Blurred boundaries: a flexible approach for segmentation applied to the car market

Laura Grigolon*

November 2017

Abstract

Prominent features of differentiated product markets are segmentation and product proliferation that blurs the boundaries between segments. I develop a tractable demand model, the Ordered Nested Logit, which allows for overlap between neighboring segments. I apply the model to the automobile market where segments are ordered from small to luxury. I find that consumers, when substituting outside their vehicle segment, are more likely to switch to a neighboring segment. Accounting for such asymmetric substitution matters when evaluating the impact of new product introduction or studying the effect of subsidies on fuel-efficient cars.

* *Laura Grigolon:* McMaster University, 1280 Main Street West Hamilton , ON L8S 4M4, Canada. (Email: Laura.Grigolon@mcmaster.ca). I am particularly grateful to Frank Verboven for his helpful comments and support. I also thank Victor Aguirregabiria, Tobias Klein, Pasquale Schiraldi, Johannes Van Biesebroeck, Patrick Van Cayseele, Mike Veall and participants at various seminars and conferences. I gratefully acknowledge financial support from SSHRC (Insight Development Grant). I declare that I have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

In most differentiated product markets, products can be partitioned into segments according to shared common features. Segmentation is not only a descriptive process, but also a practice used by firms to develop targeted marketing strategies and decide the placement of their products. Often, segments can be ordered in a natural way. Cars can be ordered from small (subcompact) to luxury according to price, size, engine performance, comfort and prestige; hotels and restaurants can be ordered on the basis of their ratings (number of stars); retail brands can be ordered in tiers according to quality and price.

In parallel with segmentation, the variety of products has also dramatically increased over time: cars, computers, printers, and smartphones are just a few examples of industries in which product proliferation is visibly prevalent. Broadening the product line has blurred the boundaries between segments, thus decreasing the distance between them: a premium subcompact car can be a potential substitute for a compact car. As a consequence, segments tend to overlap with their neighbors. Correlation between segments has important implications when we want to measure the impact of competitive events, such as the introduction of varieties combining features from different segments. Environmental policies aimed at encouraging the adoption of cleaner cars can also affect sales of upper segments differently, depending on the degree of overlap between segments.

I propose a new discrete choice model, the Ordered Nested Logit model, that captures ordered segmentation in differentiated product markets and allows for overlap between neighboring segments and, as a consequence, asymmetric substitution toward proximate neighbors. This model is a new member of the Generalized Extreme Value (GEV) model family developed by McFadden (1978). I develop the Ordered Nested Logit in the context of market level data. The GEV family is consistent with random utility theory and yields a tractable closed-form for choice probabilities. Berry (1994) has provided a framework to estimate two special members of this family with market level data: the Logit and the Nested Logit model. The Ordered Nested Logit model generalizes the Nested Logit model by incorporating an extra parameter that measures the correlation in preferences between neighboring segments: the Nested Logit model implicitly sets such correlation to zero. Hence, the Ordered Nested Logit has the Nested Logit and the Logit as special cases: it can serve as a test for the validity of the constraints imposed by the Nested Logit and, a fortiori, the Logit model. Apart from these two models, only a few other members of the GEV model family have been exploited so far with market level data: notable examples are the principle of differentiation model by Bresnahan, Stern and Trajtenberg (1997) and the flexible coefficient multinomial logit by Davis and Schiraldi (2012).

Is asymmetric substitution toward neighboring segments captured by the demand models we currently use? In the computationally simple Nested Logit model, 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, Levinsohn and Pakes (1995) has the potential to generate more flexible substitution patterns, where products tend to be closer substitutes as they share similar observed continuous characteristics. Grigolon and Verboven (2014) simulate the effect of a joint 1% price increase of all cars in a given segment and show that the random coefficients logit model tends to yield more intense substitution toward neighboring segments. But flexibility is achieved only if the parameters of the models, which determine how the random coefficients govern substitution patterns, are correctly and precisely identified. Berry and Haile (2014) provide general results for identification in differentiated product markets, showing that those parameters are identified by standard exclusion restrictions. Reynaert and Verboven (2014) and Gandhi and Houde (2016) study practical instrumentation strategies for empirical work. In practice, identification can prove difficult in complex set-ups, with four or more random coefficients, as documented by Reynaert and Verboven (2014). In addition, the random coefficients logit model does not produce a closed-form for choice probabilities and requires the use of computationally burdensome techniques. Earlier work documented sources of numerical issues (e.g. Knittel and Metaxoglou, 2008) and recent papers (Kalouptsidi, 2012, Dubé, Fox and Su, 2012, Lee and Seo, 2015) have proposed methods that improve the computational performance of random coefficients models. Avoiding the simulation of market shares altogether can alleviate the computational burden.

First I formally derive the Ordered Nested Logit model and I assess its flexibility by fitting it to datasets generated according to a Random Coefficient Logit model. I find that the Ordered Nested Logit approximates well the matrix of true elasticities. Next, I apply the Ordered Nested Logit model to a unique dataset on the car market covering three major European countries between 1998 and 2011. The process of purchasing a car is modelled as a nested sequence, with the choice between the segments (including the outside good segment) at the upper node level and the choice of the specific vehicle at the lower node. I estimate the degree of correlation in consumer preferences both within each segment, as in the Nested Logit model, and between neighboring segments. The demand estimates of the Ordered Nested Logit model clearly indicate a rejection of the simpler Nested Logit model: correlation in car choices is present not only within a segment, but also between neighboring segments.

The demand estimates have striking implications for the substitution patterns. While the Nested Logit model yields symmetric and very low substitution toward other segments, the Ordered Nested Logit model shows a large substitution effect to the neighboring segments. I look at the impact of the introduction of premium subcompact cars on sales by vehicle class. The Nested Logit model predicts that only sales of other subcompact cars are affected by the introduction of those vehicles, while the Ordered Nested Logit model shows, more plausibly, that the segment immediately above (compact cars) is affected as well. Next, I simulate a subsidy to clean vehicles: such policy is clearly asymmetric because it favours mainly subcompact and compact cars. The Nested Logit model predicts, again, that sales of non-subsidized cars do not notably change after the policy, while the Ordered Nested Logit model shows a sizeable decrease in sales of the upper segments, especially the standard segment which has cars that are just above the eligibility threshold. Green subsidies are usually temporary and naturally call for a dynamic approach to model consumers' decisions over time, which can be implemented only with additional information on the secondary market and the patterns of ownership (see Schiraldi, 2011). The Ordered Nested Logit model could be useful in a dynamic framework, as it entirely avoids the need of simulating the market share integral thus alleviating the computational burden of estimation.

The model I propose takes inspiration from the literature on Nested Logit models (Williams, 1977; Daly and Zachary, 1977; McFadden, 1978) and from the ordered generalized extreme value (OGEV) by Small (1987). The OGEV model was the first closed-form GEV model to allow for taste correlation between neighboring products. However, it 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. With market level data, such as a dataset on the car market, ordering around 100 car models in each market would prove impossible, while ordering groups of cars, the segments, is a sensible strategy to obtain a tractable model and flexible substitution patterns. Several other authors have tried to relax the hierarchical structure imposed by the Nested Logit, especially in the transportation literature; see Chu (1989); Vovsha (1997); Ben-Akiva and Bierlaire (1999). 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. Bresnahan, Stern and Trajtenberg (1997) develop a principle of differentiation model which is an example of a closed-form GEV model applied to market-level data. Davis and Schiraldi (2012) propose a fully analytic model capable of generating flexible substitution patterns.

