The demand for cars: a discrete choice model with income effects*

Jordi Jaumandreu[†] and **María J. Moral**[‡]

Universidad Carlos III de Madrid Universidad de Vigo

September 2001

Preliminary

Abstract

This paper obtains sensible estimates of the own and cross-price elasticities of demand for 164 car models by using a simple discrete choice model for market level data that combines varying coefficients with a nested logit estimation. The specification exploits the likely constraints that income imposes on preference-based consumer car choices. Segment-idiosyncratic price coefficients are shown to embody information of varying marginal utility of income according to consumers' wealth, and this information is used to estimate the patterns of substitution.

* We are grateful for the starting stimulus of the anonymous referee of another paper that required us to justify the use of varying price coefficients.

† Dpto. de Economía, c/Madrid 126, 28903 Getafe (Spain). Tel.: 34 91 6245737, Fax: 34 91 6249875.e-mail: jordij@eco.uc3m.es

‡ Dpto. de Economía Aplicada, As Lagoas Marcosende s/n, 36200 Vigo (Spain). Tel.: 34 986 812502, Fax: 34 986 812401. e-mail: mjmoral@uvigo.es

1. Introduction

This paper obtains sensible estimates of the own and cross-price elasticities of demand for 164 car models by using a simple discrete choice model for market level data that combines varying coefficients and a nested logit estimation. The specification exploits the likely constraints that income imposes on preference-based choices. Segment-idiosyncratic price coefficients are shown to embody information of varying marginal utility of income according to consumers' wealth, and this information is used to estimate the patterns of substitution.

Discrete-choice models have become a popular alternative to model and estimate the demand of markets with product differentiation (see Nevo 2000, for a list of recent works and discussion). These models, especially under the random-coefficients methodology promoted by Berry (1994), and Berry, Levinsohn and Pakes (1995) –henceforth BLP-, show a number of advantages. In particular, they allow for a reduction of the dimensionality of the problem of estimating demand parameters with many products, give ways to treat the endogeneity of prices, and permit the modelling of consumer heterogeneity.

However, income effects, one of the most important sources of consumer heterogeneity in many markets (e.g. automobiles), have only been marginally treated until now. The traditional quasilinear specification of utility has generally simply ignored these effects. Alternatively, the full random coefficient models obtain their achievements by focussing on completely unspecified, and consequently very general, forms of heterogeneity. Here, building from the ideas of these heterogeneous-consumers discrete choice models, we show that focussing on the likely effects of income heterogeneity uncovers relevant information on consumer behaviour and may be a useful way of obtaining reliable estimates.

We start by considering the constraints that income imposes on the preference-based choice through varying marginal utility of income.[1] Any consumer has a different probability of choosing a specific product according to their income, and consumers with different income levels are likely to cluster around specific products or product classes. Accordingly, product market shares are expected to be associated with different income levels, and hence to

[1] The setting may be seen as a variant of the "vertical product differentiation model", see e.g. Shaked and Sutton (1983). Bresnahan (1987) applied it to automobile, in a version of consumers heterogeneous in tastes.

idiosyncratic price coefficients which reveal different average marginal utilities of income. In particular, price coefficients of the high priced alternatives are expected to be lower. As a consequence, product elasticities are linked, in addition to consumer taste variety, to consumer clustering. Consumers of high (low) priced alternatives substitute them with other high (low) priced alternatives with a higher probability. The estimation of product or class-idiosyncratic price coefficients allows us to construct estimates of these elasticities. The empirical exercise uses the monthly sales of 164 models in the Spanish automobile market from 1990 to 1996, together with information on their price and attributes.

As a matter of fact, income effects are important in many markets, and empirical studies reveal their role. BLP and Petrin (1999), modelling the automobile market, use utility specifications log-linear in income together with data on the actual income distribution, which imply an important implicit impact of this variable on the estimated elasticities. Goldberg (1995) systematically finds lower price coefficients the higher the automobile segments. And even Nevo (2001) finds evidence on the price-coefficient impact of lower marginal utility of income of above-average income consumers in the a-priori more income-effect-free ready-to-eat cereal consumption.

The estimation of reliable own and cross-price elasticities is crucial for a series of exercises. In particular, the assessment of market power and welfare gains and loses relies heavily on these estimations. In addition, product location decisions are likely to be closely related to these elasticities and hence any exercise on firms' product strategies needs these estimations. The specification that we use in this paper gives explicit expressions which can be used in these exercises when income effects are likely to play a role, and can help to solve at a low cost the estimation of sensible elasticities.

The rest of this paper is organised as follows. Section 2 develops a simple deterministic individual framework of discrete-choice under variable marginal utility of income and then sets its stochastic counterpart. Section 3 examines the consequences of this framework of choice for the specification of market product share equations and the interpretation and measurement of elasticities. Section 4 explains the econometric estimation and results of a demand for cars that implements the previous insights. Section 5 concludes.

2. Demand with varying income marginal utility.

2.1 Consumer choice

Let $U(m,x_s)$ be the (contingent) utility derived from the consumption of a composite good m and the variant s of a unit demand good characterised by the vector of attributes x_s. Let y be income, p_s the price of good s, and normalise the price of good m to unity. Unconditional indirect utility can be written as $V(y-p_j,x_j)=\underset{s}{Max}\tilde{V}(y-p_s,x_s)$, and Roy's identity holds as $\delta_j=-\frac{\partial V}{\partial p_j}\Big/\frac{\partial V}{\partial y}=1$ for the chosen alternative and $d_s=0$ for $s\neq j$ (McFadden, 1981.)

Many empirical analyses specify V as

$$V(y-p_j,x_j)=\alpha(y-p_j)+u(x_j) \tag{1}$$

where a stands for constant marginal utility of income, and $u(.)$ is usually taken linear in attributes, i.e., $u(x_j)=x_j\beta$. This specification has the advantage of simplicity and it can be employed to derive straightforwardly an estimable system of demands for differentiated products (Berry, 1994.) It presents, however, the important drawback that it can hardly explain how consumer' income influences decisions, hindering the analysis of a basic dimension of consumer' heterogeneity.

Let us specify utility keeping separability but allowing for a varying marginal utility of income

$$V(y-p_j,x_j)=h(y-p_j)+u(x_j) \tag{2}$$

where $h(.)$ is a function with $h(0)=0$, which we will assume to be only monotonically increasing and concave. This specification encompasses expression (1) as a particular case, but shows decreasing marginal utility if $h(.)$ is strictly concave. In order to compare utilities among them, let us write (2) as

$$V(y-p_j,x_j)=h(y)-h'(y-\lambda p_j)p_j+u(x_j)=h(y)-\alpha(y,p_j)p_j+u(x_j) \tag{3}$$

where $\lambda \in (0,1)$ and depends on the values of y and p_j. The first equality is obtained applying the mean value theorem. The second establishes that marginal utility of income α is now a function of y and p_j, with $\frac{\partial \alpha}{\partial y} < 0$ and $\frac{\partial \alpha}{\partial p_j} > 0$. Expression (3) gives consumer utility derived from any consumption in terms, among other things, of the associated marginal utility of income.

