Combining Prior Information and Data to Uncover the Parameters from the Random Coefficient Discrete–Choice Demand Model

Sergio Aquino DeSouza**

Abstract

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 do 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. However, these instruments may prove to be in many instances 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 prior information into the empirical strategy in order to avoid the need for such instruments. What I propose in this work is to augment the researchers' set of prior information. I use prior information on the aggregate price elasticity to propose a two-stage methodology that is able to determine the parameters of a particular class of Random Coefficients Discrete-Choice models. I show that, provided that the prior information is valid, we 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.

Keywords- Discrete-Choice; Demand, Mixed Logit

Resumo

Estimação da demanda tem sido o objeto central em vários estudos de organização industrial. De fato, após determinar os parâmetros das preferências dos consumidores é possível analisar questões ligadas à inovação, defesa da concorrência, cálculo de índice de preços ajustados pela qualidade e previsão de efeitos competitivos de entrada e saída de produtos. No entanto, determinar preferências a partir de dados agregados em indústrias caracterizadas por produtos diferenciados impõe um sério desafio: encontrar instrumentos válidos para lidar com o problema de endogeneidade doa preços. Berry, Levinsohn, and Pakes (1995) propõem o método dos momentos generalizados (MMG) baseado em instrumentos que são funções dos regressores (exceto preço) para estimar um modelo de demanda discreta com coeficientes aleatórios. No entanto, tais instrumentos podem se mostrar fracamente correlacionados com a variável endógena (preço) em muitas aplicações, gerando problemas de inferência com respeito à estimação do coeficiente da variável preço. A principal contribuição deste artigo consiste em incorporar mais informação a priori na estratégia empírica de forma a evitar a o uso de instrumentos. O que se propõe neste trabalho é aumentar o conjunto de informações que o pesquisador impõe a priori. Especificamente, utiliza-se informação a priori sobre a elasticidade agregada para propor uma metodologia de dois estágios cuja finalidade é determinar os parâmetros de uma classe particular de modelos de demanda discreta com coeficientes aleatórios. Mostra-se que é possível determinar os parâmetros da demanda utilizando apenas os regressores exógenos (características dos produtos) como instrumentos. O que evita a necessidade de utilizar instrumentos potencialmente fracos. Para ilustrar a metodologia, aplica-se o modelo à indústria de cerais prontos para consumo e simula-se a entrada de novos produtos.

Palavras-Chave – Escolha Discrete-, Demanda, Mixed-logit

JEL - D12, L1

*Filiação Institucional: **CAEN** e Departamento de Teoria Econômica, Univ. Federal do Ceará; End.: Av. da Universidade, 2700, Segundo Andar, 60020181, Fortaleza, CE,Brazil. Fone: 55-85-33667751 ; *e-mail:* **sergiodesouza@caen.ufc.br**

I-INTRODUCTION

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 demand parameters from aggregate data on product-differentiated markets imposes several challenges: (1) number of parameters to be determined; (2) incorporation of consumer heterogeneity and (3) price endogeneity.

There are basically two categories of 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 in equilibrium may purchase 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 an over parameterized system. Furthermore, by construction, representative consumer models can not naturally deal with the presence of consumer heterogeneity. The second set of demand models is based on the theory of discrete-choice, in which a consumer is assumed to choose 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 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. An alternative to the BLP instruments was first 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 some 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'*.

The types of instruments proposed by BLP and Hausman et al. (1994) are far from being a consensus among researchers in the IO field (for a more detailed discussion see Nevo, 2001). In this paper I propose a novel methodology to uncover the demand parameters that avoids the difficult task of searching for valid instruments. By augmenting the researchers's prior information set I demonstrate that one can retrieve the demand parameters of some discrete-choice demand models using only the regressors (excluding price) as instruments.

II - MODEL

In this section, I shall describe a discrete-choice demand model with one random coefficient – henceforth ORDC. The limitations of using a mixed logit model with only one random coefficient rather than its more general version with more than one random coefficient depends on the application the researcher has in mind[1]. Song (2007) uses this class of models (ORDC) as a basis of comparison with pure characteristics models.

