac{\exp(g(\alpha, v_i, p_j) p_j + \delta_j)}{1 + \sum_{m=1}^{N} \exp(g(\alpha, v_i, p_j) p_m + \delta_m)}$$

The scalar s_{ij} is the conditional market share of product j, i.e. the market share that would prevail if all individuals had the same v_i. In the MLOGIT model this is not true, therefore, some aggregation is necessary.

Taking the expected value with respect to the distribution of v_i's yields the market share of product j implied by the model(s_j).

(4) $$s_j(\alpha, p, \delta(\beta, X, \xi)) = E_v[s_{ij}(\alpha, p, \delta(\beta, X, \xi), v_i)]$$

The theoretical market share of product j depends on the parameter α, and the $N+1$-dimensional vectors p and δ, that collect all p_j's and δ_j's respectively. Notice that, by definition, δ is an implicit function of β and X (a matrix obtained by stacking the x_j´s).

3. Augmenting the information set to uncover demand parameters

The basic idea of empirical strategies commonly adopted in structural models is to search for parameters that are able to match the shares predicted by the theoretical model $s_j(\alpha, p, \delta(\beta, X, \xi))$ to the observed shares ($\bar{s}_j$). Thus, the goal is to find the set of parameters that better explain the following relation

(5) $$\bar{s}_j = s_j(\alpha, p, \delta(\beta, X, \xi)) \;; \qquad j=1,\ldots N$$

Although traditional econometric techniques do not apply to the equation above, due to the non-linearity in the error term ξ, the main idea behind identification is standard. BLP develop an algorithm to uncover numerically the error term as function of the parameters. These error terms are combined with variables (instruments) to form moment conditions of the type $E[\xi_j \mid Z_j] = 0$, where Z_j is L-dimensional vector (L is the number of instruments). BLP propose a GMM method based on three sets of instruments. These instruments are based on the product attributes, which are assumed to be exogenous. The first set is formed by the so-called trivial instruments: the attributes themselves (excluding potentially endogenous ones, such as prices). The second is composed by the sum of the values of the same attribute across own-firm products. Finally, the third set of instruments is calculated by the sum of the values of the same attribute across rival firm products. The non-trivial instruments (those included in the second and third set of BLP instruments) are functions of the trivial ones and therefore may in many instances prove to be weakly correlated with the endogenous variable (price), leading to inference problems regarding the estimation of the coefficient on price (see Nevo, 1998).

The key contribution of this paper is to show how to incorporate external information into the empirical strategy in order to avoid the use of non-trivial instruments. The researcher brings many objects to the empirical strategy based on some belief. Indeed, structural IO models have many assumptions regarding consumer and producer behavior. Typical studies in this field assume a discrete-choice demand side and Bertrand behavior on the supply side. These assumptions constrain the data to accommodate a parametric family of functions. Obviously, the data set plays an important role, as the empirical strategy picks the parameters that better explain the observed data. However, there is one parameter of the model that is not left for the data to explain: the market size *M*. Many papers in this literature assume a particular value for this parameter. For instance, in BLP study of the U.S automobile industry, *M* is assumed to be the number of families. This assumption is based on the researcher's belief that each family is a potential consumer for an automobile in each year. A similar assumption is made by Petrin (2002) and Nevo (2001).

What I propose in this work is to go a little further and expand the set of information that is not left for the data to explain. One variable that economists and industry experts are used to dealing with is elasticity. Although any own- or cross price elasticities between any two goods could be used in the framework to be developed below, I use external information on price elasticity of the inside good l , defined as η_{ll}. The reason for this choice is that it represents a very intuitive economic magnitude: the attractiveness of the inside good l. This information could come from industry reports, other studies or market experts.

The idea of uncovering demand parameters using additional information on price elasticity is not entirely new. In another automobile study undertaken by Berry, Levinsohn, and Pakes (2004), the authors discuss possible instruments but conclude that they lead to a very imprecise estimate of the price coefficient. Then they use information on price elasticity provided by the staff at General Motors to calibrate the price coefficient. The distinctive feature of the model developed in this paper is that it addresses the estimation of an aggregate demand model, rather than a consumer-level one as in Berry, Levinsohn, and Pakes (2004), but the insight underlying the calibration of the price coefficient is the same.