The remainder of the paper is organized as follows. Section 2 puts forward the Ordered Nested Logit model and discusses its relation with other commonly used discrete choice models. A Monte Carlo study illustrates the flexibility of the model. Section 3 describes the application dataset and the econometric procedure, including the identification issues. Section 4 provides the empirical results and the implied price elasticities. Section 5 presents the policy counterfactuals. Section 6 concludes.

2 Modelling correlation between neighboring segments

The GEV family Demand is modelled within the discrete choice framework. Consider T markets, $t = 1, .., T$, with L_t potential consumers in each market. Markets are assumed to be independent, so I suppress the market subscript t to simplify notation. Each consumer i chooses a specific product j, $j = 0, \ldots J$. The outside good includes the option 'do not buy a product', $j = 0$ for which consumer i's indirect utility is $u_{i0} = \varepsilon_{i0}$. For products $j = 1, \ldots, J$, consumer i's indirect utility is:

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

where x_j is a vector of observed product characteristics and ξ_j is the unobserved product characteristic. 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.

The consumer-specific error term ε_{ij} is an individual realization of the random variable ε. The distribution of ε determines the shape of demand and the implied substitution patterns. McFadden (1978) has proposed a family of random utility models, the Generalized Extreme Value (GEV) family, in which those patterns can be modeled in different ways according to the specific behavioral circumstances. A GEV model is derived from a generating function $G = G(e^{\delta_0, \ldots, \delta_J})$, a differentiable function defined on $\mathbb{R}_+^J$: (i) which is non-negative; (ii) which is homogeneous of degree 1; (iii) tends toward $+\infty$ when any of its arguments tend toward $+\infty$; (iv) whose n^{th} cross-partial derivatives with respect to n distinct 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 product 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 is the market share of product j and G_j is the partial derivative of G with respect to e^{δ_j}.

The Ordered Nested Logit model Assume that the set of products j is partitioned into N mutually exclusive and collectively exhaustive nests. In addition, assume that those N nests are naturally ordered, with n increasing along its natural ordering: $n = 0, 1, ..., N$. The ordering may correspond to an increasing value of important characteristics such as price. I define the Ordered Nested Generalized Extreme Value model (in short, Ordered Nested Logit) as the model resulting from the following G function within the GEV class:

$$G = \sum_{r=0}^{N+M} \left(\sum_{n \in B_r} w_{r-n} \left(\sum_{j \in S_n} \exp \left(\frac{\delta_j}{1-\sigma_n} \right) \right)^{\frac{1-\sigma_n}{1-\rho_r}} \right)^{1-\rho_r}, \tag{2}$$

where n is the nest to which the products belongs; M is a positive integer; $w_m \geq 0$ and $\sum_{m=0}^{M} w_m = 1$. The weight w_m is the allocation weight of a nest to a set of nests. The parameters σ_n and ρ_r are constants satisfying $0 \leq \rho_r \leq \sigma_n < 1$: those restrictions are necessary to satisfy the four conditions for function G to belong to the GEV family; Appendix A provides the proof.[1] Finally, define the subset of N nests as $B_r = \{S_n \in \{0, ..., N\} | r - M \leq n \leq r\}$. Each of the $(N + M)$ subsets contains up to $M + 1$ contiguous nests (and *all* the alternatives in those nest). Consider a simple example with four nests (three plus the outside good nest zero), six alternatives and $M = 2$:

$$j = \underbrace{0}_{S_0}; \underbrace{1, 3}_{S_1}; \underbrace{2, 4}_{S_2}; \underbrace{5}_{S_3}$$

Alternatives within a nest need not to be ordered, but nests are. In our example the subsets of nests are: $B_0 = \{S_0\}, B_1 = \{S_0, S_1\}$, $B_2 = \{S_0, S_1, S_2\}$, $B_3 = \{S_1, S_2, S_3\}$, $B_4 = \{S_2, S_3\}$, $B_5 = \{S_3\}$, where each nest S_n belongs to $M + 1$ different subsets. The degree of proximity between neighboring nests can be modelled flexibly as each subset of nests can have its own parameter ρ_r. The shape of the demand function crucially depends on the two parameters, σ_n and ρ_r, that parameterize the cumulative distribution of the error term ε. The first one, σ_n, corresponds to a pattern of dependence in ε across products sharing the same nest (as in the Nested Logit). The second one, ρ_r, corresponds to a pattern of dependence in ε across products belonging to neighboring nests. Consider, for example, the effect of a price shock to alternative one belonging to segment S_1. The dependence in ε measured by σ_n determines that a share of consumers, who had initially chosen alternative

[1] More precisely, the condition that the sum of the weights has to add up to one ($\sum_{m=0}^{M} w_m = 1$) is not a necessary condition for function G to belong to the GEV family. However, if the condition holds, weights can be interpreted as allocation parameters of nests to subsets of nests. The condition also ensures that the generating function G of the Ordered Nested model reduces to the Nested Logit model if $\rho_r = 0$ (see next paragraph).

one, will switch to another alternative in segment S_1. The dependence in ε measured by ρ_r determines that a share of consumers will switch to the neighboring segments: in our example, with $M = 2$, the neighboring segments are S_0, S_2 and S_3.

If the random components follow the G function in (2), by the GEV postulate in (1) the choice probability of buying product j is:

$$s_j = \sum_{r=n}^{n+M} s(j|n) \cdot s(n|B_r) \cdot s(B_r), \tag{3}$$

where:

$$\begin{aligned}
s(j|n) &= \frac{\exp\left(\frac{\delta_j}{1-\sigma_n}\right)}{Z_{k\in S_n}}, \\
s(n|B_r) &= \frac{w_{r-n} Z_n^{\frac{1-\sigma_n}{1-\rho_r}}}{\exp\left(I_r\right)}, \\
s(B_r) &= \frac{\exp\left((1-\rho_r)I_r\right)}{\sum_{s=1}^{N+M} \exp\left((1-\rho_s)I_s\right)}, \\
Z_n &= \sum_{j\in S_n} \exp\left(\frac{\delta_j}{1-\sigma_n}\right), \\
I_r &= \ln \sum_{n\in B_r} w_{r-n} Z_n^{\frac{1-\sigma_n}{1-\rho_r}}.
\end{aligned}$$

2.1 The Ordered Nested Logit versus other GEV models

The Nested Logit model To clarify the logic of the modeling strategy for the Ordered Nested Logit, consider the G function associated with a traditional specification, the Nested Logit model, in which the ordering of the segments is not explicitly modeled. The model incorporates potential correlation among products only within a nest (segment), not between nests. The J alternatives are grouped into N nests labeled $S_0, ..., S_N$. The G function takes the form:

$$G = \sum_{n=0}^{N} \left(\sum_{j\in S_n} e^{\frac{\delta_j}{1-\sigma_n}} \right)^{1-\sigma_n}, \tag{4}$$

where σ_n captures correlation among products within the same nest. Consistency with random utility maximization requires σ_n to lie in the unit interval. In the Nested Logit model only alternatives belonging to the same nest have the stochastic terms that are correlated, and such correlation depends on σ_n. The condition $\sum_{m=0}^{M} w_m = 1$ ensures that the generating function G of the Ordered Nested model in (2) reduces to the Nested Logit model in (4) if

$\rho_r = 0$. If, in addition, $\sigma_n = 0$ for all nests, the model becomes the standard Logit in which each element of ε is independent.