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 (ξ_j), and unobserved idiosyncratic preferences ε_{ij}, according to the following indirect utility function:

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

where $g(\alpha, v_i)$ is a random coefficient that represents consumer *i*'s marginal utility (or disutility) of price, which is a function of the parameter α and an unobserved (by the researcher) consumer-specific term v_i. The *K*-dimensional column vector β , whose element β_k represents the marginal utility of characteristic *k*, assumed invariant across consumers.

[1] This will be made why the restriction on the number of random coefficients is necessary in the methodology developed in this paper.

Alternatively, Equation (1) can be rewritten as

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

where $\delta_j = x_j \beta + \xi_j$ and represents the mean utility o product j derived from characteristics other than prices. The utility derived form 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) $$s_{ij}(\alpha, p, \delta(\beta, X, \xi), v_i) = \frac{\exp(g(\alpha, v_i) p_j + \delta_j)}{1 + \sum_{m=1}^{N} \exp(g(\alpha, v_i) p_m + \delta_m)}$$

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 ORDC model this not true therefore, some aggregation argument has to be invoked. Indeed, 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) $$s_j(\alpha, p, \delta(\beta, X, \xi)) = E_v[s_{ji}(\alpha, p, \delta(\beta, X, \xi), v_i)]$$

The theoretical market share of product j depends on the parameter alfa, and *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 containing all observed characteristics of all products in the market).

III- AUGMENTING THE SET OF PRIOR INFORMATION 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, we try to find the set of parameters that better explain the following relation

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

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 $E[\xi_j | Z_j] = 0$, 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 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, 2001).

The key contribution of this paper is to show how to incorporate more prior information into the empirical strategy in order to avoid the use of non-trivial instruments. Although this is rarely noticed, the researcher already brings many objects to the empirical strategy based on some prior 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. However, 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*. Virtually all 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 prior 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 augment the set of prior information that is not left for the data to explain. Another variable that economists and industry experts are used to dealing with is elasticity. Although any own- or cross price elasticities between any tow goods could be used in the framework to be developed below, I use prior information on the aggregate price elasticity of the inside goods η_I, measured by the effect of an equally proportional increase in all inside goods prices on the aggregate market share of the inside goods s_I. The reason for this choice is that aggregate elasticity is easier to deal with than other types of elasticities, such as own- and cross-piece elasticities between any two goods. It represents a very intuitive economic magnitude: the attractiveness of the inside goods with respect to the outside product. This prior information could come from different sources. The researcher could use his own experience and knowledge of the industry or, alternatively, he or she could draw on industry experts as information sources.

This last type of source has been utilized in another automobile study undertaken by Berry, Levinsohn, and Pakes (2004). They report that "*based on their experience, the staff at the General Motors Corporation suggested that the aggregate price elasticity in the market for new vehicles was near one*". The aggregate elasticity η_I is given by

(6) $$\eta_I = \frac{\partial s_I(\lambda p)}{\partial \lambda} \frac{\lambda}{s_I}\bigg|_{\lambda=1}$$

For the ORDC demand model presented in section II the implied price elasticity of the aggregate demand of all inside goods is given by

(7) $$\eta_I(\alpha, p, \delta) = \frac{E_v[g(\alpha, v_i).\overline{P}_i(\alpha, p, \delta, v_i).s_{i0}(\alpha, p, \delta, v_i)]}{s_I}$$

where s_I is the aggregate demand of all inside goods, $\widetilde{P}_i = \sum_{m=1}^{N} s_{im} p_m$ is a weighted average price (the weights are given by the unobserved conditional market shares of each inside good j, given by equation (5)), and $s_{i0}(\alpha, p, \delta, v_i) = \dfrac{1}{1 + \sum_{m=1}^{N} \exp(g(\alpha, v_i) p_m + \delta_m)}$ is the conditional market share of the outside product.

Methodology to uncover the demand parameters

The methodology can be divided into two stages. In the first stage we uncover the parameter of marginal utility of price α, which is defined by Nevo(2001) as a non-linear parameter. Then, in the second stage, I show how to uncover the characteristics marginal utilities (β). These are referred to as the linear parameters. It will be clear below why this classification between linear and non-linear parameters is appropriate.