For a given consumer $V(y - p_j, x_j) > V(y - p_k, x_k)$, $\forall k \neq j$, if and only if $-\alpha(y, p_j) p_j + u(x_j) > -\alpha(y, p_k) p_k + u(x_k)$ (we will call these last terms utility contributions). As in the standard specification, the consumer can strongly prefer the attributes of k (a high quality version), but be deterred from its consumption by the counterbalancing weight of its higher price. But now this effect may crucially change according to the income that characterises a particular consumer, even if consumers share the same attributes valuation. Some consumers with higher incomes (lower α's) will choose, at the same prices, more expensive varieties.

Let us now represent in Figure 1 consumer' choices as a function of income, drawing on Bresnahan's (1987) discussion of heterogeneous tastes. To simplify the representation (and only with this purpose), assume a strictly concave function with $h''' = 0$, which implies that $\frac{\partial \alpha}{\partial y} < 0$ is a constant and $\frac{\partial^2 \alpha}{\partial y^2} = 0$. Utility contributions $V - h(y) \equiv u_s$ can be represented in the plane (y, u) by straight lines as:

$$u_s = u(x_s) - \alpha(y, p_s) p_s \tag{4}$$

with slope increasing in price ($\frac{\partial u_s}{\partial y} = -p_s \frac{\partial \alpha}{\partial y} > 0$). These lines start from the fourth quadrant[2] and cross the y-axis at the points $\bar{y}_s$ that verify $\alpha(\bar{y}_s, p_s) = \frac{u(x_s)}{p_s}$. Let us take two variants such as $u(x_k) > u(x_j)$ (i.e., the good k is superior in the attributes' valuation), and

[2] A consumer will only consider buying the good s at price p_s if he has enough income. If $y = p_s$, he gets utility $u_s = u(x_s) - h(p_s)$, presumably a negative value because of the high marginal utility of income in the absence of other consumption.

$p_k > p_j$ in a way that implies $\frac{u(x_k)}{p_k} \leq \frac{u(x_j)}{p_j}$ (i.e., the price of quality increases faster than its valuation.) Then $\alpha(\bar{y}_k, p_k) \leq \alpha(\bar{y}_j, p_j)$ and, as $\frac{\partial \alpha}{\partial p_j} > 0$, this implies $\bar{y}_k > \bar{y}_j$.

The figure illustrates the way in which the choice among the different varieties depends on consumer' income. The consumer will only decide to buy one of the product varieties when his income reaches y^j, even if he had already been able to afford the variety j before ($\bar{y}_j = y^j > p_j$). At the interval (y^j, y^k) he will buy the variety j and, if $y \in (y^k, y^l)$, he will choose the superior variety k. At higher incomes he will switch his consumption to higher price alternatives. Notice that, with the standard specification (1), this dependence of choices on income is completely missed: contributions to utility are given by horizontal straight lines.

2.2. Industry demand

Consider now an industry with differentiated products which are demanded by a population of consumers that, for simplicity, we suppose are endowed with the same *V*, but with different incomes distributed in $(0, y_{\max}]$. Varieties and their prices are given as the result of previous producers' decisions. Products can be ranked from 1 to n such as $u(x_k)/u(x_j) \leq p_k/p_j$ for any two varieties if $j<k$. It is clear that consumers with different incomes will cluster buying their most satisfactory alternative given their income. For example, consumers with income $y \in (y^j, y^k)$ will buy variety j. Therefore, the simple introduction of varying marginal utility of income in the basic model has been enough to determine an association in consumption between quality varieties and consumers' income[3].

Let us now enlarge the set of products in a direction inspired on the real structure of many markets. We allow for the existence of varieties with the same global valuation and

[3] Moreover, products present two (perfect information) boundary characteristics. Firstly, a good priced the same as another and characterised by an inferior valuation will not be demanded (its utility contribution is a downward parallel displacement of another). Secondly, a good with a valuation identical to another but with a higher price will not be sold (lines will not cross within the relevant range).

identical price (i.e., their lines are superposed), and we assume that they will share the demand of the corresponding income level. Thus, consumers with incomes belonging to the interval (y^j, y^k) will cluster buying alternative j or any one of the equivalent varieties (class-g varieties.) This agrees very well with the currently assumed consumer taste for variety. In fact, industries consisting of several income-related classes of varieties (market segments), in which several varieties compete among them more intensely than with the varieties belonging to other classes, are frequent. In fact, vertical and horizontal differentiation coexist with each other, and tastes rule the choice among the similar varieties while income has a role when comparison involves vertically differenced products. The automobile industry, with products grouped in the standard model categories of small, compact, intermediate and luxury cars, is a good example.

2.3 Stochastic set-up

Assume now that an individual's behaviour cannot be completely predicted by any of the usually alleged reasons (see Anderson *et al.*, 1992.) Then some random component must be included in the utility contributions. The simplest specification, which appends to utility an additive random term with zero mean assumed i.i.d. across varieties, is

$$u(x_s) - \alpha(y, p_s)p_s + \varepsilon_s \equiv u_s + \varepsilon_s \tag{5}$$

If the $\varepsilon's$ are distributed as a type I extreme value random variable, the probability of choosing good j for a consumer with income y is now given by the logit formula[4]

$$P(j \mid y) = \frac{e^{u_j}}{1 + \sum_s e^{u_s}} \tag{6}$$

The lines of Figure 1 and the critical income values y^s now only hold as relations for the average utility contribution (i.e., at the zero mean of the random terms.) But they allow us to easily follow the evolution of probabilities of buying different goods along incomes by looking at the relative values of the utility contributions u_s and thinking of them as arguments of

[4] The alternative of "buying nothing" is indexed by 0 and given null utility contribution.

formulas (6). For example, variety j presents the highest probability of being the one chosen by a consumer with an income that belongs to the interval (y^j, y^k), followed by the probabilities of acquiring k and then l. In particular it is clear that the higher the income, varieties with higher prices become the most probable choice.

This model is especially suited to accommodate the existence of the (income level related) classes of varieties. If varieties belonging to a class are identical in price and attributes' valuation, they will exhibit identical income conditional probabilities of choice (they will show the same average utility contribution). But now varieties can even be only "approximately" equivalent, showing minor differences in the total attributes valuation, and they will present similar conditional probabilities of choice.

