A discrete choice model for ordered nests

Laura Grigolon*

January 2013

Abstract

I introduce a new discrete choice model of demand, the Ordered Nested Generalized Extreme Value (ONGEV) model. This model allows for segmentation in markets with differentiated products where consumers are likely to substitute to neighboring segments. The model represents a tractable extension of the nested logit model, in which substitution patterns across all segments, neighboring or not, are instead symmetric. I apply the model to the automobile market where segments are ordered from small-size to luxury. The nested logit model is rejected against the ONGEV model. The implied substitution patterns illustrate the presence of relevant neighboring segment effects when consumers substitute outside their segment.

**Laura Grigolon:* University of Leuven. Email: Laura.Grigolon@econ.kuleuven.be. I am grateful to Tobias Klein, Mathias Reynaert, Pasquale Schiraldi, Johannes Van Biesebroeck, Patrick Van Cayseele and Frank Verboven for their useful comments.

1 Introduction

Discrete choice models are frequently used in empirical industrial organization to estimate demand for differentiated products. Because of their strong theoretical properties they can be flexibly adapted to capture various settings. McFadden (1978) has proposed a family of discrete choice models known as the Generalized Extreme Value (GEV) model, which is consistent with random utility theory and yields to a tractable closed-form for choice probabilities. Berry (1994) has provided a framework to estimate, using market-level data, two special members of this family: the logit and the nested logit model. Apart from these two models, only a few other members of the GEV model family have been exploited so far, especially with market-level data. One notable exception is the principle of differentiation model by Bresnahan et al. (1997).

In this paper I present a new member of the GEV model family that captures a particular feature of differentiated product markets. The starting point is the fact that often these markets present a form of segmentation which can be ordered in a natural way. I consider here the car market. It can be naturally ordered from subcompact to luxury according to important product characteristics, such as price, size, engine performance, comfort and prestige. Products belonging to the same segment tend to share similar characteristics and be closer substitutes to each other than products belonging to other segments. At the same time, these characteristics gradually increase from one segment to the other, such that a premium subcompact car can be a potential substitute for a cheap compact car. Thus, segments tend to overlap with their neighbors. A price shock to a compact car may raise the desirability, first, of other compact cars and, next, of subcompact and intermediate cars rather than luxury cars.

Is asymmetric substitution towards neighboring segments captured by the demand models we use? In the nested logit model (Williams, 1977; Daly and Zachary, 1977; McFadden, 1978), often used to estimate the degree of segmentation in a market because of its computational simplicity, neighboring segment effects are ruled out by construction. The model requires the stochastic components of utility attached to the segment choice to be independent. Therefore, while preferences can be correlated across products within the same segment (or nest), substitution outside a segment is symmetric to all other segments.

In contrast, the random coefficients logit model by Berry et al. (1995) has the potential to generate more flexible substitution patterns, where products tend to be closer substitutes as they share similar observed continuous characteristics. For example, in the first essay I simulate the effect of a joint 1% price increase of all cars in a given segment and show that the random coefficients logit model yields to more intense substitution towards neighboring segments. However, the model does not yield to a closed-form for choice probabilities and requires the use of computationally demanding techniques. Problems related to its numerical performance have been studied by several papers; see Knittel and Metaxoglou (2012); Dubé et al. (2012); Judd and Skrainka (2011).

I specify a new application of the GEV class of models denominated Ordered Nested Generalized Extreme Value (ONGEV) model. The ONGEV model explicitly recognizes the natural segment ordering by allowing correlation in unobserved utility between neighboring segments. It is a combination of the ordered generalized extreme value (OGEV) by Small (1987) and the nested logit model. Small's (1987) OGEV model has been developed in settings where a limited number of alternatives have a natural order so that correlation in unobserved utility between two alternatives depends on their proximity in the ordering. The ONGEV model extends the OGEV model to a framework with numerous alternatives where the grouping of these alternatives, rather than the alternatives as such, can be naturally ordered. In the car market, for example, ordering around 170 car models would prove impossible, while ordering grouping of cars, the segments, is a sensible strategy to obtain a tractable model and flexible substitution patterns.

The ONGEV model is appealing for three reasons. First, it provides a modeling theory that is more consistent with the particular structure of choices in some segmented markets, such as cars or hotels, than a simple nested logit model. It creates the potential for neighboring segment effects, or, more precisely, asymmetric substitution patterns across segments. Second, the ONGEV model permits the presence of overlapping nests in a closed form solution. It relaxes the hierarchical nesting structure imposed by the nested logit model while avoiding the burdensome simulation techniques and numerical problems of the random coefficients logit model. As every random utility model, the ONGEV model could be approximated by a random coefficients logit model (McFadden and Train, 2000), but doing so would imply the use of simulation techniques and a consequent reduction in tractability. Third, the ONGEV model has the nested logit and the logit as special cases. It can thus serve as a test for the validity of the constraints imposed by the nested logit and, a fortiori, the logit model.

I implement the ONGEV model using a unique dataset on the car market that covers nine European countries between 1998 and 2009. I model and estimate car demand in three ways: (i) a one-level nested logit model where consumers choose one of five segments and then the specific car; (ii) a two-level nested logit where consumers choose between 'small' and 'large' segments, then the specific segment and car; (iii) an ONGEV model where consumers choose one of five segments, as in model (i), but segments can overlap with the neighboring ones. The first model, the one-level nested logit, is the simplest model to capture market segmentation but imposes restrictions that rule out neighboring segment effects. The second model, the two-level nested logit, is an approximation of the ONGEV model where segments are grouped into two arbitrary nests (small and large), such that correlation between pairs of segment within a nest is not zero. But the segment grouping is arbitrary and can only partially solve the problem of estimating neighboring segment effects because it does not allow correlation between all neighboring segments. The third model is the ONGEV model embodying correlation between all neighboring segments.

The demand estimates of the ONGEV model clearly indicate a rejection of the one-level nested logit model: correlation in car choices is present not only within a segment, but also between neighboring segments. Robustness checks using different instrument sets indicate that the between effect is always present. The demand estimates have striking implications for the substitution patterns. I examine the effect of a 1% price increase of all cars in a given segment on the demand in the various segments. The cross-price elasticities substantially differ across the three specifications. The ONGEV model shows a large substitution effect to the neighboring segments. The one-level nested logit model yields to symmetric substitution across segments, where the amount of substitution towards other segments is very low. The two-level nested logit model only partially captures neighboring segment effects but not the full pattern of between segment correlation.

The ONGEV model is a closed-form discrete choice model that enhances the flexibility of the nested logit model. Several authors have noticed the limitation of the nested logit model consisting of unambiguous assignment of alternatives to only one nest. Small's (1987) OGEV model was probably the first closed-form GEV model that allows overlapping nests. Other closed-form models have been developed, especially in the transportation literature; see Chu (1989); Vovsha (1997); Ben-Akiva and Bierlaire (1999). Small (1994) and Bhat (1998) have extended the OGEV model to a nested logit model for the higher-level choice decision and the OGEV model formulation for the lower choice decision. The most flexible model in this literature is the generalized nested logit model by Wen and Koppelman (2001), where an alternative can be a member of more than one nest to varying degrees. Finally, Bresnahan et al. (1997) develop a principle of differentiation model which is an example of closed-form GEV model applied to market-level data and a numerous number of alternatives.