Another closely related work was developed by Ivaldi and Verboven (2005) who apply a Nested Logit model to the European truck market. After conducting estimations with different specifications, they select (ex-post) the model that provides a better match to the estimates of price elasticities provided by the European Commission[4]. The model proposed in this paper provides a more formal way to conduct this matching criterion by introducing the elasticity information directly in the empirical strategy.

For the MLOGIT demand model presented in section 2, the implied price elasticity of one of the inside goods l is given by

[4] They use this information on elasticity to pin down the market size *M*. Here, I use it to determine the price coefficient and take *M* as given.

(6) $\eta_{ll}(\alpha,p,\delta)=\frac{p_l}{s_l}E_v[\frac{\partial g(\alpha,v_i,p_j)}{\partial p_j}.s_{il}(\alpha,p,\delta,v_i)(1-s_{il}(\alpha,p,\delta,v_i)]$

Notice that instead of elasticity one could equivalently impose economic measures that are related to it. For instance, adding the assumption of profit maximization, one could use the classic formula that equates margin (Lerner index) to the inverse of elasticity. Therefore, the information on price-cost margin could be transformed into elasticity and the methodology would proceed as described originally (in terms of the elasticity).

Methodology to uncover the demand parameters

The methodology can be divided into two stages. In the first stage we uncover (calibrate) the parameter of marginal utility of price α. Then, in the second stage, we show how to uncover the characteristics marginal utilities (β).

The first stage

I begin by setting up the following system of equations:

(7) $\bar{s}_j = s_j(\alpha,p,\delta)$; j=1,….N

(8) $\bar{\eta}_{ll}=\eta_{ll}(\alpha,p,\delta)$

The first equation in this system is simply the reproduction of equation (5), while the second equation is a consequence of the new information brought to the empirical method (equation 6). In addition to matching the observed market shares $\bar{s}_j$, the parameters of the theoretical model also have to match the elasticity of the inside good l ($\bar{\eta}_{ll}$).

Notice that, the system of equations above has *N+1* equations and, since *p* represents data (prices), there are *N+1* unknowns (*N*-dimensional vector δ plus the scalar α)[5]. Therefore, one can apply commonly available methods for solving non-linear equations to find the solution for the *N+1*- dimensional vector (δ,α).

The second stage

Once we have δ^*, obtained from the first part of the methodology, one is able to project this vector onto the space of product characteristics (except price) and estimate the parameters of the corresponding regression equation, which is given by

(9) $\delta_j = x_j\beta+\xi_j$

This equation can be estimated by OLS since characteristics are assumed to be exogenous, an assumption that, to the best of my knowledge, is shared by all papers in this literature. Notice also that we do not need to search for non-trivial instruments, i.e. instruments other than non-price characteristics (the trivial instruments), avoiding the problems associated with BLP instruments, that are likely to be weak in many instances,

[5] If α is vector of dimension greater than one, and not a scalar as assumed here, or if we had more than one random coefficient, the system would certainly be under identified. For this reason, I have to impose a mixed logit model with only one random coefficient with only one parameter. Whether this is a plausible model is largely an empirical question. Notice also that α is deterministic (calibrated) and therefore it does not have a standard error.

and Hausman price instruments, that places greater demands on the data set[6] and may be invalid in some situations.

The Simple Logit

In this subsection I present the simplest discrete-choice model: the Logit. This exposition serves the purpose of highlighting the contribution of bringing more external information (price elasticity) to the empirical strategy without having to deal with the lack of analytical formulas and the consequent numerical and computational issues. However, this is done for expositional purposes only. As well documented in the discrete-choice literature (Berry,1994), the Logit demand model places very restrictive limitations on own and cross price elasticities, which constitute critical parameters in the economic evaluation of innovation, mergers and entry of new products.