Following Berry (1994), I can write the choice probability of a product j for the Nested Logit model as follows:

$$s_j = s(j|n) \cdot s(n) \tag{5}$$

where:

$$\begin{aligned} s(j|n) &= \frac{\exp\left(\frac{\delta_j}{1-\sigma_n}\right)}{Z_{k\in S_n}}, \\ s(n) &= \frac{Z_n^{1-\sigma_n}}{\exp(I_n)}, \\ Z_n &= \sum_{j\in S_n} \exp\left(\frac{\delta_j}{1-\sigma_n}\right), \\ I_n &= \ln \sum_{n=0}^{N} Z_n^{1-\sigma_n}. \end{aligned}$$

Compare the market shares of the Ordered Nested Logit model in (3) with the market shares of the one-level Nested Logit model in (5): similarly to the one-level Nested Logit model, in the Ordered Nested Logit model s_j is diminished by the presence of attractive alternatives within a nest n. Differently from the Nested Logit model, s_j is also diminished by the presence of attractive alternatives in neighboring nests B_r. Ceteris paribus, this effect is increasing in ρ_r: one may expect that if the values of σ_n and ρ_r are sufficiently high, products belonging to the same segment or to neighboring segments will be closer substitutes compared to products belonging to further segments.

The OGEV model The OGEV model derived by Small (1987) is based on the following G function (see Definition 1 in Small, 1987):

$$G = \sum_{r=0}^{J+M} \left(\sum_{j\in B_r} w_{r-j} \exp\left(\frac{\delta_j}{1-\rho_r}\right)\right)^{\rho_r},$$

where M is a positive integer; the weights w_m are overlapping parameters for alternatives; the parameter ρ_r indicates the covariance between alternatives, rather than nests as in our model, and B_r is a subset of alternatives, not nests.

The OGEV model responds to different modelling needs with respect to the Ordered Nested Logit: the OGEV is designed when individual-level data are available, in which a limited number of alternatives can be naturally ordered. The Ordered Nested Logit model is designed for situations in which numerous alternatives are present. Groups of those alternative can be naturally ordered, while alternatives in each group need not to be ordered.[2]

The Generalized Nested Logit model The Ordered Nested Logit model can be viewed as a special case of the Generalized Nested Logit (GNL) by Wen and Koppelman (2001). Recall the generating function of the GNL model:

$$G = \sum_{k=0}^{K} \left(\sum_{j \in S_k} \left(\alpha_{jk} \exp\left(\delta_j\right)\right)^{\frac{1}{1-\rho_k}} \right)^{1-\rho_k},$$

where S_k is the set of all alternatives included in nest k, α_{jk} is the allocation parameter which is the portion of alternative j assigned to nest k.

The Ordered Nested Logit model can be written as a special case of the GNL if (i) alternatives are positioned in the nest to which they originally belong, so $S_n = \{j \in S_n\}$; (ii) all the alternatives in neighboring nests are put together in a nest B_r formed by combinations of nests in ordered position: $B_r = \{S_n \in \{0, ..., N\} | r - M \leq n \leq r\}$; (iii) the weights or allocation parameters α_{jk} are equal for all alternatives in one nest B_r. Hence, weights are associated to the nest B_r rather than its alternatives.

Summary The Ordered Nested Logit model generalizes the Nested Logit model by capturing asymmetric interactions across nests. It differs from the OGEV model by Small (1987) because it is designed to capture asymmetric interactions across *nests,* not across *alternatives.* Hence, it does not impose an order across alternative, but across groups of alternatives (nests). The Generalized Nested Logit model by Wen and Koppelman (2001) is the most general instance of GEV model, but the complexity of the model requires normalization assumptions to identify the parameters and constraints to make the estimation feasible: see Bierlaire (2006). The Ordered Nested Logit includes an ordered nesting structure motivated by features commonly found in differentiated product markets: those restrictions render the model easy to handle for estimation.

[2]The Ordered Nested Logit model also differs with respect to the nested version of the OGEV model described by Small (1994) and Bhat (1998), which is similar to a nested logit except that at the lower node the alternatives (not segments) are grouped according to the OGEV model rather than the standard logit.

2.2 Monte Carlo experiment

The Ordered Nested Logit model is appealing for its closed form formulation and for its ability to capture more complicated substitution patterns than the Nested Logit. As a first step to test the benefits of the Ordered Nested Logit model, I perform a Monte Carlo experiment in which I generate datasets according to a Random Coefficient Logit model and fit the Logit, Nested Logit and Ordered Nested Logit models. I then assess the flexibility of the misspecified models by checking their ability to approximate the true elasticities.

Specification I generate a dataset for $T = 50$ independent markets consisting of $J = 100$ products and one outside good. Each product j is described by seven vectors of observed characteristics consisting of a constant, four dummies, and two exogenous continuous characteristics: $x_{jt} = (1, d^1_{jt}, d^2_{jt}, d^3_{jt}, d^4_{jt}, x^S_{jt}, x^P_{jt})$.The continuous variable x^S_{jt} is the basis upon which products are assigned into segments: think of the variable size in a car dataset. The variable x^P_{jt} intends to mimic the variable price in a non-simulated dataset. The vector ξ_{jt} of unobserved product characteristics is drawn from a standard normal distribution.[3] Products are partitioned into five nests on the basis of the size variable x^S_{jt}; the four dummies correspond to each product's nest (except one). In most markets, low nests with cheaper products tend to be more crowded by a larger number of competitors than higher nests. To mimic this feature, the lowest nest (grouping products with lower values of the continuous characteristic) contains twice as many products with respect to the contiguous nest and so on. I assume that price (x^P_{jt}) and size are drawn from a multivariate normal truncated below zero (so that all elements are positive). The correlation between price and size can be set at different levels to vary the degree of overlap between segments: while size is always informative of the segment, a lower nest can contain some items with a higher price, depending on the level of correlation. Finally, I specify a random coefficient for the size characteristic x^S_{jt}. The mean valuations for the constant, the four dummies and the two continuous characteristics are $\beta = (-5, 0, 0, 0, 0, 1, -2)$.[4]

I assume that data is generated by a Random Coefficient Logit process. The random coefficient for the characteristic x^S_{jt} is $\beta^S_i = \beta^S + \sigma\nu^S_i$,where ν^S_i is a random variable vector drawn from a normal distribution with zero mean and unit variance and σ is the standard deviation around the mean valuation β^S. In the Monte Carlo, the standard deviation of the characteristic x^S_{jt} is set at 2.

[3] For simplicity I discuss a data generating process in which there is no endogenous product characteristic. Extending the set-up to the case of one endogenous characteristic would imply an increase in the number of instrumental variables, but would not add to the substantial findings.

[4] Coefficients are chosen to ensure that the outside good is between 0.5 and 0.7.