The remainder of the paper is organized as follows. Section 2 discusses the ONGEV model. Section 3 describes the dataset and the econometric procedure, including the identification issues. Section 4 shows the empirical results and the implied price elasticities. Section 5 concludes.

2 The ordered nested generalized extreme value model

I estimate the demand for cars at product level within the random utility framework as developed in the GEV class of models proposed by McFadden (1978).

There are T markets, $t = 1, .., T$. In each market t there are L_t potential consumers. I suppress the market subscript t for the moment, since consumers are assumed to purchase the car only in the market where they are located. Consumer i chooses the specific car j, $j = 0, \ldots J$. The outside good includes only the option 'do not buy a car', $j = 0$, for which consumer i's indirect utility is $u_{i0} = \varepsilon_{i0}$. For cars $j = 1, \ldots, J$, consumer i's indirect utility is:

$$\begin{aligned} U_{ij} &= x_j \beta + \alpha p_j + \xi_j + \varepsilon_{ij} \\ &\equiv \delta_j + \varepsilon_{ij}, \end{aligned}$$

where x_j is a vector of observed product characteristics, including size, horsepower, fuel efficiency and country of origin; p_j is the price; ξ_j is the unobserved product characteristic, including style, image or comfort.[1] The coefficient of price, α, is specified as $\overline{\alpha}/y$, where y is equal to income per capita.[2] Following Berry (1994), I decompose U_{ij} into two terms: δ_j, the mean utility term common to all consumers and ε_{ij}, the utility term specific to each consumer.

In particular, ε_{ij} is an individual realization of the random variable ε. The distribution of ε is determinant for the shape of demand and the implied substitution pattern. I assume that segments are ordered from the outside good segment (the 'inferior quality' good) to the luxury segment according to observable characteristics, such as price, size, engine performance and fuel efficiency, and other unobservable characteristics, such as comfort, prestige or safety. The modeling strategy takes into account that if cars j and k belong to the same segment or to a neighboring segment, a consumer's draw of ε_{ij} can be correlated with ε_{ik}.

McFadden (1978) has proposed a GEV class of random utility model in which such correlation can be modeled in different ways. A GEV model is derived from a function $G = G(e^{\delta_0, \ldots, \delta_J})$ which is a differentiable function defined on $\mathbb{R}_+^J$ which is: (i) non-negative; (ii) homogeneous of degree 1; (iii) tending toward $+\infty$ when any of its arguments tend toward $+\infty$; (iv) whose n^{th} cross-partial derivatives with respect with respect to n distinct

[1] Since my data are at the annual level, I also add controls for the number of months each car is available in a country within a given year (for cars introduced or dropped within a year).

[2] This utility specification approximates Berry et al.'s (1995) Cobb-Douglas specification $\alpha \ln(y - p_j)$ when the price is small relative to (capitalized) income. It is convenient when studying countries with different exchange rates, since local price is simply expressed relative to local income; see Goldberg and Verboven (2001).

e^{δ_j} are non-negative for odd n and non-positive for even n.

According to the GEV postulate, the choice probability of buying car j is:

$$s_j = \frac{e^{\delta_j} \cdot G_j(e^{\delta_0,\ldots,\delta_J})}{G(e^{\delta_0,\ldots,\delta_J})}, \tag{1}$$

where s_j if the equation for the market share of car j and G_j is the partial derivative of G with respect to e^{δ_j}.

To make the logic of the modeling strategy more clear, consider the G function associated with a traditional specification, the nested logit model, in which the ordering of the segments is not explicitly modeled. I present first a one-level nested logit model, in which correlation between segments is ruled out by construction. Then I present a two-level nested logit model, in which correlation between segments is partially modeled. Finally I turn to the ONGEV model. The three models are represented in Figure 1.

One-level nested logit model In the one-level nested logit model, the segment ordering does not matter. The model incorporates potential correlation among products only within a nest (segment), not between (panel a of Figure 1). The J alternatives are grouped into S nests (segments) labeled $S_1, ..., S_S$. The G function takes the form:

$$G = e^{\delta_0} + \sum_{r=1}^{S} \left(\sum_{j \in S_r} e^{\frac{\delta_j}{1-\sigma_s}} \right)^{1-\sigma_s}, \tag{2}$$

where σ_s captures correlation among products within the same nest. Consistency with random utility maximization requires σ_s to lie in the unit interval. As σ_s goes to 0, the expression in (2) simplifies to the G function associated with a logit model, where each element of ε is independent. As σ_s goes to 1, the error term becomes perfectly correlated, so that the probability of choosing an alternative dominated by another alternative ($|\delta_j| > |\delta_k|$) in the same nest is 0.

Two-level nested logit model One could think about more complicated nesting structures to capture neighboring substitution effects within the modeling framework of traditional nested logit models. For example, a two-level nested structure could combine the segments into two nests: a 'small' nest, including subcompact and compact segments and a 'large' nest, including intermediate, standard and luxury segments (panel b of Figure 1). In practice, the J alternatives are grouped into 2 nests, small and large, labeled B_1(small) and B_2,(large) and S sub-nests (segments), labeled $S_1, ...S_S$.The two-level nested logit formula of the G function is the following:

$$G = e^{\delta_0} + \sum_{r=1}^{S} \left\{ \left[\left(\sum_{j \in B_1, S_r} e^{\frac{\delta_j}{1-\sigma_g}} \right)^{\frac{1-\sigma_g}{1-\sigma_s}} + \left(\sum_{j \in B_2, S_r} e^{\frac{\delta_j}{1-\sigma_g}} \right)^{\frac{1-\sigma_g}{1-\sigma_s}} \right]^{1-\sigma_s} \right\}, \tag{3}$$

where σ_s captures the correlation of preferences within a specific sub-nest (segment) and σ_g the correlation of preferences within a grouping of segments (B_1, small and B_2, large). A price shock to a compact car would determine, first, substitution within the compact segment and, next, within the small group, towards the subcompact segment. But this specification assumes away potential correlation in ε towards the other neighboring segment, the intermediate one. Segments could be grouped in different ways, but no two-level nested logit specification can fully parameterize the neighboring segment effects. I turn to the ONGEV model to better accommodate this effect.

The ONGEV model Assume that the 5 + 1 segments are ordered as follows: S_0, the outside good; S_1, subcompact; S_2, compact; S_3, standard; S_4, intermediate; S_5, luxury. The ordering corresponds to an increasing value of important characteristics such as price. The outside good segment is, thus, the segment with the 'inferior quality' good.