The Logit is a particular case of the MLOGIT. Indeed, if $g(\alpha, v_i, p_j) = -\alpha p_j$, the shares are given by $s_j(\alpha, p, \delta) = \dfrac{\exp(-\alpha p_j + \delta_j)}{1 + \sum_{m=1}^{N} \exp(-\alpha p_m + \delta_m)}$

Log-linearizing this equation we have $\ln s_j - \ln s_0 = -\alpha p_j + \delta_j$. The Logit also implies an analytical formula for the own price elasticity of a given good l.

Indeed, $\eta_{jj}(\alpha, p, \delta) = -\alpha p_l (1 - s_l(\alpha, p, \delta))$. The system of equation - equations (7) and (8) - simplifies to the following system of linear equations[7]:

(10) $\ln \bar{s}_j - \ln \bar{s}_0 = -\alpha p_j + \delta_j$; *j=1,...N*

(11) $\bar{\eta}_{jj}(\alpha, p, \delta) = -\alpha p_l (1 - s_l)$

This system is much simpler than its version for the more general MLOGIT model. We can directly solve for α from Equation (11), giving $\alpha = -\dfrac{\bar{\eta}_{ll}}{p_l(1 - s_l)}$. Once α is determined, we can find the corresponding δ_j's ($\delta_j = \ln \bar{s}_j - \ln \bar{s}_0 + \alpha p_j$) from Equation (10). The second part of the methodology is the same as in the MLOGIT. With the δ_j's we are able to run the regression $\delta_j = x_j \beta + \xi_j$ using OLS. The logit version of the model bears a resemblance with the so-called Antitrust Logit Model, a methodology developed by Werden and Froeb (1994). Indeed, these authors use an equivalent set of equations to determine α and the δ_j's.

It is important to notice that the MLOGIT model presented in this paper provides a generalization of their idea as it accommodates consumer heterogeneity, a crucial element if we want to generate reasonable patterns for the elasticities between any two products.

Additional comments on the methodology

Since the key element of the methodology relies on external information (elasticity), one legitimate concern is how it affects the main outputs, namely α and mean utilities (δ_j´s). One way to assess the relationship between the key input (elasticity) and the main outputs of

[6] we need to observe at least one cross-section of markets

[7] The system is linear in the unknowns (δ, α)

the model is through a formal (analytical) approach. This can be easily achieved for the simple logit model. Indeed, from equation (11), $\alpha = |\overline{\eta}_{ll}| / (p_l(1-s_l))$. Since the denominator is data, a higher (lower) elasticity in absolute value $|\overline{\eta}_{ll}|$ implies a higher (lower) α. In turn, solving equation (10) for δ_j leads to the following equation $\delta_j = \ln \overline{s}_j - \ln \overline{s}_0 + \alpha p_j$. Since the first two right hand side terms of this equation are data, one can easily verify that δ_j and α move in the same direction. Therefore, a higher (lower) elasticity in absolute value $|\overline{\eta}_{ll}|$ implies a higher (lower) δ_j. Both results can be summarized in the following statement: α and mean utilities (δ_j´s) move in the same direction as the absolute value of the elasticity $|\overline{\eta}_{ll}|$.

However, due to the high nonlinearity of the MLOGIT system of equations (7 and 8), one can only conjecture that the same analytical results carry over to the more general model proposed in this paper. One way to improve confidence in this conjecture is to perform sensitivity analysis by running the model for different values of own-price elasticity. Such exercises are conducted in the appendix and the results are consistent with this conjecture.

Another key element of the methodology is the specification of the price term in the utility function. Notice that the MLOGIT model presented in sections 2 and 3 accommodate different functional forms $g(\alpha, v_i, p_j)$. For instance, it can accommodate the functional forms found in seminal papers in the literature (Berry, Levinsohn and Pakes, 1995, 1999). In the empirical example below, their 1999 formulation is used $g(\alpha, v_i, p_j) = -(\alpha / v_i) p_j$. Alternatively, the price term in the utility function could take the following functional form assumed in their 1995 paper $g(\alpha, v_i, p_j) = \alpha \log(v_i - p_j)$. As noted in Berry, Levinsohn and Pakes (1999), the former functional form is a first order Taylor series approximation of the latter. Which functional form is the most appropriate is an open question in the literature and is beyond the scope of this work.