As far as the probabilities' evolution is concerned, it is easy to show the following

Lemma. $P(0 | y)$ is continuously decreasing in y and $P(j | y)$ is, for each j, either a continuously increasing function of y or reaches a maximum and then decreases. The alternatives whose probability reaches a maximum are the lowest priced ones, and they reach it at a sequence of y values that reproduce the price ordering.

Proof. See Appendix.

3. Product shares under income effects

3.1. Product shares

Assume a large sample of M consumers, endowed with income-conditional probabilities of buying each alternative of the J product varieties of the industry. Consumer i-th has income y_i and the probability density function $f(y)$ characterises the income distribution among consumers. Given product prices, it turns out that observed market shares will converge in probability to expected market shares that can be approximated by the logit expressions evaluated at sample average utility levels. The average utility level of alternative j

is given by u_j valuated at an α equal to consumers' mean marginal utility of income conditional on choosing alternative j.

To see why, note firstly that the number of buyers of product j is the sum of a large sample of Bernoulli independent variables χ_{ij} with means $E(\chi_{ij}) = P(j \mid y_i)$ and variances $V(\chi_{ij}) = P(j \mid y_i)(1 - P(j \mid y_i))$. The observed share $\hat{P}(j) = \sum_i \chi_{ij} / M$ will converge in probability, by the law of large numbers, to $P(j) = \sum_i E(\chi_{ij})/M = \sum_i P(j/y_i)/M$, which, using $f(y)$, we will write as $P(j) = \sum_y P(j \mid y) f(y)$.

Associated to this share there is an expected income, or mean income, that buyers of j are expected to exhibit, obtainable by weighting each income by its probability. That is, $y_j = \sum_i y_i \left[\frac{P(\chi_{ij} = 1)}{\sum_i P(\chi_{ij} = 1)} \right] = \sum_y \frac{y_i P(j \mid y) f(y)}{\sum_j P(j \mid y) f(y)}$. Applying Bayes's rule we can write $y_j = \sum_y y\, f(y \mid j)$, where we will call $f(y \mid j)$ the probability density function of income conditional on choosing product j. See Figure 2 for an example of the typical pattern expected for conditional densities. Similarly, $P(j)$ also has an associated marginal utility of income, or mean marginal utility, that buyers of j are expected to experience, which we can write, by the same reasoning, as $\alpha_j = \sum_y \alpha(y, p_j) f(y \mid j)$.

It is natural to approximate shares as a function of these values. To see how, firstly note that, given prices, (individual) marginal utilities of income $\alpha(y, p_s) = \alpha_s(y)$ can be taken as variables with probability density functions $g_s(\alpha) = f(\alpha_s^{-1}(\alpha)) \mid \alpha_s^{-1}{}'(\alpha) \mid$ and conditional densities $g(\alpha \mid s)$. Then recall that $P(j) = \sum_y P(j \mid y) f(y)$ implies an average across income densities of all marginal utilities of income. Then replace this expectation by the function $P(j \mid .)$ evaluated at the vector of expected income marginal utilities $\alpha = (\alpha_1, ..., \alpha_J)$. Calling this approximation $P(j; \alpha)$, we have

$$P(j) \approx P(j;\alpha) = \frac{e^{u_j}}{1+\sum_s e^{u_s}} \tag{7}$$

where $u_s = u(x_s) - \alpha_s p_s$ and $\alpha_s = \sum_y \alpha(y, p_s) f(y \mid s) = \sum_a \alpha g(\alpha \mid s)$.

We had already established that $\hat{P}(j) \xrightarrow{p} P(j) \approx P(j;\alpha)$. Now it is clear what the contents of the α_s specific price coefficients at expected "aggregated" logit formulas are: expected average utilities of income associated with choosing each alternative. Moreover, given prices and product attributes (and normalising u_0), a unique vector of shares P will correspond to each vector α, and there is only one vector α corresponding to a vector of shares P (see Berry, 1994.) Then, it seems natural to interpret observed shares $\hat{P}$ as corresponding to average marginal utilities $\hat{\alpha}$ converging to α, that is $p\lim \hat{P}(j) = p\lim P(j;\hat{\alpha}) = P(j;\alpha)$.

In addition, expected mean marginal utilities will show a very definite pattern with respect to product prices/quality, as the following proposition establishes.

Proposition. With decreasing marginal utility of income, products with higher prices have lower expected average marginal utilities.

Proof. See Appendix.

3.2. Elasticities

Replacing sums by the integral sign, in order to simplify notation, the own-price and cross-price elasticities of the expected product j market share $P(j)$ can be written as

$$\eta_j = -\frac{p_j}{P(j)} \int \tilde{\alpha} P(j \mid y)(1 - P(j \mid y)) f(y) dy \qquad \forall j = 1,\ldots,J \tag{8.a}$$

$$\eta_{jk} = \frac{p_k}{P(j)} \int \tilde{\alpha} P(k \mid y) P(j \mid y) f(y) dy \qquad \forall j, k \neq j \qquad (8.b)$$

where $\tilde{\alpha} = \alpha(y, p_s)\left[1 + \frac{p_s}{\alpha(.)}\frac{\partial \alpha(.)}{\partial p_s}\right] = \alpha(y, p_s)(1+\gamma_s)$. The elasticity γ_s of marginal utility with respect to the product price may be considered small if price is not too big with respect to the relevant income. Moreover, in many cases these elasticities can be sensibly assumed to be independent from income. We will assume this henceforth, without loss of generality, in order to simplify notation.

Expressions (8) present the general form corresponding to logit elasticities under consumer' attribute-related heterogeneity (see, for example, Nevo 2000). But the focus on income heterogeneity allows us to be more specific. These elasticities can be rewritten as

$$\eta_j = -\frac{p_j}{P(j)}\left[\tilde{\alpha}_j P(j)(1-P(j)) + Cov(\tilde{\alpha} P(j \mid y), P(j \mid y))\right] = -\tilde{\alpha}_j p_j \left[1 - P(j)(1+\omega_j)\right] \qquad \forall\ j \qquad (9.a)$$

$$\eta_{jk} = \frac{p_k}{P(j)}\left[\tilde{\alpha}_k P(k) P(j) + Cov(\tilde{\alpha} P(k \mid y), P(j \mid y))\right] = \tilde{\alpha}_k p_k P(k)(1+\omega_{jk}) \qquad \forall j, k \neq j \qquad (9.b)$$

where $\omega_{jk} = \int (\frac{\tilde{\alpha} P(k \mid y)}{\tilde{\alpha}_k P(k)} - 1)(\frac{P(j \mid y)}{P(j)} - 1) f(y) dy$, ω_j is the corresponding expression for j, the ω factors fulfil the constraint $\omega_j = -\sum_s (P(s)/P(j))\omega_{sj}$, and $\tilde{\alpha}_k$ stands for the mean marginal utility of income associated to product k augmented in the price variation effect, that is $\tilde{\alpha}_k = (\int \alpha f(y \mid k) dy)(1+\gamma_k) = \alpha_k (1+\gamma_k)$.