Assume the following G function within the GEV class:

$$G = \sum_{r=0}^{S+1} \left[\frac{1}{2} \left(\sum_{j \in S_{r-1}} e^{\frac{\delta_j}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} \left(\sum_{j \in S_r} e^{\frac{\delta_j}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} \right]^{1-\sigma_n}, \tag{4}$$

where σ_s captures correlation of preferences within a specific nest (segment) and σ_n correlation of preferences between neighboring segments. Consistency with random utility maximization requires $0 \leq \sigma_n \leq \sigma_s < 1$. These restrictions on the values of σ_n and σ_s are necessary to satisfy the four conditions for function G to belong to the GEV family. The Appendix provides the proof in paragraph A.1.

The shape of the demand function crucially depends on the two parameters, σ_s and σ_n, that parameterize the cumulative distribution of the error term ε. The first one, σ_s, corresponds to a pattern of dependence in ε across products sharing the same segment. The second one, σ_n, corresponds to a pattern of dependence in ε across products belonging to neighboring segments. Consider the effect of a price shock to a specific compact car. The dependence in ε measured by σ_s determines that a share of consumers, who had initially chosen a compact car, will switch to another compact car. The dependence in ε measured by

σ_n determines that a share of consumers will switch to a subcompact or an intermediate car (the neighboring segments). In other words, if the values of σ_s and σ_n are sufficiently high, products belonging to the same segment or to neighboring segments will be closer substitutes compared to products belonging to further segments (panel c of Figure 1). I formalize the argument by looking at the correlation between ε's for two cars that (i) belong to the same segment; (ii) belong to neighboring segments; (iii) belong neither to the same segment nor to neighboring segments. The Appendix in paragraph A.2 provides these expressions.

If the random components follow the G function in (4), by GEV postulate in (1) the choice probability of buying car j is:

$$s_j = \frac{e^{\frac{\delta_j}{1-\sigma_s}} \cdot \left(\sum_{j \in S_r} e^{\frac{\delta_j}{1-\sigma_s}}\right)^{\frac{\sigma_n-\sigma_s}{1-\sigma_n}} \cdot \left[\left(\sum_{j \in S_{r-1}} e^{\frac{\delta_j}{1-\sigma_s}}\right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \left(\sum_{j \in S_r} e^{\frac{\delta_j}{1-\sigma_s}}\right)^{\frac{1-\sigma_s}{1-\sigma_n}}\right]^{-\sigma_n}}{\sum_{r=0}^{S+1}\left[\left(\sum_{j \in S_{r-1}} e^{\frac{\delta_j}{1-\sigma_s}}\right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \left(\sum_{j \in S_r} e^{\frac{\delta_j}{1-\sigma_s}}\right)^{\frac{1-\sigma_s}{1-\sigma_n}}\right]^{1-\sigma_n}}. \quad (5)$$

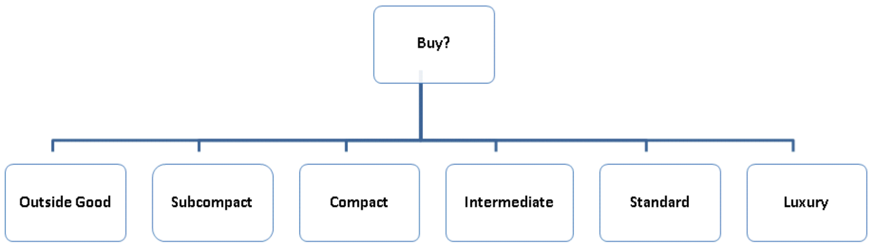
If $\sigma_n = 0$, the expression boils down to the one-level nested logit choice probability:

$$s_j = \frac{e^{\frac{\delta_j}{1-\sigma_s}} \cdot \left(\sum_{j \in S_r} e^{\frac{\delta_j}{1-\sigma_s}}\right)^{-\sigma_s}}{\sum_{r=0}^{S}\left[\left(\sum_{j \in S_r} e^{\frac{\delta_j}{1-\sigma_s}}\right)^{1-\sigma_s}\right]}. \quad (6)$$

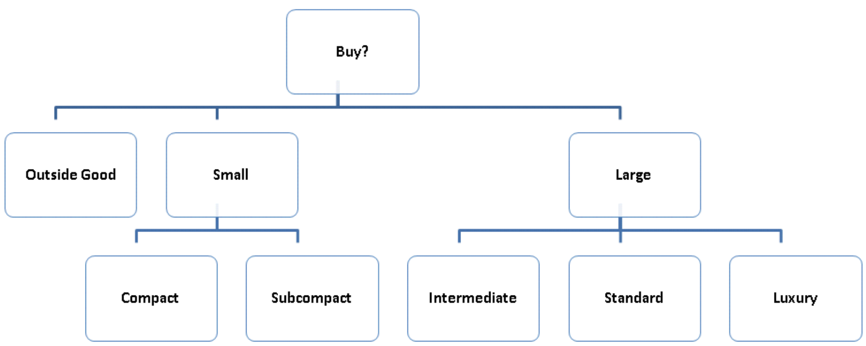
Compare the market shares of the ONGEV model in (5) with the corresponding shares of the one-level nested logit model in (6). Similarly to the one-level nested logit model, in the ONGEV model s_j is diminished by the presence of attractive alternatives within a nest. Differently from the nested logit model, s_j is also diminished by the presence of attractive alternatives in neighboring nests, because the denominator will increase. Ceteris paribus, this effect is increasing in σ_n.

Compare the ONGEV model with the two-level nested logit model. In the two-level nested logit model, segments can be paired in different ways, but no specification can fully parameterize the neighboring segment effects as in the ONGEV model. Note that the ONGEV model does not imply the estimation of a greater number of parameters with respect to the two-level nested logit, but simply accommodates neighboring effects in a more appropriate way.

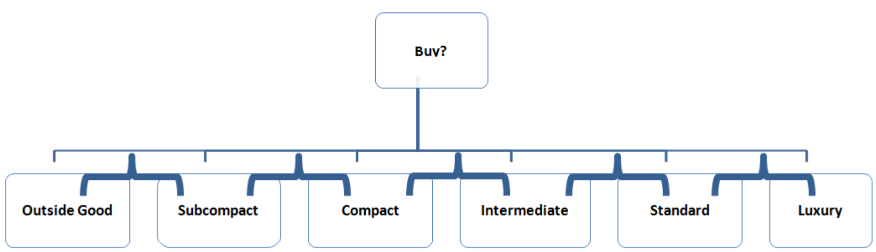
Figure 1: The nested logit models versus the ONGEV model



(a) One-level nested logit



(b) Two-level nested logit



(c) ONGEV

This figure represents the three model specifications based on McFadden's (1978) GEV model: the one-level nested logit that ignores neighboring segments, the two-level nested logit that partially models neighboring segments, and the ONGEV model, that flexibly models neighboring segments. Note that in the ONGEV model each segment overlaps with two neighbors, apart from the extreme segments (outside good and luxury).