4. An empirical example

In order to illustrate the methodology, I use data on the ready-to-eat cereal industry. However, it should be noticed that the objective of this section is to illustrate the methodology proposed in this paper rather than providing a detailed study of the ready-to-eat cereal industry. Nonetheless, an application of this methodology that takes into consideration all or most of the idiosyncrasies of this industry would be an interesting extension of this work.

The data set is a cross-section of the fifty top selling brands in the U.S in 1992. The summary statistics are presented below[8]. The data set reports information on shares, prices, fat, sugar, advertising exposure and two dummies: DKIDS assumes the value 1 if the brand belongs to the kids segment and 0 otherwise, and DKG, which takes on the value 1 if the brand belongs to Kelloggs (the market leader) and 0 otherwise. To construct the shares it is assumed that *M* is the total cereal purchases observed in the dataset. This implies that the outside good is representative of all other brands not included in the top fifty best selling list[9].

---

[8] This data was collected by Matt Shum and is publicly available in his personal webpage.(Acessed August 2011). http://www.hss.caltech.edu/mshum/gradio/ioclass.html.

[9] This implies that not purchasing the product is not an option, which may constitute a restrictive assumption in many setups. However, according to Schum's data, for the cereal industry this is could be a good approximation since, in 1992, 97.1% of American households purchased some cereal during the year. Furthermore, notice that

Table I- Summary statistics for Ready-To-Eat Cereal Industry in the U.S – 1992

	Mean	*Std Dev*	*Variance*	*Min*	*Max*
Share	0.0152	0.0102	0.0001	0.0067	0.0567
Price ($/lb)	2.9830	0.4916	0.2416	1.7700	3.9600
Fat(cal)	1.6080	1.6884	2.8505	0	8.0000
Sugar(g)	10.1080	5.4177	29.3514	0	20.000
Advert. ($millions)	2.8643	1.9049	3.6287	0	7.8670
DKIDS	0.24	0.4314	0.1861	0	1.000
DKG	0.34	0.4785	0.229	0	1.000

Source: Descriptive statistics for variables available in the data set mentioned above.

I follow Berry, Levinsohn, and Pakes (1999) and parameterize the consumer marginal utility for price according to the functional form given by $g(\alpha, v_i, p_j) = -\frac{\alpha}{v_i} p_j$, where the consumer-specific term v_i represents household income, whose distribution is obtained from the 1992 Current Population Survey (CPS). In order to simplify the computation of the MLOGIT model, I made a few simplifications regarding this distribution. I have divided the income space into intervals of the same size (2500 USD) and computed the frequencies of each interval. Then, I discretize the distribution assuming that the average income in each interval is representative of all individuals included in this interval. In the end, we have 21 income levels and thus 21 consumer types. The discretization is a non-parametric approach, which avoids the need to impose distributional assumptions on income. Notice that if the researcher is not willing to make these simplifications, the methodology outlined in section 3 can accommodate different parametric distributions for income, in which case numerical integration methods should be used to calculate the shares and the elasticity according to equations (7) and (8).

In the first stage of the methodology, I pick the brand Apple Cinnamon Cheerios (ACC) from General Mills and assume its elasticity to be $\overline{\eta}_{ll}$ =-3. And, as mentioned before, *M* is the total cereal purchases observed in the dataset[10]. Then, one is able to uncover *N+1*-dimensional vector (δ, α). I find that α is 36482.18, from which we can derive the distribution of the price coefficients (in absolute values) across consumers. This distribution is given by the ratio (α/v_i). We can also construct descriptive statistics for the δ_j's. These results are summarized in Table II below.

the methodology developed in this paper can accommodate any other value for *M*, and therefore any other value of the market size could have been used to illustrate the methodology.

[10] These values compose the information set the researcher brings to the empirical strategy. I could have used other values for the price elasticity and market size to illustrate the methodology.

Table II- Summary statistics of stage 1 results (MLOGIT model)

	Mean	Median	Max	Min
Price coefficient	1.739	0.694	14.593	0.347
Mean utilities (δ_j's)	3.223	3.258	4.647	1.051

The distribution of the price coefficient has mean 1.739 and median 0.694, implying that the distribution is not symmetric around its mean. The mean utilities do not exhibit much variation across brands and the distribution is approximately symmetric around the mean since the mean and the median are approximately equal.