The first stage

I begin by setting up the following system of equations:

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

(9) $\bar{\eta}_I = \eta_I(\alpha, p, \delta)$

The first equation in this system is simply the reproduction of Equation (5), while the second equation comes from the new information brought to the empirical method. In addition to matching the observed market shares, the parameters of the theoretical model are also asked to match the aggregate elasticity of the inside goods. 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)[2]. Therefore, we can solve for the *N+1*- dimensional vector (δ, α). One possible method to find the solution of the system is to employ commonly applied algorithms that search for the solution directly in the (δ, α) space. However, this would be computationally inefficient. Recall that one of the main motivations of the discrete-choice model relies on its the ability to deal with markets characterized by the presence of many brands. If we had 40 brands, for example, the algorithm would be searching directly in a space with dimension 41.

In order to deal with this dimensionality problem, we can take advantage of an important result derived in BLP. Given the parameter α and *p* the mapping defined pointwise by

$$T(s, \alpha, p)[\delta_j] = \delta_j + \ln(\bar{s}_j) - \ln(s_j(\alpha, p, \delta))$$

is a contraction mapping with modulus less than one. Therefore, we can improve computational efficiency by concentrating the search. Shortly, the algorithm goes as follows. The first step initiates the outer loop, which begins with a value of α', solve for the implied $\delta'(\alpha')$ by applying the contraction mapping algorithm (inner loop) to the sub-system formed by the *N* equations in (8). Then we calculate the implied aggregate elasticity of the inside goods $\eta_I(\alpha', p, \delta')$ and then check whether equation (9) is satisfied. In this last step we verify how large is the distance between the prior information on the elasticity $\bar{\eta}_I$ and the implied $\eta_I(\alpha', p, \delta')$. If this α' does not imply a close enough distance, measured by $|\bar{\eta}_I - \eta_I(\alpha', p, \delta')|$, we repeat this process, by reinitiating the outer loop, until convergence has been attained[3].

---

[2] 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 we have to posit 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 and therefore it does not have a standard error.

[3] Thus, no matter how large is *N* (number of brands) the algorithm searches directly in a one-dimensional space.

The second stage

Once we have δ^*, obtained from the first part of the methodology, we are 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

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

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 no 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, and Hausman price instruments, that places greater demands on the data set[4] and may be invalid in some situations (see Nevo,2001).

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 prior information (aggregate elasticity) to the model 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 (see BLP), 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.
In the Logit case, we can assume without loss of generality that $g(\alpha, v_i) = \alpha$.

Then 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 $\ln s_j(\alpha, p, \delta) - \ln s_0(\alpha, p, \delta) = \alpha p_j + \delta_j$.

The Logit also implies an analytical formula for the aggregate elasticity.

Indeed, $\eta_I(\alpha, p, \delta) = \dfrac{\alpha.\overline{P}.s_0}{s_I}$, where , $\tilde{P} = \sum_{m=1}^{N} s_m p_m$ is a weighted average price (the weights are given by the observed market shares of each inside

---

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

good *j*). Notice that s_0, s_I and $\tilde{P}$ are observed. The system of equation - Equations (8) and (9) - simplifies to the following system of linear equations[5]:

(11) $$\ln \bar{s}_j - \ln \bar{s}_0 = \alpha p_j + \delta_j \quad ; \quad j=1,\ldots N$$

(12) $$\bar{\eta}_I = \frac{\alpha.\bar{\tilde{P}}.\bar{s}_0}{\bar{s}_I}$$

This system is much simpler than its version for the more general ORDC model. We can directly solve for α from Equation (12), giving $\alpha = \frac{\bar{\eta}_I.\bar{s}_I}{\bar{\tilde{P}}.\tilde{s}_0}$. 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 (11). The second part of the methodology is the same as in the ORDC model. With the δ_j's we are able to run the regression $\delta_j = x_j \beta + \xi_j$ using OLS. For those who know the so-called antitrust model, a methodology developed by Werden and Froeb (1994), the simple logit version of the methodology presented above may sound familiar. Indeed, these authors use the same set equations to determine α and the δ_j's. The improvement presented here is, provided that we have enough data, to project the δ_j' onto the space of characteristics using simple OLS.

It is also important to notice that the ORDC 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. The model also has, both in its logit and ORDC version, the additional advantage, when compared to the antitrust logit model, of proposing a method to determine the marginal effects of characteristics (β). Calculating this marginal effects vector is important to measure the welfare effects of the entry of new products.