3 Data and identification

3.1 Data

I use a unique dataset on the automobile market maintained by JATO. The data are at the level of the car model (e.g. VW Golf) and include essentially all passenger cars sold between 1998 and 2009 in nine West-European countries, which include Belgium, France, Germany, Greece, Italy, the Netherlands, Portugal, Spain and the UK. For each model/country/year, I build a dataset including information on sales and list prices and various characteristics such as vehicle size (curb weight, width and height), engine attributes (horsepower and displacement) and fuel consumption (liter/100 km or €/100 km). I assign each model to the brands' specific perceived country of origin. Models sold under the brand of Citroën, Peugeot and Renault are perceived as French cars, even though the production can take place in different locations. The dataset is augmented with macro-economic variables including the number of households for each country, the population size and GDP. The resulting dataset consists of 18,624 model/country/year observations or, on average, about 172 models per country/year. A more detailed description of the dataset for a shorter number of years (1998-2006) is provided by Grigolon and Verboven (2011). Table 1 provides summary statistics for sales, price and the other product characteristics used in our empirical demand model. In my counterfactual I report substitution patterns for one country, Germany, so I report summary statistics for that country as well.

Starting from JATO's classification, I attribute each model to a marketing segment. There are five marketing segments: subcompact, compact, standard, intermediate, luxury.[3] Since my empirical analysis focuses on the ordering of the segments from subcompact to luxury, I provide more details on how the characteristics relate to neighboring segments. Table 2 presents, on the top panel, the mean for each characteristic per segment. The mean values of price, horsepower, fuel consumption and width increase from subcompact to luxury. Height shows neither a clearly increasing nor a clearly decreasing trend. Segment averages of price, horsepower and fuel consumption vary more widely with respect to width and height between segments. On the bottom panel, the table presents a non-parametric test on the probability that a randomly-chosen value of, say, horsepower from the subcompact segment has a higher value than a randomly-chosen value of horsepower of the compact segment. The test is useful to illustrate both the segment ordering and the extent of overlap of the characteristics between neighboring segments. Results confirm that all characteristics tend to increase when switching from one segment to the neighboring one (probability $< 50\%$)

[3] Other segments that are present in the car market, such as SUV, sports and minivan, are not considered here because different characteristics would suggest different ordering of the segments.

Table 1: Summary Statistics

	All countries		Germany	
	Mean	St. Dev.	Mean	St. Dev.
Sales (units)	6,759	16,048	13,872	23,299
Price/Income	1.01	0.76	0.79	0.47
Horsepower (in kW)	81.3	35.4	82.7	36.6
Fuel efficiency (€/100 km)	8.8	2.5	9.0	2.2
Width (cm)	172.4	8.2	172.6	8.2
Height (cm)	146.2	6.8	146.4	6.9
Foreign (0-1)	0.9	0.3	0.7	0.4
Months present (1-12)	9.7	2.6	9.7	2.5

The table reports means and standard deviations of the main variables. The total number of observations (models/markets) is 18,624, where markets refer to the 9 countries and 12 years.

with the exception of height. This partial measure of size should not be taken into account in the ordering because smaller cars often tend to compensate their reduced size in terms of width and length with height.

Neighboring segments also present different degrees of overlap depending on the characteristic. Fuel efficiency tends to overlap across neighboring segments more than price and horsepower. For example, while the average fuel efficiency increases from the intermediate (9.2) to the standard segment (9.9), the probability that a randomly chosen value of fuel efficiency from the intermediate segment has a higher value than a randomly-chosen value of fuel efficiency of the standard segment is very high (42%). Probabilities across all characteristics suggest that the border between compact-intermediate and intermediate-standard is less defined compared to the border between subcompact-compact and standard-luxury.

3.2 The estimation procedure

The estimation procedure for the ONGEV model follows the methodological lines of Berry (1994), Berry et al. (1995) and the subsequent literature. Following Grigolon and Verboven (2011), I exploit the panel features of the dataset to specify the error term capturing unobserved product characteristics. For this purpose, I reintroduce the market subscript t. Specifically, I model the product-related error term as follows: $\xi_{jt} = \xi_j + \xi_t + \Delta\xi_{jt}$, where ξ_j is a fixed-effect for each car model, ξ_t is a fixed-effect for each country interacted with a time trend and squared time trend. $\Delta\xi_{jt}$ is the remaining product-related error term.

I follow a two-step procedure. First, I numerically solve for the error term $\Delta\xi_{jt}$ as a

Table 2: Summary Statistics by Segment

Segment	Subc	Comp	Interm	Stand	Lux
Mean of the characteristics x_{jt}					
Price/Income	0.528	0.792	1.023	1.378	2.162
Horsepower (in kW)	50.1	71.7	87.5	102.8	142.1
Fuel efficiency (€/100 km)	7.2	8.3	9.2	9.9	12.0
Width (cm)	163.2	172.5	176.5	176.0	183.1
Height (cm)	149.4	145.2	145.6	142.9	145.8
% overlap between neighboring segments					
Price/Income	-	20	26	25	24
Horsepower (in kW)	-	8	17	25	16
Fuel efficiency (€/100 km)	-	33	36	42	29
Width (cm)	-	7	23	54	9
Height (cm)	-	65	46	70	26
Number of observations	5,165	5,352	3,572	2,159	2,376

The top panel of the table reports means of the five continuous characteristics (i.e. price/income, horsepower, fuel efficiency, width and height) by segment. The bottom panel of the table reports a non-parametric test on the probability that a randomly-chosen value of a characteristic from segment $r-1$ has a higher value than a randomly-chosen value of the characteristic from segment r. Subc=subcompact, Comp=compact, Interm=intermediate, Stand=standard, Lux=Luxury.

function of the vector of parameters. Second, I interact $\Delta\xi_{jt}$ with a set of instruments to form a generalized method of moments (GMM) estimator.

Consider the solution of $\Delta\xi_{jt}$ first. In the nested logit model $\Delta\xi_{jt}$ has an analytic solution. In the ONGEV model $\Delta\xi_{jt}$ is the numerical solution of the system $s = s(\delta(\alpha, \sigma_s, \sigma_n), \alpha, \sigma_s, \sigma_n)$. In contrast with the random coefficients logit model, the numerical procedure is well-behaved because the market share function is expressed in closed-form.[4]

Let $\widehat{\Delta\xi}$ be the sample analogue of the vector $\Delta\xi$, and Z the matrix of instruments. The GMM estimator is defined as:

$$\min_{\alpha,\sigma_s,\sigma_n} \widehat{\Delta\xi}'(Z\Omega Z')\widehat{\Delta\xi},$$

where Ω is the weighting matrix. To minimize the GMM objective function with respect to the parameters $\alpha, \sigma_s, \sigma_n$, I first concentrate out the linear parameters β. Also, I do not directly estimate the more than 150 car model fixed effects ξ_j, but instead use a within transformation of the data (Baltagi, 1995). Standard errors are computed using the standard GMM formulas for asymptotic standard errors.