In the second stage of the MLOGIT model, we are able to estimate the characteristics coefficients using OLS. The results for the MLOGIT model can be found in Table III below. All coefficients are statistically significant at the 10% confidence level. However, only the coefficients on fat, sugar and advertising are significant at the 5% confidence level.

Table III- Stage 2 results (MLOGIT model)

	Coef. (β)	Stand. error	t-value	Prob>ltl
Fat	0.223	0.108	2.069	0.044
Sugar	0.080	0.029	2.726	0.009
Advert.	0.463	0.072	6.426	0.000
DKIDS	0.764	0.411	1.859	0.070
DKG	0.698	0.363	1.923	0.061

Counterfactual experiment

An advantage of structural estimation is that, once the parameters of interest are determined, one can simulate the effect of different market environments using the usual welfare metrics. The framework for counterfactual simulations laid out in this section is standard in discrete-choice demand models. The counterfactual experiment goes as follows. Determine the demand parameters. Next, simulate the entry of a new good with a given price (p_*), a k-dimensional row vector of characteristics (x_*) and a value for quality that is not captured by these characteristics (ξ_*). Then, calculate the market penetration of the new good (share) and consumer surplus variation. For the MLOGIT model described in this section, McFadden (1981) and Nevo (200b) show that surplus variation (ΔCS) of consumer *i* is given by

$$(12) \qquad \Delta CS_i = \frac{1}{(\alpha / v_i)} \ln \left\{ \frac{1 + \left(\sum_{m=1}^{N} \exp[-(\alpha/v_i) p_m + x_m \beta + \xi_m] \right) + \exp[-(\alpha/v_i) p_* + x_* \beta + \xi_*]}{1 + \sum_{m=1}^{N} \exp[-(\alpha/v_i) p_m + x_m \beta + \xi_m]} \right\}$$

In order to obtain the average of consumer welfare variation we have to integrate out the consumer specific term v_i. This measure is given by

$$(13) \qquad \Delta CS = E_v \{\Delta CS_i\}$$

Tables IV and V show the results from different simulations. The first columns describe the characteristics of the new good (indexed in the first column). The last 2 columns present the simulation results in terms of market shares the new product is able to gain and average per consumer surplus in 1992 USD. Each row of this table defines the characteristics of the new good introduced in the market. For instance, in the experiment indexed by 1, I simulate the introduction of a product with the following characteristics. It is the destination of 2.86 million USD spent on advertising and contains zero fat and 20 g of sugar. Also, it does not belong to the kids segment and is not produced by Kelloggs (the market leader). From table IV below we verify that this new product gains a market share of 1.24% and implies a positive per consumer surplus variation of 3.94 USD. In the other entries of this table, I reduce the sugar content and verify that market shares and consumer gains decrease. In each experiment, I simulate the introduction of a different good. This process is non-cumulative.

In addition, I conduct the same sequence of experiments but assume that the introduced product belongs to Kelloggs (see table V). The results are superior for market shares and consumer gains, due to the fact that Kelloggs' products are in average more attractive than non-kelloggs' products (see regression results in table III).

Table IV- First set of Simulation results

Experiment Index	Fat	Sugar	Adv	DKIDS	DKG	Mkt.Share (%)	ΔCS (1992 USD)
1	0	20	2.86	0	0	1.248	3.941
2	0	15	2.86	0	0	0.842	2.645
3	0	10	2.86	0	0	0.567	1.774
4	0	5	2.86	0	0	0.381	1.191

Note: Only sugar content varies across experiments

Table V-Second set of Simulation results

Experiment Index	Fat	Sugar	Adv	DKIDS	DKG	Mkt.Share (%)	ΔCS (1992 USD)
5	0	20	2.86	0	1	2.467	7.920
6	0	15	2.86	0	1	1.673	5.314
7	0	10	2.86	0	1	1.131	3.566
8	0	5	2.86	0	1	0.762	2.393