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

IV - 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 reason for the choice of this industry is mainly methodological. Indeed, the BLP instruments, constructed from typical data sets available for this industry, are likely to be weak. Indeed, unlike the automobile industry, there is not much variation in these instruments over time, and even less so between geographic markets (Nevo, 2001). Therefore, unless we are willing to exploit the panel structure and use the prices in other geographic markets as instruments, we are stuck with a cross-section and the weak BLP instruments. This is the scenario for which the methodology presented in this paper is most appealing. The data set is a cross-section of fifty top selling brands in the U.S in 1992. The summary statistics are presented below[6]. 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 DKG, which takes on the value 1 if the brand belongs to Kelloggs (the market leader). To construct the shares it is assumed that M is the total cereal purchases observed in the dataset. Thus, this implies that the outside good is representative of all other brands not included in the top fifty best selling list[7].

[6] This data was collected by Matt Shum and is publicly available in his personal webpage.(Acessed December 2007).http://www.econ.jhu.edu/people/shum.

[7] 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 the methodology developed in this paper can accommodate any prior information on M, and therefore any other value of the market size could have been used to illustrate the methodology.

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

* Descriptive statistics for variables available in the data set mentioned in the text.

I follow Berry, Levinsohn, and Pakes (1999) and parameterize the consumer marginal utility for price according to the functional form given by $g(\alpha, v_i) = -\frac{\alpha}{v_i}$, 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 ORDC 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 avoids the need for numerical integration (e.g. quadrature methods) or simulation methods (as employed by BLP) to compute the markets shares in Equation (4). This is done to reduce the computational burden. Notice that if the researcher is not willing to make these simplifications, the methodology model outlined in section III can certainly accommodate different distributional assumptions for income such that quadrature or simulation methods can be used.

In the first stage of the ORDC model, I posit that $\overline{\eta}_I$ =-3 and, as mentioned before, *M* is the total cereal purchases observed in the dataset[8]. Then we are able to uncover *N+1*-dimensional vector (δ, α). I find that α is 41567.92, from which we can derive the distribution of the price coefficients

[8] These values compose the prior information set. I could have used other values for the aggregate elasticity and market size to perform robustness checks, especially by changing the $\overline{\eta}_I$'s. This is left for future developments of this work.

(in absolute values) across consumers. This distribution is given by the distribution of the ratio $\frac{\alpha}{v_i}$. We can also construct descriptive statistics for the δ_j's. These results are summarized in Table II below.

Table II

Summary statistics of stage 1 results (ORDC model)

	Mean	Median	Max	Min
Price coefficient	1.982	0.791	16.627	0.396
Mean utilities (δ_j's)	4.087	4.099	5.541	1.705

The distribution of the price coefficient has mean 1.982 and median 0.791, implying that the distribution is not symmetric around its mean. The mean utilities dos 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 ORDC model, we are able to estimate using OLS the characteristics coefficients. The results for the ORDC 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 (ORDC model)

	Coef. (β)	Stand. error	t-value	Prob>|t|
Fat	0.280	0.135	2.072	0.044
Sugar	0.110	0.0368	3.000	0.004
Advert.	0.556	0.090	6.150	0.000
DKIDS	0.975	0.5154	1.892	0.065
DKG	0.858	0.455	1.885	0.066

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 distinctive difference is that the entries on the welfare metric are obtained by the method described in section III that shows how to incorporate prior information to uncover the demand parameters without the need to search for instruments. 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 consumer surplus variation.

For the ORDC model described in section II, McFadden (1981) shows that surplus variation (ΔCS) of consumer *i* is given by

$$(13) \quad \Delta CS_i = \left\{ \frac{\ln\left(1+\left(\sum_{m=1}^{N}\exp[g(\alpha,v_i)p_m + x_m\beta+\xi_m]\right)+\exp[g(\alpha,v_i)p^* + x_*\beta+\xi_*]\right)}{g(\alpha,v_i)\ln\left(1+\sum_{m=1}^{N}\exp(g(\alpha,v_i)p_m+\delta_m)\right)} \right\}$$

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

$$(14) \quad \Delta CS = E_v \left\{ \frac{\ln\left(1+\left(\sum_{m=1}^{N}\exp[g(\alpha,v_i)p_m + x_m\beta+\xi_m]\right)+\exp[g(\alpha,v_i)p^* + x_*\beta+\xi_*]\right)}{g(\alpha,v_i)\ln\left(1+\sum_{m=1}^{N}\exp(g(\alpha,v_i)p_m+\delta_m)\right)} \right\}$$