3.3 Identification

The GMM estimator requires an instrumental variable vector Z with a rank of at least $K+3$ (K is the dimension of the β vector; the price parameter α; the two nesting parameters σ_s and σ_n). The interpretation of $\Delta\xi_{jt}$ as unobserved product quality disqualifies price p_{jt} as an instrument since it could imply a positive correlation with $\Delta\xi_{jt}$. There are two main reasons for this positive correlation. First, if an unobservable characteristic, for example comfort, rises with price, consumers will avoid expensive cars less than they would without that characteristic. Second, if adding comfort is costly for the manufacturer, the price of the car is expected to reflect this cost. A similar argument holds for the correlation between the shares within a segment (or within neighboring segments) and $\Delta\xi_{jt}$. This calls for instrumentation of the share terms to avoid an upward bias on the parameters σ_s and σ_n.

Following Berry et al. (1995), I assume that the observed product characteristics x_{jt} are uncorrelated with the unobserved product characteristics $\Delta\xi_{jt}$. Note that this assumption is weaker than the often adopted assumption that x_{jt} is uncorrelated with ξ_{jt}. We can therefore include x_{jt} in the matrix of instruments. I also use functions of these characteristics as instruments to estimate the $K+3$ parameters. Specifically, I include: (i) counts and sum of

[4] I use a modified version of Berry et al.'s (1995) contraction mapping: $\delta^{k+1} = \delta^k + [1 - \max(\widehat{\sigma_s}, \widehat{\sigma_n})] \cdot [\ln(s) - \ln(s(\delta^k))]$. If one does not weigh the second term by $[1 - \max(\widehat{\sigma_s}, \widehat{\sigma_n})]$ the procedure may not lead to convergence; see Brenkers and Verboven (2006). Following the insights of the literature on random coefficients logit models, particularly by Dubé et al. (2012), I use a tight tolerance level of $1e-14$ to invert the shares using Berry et al.'s (1995) contraction mapping.

the characteristics of other products of competing firms by segment; (ii) counts and sum of the characteristics of other products of the same firm by segment; (iii) counts and sum of the characteristics of other products of competing firms by neighboring segments; (iv) counts and sum of the characteristics of other products of the same firm by neighboring segments. These instruments originate from supply side considerations, where I assume that firms set prices according to a Bertrand-Nash game. When the number of products in one segment, or in the neighboring segments increases, demand should become more elastic and this should affect prices and shares. Similarly, if one firm produces a large share of the products in one segment or in neighboring segments, sales and prices for each product of that particular firm should be higher.

4 Results

4.1 Demand estimates

Table 3 shows the parameter estimates for the three alternative demand models. The first one is the one-level nested logit model, which imposes $\sigma_n = 0$. The second one is the two-level nested logit model, which is an approximation of the ONGEV model which assumes away potential correlation between ε among compact and intermediate segments. Again in this model $\sigma_n = 0$, but two σ's are estimated: σ_s corresponding to the specific segment (sub-nest) and σ_g corresponding to the higher level in which segments are grouped into two nests, small and large. The third one is the ONGEV model, where both σ_s and σ_n are estimated.

In all three models, the price parameter ($\overline{\alpha}$) and the parameters of the characteristics (β) have the expected sign and are all significantly different from zero. Most parameter estimates have also roughly the same magnitude.

I now examine the nesting parameters. In the first model, the nesting parameter for the segment is estimated very precisely: $\sigma_s = 0.84$. The magnitude is consistent with random utility maximization ($0 \leq \sigma_s < 1$) and implies that consumer preferences are strongly correlated within a segment. This is consistent with earlier findings by Goldberg and Verboven (2001) and Brenkers and Verboven (2006).

In the second model, the two nesting parameters are again very precisely estimated: $\sigma_s = 0.89$ and $\sigma_g = 0.69$. Their magnitude is also consistent with random utility maximization ($0 \leq \sigma_g \leq \sigma_s < 1$). They imply that consumer preferences are more strongly correlated across cars from both the same segment grouping (small and large) and sub-grouping (segment) rather than merely from the same grouping. However, correlation within a grouping of segments (small and large) is still important. Neighboring segment effects are therefore present.

In the third model, the ONGEV model, we estimate both σ_s and σ_n significantly different from zero. The parameter measuring correlation within segments, σ_s, is very precisely estimated, $\sigma_s = 0.85$. The parameter measuring correlation between neighboring segments, σ_n, is less precisely estimated, $\sigma_n = 0.72$. Nevertheless, correlation between neighboring segments is strongly supported by the data since σ_n is significantly higher than zero and has a high magnitude: the one-level nested logit assuming $\sigma_n = 0$ is rejected against the ONGEV model. The magnitude of the parameters is consistent with random utility maximization ($0 \leq \sigma_n \leq \sigma_s < 1$).

Table 4 presents alternative estimations on the ONGEV model with different instrument sets to check the robustness of the parameter estimates. The first column reports the baseline ONGEV. In the second column I remove half of the instruments, specifically the ones based on the neighboring segments. In the third column I remove the other half of the instruments, the ones based on the segments. The magnitude of the parameters is stable, where the hypothesis that $\sigma_n = 0$ is always rejected. Note a reduction in the measured precision of the important parameters, $\overline{\alpha}$, σ_s and especially σ_n, which is, however, not dramatic.

Table 3: Parameter Estimates for Alternative Demand Models

	One-level Nested Logit		Two-level Nested Logit		Ordered Nested Logit	
	Param.	St. Er.	Param.	St. Er.	Param.	St. Er.
	Mean valuations for the characteristics in x_{jt} (β)					
Price/income	-1.26	0.02	-0.94	0.02	-1.08	0.05
Horsepower (kW/100)	1.24	0.04	0.91	0.03	1.04	0.06
Fuel (€/10,000 km)	-3.06	0.34	-1.93	0.25	-2.28	0.29
Width (cm/100)	7.01	1.78	5.42	1.30	7.16	1.49
Height (cm/100)	11.70	1.32	9.39	0.97	10.24	1.11
Foreign (0/1)	-0.27	0.03	-0.16	0.02	-0.24	0.03
	Nesting parameters ($\sigma_s, \sigma_n, \sigma_g$)					
Segment σ_s	0.84	0.02	0.89	0.02	0.85	0.02
Neighboring segment σ_n	n/a		n/a		0.72	0.21
Group of segments σ_g	n/a		0.69	0.02	n/a	
Model fixed effects	Yes		Yes		Yes	
Market fixed effects	Yes		Yes		Yes	

The table shows the parameter estimates and standard errors for the three different demand models. The nested logit models assume $\sigma_n = 0$, while the ONGEV model estimates σ_n. The total number of observations (models/markets) is 18,624, where markets refer to the 9 countries and 12 years.