Elasticities in (9) show two important changes with respect to the standard logit elasticities. Firstly, marginal utility parameters α are now product specific. This reflects, as we have shown above, the different average marginal utilities associated with each product. Secondly, cross-price elasticities with respect to the price of product k vary across the substitutes j, in contrast with the uniform effects which characterise the standard logit model. This variation reflects that the model is now distinguishing among products "closer" and "farther" to product k in a very precise meaning, cast in the covariance term. This term shows that a price increase of product k will induce a high substitution by the products that exhibit a high probability of being chosen at the same income levels, and a low substitution by the products that are the typical choices at other income levels.

The differences between the formulas (9) and the standard ones is easy to interpret. Let us first examine the conventional logit price effects, given by the first term of the sum in the first part of the equalities 9.a and 9.b. In 9.a we find the share reduction provoked by a price increase of product j, $-\tilde{\alpha}_j P(j)(1-P(j)) = -\tilde{\alpha}_j P(j)\left[P(0)+\sum_s P(s)\right]$, which in case of a price increase of product k will be $-\tilde{\alpha}_k P(k)\left[P(0)+\sum_j P(j)\right]$. In 9.b we find the corresponding increment of all the other market shares $\tilde{\alpha}_k P(k)P(0)+\sum_j \tilde{\alpha}_k P(k)P(j)$, which turns out to be proportional to their relative importance.

It is simply natural to expect and specify such a proportional effect in an individual that we are only able to characterise up to a set of probabilities of buying the alternative products. If one probability is going to be reduced, we expect the probabilities of all the other alternatives to benefit proportionally from this reduction. But, when we look at the distribution of buyers, we must take into account that this simple change modifies the relative individual probabilities of all buyers. If buyers cluster around alternatives, in the sense that some consumers have higher probabilities of buying some alternatives than others, individual changes will induce expected aggregated changes according to the patterns of substitution given by the clustering. This is what is reflected by the covariance term, raising what can be called "aggregated" elasticities (Ben-Akiva and Lerman, 1985).

3.3. Nested probabilities

The present formulation can be combined with the specification of nested logit probabilities. In this case, the nesting must be intended to reflect the structure of the individual preferences and hence probabilities of choice given income. That is, as products to be chosen by any consumer present among them different degrees of "similarity" and "dissimilarity" in attributes, the model avoids the IIA property by grouping together similar alternatives. Then the rest of the model applies.

One particular version of this combination is obtained by assuming that income does not influence the choice among alternatives inside a nest. That is, all products grouped at a nest present conditional probability choices that evolve with income at the same pace as the probability of choosing the group (segment). This corresponds to the simple graphical example of sets of products represented for the same line. It can be regarded as a simplified situation in which income and tastes influence global choice, but only tastes determine the choice inside the segment. This situation is likely to emerge when producers choose to locate products slightly differentiated in nests characterised by the same price.

Formally, let $g = 1,...,G$ index the groups of products and let J_g represent the set of products in group g. We will use the notation $P_g(k \mid y) = P(k \mid y) / P(J_g \mid y)$ for the intra-group share of product k, where $P(J_g \mid y) = \sum_{s \in J_g} P(s \mid y)$. According to the above suggestion assume that $P_g(k \mid y) = P_g(k)$, which is equivalent to $f(y \mid k) = f(y \mid J_g)$. Then, it is easy to check that nested elasticities with consumer income heterogeneity are given by

$$\eta_j = -\tilde{\alpha}_j p_j \left[1 - P(j)(1 + \omega_j) + \frac{\sigma}{1-\sigma}(1 - P_g(j)) \right] \qquad \forall j \qquad (10.a)$$

$$\eta_{jk} = \tilde{\alpha}_k p_k \left[P(k)(1 + \omega_{jk}) + \frac{\sigma}{1-\sigma} P_g(k) \right] \qquad \forall j, k \neq j \quad \text{if} \quad j, k \in J_g \qquad (10.b)$$

or $$\eta_{jk} = \tilde{\alpha}_k p_k P(k)(1 + \omega_{jk}) \qquad \forall j, k \neq j \quad \text{otherwise} \qquad (10.c)$$

3.4. Measuring elasticities

Formulas (9) or (10) have two advantages. Firstly, they provide explicit expressions for the relevant elasticities (which apply under a particular form of consumer heterogeneity). This is important, for example, at the time of specifying price equations (see Moral and Jaumandreu, 2000). Secondly, elasticities can be estimated (up to the factors γ) given estimates of the α parameters. Covariance terms can be approximated in different ways (see Appendix).

4. Data, econometric estimation and results

Over the seven years 1990-96, we observe the monthly registrations (sales) in the Spanish market of a total of 182 car models. These unbalanced panel data have been elaborated and matched to a database on prices and technical and physical characteristics of the models.[5]

Applying the logit linear transformation to equations like (7), and allowing explicitly for a time subscript, our estimating equation can be written as

$$\ln \hat{P}_{jt} - \ln \hat{P}_{0t} = x_{jt}\beta^{*} - \alpha_{g}^{*} p_{jt} + \eta_{g}^{*} + \xi_{j} + \xi_{jt} \qquad \forall j,t \quad with \quad j \in J_{g} \quad (11)$$

where we apply the nested type specification, with segment-idiosyncratic coefficients (the asterisk indicates that the coefficients must be understood to be divided by the factor $(1-\sigma)$). This specification can be seen as splitting unobserved utility effects in three components: η_g, a segment effect that we will take as time invariant and in which we will condition the estimation (see below); ξ_j, a time invariant component related to the stable unobservable characteristics of model j; and ξ_{jt}, the remaining time varying zero-mean unobserved effects on mean utility of model j. We will estimate the η_g's by means of segment dummies, interpreting their values as the realizations of the random variables conjugated to the extreme value errors that raise nested probabilities (Cardell, 1997).