The market share equation is given by the logit choice probability integrated over the individual-specific valuations for the continuos characteristic x_{jt}^S :

$$s_{jt}(\delta_{jt}, \sigma) = \int \frac{\exp(\delta_{jt} + x_{jt}^S \sigma \nu_i^S)}{1 + \sum_{l=1}^J \exp(\delta_{lt} + x_{lt}^S \sigma \nu_i^S)} dP_\nu(\nu), \tag{6}$$

where $\delta_{jt} \equiv x_{jt}\beta + \xi_{jt}$ identifies the mean utility term. Following Heiss and Winschel (2008), the market shares in (6) are approximated using a polynomial-based Sparse Grid quadrature rule as follows:

$$s_{jt}(\delta_{jt}, \sigma) \approx \sum_{i=1}^m \phi_i \frac{\exp(\delta_{jt} + x_{jt}^S \sigma \nu_i^S)}{1 + \sum_{l=1}^J \exp(\delta_{lt} + x_{lt}^S \sigma \nu_i^S)},$$

where m is the number of nodes for ν and ϕ_i is the weight associated with the node ν_i^S.

I use the simulated data to estimate a Logit, a Nested Logit, an Ordered Nested Logit and a Random Coefficient Logit model. In all models I use the same set of optimal instruments generated from within the model, following the approach of Chamberlain (1987) and Reynaert and Verboven (2014). Both in the Nested Logit and in the Ordered Nested Logit, the nesting parameters (σ_d) are constrained to be the same for all nests. In addition, for the Ordered Nested Logit I assume $M = 2$ (the two immediately proximate segments are neighbors), $\rho_r = \rho$ for all r, and $w_m = 1/(M + 1) \equiv 1/3$.

Finally, for each simulation, I minimize the GMM objective function using Knitro 10.3 in Matlab and tight convergence criteria for the contraction mapping (1e-12) and the gradient (1e-6).

Results Table B.1 in the Appendix shows the estimated demand parameters for a set-up in which the correlation between the price variable x_{jt}^P and the size variable x_{jt}^S is set at 0.9, which sets the correlation between the price variable and the nest identifier at around 0.7. The parameter estimates of the correctly specified model, the Random Coefficients Logit, are, not surprisingly, very close to the true values. In the Nested Logit model, the nesting parameter is estimated at $\sigma_d = 0.7$: correlation in consumer preferences within each nest is therefore sizeable. The Ordered Nested Logit shows that correlation between neighboring nests is also important ($\sigma_d = 0.9; \rho = 0.89$).

The implications of the parameter estimates are most clearly illustrated by looking at the nest (segment) level price elasticities. Table 1 represents the effect of a 1% increase in price (the continuous variable x_{jt}^1) of all products in one nest on the market shares of the other nests. For example, under the correctly specified Random Coefficient Logit model, the model predicts that, if the price of goods in nest 4 increases by 1%, consumers will be more likely to buy a product from a contiguous nest (sales in nests 3 increase by 0.2%) rather than buying a cheap product (sales in nest 1 increase by only 0.08%). By construction, both the Logit and the Nested Logit models imply fully symmetric substitution patterns, namely identical cross-price elasticities. In addition, the Nested Logit model underestimates the cross-price elasticities. In contrast, the Ordered Nested Logit does a much better job in approximating such asymmetric substitution patterns. In conclusion, the Nested Logit implies biased cross-price elasticities because the model does not have the flexibility to deal with overlapping nests as well as the Ordered Nested Logit. This result is robust to a variety of designs for the Monte Carlo experiment: the level of the elasticities may change, but the Ordered Nested Logit is always better than the Nested Logit in approximating the flexibility in substitution patterns implied by Random Coefficient Logit model.

Table 1: Monte Carlo Segment Elasticities: Ordered Nested vs Random Coefficients Logit

True	Nest 1	Nest 2	Nest 3	Nest 4	Nest 5
Nest 1	-0.434	0.069	0.060	0.047	0.021
Nest 2	0.095	-0.771	0.120	0.113	0.065
Nest 3	0.118	0.172	-1.041	0.222	0.155
Nest 4	0.083	0.144	0.200	-1.837	0.263
Nest 5	0.135	0.286	0.496	0.889	-1.623
Logit					
Nest 1	-0.431	0.063	0.063	0.063	0.063
Nest 2	0.088	-0.766	0.088	0.088	0.088
Nest 3	0.128	0.128	-1.076	0.128	0.128
Nest 4	0.119	0.119	0.119	-1.904	0.119
Nest 5	0.481	0.481	0.481	0.481	-3.411
Nested Logit					
Nest 1	-0.140	0.020	0.020	0.020	0.020
Nest 2	0.029	-0.251	0.029	0.029	0.029
Nest 3	0.042	0.042	-0.353	0.042	0.042
Nest 4	0.039	0.039	0.039	-0.623	0.039
Nest 5	0.158	0.158	0.158	0.158	-1.117
Ordered Nested Logit					
Nest 1	-0.222	0.073	0.032	0.007	0.007
Nest 2	0.110	-0.416	0.136	0.046	0.010
Nest 3	0.076	0.231	-0.559	0.170	0.070
Nest 4	0.014	0.085	0.195	-1.219	0.287
Nest 5	0.053	0.053	0.305	1.409	-1.082
Random Coefficient Logit					
Nest 1	-0.434	0.068	0.059	0.047	0.021
Nest 2	0.095	-0.767	0.121	0.113	0.064
Nest 3	0.118	0.172	-1.036	0.222	0.153
Nest 4	0.083	0.143	0.199	-1.834	0.258
Nest 5	0.133	0.283	0.492	0.882	-1.608

The table reports the segment-level own- and cross-price elasticities (when all products in the same segment raise their price by 1%). The segment-level elasticities are based on the parameter estimates in Table B.1 from 100 random samples of 50 markets and 100 products for each DGP.

3 Empirical study

3.1 Data

We now turn to the application of the Ordered Nested Logit to the automobile market. For the empirical study, I combine different datasets. The main one is a dataset on the automobile market provided by a marketing research firm, JATO: it includes essentially all transactions of passenger cars sold between 1998 and 2011 in the three largest European car markets: France, Germany, and Italy. The data is highly disaggregated, and I aggregate it at the level of the car model (nameplate), e.g. Volkswagen Golf. For each car model/country/year, I have information on sales, prices and various characteristics such as vehicle size (curb weight, width and height), engine attributes (horsepower and displacement), fuel consumption (liter/100 km or €/100 km), emissions, the brands' specific perceived country of origin, and, for models introduced or eliminated within a given year, the number of months with positive sales. The dataset is augmented with macro-economic variables including the number of households for each country, fuel prices and GDP. Low-sold car models, which are more susceptible to recording or measurement errors, as well as non-passenger cars, such as pickups and large vans, are removed. I also exclude minivans, sports cars and sport utility vehicles because they do not naturally fit in a univocal ordering of the segments: for example, sports cars are on average more powerful but not more expensive than luxury cars. The resulting dataset consists of 5,788 model/country/year observations or, on average, about 138 models per country/year.

Prices are list prices including value added taxes and registration taxes which differ across countries and engines: such information comes from the European Automobile Manufacturers Association. Prices are also corrected to account for active scrapping schemes and feebate programs according to the eligibility criteria for each vehicle: information on those programs comes from IHS Global Insight (an automotive consultant) and the European Automobile Manufacturers Association. Finally, the dataset is augmented with information on the location of the main production plant for each car model (from PWC Autofacts), and three input prices by country of production: unit labor costs, steel prices, and a producer price index. Table 2 presents summary statistics for sales, price, and vehicle characteristics used in demand estimation.