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 that is introduced. 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 1.37% and implies a positive per consumer surplus variation of 7.40 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, we 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	1.37	7.40
2	0	15	2.86	0	0	0.79	4.26
3	0	10	2.86	0	0	0.46	2.45
4	0	10	2.86	0	0	0.27	1.42

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	3.16	17.43
6	0	15	2.86	0	1	1.85	10.04
7	0	10	2.86	0	1	1.07	5.79
8	0	10	2.86	0	1	0.62	3.34

Note: Only sugar content varies across experiments

V. FINAL REMARKS

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 do 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. Therefore these instruments in many instances may prove 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 prior information into the empirical strategy in order to avoid the need for such instruments. What I propose in this work is to augment the researchers' set of prior information. I use prior information on the aggregate price elasticity, measured by the effect of equally proportional increase in all inside goods prices on the aggregate market share of the inside goods, to propose a two-stage methodology to determine the parameters of a particular class of Random Coefficients Discrete-Choice models. I show that, provided that the prior information valid, we can determine the demand parameters using only the exogenous regressores (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

Ackerberg, D. and Rysman, M. (2002), "Unobserved Product Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects," NBER Working Paper No. 8798.

Anderson, S., De Palma, A., and Thisse, J.-F. (1992), *Discrete Choice Theory of Product Differentiation*, MIT Press, Cambridge, Massachusetts.

Berry, S. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *Rand Journal*, 25(2), pp. 242-262.

Berry, S., Levinsohn, J., and Pakes, A. (1995), "Automobile Prices in Market Equilibrium" *Econometrica*, 63(4), pp. 841-890.

____________________ (1999), "Voluntary Export Restraints in Automobiles," *American Economic Review*, 89(3), pp. 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 (1), pp. 68-105.

Caplin, A. and Nalebuff, B. (1991), "Aggregation and Imperfect Competition: On the Existence of Equilibrium," *Econometrica*, 59(1), pp. 25-59.

Cardell, S. (1997), "Variance Components Structures for the Extreme Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, 13(2), pp. 185-213.

Goldberg, Pinelopi (1995), "Product differentiation and Oligopoly in International Markets: The Case of the U.S Automobile Industry" Econometrica, 63 (4), pp.891-952.

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

Imbens, G., and Lancaster, T. (1994) "Combining Micro and Macro Data in Microeconometric Models" *Review of Economic Studies*, 61(4), pp. 655-680.

Katayama, H., Lu, S. and Tybout, J. (2003) "Why Plant-Level Productivity Studies Are Often misleading, and an Alternative Approach to Inference". NBER Working Paper No.9617.

Lahiri, K. and Gao, J. (2001), "Bayesian Analysis of Nested Logit Model by Markov Chain Monte Carlo," mimeo.

McFadden, D. (1981) “Econometric Models of Probabilistic Choice’’, in C. Manski and D. McFadden (Eds), *Structural Analysis of Discrete Data.*

Marshak, J. and Andrews, W. H. (1944), “Random Simultaneous Equations and the Theory of Production,” *Econometrica*, 12, pp. 143-205.

Nevo, A.(2000) “A Practitioner’s Guide to Estimation of Random-Coefficients Logit Models of Demand”. *Journal of Economics & Management Strategy*, 9(4), pp.513–548

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

Pakes, A. and McGuire, P. (1994), “Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model,” *RAND Journal of Economics*, 25(4), pp. 555-589.

Petrin, A. (2002) “Quantifying the benefits of New Products: The Case of the Minivan”. *Journal of Political Economy*. 110 (4), pp. 705-729.

Poirier, D. (1996), “A Bayesian Analysis of Nested Logit Models,” *Journal of Econometrics*, 75(1), pp. 163-181.

Slade, M. (2004), “"Market Power and Joint Dominance in U.K. Brewing", *Journal of Industrial Economics,* 52, pp. 133-163

Werden, Gregory J. and Luke M. Froeb (1994). “The Effects of Mergers in differentiated Products Industries: Logit Demand and Merger Policy.” *Journal of Law, Economics, & Organization* 10, 407–26.