Estimation of (11), using the constraint $\sum_{g}(\eta_{g} - \eta) = 0$ to specify all the segment effects (η represents the average of these effects), gives coefficient estimates up to the scale factor $(1-\sigma)$ and mixes two unidentifiable components in the regression constant. Then, to

[5] There is an average of 110 models marketed per month and an average of 50 monthly observations per model. Details on the database can be found in Moral and Jaumandreu (2001).

estimate the σ parameter, we construct estimates of the "inclusive values" $D_g = \sum_{j \in J_g} e^{\frac{u_j}{(1-\sigma)}}$ up to a multiplicative constant and perform the regression[6]

$$\ln \hat{P}(g) - \ln \hat{P}(0) = c + (1-\sigma) \ln \hat{D}_g^* \tag{12}$$

The dependent variable of (11) consists of the (log of the) model monthly share observations minus the (log of the) monthly shares of the outside good.[7] Both shares are computed taking the current number of households as the market size. To control for seasonality and unspecified time effects (for example, a fall in demand happened at mid-period), we include a set of monthly dummies and another of yearly time dummies, respectively. In addition, we control for age-of-the-model effects (the time the model has been marketed) by including a polynomial of order three with age measured in months and a set of dummies interacted with price (marginal utility effects). Moral and Jaumandreu (2001) develop in detail the meaning and results of the age effects.

The specification employs almost the same attributes as BLP, replacing the air conditioning "luxury" proxy for the *maximum speed in km/h (Maxspeed*), probably a better proxy for quality in Spain during the nineties. The other employed attributes are the power measure *ratio cubic centimetres to weight* (*CC/Weight*), the fuel efficiency *ratio km to litre* (*km/l*), and the measure of size and safety *length times width* (*Size*). The use of other characteristics or a more complete list does not change the main results.

We group models into 6 categories that closely resemble common industry and marketing classifications. The main classes of cars considered are: small-mini, small-domestic, compact, intermediate, and luxury. We separately group minivans, which were at that moment a product beginning their market penetration. The consideration of two different subgroups in the small cars, one constituted mainly for the very popular non-mini domestic brand cars, proved to be important for the nesting in practice. We estimate, however, only 5 different

[6] To avoid simultaneity biases we construct the "inclusive values" with the price values predicted using the instruments.

[7] Montly shares are previously multiplied by 12 in order to facilitate comparability with elasticities obtained with yearly data.

segment price effects: small, compact, intermediate, luxury and minivan. The number of models in each segment are, respectively, 33, 37, 56, 47 and 9.

Let us detail our identification framework. Prices are likely to be correlated with the x_j and x_{jt} components of the disturbance (the impact of model unobserved characteristics and the shocks.) In the first case this happens because there are presumably many unobserved characteristics that enter the determination of the models' marginal cost, and hence their prices, which simultaneously influence consumers' utility. In the second case, it occurs because prices are determined at the same time as consumers' demand, and both variables are likely to be influenced by common market shocks.[8] Accordingly we will use as instruments, in a GMM framework, the differences of the prices with respect to their individual time means, $\tilde{p}_{jt} = p_{jt} - (1/T)\sum_s p_{js}$, lagged a number of periods. This instruments prices with their past time variations, avoiding the use of their level variations across models. In addition, to test the validity of the employed instruments, we employ the Sargan test statistic of the overidentifying restrictions.[9]

Several instrument sets were tested, mainly using price differences with respect to the individual time means with different lags, obtaining very similar results.[10] The reported estimate uses as instruments the sixth and twelfth lags of the (segment) price variables in differences (see details in Table 1). The number of overidentifying restrictions is 25, although very similar results are obtained with a smaller number of instruments.

The reported coefficient estimates are one step GMM estimates, obtained by employing the standard weighting matrix (the inverse of $E(Z_j^{'}\bar{x}_j\bar{x}_j^{'}Z_j)$, where $\bar{x}_j = (x_j + x_{j1},...,x_j + x_{jT})'$ and Z_j represents the set of instruments for individual *j*,

[8] The most standard way of treating such a setting is the estimation of the equation taking first differences in order to difference out the individual correlated component, and the use of lags of the endogenous variable to set valid moment restrictions (see, for example, Arellano and Honore (2000)). In our case, this is an undesirable alternative because T is short in relation to the pace of variation of attributes (many attributes change very little or not at all in the seven years). The differentiation of the attributes would eliminate crucial information contained in the levels equation and would exacerbate the variance of the disturbances.

[9] A more detailed justification of the estimation procedure may be found in Moral and Jaumandreu (2001).

estimated by $(\sum_j Z_j^{'} Z_j)^{-1}$). All the statistics are then computed using the robust to heteroskedasticity and serial autocorrelation "two-step" weighting matrix.[11] The reported Sargan test is also a two-step statistic. To estimate a robust inverse of $E(Z_j^{'} \bar{x}_j \bar{x}_j^{'} Z_j)$, we assume that $\bar{x}_j \bar{x}_j^{'} = \Omega_j$ are matrices corresponding to conditional homoskedastic errors, and we obtain $\hat{\Omega}_j$ values using the Newey-West Bartlett kernel computations for the autocovariances of individual j. Then we employ the usual "two-step" estimate $(\sum_j Z_j^{'} \hat{\Omega}_j Z_j)^{-1}$. We use 72 time observations as the maximum lag that we take into account in the Bartlett kernel.

Table 1 presents the results of our preferred estimation. The statistics and estimated coefficients look sensible. The Sargan test confirms the validity of the employed instruments. Regression residuals show a strong autocorrelation, but the use of a robust covariance matrix makes the inferences reliable. Control variables present reasonable patterns. Attributes are significant and show the expected sign. And the price effects clearly exhibit the expected pattern: the higher the segment, the lower average marginal utility is. All own-price implicit estimated elasticities are, as expected from theory, higher than one in absolute value.

Let us focus on the estimated elasticities. Estimated elasticities are no longer exclusively ruled by the product prices and shares, and they reveal patterns of substitution which deserve detailed comment. Table 2 reports a sample of these elasticities, which includes three models for each segment. Firstly, own-price elasticities of intermediate and luxury cars, mostly preferred at higher income levels, are clearly lower than the elasticities shown by the small and compact cars. However, cars at the bottom of the attributes/price scale (small-mini) also tend to show slightly lower own-price elasticities than the average of small-compact classes of cars.

---

[10] We also experimented with sums of characteristics across own-firm products and rival firm products, in their totality and by segments. In general they revealed to be poorer instruments than the lagged price differences and tended to produce worse values for the Sargan statistics.

[11] The use of a robust to serial autocorrelation weighting matrix is especially important because monthly shares data tend to induce strongly autocorrelated disturbances.

Secondly, intra-segment cross-price elasticities tend to be lower the higher the segment, while the ratios of the cross-price elasticities shown by one model to its own-price elasticity always remain roughly within the same range. This simply means that cross intra-segment elasticities tend to follow the same income-related pattern as the own-price elasticities.