Starting from JATO's classification, I attribute each model to a marketing segment. I define five segments: subcompact, compact, standard, intermediate, and luxury.[5] Cars belonging to the same segment share similar characteristics in terms of price, horsepower,

[5] For example, a Volkswagen Golf belongs to the compact segment. The smaller Polo belongs one segment below the Golf (subcompact), while the bigger Passat is located one segment above (intermediate).

fuel consumption and size. Segmentation is used by carmakers to position their vehicle in the market place: they often advertise their vehicle as the cheapest or best performing in its class. Leading automotive magazines, such as *Auto motor und sport*, award a 'best car' prize for each segment. Comparison websites and consumer reports also feature the classification into segments as a prominent search tool. But the boundaries between segments are blurred by the presence of cars with some characteristics, including price, image and extra accessories, which would position those cars in an upper segment. Audi A1 or BMW Mini are examples of 'luxury subcompacts' designed to compete across segments. Table 3 and Figure 1 provide a descriptive illustration of segmentation in the car market. The top panel of the table presents the mean and standard deviation of price, horsepower, fuel consumption, and size by segment. Figure 1 represents also the median, the minimum and maximum values, and the values of the lower and upper quartiles of those characteristics. The table and the figure illustrate that the mean and median values of all characteristics increase from subcompact to luxury (with the exception of size from the intermediate to standard segment). At the same time, the large variability displayed by those characteristics within a segment suggest that overlap across segments is plausible and depends on the proximity of the ordering. The bottom panel of Table 3 shows how well characteristics predict to which segment each car model belongs. Classifications are reasonably accurate (always above 80% with one exception), but the prediction power is not perfect and confirms the need to quantify the presence of neighboring segment effects.

Table 2: Summary Statistics

	Mean	Std. Dev.
Sales (units)	13,821	24,312
Price/Income	0.84	0.50
Power (in kW)	82.19	35.72
Fuel efficiency (€/100 km)	7.27	1.46
Size (m^2)	7.46	1.14
Foreign (0-1)	0.78	0.42
Months present (1-12)	9.66	2.61

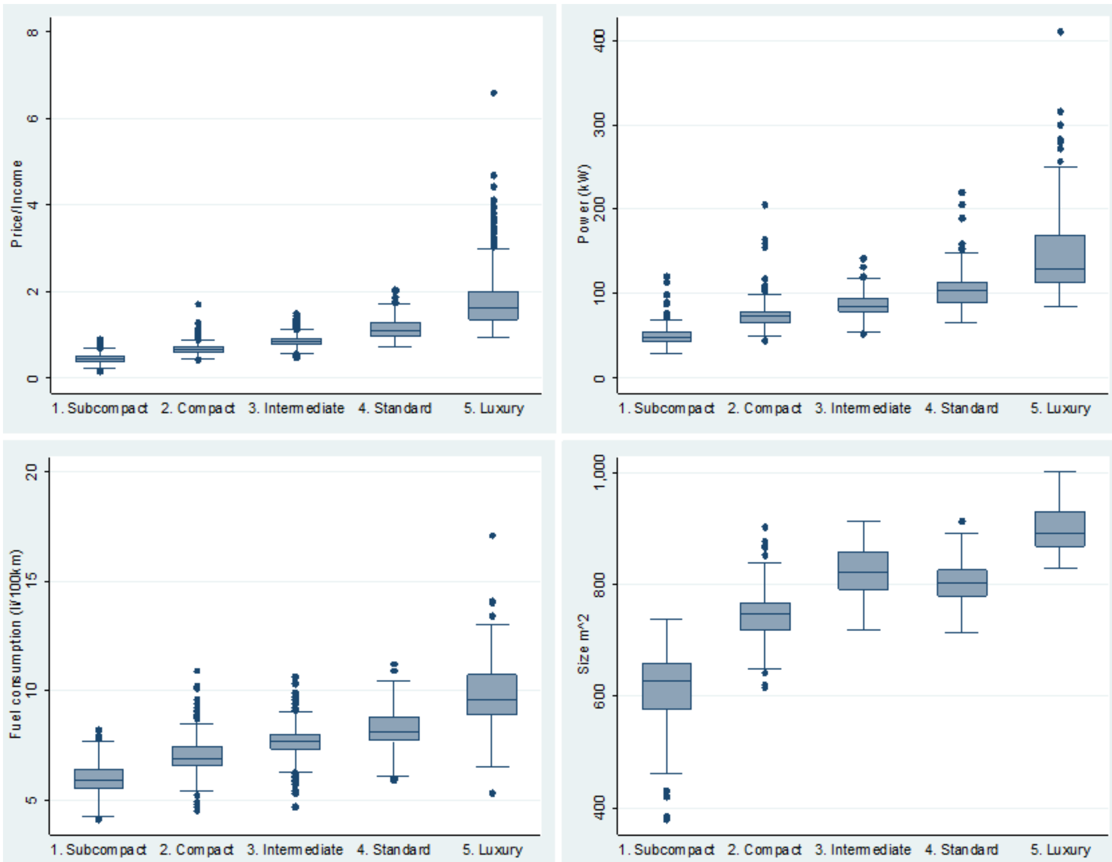
The table reports means and standard deviations of the main variables. The total number of observations (models/markets) is 5,788, where markets refer to the 3 countries and 14 years.

Table 3: Summary Statistics by Segment

		Subcomp.	Compact	Interm.	Standard	Luxury
Price/Income	Mean	0.45	0.68	0.86	1.13	1.80
	Std. Dev.	(0.10)	(0.12)	(0.13)	(0.22)	(0.67)
Power	Mean	50.37	73.84	88.52	104.91	145.47
(kW)	Std. Dev.	(10.83)	(14.16)	(12.96)	(19.48)	(42.44)
Fuel consumption	Mean	5.90	6.98	7.69	8.21	9.65
(li/100km)	Std. Dev.	(0.71)	(0.72)	(0.63)	(0.88)	(1.34)
Size	Mean	6.10	7.46	8.23	8.06	9.00
(m2)	Std. Dev.	(0.71)	(0.43)	(0.41)	(0.34)	(0.38)
Number of obs.		1,802	1,409	1,131	716	730
		Correct classifications into segments (percent)				
Subcompact		-	92.37	97.28	98.89	100.00
Compact			-	74.59	89.80	96.27
Intermediate				-	81.20	90.74
Standard					-	84.72
Luxury						-

The top panel of the table reports means of the main variables per segment in the top panel. The bottom panel of the table reports the percentage of correctly classified car models, based on binary logit of a segment dummy per pair on four continuous characteristics (i.e. power, fuel efficiency, width and height). Subcomp.=subcompact, Interm=intermediate.

Figure 1: Characteristics by Segment



The figure reports the median, the minimum and maximum values, and the values of the lower and upper quartiles by segment of the following vehicle characteristics: price/income, power, fuel consumption, size.

3.2 Specification

To estimate the demand for cars in France, Germany, and Italy, I modify the Ordered Nested Logit specified above. In each period (year) and country t, L_t potential consumers choose one alternative, either the outside good $j = 0$ or one of the J cars. Following Berry (1994) and the subsequent literature, price is treated separately because it is an endogenous characteristic. Hence, the utility specification becomes:

$$U_{ijt} = x_{jt}\beta - \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt} \equiv \delta_{jt} - \alpha_i p_{jt} + \varepsilon_{ijt},$$

where x_{jt} is a $1 \times K$ vector of characteristics including price, horsepower, fuel consumption, size measures (width and height), and a dummy variable for the country of origin. For the potential market size (L_t), I follow the literature and use the total number of households in each year and market.