Note: Only sugar content varies across experiments

5. Final comments

Demand estimation in product-differentiated industries has been the central object in many studies in the industrial organization field. Indeed, after pinning down the preference parameters it is possible to analyze issues related to innovation, antitrust (mergers and divestitures), calculation of quality adjusted price-indices and prediction of the competitive effect of entry and exit of products. However, uncovering consumers' preferences using aggregate data on product-differentiated markets imposes a serious challenge: find instruments to deal with price endogeneity. Berry, Levinsohn, and Pakes (1995) propose a GMM method based on instruments that are functions of the regressors (except price) to estimate general Random Coefficients Discrete-Choice models. These instruments may prove in many instances to be weakly correlated with the endogenous variable (price), leading to inference problems regarding the estimation of the coefficient on price. The key contribution of this paper is to show how to incorporate more information into the empirical strategy in order to avoid the need for such instruments. I use external information on price elasticity to propose a methodology to determine the parameters of a particular class of Random Coefficients Discrete-Choice models. I show that, provided that the external information is valid, one can determine the demand parameters using only the exogenous regressors (characteristics other than prices) as instruments, avoiding then the need to use potentially weak instruments. Finally, for illustrative purposes, I apply this methodology to the ready-to-eat cereal industry and simulate the entry of new products.

References

Arellano M, Bover O. (1995) Another look at the instrumental variables estimation of error components models. *Jornal of Econometrics* ;68;29–51

Berry, S. (1994) Estimating Discrete-Choice Models of Product Differentiation. *Rand Journal*; 25; 242-262.

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

____________________(1999) Voluntary Export Restraints in Automobiles. *American Economic Review*; 89; 400-430.

_____________________(2004) Estimating Differentiated Product Demand Systems from a Combination of Micro and Macro Data: The Market for New Vehicles. Journal of Political Economy; 112; 68-105.

Blundell R, Bond S. (1998) Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*; 87; 115–143

Bresnahan, T. (1996) Comments on "Valuation of New Goods Under Perfect and Imperfect Competition" by Jerry Hausman, in T. Bresnahan and R. Gordon , eds., The Economics of New Goods, Studies in Income and Wealth 58; Chicago: National Bureau of Economic Research.

Di Giacomo, M.(2008) GMM estimation of a structural demand model for yogurt and the effects of the introduction of new brands. *Empirical Economics* 2008; 34;537–565

Hausman, J., Leonard, G., Zona J. (1994) Competitive Analysis with Differentiated Products. *Annales d'Economie et de Statistique*; 34; 159-180.

Ivaldi, M. and Verboven, F. (2005) Quantifying the Effects of Horizontal Mergers in European Competition Policy. International Journal of Industrial Organization 2005;23;669-691

McFadden, D. Econometric Models of Probabilistic Choice., in C. Manski and D. McFadden (Eds), (1981) *Structural Analysis of Discrete Data.*. University of Chicago Press: Chicago.

Nevo, A. (1998) Measuring Market Power in the Ready-to-Eat Cereal Industry". NBER *Working Paper*, N^0 6387, 1998.

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

_______(2000b). "Mergers with Differentiated Products: The Case of the Ready-to- Eat Cereal Industry," *Rand Journal of Economics*, 31, 395-421

______ (2001) "Measuring Market Power in the Ready-to-Eat Cereal Industry". *Econometrica*, 69(2) pp.307-342.

_______ (2001) Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*;69;307-342.

Petrin, A.(2002) Quantifying the benefits of New Products: The Case of the Minivan. *Journal of Political Economy*;110; 705-729.

Song, M.(2007) Measuring Consumer Welfare in the CPU Market: An Application of the Pure Characteristics Demand Model. *RAND Journal of Economics*;38; 429-446

Villas-Boas, S. (2007) Vertical Relationships Between Manufacturers and Retailers: Inference With Limited Data The Review of Economic Studies; 74(2); 625-652

Werden, G. and Luke, F. (1994) The Effects of Mergers in differentiated Products Industries: Logit Demand and Merger Policy. *Journal of Law, Economics, & Organization*;10; 407–26.