Table 4: Alternative Treatment of the Instrument Matrix

	Baseline ONGEV		No neighboring seg. IV		No segment IV	
	Param.	St. Er.	Param.	St. Er.	Param.	St. Er.
	Mean valuations for the characteristics in x_{jt} (β)					
Price/income	-1.08	0.05	-1.02	0.06	-1.24	0.05
Horsepower (kW/100)	1.04	0.06	0.98	0.08	1.23	0.07
Fuel (€/10,000 km)	-2.28	0.29	-2.14	0.30	-3.00	0.32
Width (cm/100)	7.16	1.49	9.27	1.73	7.82	1.74
Height (cm/100)	10.24	1.11	11.04	1.20	11.00	1.25
Foreign (0/1)	-0.24	0.03	-0.31	0.04	-0.24	0.03
	Nesting parameters (σ_s, σ_n)					
Segment σ_s	0.85	0.02	0.80	0.03	0.84	0.02
Neighboring segment σ_n	0.72	0.21	0.73	0.24	0.71	0.26
Model fixed effects	Yes		Yes		Yes	
Market fixed effects	Yes		Yes		Yes	

The table shows the parameter estimates and standard errors for the ONGEV model with different instrument sets. The first model is the baseline one. The second model excludes: (i) counts and sum of the characteristics of other products of competing firms by neighboring segments; (ii) counts and sum of the characteristics of other products of the same firm by neighboring segments. The third model excludes: (i) counts and sum of the characteristics of other products of competing firms by segment; (ii) counts and sum of the characteristics of other products of the same firm by segment. The total number of observations (models/markets) is 18,624, where markets refer to the 9 countries and 12 years.

4.2 Substitution patterns: segment-level price elasticities

The implications of the ONGEV model are clearly illustrated by the substitution patterns. I consider own- and cross-price elasticities at the level of an entire segment. These elasticities represent the effect of a joint 1% price increase of all cars in a given segment on the demand in the various segments. Table 5 shows the segment-level own- and cross-price elasticities. The table reports both the point estimates and the bootstrapped 95% confidence interval.

The own-price elasticities across the three models are very similar in terms of magnitude and often overlap in the confidence intervals. The own-price elasticities of the second model, the two-level nested logit model, tend to be higher, especially for the most-expensive classes. This proportional relationship between own-price elasticity and price is a result of the functional form assumption of the nested logit model, since price enters utility linearly. The proportionality is a common feature of all three models, but the higher values of the nesting parameters for the second model make it more evident.

The cross-price elasticities are the most interesting. By construction, the one-level nested logit model implies a fully symmetric substitution pattern, namely identical cross-price elasticities in each row. The two-level nested logit model retains the same feature, but within each sub-nest. Thus, a 1% price increase to the intermediate segment raises the demand in the standard and luxury segments by the same amount, 0.73%. In addition, a price increase in the intermediate segment implies a very small increase in the demand of compact cars, a segment that belongs to another grouping (small). This counterintuitive result is the consequence of the inability of the model to represent the overlapping segments. Similarly, a price increase in the compact segment implies a strong substitution effect to the subcompact segment, but not to the intermediate one, which attracts demand in the same proportion of the luxury segment.

By contrast, the ONGEV model seems to capture substitution to neighboring segments well. Taking the same example, i.e. a price increase in the compact segment, note the high effect on the demand of the two neighboring segments, subcompact (+0.14%) and intermediate (+0.13%). These numbers are lower, but comparable to the ones reported by Grigolon and Verboven (2011) in the analysis of the segment-level price elasticities for the random coefficients logit model. The fact that between segment correlation, σ_n, is less precisely estimated than within segment correlation, σ_s, leads to larger confidence intervals for the point estimates of the cross-price elasticities compared to the own-price elasticities. Note that the distribution of the elasticities tends to be skewed to the left, so that the point estimates are higher than their expected value.

The ONGEV model is parsimonious in the number of parameters, so that only the immediately proximate segments (on the left and on the right) are the neighboring ones. Outside the neighboring segments, the ONGEV model still retains the modeling assumptions of the nested logit model. Thus, substitution patterns are symmetric outside the neighboring segments. This constitutes a limitation of the model which is especially evident for the segments allocated at the beginning and at the end of the ordering. For example, one could imagine that a price increase to luxury cars would imply higher substitution effects towards standard cars and, to a lesser extent, towards intermediate cars. Now substitution towards the intermediate segment is very small and identical to the one towards the compact and subcompact segments. Modeling further neighboring segments could be an interesting extension of this work.

5 Conclusion

I present a new member of the GEV model family denominated ordered nested generalized extreme value (ONGEV) model. The ONGEV model is particularly suitable for markets of

Table 5: Segment-level Price Elasticities in Germany for Alternative Demand Models

	One-level Nested Logit				
	Subcompact	Compact	Intermediate	Standard	Luxury
Subcompact	-0.50	0.02	0.02	0.02	0.02
	-0.51;-0.48	0.02;0.02	0.02;0.02	0.02;0.02	0.02;0.02
Compact	0.02	-0.76	0.02	0.02	0.02
	0.02;0.02	-0.79;-0.74	0.02;0.02	0.02;0.02	0.02;0.02
Intermediate	0.01	0.01	-1.07	0.01	0.01
	0.01;0.01	0.01;0.01	-1.11;-1.04	0.01;0.01	0.01;0.01
Standard	0.01	0.01	0.01	-1.36	0.01
	0.01;0.01	0.01;0.01	0.01;0.01	-1.40;-1.32	0.01;0.01
Luxury	0.01	0.01	0.01	0.01	-1.99
	0.01;0.01	0.01;0.01	0.01;0.01	0.01;0.01	-2.05;-1.92
	Two-level Nested Logit				
Subcompact	-0.70	0.53	0.01	0.01	0.01
	-0.76;-0.65	0.46;0.61	0.01;0.01	0.01;0.01	0.01;0.01
Compact	0.56	-1.35	0.01	0.01	0.01
	0.49;0.64	-1.47;-1.23	0.01;0.01	0.01;0.01	0.01;0.01
Intermediate	0.03	0.01	-1.89	0.73	0.73
	0.02;0.03	0.01;0.01	-2.06;-1.73	0.63;0.84	0.63;0.84
Standard	0.03	0.01	0.86	-2.46	0.86
	0.02;0.03	0.01;0.01	0.74;0.99	-2.68;-2.25	0.74;0.99
Luxury	0.03	0.01	0.77	0.77	-4.05
	0.02;0.03	0.01;0.01	0.66;0.88	0.66;0.88	-4.43;-3.68
	Ordered Nested Logit				
Subcompact	-0.52	0.14	0.01	0.01	0.01
	-0.56;-0.44	0.02;0.19	0.01;0.01	0.01;0.01	0.01;0.01
Compact	0.14	-0.86	0.13	0.01	0.01
	0.02;0.19	-0.96;-0.68	0.01;0.21	0.01;0.02	0.01;0.02
Intermediate	0.01	0.05	-1.49	0.48	0.01
	0.01;0.01	0.01;0.08	-1.75;-1.07	0.11;0.85	0.01;0.01
Standard	0.01	0.01	0.53	-1.94	0.37
	0.01;0.01	0.01;0.01	0.13;0.87	-2.36;-1.28	0.01;0.56
Luxury	0.01	0.01	0.01	0.32	-2.19
	0.01;0.01	0.01;0.01	0.01;0.01	0.01;0.48	-2.52;-1.74