In estimation, the coefficient of price, α_i, is specified in two ways: (i) $\alpha_i = \alpha/y$, where y is equal to income per capita; (ii) $\alpha_i = \alpha/y_i$, a specification in which I exploit information on income distribution.

The error term ε_{ij} is the individual realization of the random variable ε : as discussed above, its distribution determines the substitution patterns. From the insights offered by industry sources, I assume that the $5 + 1$ nests (segments) are ordered ordered as follows: S_0, the outside good; S_1, subcompact; S_2, compact; S_3, standard; S_4, intermediate; and S_5, luxury. The ordering corresponds to an increasing value of observable and unobservable characteristics such as price, size or comfort. The outside good nest is the nest with the 'inferior quality' good. The industry and the European Commission[6] have at times used more detailed classifications, for example by distinguishing the subcompact segment between city/mini cars and small cars (segment A and B). When using more detailed classifications, I found that the model was always not supported in the data ($\rho_r > \sigma_n$).

The distribution of the error term ε_{ij} thus follows the assumptions of the Ordered Nested Logit as defined in equation (2). In particular, I assume that: (i) $M = 2$: each nest has the two contiguous nests as neighbors, or, in other words, each nest belongs to 3 different subsets of nests; (ii) all nests have the same weight $1/(M + 1) = 1/3$; the nesting parameter σ_n is allowed to differ across nests; the parameter determining the degree of proximity between neighboring nests is constrained to be the same across subsets of nests: $\rho_r = \rho$.

I experimented with different assumptions. The assumption $M = 1$ (correlation only between pairs of nests) did not find support in the data because it resulted in a neighboring nest parameter (ρ) significantly higher than σ_5, the nesting parameter of the luxury nest. I also tried different methods to attribute the weights with robust results.

3.3 The estimation procedure

The estimation procedure for the Ordered Nested Logit model follows the methodological lines of Berry (1994), Berry, Levinsohn and Pakes (1995) and the subsequent literature. I exploit the panel features of the dataset to specify 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 full set of country/year fixed effects and a set of dummy variables for the number of months each model was available

[6] Case No COMP/M.1406 -HYUNDAI/KIA, available at mergers decision case 1406

in a country within a given year (for models introduced or dropped within a year). $\Delta\xi_{jt}$ is the remaining product-related error term.

The estimation procedure is standard in the literature. First, I numerically solve for the error term $\Delta\xi_{jt}$ as a 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 Ordered Nested Logit model $\Delta\xi_{jt}$ is the numerical solution of the system $s = s(\delta(\alpha_i, \sigma_s, \sigma_n), \alpha, \sigma_n, \rho)$. I use a modified version of Berry, Levinsohn and Pakes's (1995) contraction mapping: $\delta^{k+1} = \delta^k + [1 - \max(\widehat{\sigma_n}, \widehat{\rho})] \cdot [\ln(s_t) - \ln(s_t(\delta_t^k))]$. If one does not weigh the second term by $[1 - \max(\widehat{\sigma}_s, \widehat{\rho})]$ the procedure may not lead to convergence; see Appendix A in Grigolon and Verboven (2014).

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:

$$\min_{\alpha_i, \sigma_n, \rho} \widehat{\Delta\xi}'(Z\Omega Z')\widehat{\Delta\xi},$$

where Ω is the weighting matrix. I follow a two step-procedure: first I use the weighting matrix $\Omega = (Z'Z)^{-1}$. Then I re-estimate the model with the optimal weighting matrix. To minimize the GMM objective function with respect to the parameters α_i, σ_n, ρ, I first concentrate out the linear parameters β. Also, I do not directly estimate 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. Following Dubé, Fox and Su (2012), I use a tight tolerance level to invert the shares using the contraction mapping ($1e - 12$), check convergence for 10 starting values at each step, and check that the first order conditions are satisfied at convergence.

3.4 Identification

The GMM estimator requires an instrumental variable vector Z with a rank of at least $K + 7$ (K is the dimension of the β vector; the price parameter α; the five nesting parameters σ_n and the parameter characterizing correlation between neighboring nests ρ). 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 such 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 $\Delta\xi_{jt}$: parameters σ_n and ρ are special kinds of random coefficients

(Cardell, 1997). Berry and Haile (2014) clarify that, even abstracting from price endogeneity, identification of random coefficients requires instrumentation for the endogenous market shares: this calls for instrumentation of the share terms to avoid an upward bias on the parameters σ_n and ρ.

Following Berry, Levinsohn and Pakes (1995), I assume that the observed product characteristics x_{jt} are uncorrelated with the unobserved product characteristics $\Delta\xi_{jt}$, so product characteristics x_{jt} are included in the matrix of instruments. Note that this assumption is weaker than the often adopted assumption that x_{jt} is uncorrelated with ξ_{jt}.

I include three sets of moment conditions. The first set focuses on the identification of the price coefficient. Armstrong (2016) suggests the use of cost-shifters, especially when the number of products is large, to identify price effects. I use input prices derived from the country of production of each car: a steel price index interacted with car weight (as a proxy for material costs) and unit labor costs in the country of production.

The second set of instruments, often used in the literature, includes interactions of the exogenous characteristics. In particular, I use (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; (iii) counts and sum of the characteristics of other products of competing firms by a subset of segments B_r; (iv) counts and sum of the characteristics of other products of the same firm by a subset of segments B_r. 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.

Following Gandhi and Houde (2016), the third set of instruments is the difference in car attributes to capture the relative position of each product in the characteristic space. Those instruments approximate the optimal instrumental variables I used in the Monte Carlo without requiring initial estimates.[7] In particular I construct the sum of square of characteristic differences within each segment and within each subset of segments, B_r.

[7] In the Monte Carlo, I did not use approximation because I constructed the optimal instruments from the parameters and the functional form assumptions of the true data generating process.

4 Results

4.1 Demand estimates

Table 4 shows the parameter estimates for the three specifications. The first one is the one-level Nested Logit model, which imposes $\rho = 0$. The second specification is an Ordered Nested Logit with $M = 2$; both σ_s and ρ are estimated and the coefficient of price, α_i, is specified as α/y, where y is equal to income per capita of each country. The third specification is identical to the second one, except that it incorporates information on the empirical distribution of income within each country, so $\alpha_i = \alpha/y_i$. The specification shows that it is possible to incorporate random coefficients in the Ordered Nested Logit model. This strategy comes at the cost of losing the tractable closed-form solution for market shares, but can be reasonable to capture the features of the market under study.

In all three models, the price parameter (α_i) 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.

In the Nested Logit model, the nesting parameters are all precisely estimated; their magnitude is consistent with random utility maximization ($0 \leq \sigma_n < 1$) and (non-monotonically) decreases from subcompact to luxury: consumer preferences are more homogeneous for subcompact cars ($\sigma_1 = 0.95$) with respect to luxury cars ($\sigma_5 = 0.35$). This is consistent with earlier findings by Goldberg and Verboven (2001) and Brenkers and Verboven (2006).