Thirdly, cross-segment cross-price elasticities are significantly lower, presenting a fall that is highly influenced by a high estimated level of similarity among models inside the nests (0.84, which notwithstanding turns out to be a low value when compared with the s obtained in standard nested estimations). Cross-segment cross-prices, however, show a very definite pattern. Price changes of small-mini and luxury cars, the two extremes of the scale, have very small impacts on the demands for cars of all the other segments. Price changes of small-domestic and compact cars, however, have higher effects more or less disseminated among all the other classes. And the price of intermediate cars has mainly significant impacts on the demand for luxury cars. Reading the other way, the smallest cars are relatively good substitutes for other small and compact cars, and luxury cars especially for the intermediate cars, while price changes of the smallest and luxury cars promote substitution relatively more intensely inside the segment.

5. Conclusion

This paper has obtained sensible own-price and cross-price elasticity estimates for the high number of car models that constitute a market (the Spanish market during a six-year period) by using a simple discrete choice model, applicable to market level data, which combines segment-idiosyncratic coefficients with a nested logit estimation.

Our starting theoretical development has shown the effects that consumer income heterogeneity is likely to have on consumer choices, product market shares and demand elasticities, and hence the consequences for the specification estimation of market equations based on these shares. Estimation has shown that segment coefficients embody consistent information on marginal utilities of income conditional on choosing the cars that belong to the segment. The use of this information gives reasonable patterns of substitution: luxury, and to some extent the bottom models of the small cars category, show relatively less elastic demands; intra-segment substitution is very intense and also tends to decrease with the car range; the smallest and luxury cars are relatively good substitutes for the intermediate categories but their price changes have a very limited impact on other segments.

A market structure consisting of consumers clustering at classes of cars according to income levels is confirmed. The recognition (when relevant) of this structure allows for the derivation of useful market product elasticity-explicit formulas and easy ways to estimate them. Future work must improve the estimation of elasticities by using more complete specifications of the varying price coefficients, but we intend this approach to be especially suited for the simultaneous specification and estimation of price determination or product location equations, in which jointly determined market elasticities are likely to play a crucial role.

Appendix

Proof of the Lemma

Note that the e^{u_s} are increasing in y. Hence $P(0\mid y)=1/(1+\sum_s e^{u_s})$ is continuously decreasing in y. On the other hand, each $P(j\mid y)$ will increase if $\frac{\partial P(j\mid y)}{\partial y}=-\frac{\partial \alpha(y,p_j)}{\partial y}P(j\mid y)\left[p_j-\sum_s \theta_{sj}P(s\mid y)p_s\right]>0$, where the terms θ_{sj} have the form $\theta_{sj}=\frac{\partial \alpha(y,p_s)}{\partial y}\Big/\frac{\partial \alpha(y,p_j)}{\partial y}>0$ and fulfil $\theta_{sj}>\theta_{sk}$ *if* $p_k>p_j$. What we need is the positiveness of the term between brackets. With J varieties ordered according to prices (from the lowest to the highest), the positiveness of conditions from 1 to J form a system of inequalities in which the constraints will be violated in turn as y grows (the constraints are easier to be fulfilled the bigger the price). Then the corresponding probabilities will begin to decrease.

Proof of the Proposition.

We want to show that, if $p_k>p_j$, then $\alpha_k<\alpha_j$. We have $\alpha_j-\alpha_k=\int \alpha(y,p_j)f(y\mid j)dy-\int \alpha(y,p_k)f(k\mid y)dy=\int(\alpha(y,p_j)-\alpha(y,p_k))f(y\mid j)-$
$-\int \alpha(y,p_k)(f(y\mid j)-f(y\mid k))=-\int \frac{\partial \alpha}{\partial p_j}(p_k-p_j)f(y\mid j)-\int \frac{\partial \alpha}{\partial y}(F(y\mid j)-F(y\mid k))$
where the first term of the last equality constitutes a first order approximation and the second is obtained by integration. This expression can be approximated by the integral $-\int \frac{\partial \alpha}{\partial y}(F(y-\Delta p\mid y)-F(y\mid k))$. A sufficient condition for this integral to be positive is the stochastic dominance of the distribution $F(y-\Delta p\mid y)$ over the distribution $F(y\mid k)$. As we are going to see, this is what we must expect when $p_k>p_j$ and prices are not too big with respect to income.

Given the previous Lemma, it is easy to check that the ratio of probabilities of alternative k to j is increasing, that is $\frac{\partial (P(k \mid y)/P(j \mid y))}{\partial y} > 0$. This implies

$$\int \frac{P(k \mid y)}{P(j \mid y)} f(y \mid j) dy < \int_{y}^{y^M} \frac{P(k \mid y)}{P(j \mid y)} \frac{P(j \mid y) f(y)}{\int_{y}^{y^M} P(j \mid y) f(y)} dy \quad \text{or}$$

$$\frac{P(k)}{P(j)} = \frac{\int P(k \mid y) f(y) dy}{\int P(j \mid y) f(y) dy} > \frac{\int_0^y P(k \mid y) f(y) dy}{\int_0^y P(j \mid y) f(y) dy}.$$ Hence, $\int_0^y f(y \mid j) dy > \int_0^y f(y \mid k) dy$ or $F(y \mid j) > F(y \mid k)$. If prices are not too big with respect to income, this will imply the required dominance.

Approximating the covariance terms

Assume that the g factors are negligible, and that the terms $\boldsymbol{a}(y, p_s)$ can be sufficiently approximated at $\boldsymbol{a}(y,0) = \boldsymbol{a}(y)$. Then conditional probabilities may be seen as a function of only the variable y or, more conveniently, $\boldsymbol{a}$. Define $\boldsymbol{a}_k^*$ as the unique $\boldsymbol{a}$ value for which $P(k \mid \boldsymbol{a}_k^*) = P(k)$ holds. The differences under the integral sign may then be approximated as functions of $(\boldsymbol{a} - \boldsymbol{a}_k^*)$. Thus we have

$$\begin{aligned} Cov(\boldsymbol{a}P(k \mid y), P(j \mid y)) &\approx -\int \frac{\partial \boldsymbol{a}P(k \mid y)}{\partial \boldsymbol{a}} \frac{\partial P(j \mid y)}{\partial \boldsymbol{a}} (\boldsymbol{a} - \boldsymbol{a}_k^*)(\boldsymbol{a} - \boldsymbol{a}_j^*) g(\boldsymbol{a}) d\boldsymbol{a} = \\ &= P(k)P(y)(1 - \boldsymbol{a}_k p_k^*) p_j^* \left[V(\boldsymbol{a}) + (\boldsymbol{a} - \boldsymbol{a}_k^*)(\boldsymbol{a} - \boldsymbol{a}_j^*) \right] \end{aligned}$$