The table reports the segment-level own- and cross-price elasticities (when all products in the same segment raise their price by 1%), together with the percent confidence intervals, based on a bootstrapping procedure. The elasticities are based on the parameter estimates in Table 3. They refer to Germany in 2009.

differentiated products that present a form of segmentation which can be ordered in a natural way. The model relaxes the assumption of independent and identically distributed random errors between segments of the nested logit. In the nested logit framework, the assumption implies symmetric elasticities outside a segment. In contrast, the ONGEV model allows for asymmetric elasticities, where consumers subject to a price increase for alternatives in one segment favor alternatives in neighboring segments rather than in further segments. Flexibility in substitution patterns is accomplished with a closed-form solution that avoids the numerical problems of the random coefficients logit model.

I apply the ONGEV model to the car market which is classified into segments that are naturally ordered from subcompact to luxury. Results show that neighboring segment effects are strongly supported in the data.

The specific modeling strategy I adopt here seems to be a promising starting point to capture neighboring segment effects. One could extend the model to capture further neighbors, rather than just the proximate ones. Also, other industries could benefit from this modeling strategy when ordering a high number of alternatives would prove impossible, but ordering grouping of these alternatives represents a sensible way to obtain flexible substitution patterns in a tractable setting. Take for example the hotel market, where hotel categories are naturally ordered from motel to luxury according to several characteristics such as price, comfort or number of available facilities (Venkataraman and Kadiyali, 2005).

A Appendix

A.1 Proof GEV

Proof. I show that under the assumption $0 \leq \sigma_n \leq \sigma_s < 1$, the G function in (4) verifies the four properties of GEV generating functions. To simplify the notation, let $e^{\delta_j} = Y_j$.

1. G is obviously non-negative since $Y_j \in \mathbb{R}_+ \forall j$.

2. G is homogeneous of degree 1, that is $G(\lambda Y_0, ..., \lambda Y_J) = \lambda G(Y_0, ..., Y_J)$. Indeed:

$$
\begin{aligned}
G(\lambda Y_0, ..., \lambda Y_J) &= \sum_{r=0}^{S+1} \left[\frac{1}{2} \left(\sum_{j \in S_{r-1}} \lambda^{\frac{1}{1-\sigma_s}} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} \left(\sum_{j \in S_r} \lambda^{\frac{1}{1-\sigma_s}} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} \right]^{1-\sigma_n}, \\
&= \sum_{r=0}^{S+1} \left[\frac{1}{2} \left(\lambda^{\frac{1}{1-\sigma_s}} \sum_{j \in S_{r-1}} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} \left(\lambda^{\frac{1}{1-\sigma_s}} \sum_{j \in S_r} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} \right]^{1-\sigma_n}, \\
&= \sum_{r=0}^{S+1} \left[\frac{1}{2} \lambda^{\frac{1}{1-\sigma_n}} \left(\sum_{j \in S_{r-1}} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} \lambda^{\frac{1}{1-\sigma_n}} \left(\sum_{j \in S_r} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} \right]^{1-\sigma_n}, \\
&= \lambda \sum_{r=0}^{S+1} \left[\frac{1}{2} \left(\sum_{j \in S_{r-1}} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} \left(\sum_{j \in S_r} Y_j^{\frac{1}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} \right]^{1-\sigma_n}, \\
&= \lambda G(Y_0, ..., Y_J).
\end{aligned}
$$

3. The limit property obviously holds.

4. The property of the sign of the derivatives holds if $0 \leq \sigma_n \leq \sigma_s < 1$. The first cross-derivative G_j is given by:

$$
G_j = \underbrace{\frac{1}{2} \underbrace{Y_j^{\frac{\sigma_s}{1-\sigma_s}}}_{\geq 0} \cdot \underbrace{A_r^{\frac{1-\sigma_s}{1-\sigma_n}-1}}_{\geq 0} \cdot \underbrace{\Theta}_{\geq 0}}_{\geq 0},
$$

where A_r and Θ are defined as follows:

$A_r = \sum_{j \in S_r} Y_j^{\frac{1}{1-\sigma_s}}$,

$\Theta = \left(\frac{1}{2} A_{r-1}^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} A_r^{\frac{1-\sigma_s}{1-\sigma_n}} \right)^{-\sigma_n} + \left(\frac{1}{2} A_r^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} A_{r+1}^{\frac{1-\sigma_s}{1-\sigma_n}} \right)^{-\sigma_n}$.

$G_j \geq 0$ as required.

The second cross-derivative is given by:

$$G_{ji} = \underbrace{\frac{1}{2} \cdot \underbrace{(Y_i Y_j)^{-\frac{\sigma_n}{1-\sigma_n}} \cdot A_r^{\frac{1-\sigma_s}{1-\sigma_n}-2}}_{\geq 0} \cdot \left(\underbrace{\underbrace{-\frac{\sigma_n}{1-\sigma_n}}_{\leq 0} \cdot \underbrace{\Gamma}_{\geq 0} + \underbrace{\frac{\sigma_n - \sigma_s}{(1-\sigma_n)\cdot(1-\sigma_s)}}_{\leq 0 \; iff \; \sigma_n \leq \sigma_s} \cdot \underbrace{\Theta}_{\geq 0}}_{\leq 0} \right)}_{\leq 0},$$

where Γ is defined as follows:

$$\Gamma = A_r^{\frac{1-\sigma_s}{1-\sigma_n}} \cdot \left(\tfrac{1}{2} A_{r-1}^{\frac{1-\sigma_s}{1-\sigma_n}} + \tfrac{1}{2} A_r^{\frac{1-\sigma_s}{1-\sigma_n}} \right)^{-1-\sigma_n} + \left(\tfrac{1}{2} A_r^{\frac{1-\sigma_s}{1-\sigma_n}} + \tfrac{1}{2} A_{r+1}^{\frac{1-\sigma_s}{1-\sigma_n}} \right)^{-1-\sigma_n}.$$

$G_{ji} \leq 0$ as required. Higher cross-partials are calculated similarly.

■

A.2 The error term ε

I formally show that the ONGEV model implies correlation among unobserved random utility components for segments that are close on the ordering. I look at the bivariate marginal cumulative distribution function (CDF) and, if possible, at the correlation of any two stochastic elements (ε) for two cars that (i) belong to the same segment; (ii) belong to neighboring segments; (iii) belong neither to the same segment nor to neighboring segments.