In the second specification, the Ordered Nested Logit I, parameters σ_n are again precisely estimated and non-monotonically decreasing. The parameter capturing correlation between proximate nests is also precisely estimated and it indicates that correlation between neighboring segments is strongly supported by the data: $\rho = 0.61$ with a standard error of 0.08. Its magnitude is also consistent with random utility maximization ($0 \leq \sigma_n \leq \rho < 1$) with the exception of $\sigma_5 = 0.47$. However, the hypothesis that $\sigma_5 = \rho$ cannot be rejected (p-value 0.21), so in our counterfactual analysis I will set $\rho = \sigma_5$. In conclusion, the null hypothesis of $\rho = 0$ assumed by the Nested Logit is rejected against the alternative hypothesis of a more general Ordered Nested Logit model; in other words, the Nested Logit model is rejected against the more general Ordered Nested Logit model.

The third specification, the Ordered Nested Logit II, incorporating income distribution, presents parameter estimates that are very similar to the Ordered Nested Logit I. Again, the estimate of both σ_n and ρ are significantly different from zero and their magnitude is consistent with random utility maximization ($0 \leq \sigma_n \leq \rho < 1$). In sum, it is feasible to combine the nesting structure of the Ordered Nested Logit with random coefficients to obtain additional flexibility.

All models imply similar own-price elasticities; demand is always elastic, which is consistent with oligopolistic profit maximization.

Table 4: Parameter Estimates for Alternative Demand Models

	Nested Logit		Ordered NL I		Ordered NL II	
	Estimate	St.Error	Estimate	St.Error	Estimate	St.Error
	Mean valuations for the characteristics in $x_{jt}(\beta)$					
Price/Income	-1.43	0.17	-1.23	0.13	-1.06	0.21
Power (kW/100)	0.80	0.12	0.64	0.10	0.35	0.08
Fuel consumption (€/10,00km)	-0.72	0.10	-0.47	0.08	-0.56	0.08
Width (cm/100)	0.52	0.18	0.42	0.15	0.54	0.15
Height (cm/100)	1.13	0.16	0.89	0.12	0.97	0.13
Foreign (0/1)	-0.44	0.02	-0.34	0.02	-0.37	0.02
	Nesting parameters (σ_n)					
Subcompact	0.95	0.02	0.95	0.02	0.92	0.02
Compact	0.77	0.02	0.81	0.01	0.81	0.01
Intermediate	0.80	0.02	0.84	0.02	0.83	0.02
Standard	0.78	0.03	0.87	0.02	0.87	0.02
Luxury	0.35	0.07	0.48	0.06	0.47	0.06
	Neighboring Nesting Parameter (ρ)					
Neighboring Nests ρ	-		0.61	0.08	0.62	0.08
Model fixed effects	Yes		Yes		Yes	
Year*Country fixed effects	Yes		Yes		Yes	
Income distribution	No		No		Yes	
Own Elasticity	-6.931		-7.415		-4.980	

The table shows the parameter estimates and standard errors for the three demand models: (i) the Nested Logit model, which assumes homogenous income distribution ($\alpha_i = \alpha/y$) and set the neighboring segmentation parameter at zero ($\rho = 0$); (ii) the Ordered Nested Logit I with homogenous income distribution ($\alpha_i = \alpha/y$); (iii) the Ordered Nested Logit with heterogeneous income distribution ($\alpha_i = \alpha/y_i$). The total number of observations (models/markets) is 5,788, where markets refer to the 3 countries and 14 years. NL=Nested Logit.

4.2 Substitution patterns: segment-level price elasticities

As shown in the Monte Carlo simulation, the implications of rejecting the Nested Logit in favour of the Ordered Nested Logit model are most clearly illustrated by the implied substitution patterns at segment level. Table 5 presents own- and cross-price elasticities constructed by simulating the effect on demand of a joint 1% price increase of all cars in a given segment.

The own-price elasticities across the three models are similar in terms of magnitude and tend to be higher for the most expensive classes. This monotonic relationship between own-price elasticity and price is the result of the assumption that price enters utility linearly and is partially mitigated by modelling heterogeneity in consumer preferences for segments (σ_n and ρ) and income ($\alpha_i = \alpha/y_i$).

The cross-price elasticities are the most interesting results. By construction, the one-level Nested Logit model implies a fully symmetric substitution pattern, namely identical cross-price elasticities in each row. Thus, a 1% price increase to all subcompact cars raises demand in the compact and luxury segments by the same amount, 0.01%. By contrast, the Ordered Nested Logit model delivers more plausible substitution patterns. A 1% price increase in the subcompact segment has a stronger effect on demand of the two proximate segments: compact (+0.13%) and intermediate (+0.06%) compared to luxury (+0.01%). These numbers are comparable to the ones reported by Grigolon and Verboven (2014) in the analysis of the segment-level price elasticities for the random coefficients logit model. The Ordered Nested Logit model I estimated is rather flexible, but still parsimonious in the number of parameters, so that only the two immediately proximate segments (on the left and on the right) are the neighboring ones. Outside the neighboring segments, the Ordered Nested Logit model still retains the modeling assumptions of the Nested Logit model. Thus, substitution patterns are symmetric outside the neighboring segments.

The symmetry outside proximate segments does not hold in the third model, in which the Ordered Nested Logit model incorporates also a random coefficient on price. However, cross-price elasticities are still rather symmetric and similar to the ones implied by the Ordered Nested Logit I.

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

Nested Logit	Outside	Subcompact	Compact	Intermediate	Standard	Luxury
Subcompact	0.010	-0.593	0.010	0.010	0.010	0.010
Compact	0.015	0.015	-0.872	0.015	0.015	0.015
Intermediate	0.006	0.006	0.006	-1.204	0.006	0.006
Standard	0.009	0.009	0.009	0.009	-1.417	0.009
Luxury	0.011	0.011	0.011	0.011	0.011	-2.092
Ordered Nested Logit I						
Subcompact	0.010	-0.719	0.127	0.064	0.009	0.009
Compact	0.014	0.191	-1.046	0.208	0.114	0.013
Intermediate	0.006	0.041	0.087	-1.716	0.183	0.103
Standard	0.008	0.008	0.067	0.255	-1.843	0.316
Luxury	0.009	0.009	0.009	0.173	0.381	-2.424
Ordered Nested Logit II						
Subcompact	0.009	-0.666	0.109	0.052	0.008	0.007
Compact	0.012	0.165	-0.920	0.169	0.091	0.011
Intermediate	0.005	0.033	0.071	-1.419	0.151	0.073
Standard	0.006	0.007	0.053	0.208	-1.469	0.223
Luxury	0.007	0.007	0.008	0.119	0.264	-1.720

The table reports the segment-level own- and cross-price elasticities (when all products in the same segment raise their price by 1%). The elasticities are based on the parameter estimates in Table 4. They refer to Germany in 2011.

5 Counterfactuals

Entry of premium subcompact Since the early 2000s, luxury brands have entered the lower segments of the car market, such as subcompacts and compacts. The vehicles launched by those brands feature distinctive characteristics with respect to the incumbents: for their power, accessories, image, and, of course, price they resemble a vehicle from a higher segment. This trend has diluted the traditional borders between segments in the automobile market. I consider in particular three premium subcompacts: Audi A1, BMW Mini (both the hatchback and wagon versions) and the Fiat 500 Abarth, an upgraded version of the Fiat 500. Table B.2 in the Appendix presents summary statistics of the characteristics of those three vehicles compared to the average subcompact and compact car. Their price and horsepower are significantly higher, while there is no statistically significant difference in fuel consumption and size with respect to the average subcompact car. In contrast, with respect to the average compact car, only size is significantly lower.