where $p_j^* = p_j - \sum_s P(s) p_s$. The values $\boldsymbol{a}_s^*$ are unknown, but using the equality $P(j;\boldsymbol{a}) = P(j \mid \boldsymbol{a}_s^*)$, they can be approximated as $\boldsymbol{a}_j^* = \boldsymbol{a}_j \frac{p_s - \sum P(s) p_s \frac{\boldsymbol{a}_s}{\boldsymbol{a}_j}}{p_s - \sum P(s) p_s}$. However, in practice, there can be almost no change from using $\boldsymbol{a}_j$ or $\boldsymbol{a}_j^*$. In any case, $\boldsymbol{w}_{jk} \approx (1/\boldsymbol{a}_k - p_k^*) p_j^* \boldsymbol{j}(\boldsymbol{a})$, which is an expression that can be evaluated using estimates of $\boldsymbol{a}$.

References

Arellano, M. and B. Honorè (2000), *Panel Data Models: Some Recent Developments*, forthcoming in Handbook of Econometrics.

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

Ben-Akiva, M. and S.R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press.

Berry, S.T. (1994), "Estimating discrete-choice models of product differentiation," *RAND Journal of Economics*, 25 (2), 242-262.

Berry, S.T., Levinsohn, J., and A. Pakes (1995), "Automobile prices in market equilibrium," *Econometrica,* 63 (4), 841-890.

Bresnahan, T.F. (1987), "Competition and collusion in the American automobile industry: The 1955 price war," *Journal of Industrial Economics*, 35 (4), 457-82.

Cardell, N.S. (1997), "Variance components structures for the extreme-value and logistic distributions with applications to models of heterogeneity," *Econometric Theory*, 13 (2), 185-213.

Goldberg, P.K. (1995), "Product differentiation and oligopoly in international markets: the case of the U.S. automobile industry," *Econometrica*, 63(4), 891-951.

McFadden, D. (1981), "Econometric models of probabilistic choice," in *Structural Analysis of Discrete Data with Econometric Applications*, C. Manski and D. McFadden Eds., MIT Press.

Moral, M.J., and Jaumandreu, J. (2000), "Optimal multiproduct prices. An application in the automobile market," mimeo.

Moral, M.J., and Jaumandreu, J. (2001), "Automobile demand, model cycle and age effects," mimeo.

Nevo, A. (2000), "A practitioner's guide to estimation of random-coefficients logit models of demand," *Journal of Economics and Management Strategy*, 9 (4), 513-548.

Nevo, A. (2001), "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69(2), 307-342.

Petrin, A. (1999), "Quantifying the benefits of new products: The case of the minivan," University of Chicago, mimeo.

Shaked, A., and J.Sutton (1982), "Natural oligopolies," *Econometrica*, 51(1), 1469-1483.

Table 1
Results from model estimation

Dependent variable: $\ln \hat{P}_j - \ln \hat{P}_0$
Sample period[1]: I-1991 to XII-1996
Observations[1]: 7,122
Nº of models[1]: 164
Estimation method: GMM[2]

Variable	Coefficient	t-ratio[3]
Constant	-15.840	-6.70
Attributes:		
CC/Weight	1.332	2.46
Maxspeed	0.034	2.92
Km/l	0.071	1.61
Size	0.651	3.42
Prices:		
Small	-4.916	-2.67
Compact	-3.374	-2.65
Intermediate	-0.931	-3.53
Luxury	-0.593	-2.97
Minivan	-2.575	-3.12
Segment effects[4]*:*		
Small-domestic	5.152	3.49
Intermediate	-2.831	-1.97
Luxury	-4.969	-3.57
Seasonal effects	included	
Time dummies	included	
Age polynomial	included	
Age-price interactions	included	
s estimate	0.842	7.51
Sargan test[5] (25 degrees of freedom)	35.86	

Notes:
1. Instruments lagged 12 months imply that models with 12 and fewer observations must be removed.
2. Instruments: differences of segment-prices with respect to their time mean lagged 6 and 12 months, 20 age dummies (years) and interactions of the age dummies with the price differences lagged 12 months.
3. Standard errors are robust to heteroskedasticity and serial correlation.
4. Small-mini, compact and minivan coefficients constrained to be equal to the average effect.
5. Two-step statistic.

Table 2
A sample of own and cross-price elasticities[1]
(x100 cross-price cross-segment elasticities[2])

	Small mini			Small			Compact			Intermediate			Luxury		
	Fiat Uno	Seat Marbella	Rover 114	Ford Fiesta	Seat Ibiza	Peugeot 205	Ford Escort	Opel Astra	VW Golf	Citroen Xantia	Ford Mondeo	Opel Vectra	BMW 525	Mercedes 300	Volvo 850
Fiat Uno	**-4.033**	**1.185**	**0.169**	0.421	0.390	0.203	0.279	0.420	0.278	0.019	0.024	0.021	0.007	0.017	0.014
Seat Marbella	**1.396**	**-3.235**	**0.144**	0.417	0.399	0.192	0.278	0.394	0.277	0.037	0.042	0.034	0.007	0.017	0.012
Rover 114	**1.215**	**1.085**	**-6.227**	0.410	0.381	0.167	0.282	0.400	0.266	0.014	0.017	0.038	0.007	0.016	0.013
Ford Fiesta	0.081	0.060	0.010	**-5.776**	**0.913**	**0.455**	0.280	0.405	0.267	0.010	0.012	0.019	0.007	0.016	0.013
Seat Ibiza	0.081	0.060	0.010	**0.949**	**-5.820**	**0.455**	0.280	0.404	0.268	0.010	0.012	0.019	0.007	0.017	0.013
Peugeot 205	0.081	0.059	0.010	**0.949**	**0.913**	**-5.899**	0.280	0.400	0.269	0.009	0.011	0.021	0.007	0.017	0.013
Ford Escort	0.080	0.058	0.010	0.415	0.399	0.193	**-5.294**	**0.909**	**0.616**	0.207	0.231	0.139	0.009	0.018	0.006
Opel Astra	0.044	0.049	0.007	0.387	0.389	0.140	**0.721**	**-5.764**	**0.586**	0.210	0.234	0.081	0.009	0.016	0.005
VW Golf	0.080	0.059	0.010	0.414	0.399	0.193	**0.740**	**0.909**	**-7.030**	0.254	0.285	0.155	0.010	0.018	0.004
Citroen Xantia	0.017	0.045	0.005	0.417	0.430	0.140	0.297	0.361	0.256	**-2.449**	**0.643**	**0.440**	0.011	0.013	0.000
Ford Mondeo	0.019	0.044	0.005	0.414	0.424	0.139	0.294	0.359	0.255	**0.580**	**-2.410**	**0.438**	0.011	0.014	0.000
Opel Vectra	0.076	0.051	0.009	0.422	0.402	0.189	0.280	0.403	0.267	**0.581**	**0.651**	**-2.477**	0.017	0.023	0.000
BMW 525	0.074	0.046	0.008	0.425	0.403	0.186	0.284	0.426	0.253	1.567	1.748	1.021	**-2.631**	**0.545**	**0.433**
Mercedes 300	0.073	0.044	0.008	0.427	0.404	0.185	0.286	0.440	0.243	2.012	2.239	1.201	**0.321**	**-3.258**	**0.433**
Volvo 850	0.032	0.040	0.006	0.403	0.407	0.134	0.273	0.382	0.248	1.206	1.349	0.301	**0.350**	**0.515**	**-2.797**