The following function represents the cumulative extreme-value distribution of ε:

$$F(\varepsilon_{11}, ..., \varepsilon_{1J}, ..., \varepsilon_{I1}, ..., \varepsilon_{IJ}) = \exp\left\{ \sum_{r=0}^{S+1} \left[\frac{1}{2} \left(\sum_{j \in S_{r-1}} e^{-\frac{\varepsilon_{ij}}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \frac{1}{2} \left(\sum_{j \in S_r} e^{-\frac{\varepsilon_{ij}}{1-\sigma_s}} \right)^{\frac{1-\sigma_s}{1-\sigma_n}} \right]^{1-\sigma_n} \right\}.$$

The marginal CDF of each stochastic element ε_{ij} is univariate extreme-value as follows:

$$F(\varepsilon_{ij}) = \exp\left[-2^{\sigma_n} \left(e^{-\frac{\varepsilon_{ij}}{1-\sigma_s}} \right)^{1-\sigma_s} \right].$$

The bivariate marginal CDF for two cars j and k belonging the same segment is the following:

$$H(\varepsilon_{ij}, \varepsilon_{ik}) = \exp\left[-2^{\sigma_n} \left(e^{-\frac{\varepsilon_{ij}}{1-\sigma_s}} + e^{-\frac{\varepsilon_{ik}}{1-\sigma_s}} \right)^{1-\sigma_s} \right]. \tag{7}$$

CDF (7) generates correlation between utilities that corresponds to the nested logit model:

$$Cor(\varepsilon_{ij}, \varepsilon_{ik}) = \sigma_s$$

The bivariate marginal CDF for two cars j and k' belonging to neighboring segments but not to the same segment depends on σ_s and σ_n as follows:

$$H(\varepsilon_{ij}, \varepsilon_{ik'}) = \exp\left\{-2^{\sigma_n+1}\left[\left(e^{-\frac{\varepsilon_{ij}}{1-\sigma_s}}\right)^{1-\sigma_s} + \left(e^{-\frac{\varepsilon_{ik'}}{1-\sigma_s}}\right)^{1-\sigma_s} + \left(\left(e^{-\frac{\varepsilon_{ij}}{1-\sigma_s}}\right)^{\frac{1-\sigma_s}{1-\sigma_n}} + \left(e^{-\frac{\varepsilon_{ik'}}{1-\sigma_s}}\right)^{\frac{1-\sigma_s}{1-\sigma_n}}\right)^{1-\sigma_n}\right]\right\}. \tag{8}$$

Correlation between cars belonging to neighboring segments computed from the CDF (8) cannot be written in closed-form, but it is different from 0 and intuitively increasing in σ_n.

The bivariate marginal CDF for two cars j and k'' that do not belong to the same segment or to neighboring segments is the product of the corresponding univariate CDF's because the random elements attached to different segments are independent:

$$H(\varepsilon_{ij}, \varepsilon_{ik''}) = \exp\left\{-2^{\sigma_n}\left[\left(e^{-\frac{\varepsilon_{ij}}{1-\sigma_s}}\right)^{1-\sigma_s} + \left(e^{-\frac{\varepsilon_{ik''}}{1-\sigma_s}}\right)^{1-\sigma_s}\right]\right\}. \tag{9}$$

Correlation between cars that do not belong to the same segment or to neighboring segments computed from CDF (9) is simply 0:

$$Cor(\varepsilon_{ij}, \varepsilon_{ik''}) = 0.$$

References

Baltagi, Badi H., *Econometric Analysis of Panel Data*, John Wiley & Sons, 1995.

Ben-Akiva, Moshe and Michel Bierlaire, "Discrete choice methods and their applications to short-term travel decisions," in R. Hall, ed., *Handbook of Transportation Science*, Operations Research and Management Science Kluwer 1999, pp. 5–34. ISBN:0-7923-8587-X.

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25* (2), 242–262.

_ **, James Levinsohn, and Ariel Pakes**, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

Bhat, Chandra R., "An Analysis of Travel Mode and Departure Time Choice for Urban Shopping Trips," *Transportation Research Part B: Methodological*, 1998, *32* (6), 361 – 371.

Brenkers, Randy and Frank Verboven, "Liberalizing A Distribution System: The European Car Market," *Journal of the European Economic Association*, 03 2006, *4* (1), 216–251.

Bresnahan, Timothy F., Scott Stern, and Manuel Trajtenberg, "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *The RAND Journal of Economics*, 1997, *28*, S17–S44.

Chu, C., "A Paired Combinational Logit Model for Travel Demand Analysis," in "Proceedings of Fifth World Conference on Transportation Research," Vol. 4 Venturea 1989, pp. 295–309.

Daly, Andrew and S. Zachary, "Improved multiple choicemodels," in Q. Dalvi eds. D.A. Hensher, ed., *Determinants of Travel Choice*, Saxon House Farnborough, 1977, pp. 335–357.

Dubé, Jean-Pierre H., Jeremy T. Fox, and Che-Lin Su, "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," *Econometrica*, September 2012, *80* (5), 2231 – 2267.

Goldberg, Pinelopi Koujianou and Frank Verboven, "The Evolution of Price Dispersion in the European Car Market," *Review of Economic Studies*, October 2001, *68* (4), 811–48.

Grigolon, Laura and Frank Verboven, "Nested logit or random coefficients logit? A comparison of alternative discrete choice models of product differentiation," CEPR Discussion Papers 8584, C.E.P.R. Discussion Papers September 2011.

Judd, Kenneth L. and Ben Skrainka, "High performance quadrature rules: how numerical integration affects a popular model of product differentiation," CeMMAP working papers CWP03/11, Centre for Microdata Methods and Practice, Institute for Fiscal Studies February 2011.

Knittel, Christopher R. and Konstantinos Metaxoglou, "Estimation of Random Coefficient Demand Models: Two Empiricists' Perspective," In Mimeo 2012.

McFadden, Daniel L., "Modelling of Choice of Residential Location," in F. Snickers A. Karlquist, L. Lundquist and J. Weibull, eds., *Spatial Interaction Theory and Residential Location*, 1978.

_ **and Kenneth Train**, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, 2000, *15* (5), 447–470.

Small, Kenneth A., "A Discrete Choice Model for Ordered Alternatives," *Econometrica*, March 1987, *55* (2), 409–24.

_ , "Approximate generalized extreme value models of discrete choice," *Journal of Econometrics*, 1994, *62* (2), 351 – 382.

Venkataraman, Sriram and Vrinda Kadiyali, "An Aggregate Generalized Nested Logit Model of Consumer Choices: An Application to the Lodging Industry," *SSRN eLibrary*, 2005.

Vovsha, Peter, "Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area," *Transportation Research Record: Journal of the Transportation Research Board*, 1997, *1607*, 6–15.

Wen, Chieh-Hua and Frank S Koppelman, "The generalized nested logit model," *Transportation Research Part B: Methodological*, 2001, *35* (7), 627 – 641.

Williams, H., "On the formation of travel demand models and economic evaluation measures of user benefit," *Environment and Planning - Part A*, 1977, *9* (3), 285 – 344.