1. Cell entries *j*,*k* where *j* indexes row and *k* column, give the percent change in sales (or market share) of model *j* with a one percent change in price of model *k*.
2. Cross- price elasticities between models of different segments are multiplied by 100 (the sample includes an average of 110 models/year).

Figure 1

Utility and choices as a function of consumer' income.

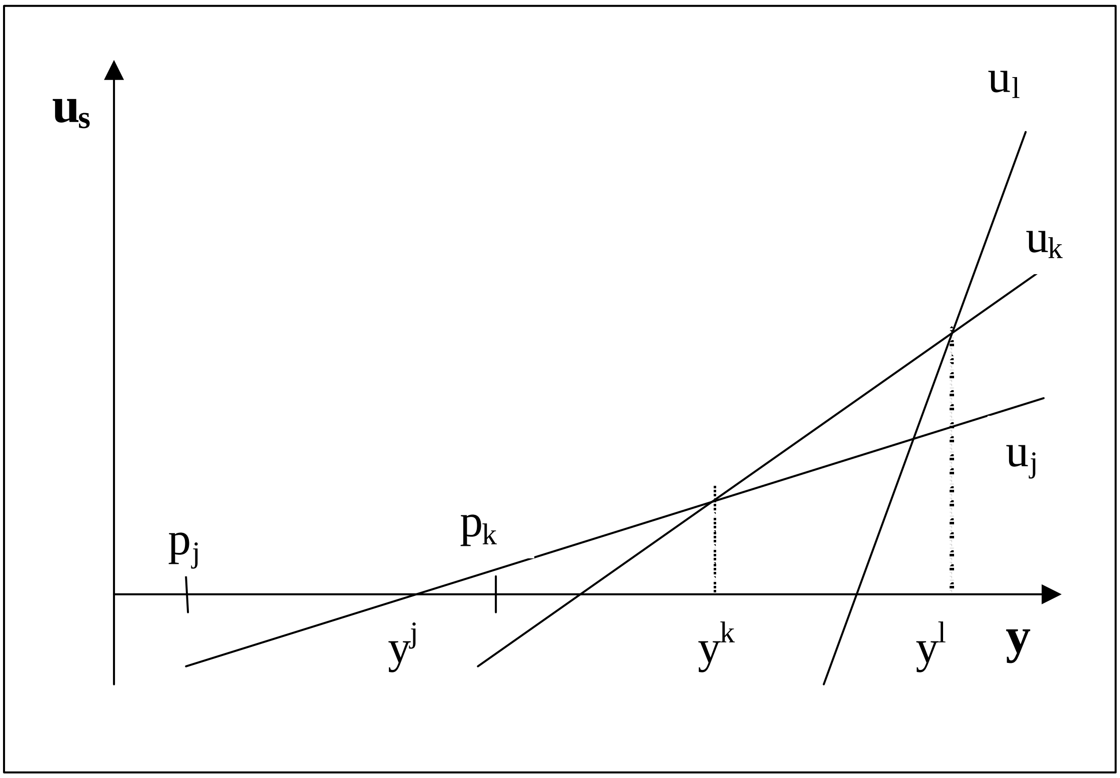
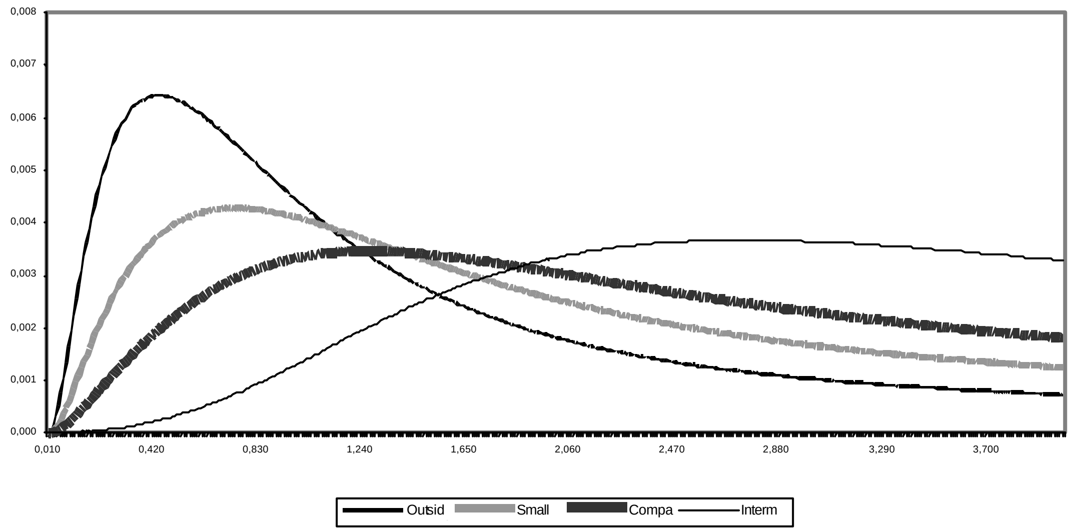


Figure 2: Income densities given consumer car model choices



Note: The figure depicts the $f(y\,|\,s)$ distributions for the outside good and three choice alternatives in a stylised case based on some real data. The function $h(y)$ is simply taken as $1-e^{-x}$, the vector of valuations as $u=(1,1.2,1.3)$, and prices as $p=(1,1.6,2.3)$. Income is assumed to be distributed lognormal with parameters (0,1), which implies E(y)=1.6 and V(y)=4.7. Valuations and prices are proportional to the average maximum speeds and prices of the car models in three segments of the automobile market (small, compact and intermediate).