I simulate a counterfactual scenario without those three premium subcompacts. Table 6 summarizes the implied diversion ratios by segment. Those ratios measure the fraction of sales diverted to other products, in the same segment or other segments, when the premium subcompacts are removed. In the simulation I account for the response of other carmakers by solving the differentiated product model for the change in equilibrium prices induced by the removal of the products. The Nested Logit model suggests that, absent the choice of premium subcompacts, 95% of sales would be diverted to other subcompact cars, while sales of upper segments would practically not be affected. The Ordered Nested Logit I, which allows for the possibility of asymmetric correlation between neighboring nests, still predicts that most substitution (93%) happens within the subcompact segment, but now 1.2% of sales would be diverted to compact cars. In both cases, the diversion ratio to the outside good is around 5%. The Ordered Nested Logit II, which incorporates income heterogeneity as well, predicts that 85% of substitution happens within the subcompact segment, 2.6% of sales are diverted to the compact segment, and 12% to the outside good.

Table 6: Diversion Ratios After the Removal of Premium Subcompact Cars

Diversion ratios (%)	Nested Logit	Ordered Nested Logit I	Ordered Nested Logit II
Outside	5.17	5.12	11.70
Subcompact	94.60	93.42	85.10
Compact	0.10	1.21	2.56
Intermediate	0.03	0.19	0.45
Standard	0.04	0.03	0.10
Luxury	0.06	0.03	0.08

The table reports the diversion ratios (in percent) by segment after removing three premium subcompact car models: Audi A1, BMW Mini (both the hatchback and wagon versions) and the Fiat 500. Diversion ratios: share of fraction of sales diverted to other products in the same segment or other segments. The simulations are based on the parameter estimates in Table 4. They refer to Germany in 2011.

The effects of targeted environmental policies Asymmetric substitution patterns across segments are particularly important when looking at asymmetric policies. An example is a targeted scrapping scheme, which encourages consumers to scrap an old vehicle and purchase a cleaner one. The dataset comprises: (i) the 2009 German scrapping scheme, which was not targeted (it provided an incentive to purchase a new car regardless of its fuel efficiency); (ii) the 2008-2011 French scrapping scheme, which was targeted, and the feebate program (Bonus/Malus); (iii) various Italian scrapping schemes, which are mostly targeted but not sizeable.[8] The French scrapping scheme in combination with the feebate program is the only notably asymmetric policy, so I tested the predictions of the three models to the French environmental policy. In particular, I compare the market shares observed in 2007 (before the policy) and the simulated market shares of 2008 setting the environmental policy to zero and the fuel prices at the level of 2007. Table B.3 in the Appendix shows that the three models, though suffering from the limitations of a static framework, predict counterfactual shares that are very close to the observed ones. The Ordered Nested Logit models implies counterfactual shares similar to the ones produced by Nested Logit model because the asymmetry in the policy is actually rather limited. In practice, cleaner cars in the dataset mostly received only a modest rebate (€200), while polluting cars were mostly subject to a modest fee ranging from €200 to 750. Cars emitting more than 160g of CO_2 per kilometer would be subject to the sizeable fee of €2,600, but only two cars in the dataset meet the requirement.

What would be the effect of a bolder environmental policy? I simulate the impact of a €5,000 subsidy to cars emitting less than 140g of CO_2 per kilometer. The first column of Table 7 illustrates the asymmetry of the policy as it mostly benefits subcompact and compact cars. The other columns simulate the effect of the subsidy. As in the previous counterfactual, I account for the pricing responses of manufacturers. Under the Nested Logit model, subcompact cars gain a significant amount of sales (+ 24%). Sales increase, by a smaller amount, also for the compact and intermediate segments. Most importantly, standard and luxury cars are unaffected by the policy. The Ordered Nested Logit I model tells another story: sales of non-eligible cars, especially in the standard segment, are affected by the policy and decrease by 2.4%. The Ordered Nested Logit II predicts a similar decrease (2%).

[8] For more information, see Table A1 of Grigolon, Leheyda and Verboven (2016) and Table 1 of D'Haultfoeuille, Givord and Boutin (2014).

Table 7: The Effect of a Subsidy to Clean Cars on Market Shares

	Eligible cars %	% Change in Sales Nested Logit	Ordered Nested Logit I	Ordered Nested Logit II
Outside	-	-0.61	-0.59	-0.60
Subcompact	93.02	24.28	27.56	27.83
Compact	39.39	8.57	5.59	6.00
Intermediate	8.33	5.40	3.94	2.58
Standard	0.00	-0.61	-2.40	-2.06
Luxury	0.00	-0.65	-0.99	-0.82

The table reports the effect of €5,000 subsidy to cars emitting less than 140g of CO_2. The simulations are based on the parameter estimates in Table 4. They refer to Germany in 2011.

6 Conclusion

I present a new member of the GEV model family denominated Ordered Nested Logit model. The Ordered Nested GEV 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, than a simple Nested Logit model. It creates the potential for neighboring segment effects, or, more precisely, asymmetric substitution patterns across segments. Second, the 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 of the random coefficients logit model. Third, the Ordered Nested GEV 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 apply the Ordered Nested Logit 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. I show that asymmetry in substitution matters when simulating the introduction of vehicles combining features from different segments, such as premium subcompacts, or when studying the consequence of asymmetric policies, such as targeted subsidies.

The model I propose here can be a promising starting point to capture neighboring segment effects. Future research on other industries could benefit from this modeling strategy: ordering a high number of alternatives can prove impossible, but ordering groups of these alternatives may represent a sensible way to obtain flexible substitution patterns in a tractable setting.

References

Armstrong, Timothy B. 2016. "Large Market Asymptotics for Differentiated Product Demand Estimators With Economic Models of Supply." *Econometrica*, 84(5): 1961–1980.

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

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

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

Berry, Steven T., and Philip A. Haile. 2014. "Identification in Differentiated Products Markets Using Market Level Data." *Econometrica*, 82: 1749–1797.

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

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

Bierlaire, Michel. 2006. "A theoretical analysis of the cross-nested logit model." *Annals of Operations Research*, 144(1): 287–300.

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

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

Cardell, Scott N. 1997. "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity." *Econometric Theory*, 13(02): 185–213.

Chamberlain, Gary. 1987. "Asymptotic efficiency in estimation with conditional moment restrictions." *Journal of Econometrics*, 34(3): 305 – 334.

Chu, C. 1989. "A Paired Combinational Logit Model for Travel Demand Analysis." Vol. 4, 295–309. Venturea.

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

Davis, Peter, and Pasquale Schiraldi. 2012. "The Flexible Coefficient Multinomilal Logit (FC-MNL) Model of Demand for Differentiated Products." Mimeo.

D'Haultfoeuille, Xavier, Pauline Givord, and Xavier Boutin. 2014. "The Environmental Effect of Green Taxation: The Case of the French Bonus/Malus." *The Economic Journal*, 124(578): F444–F480.

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

Gandhi, Amit, and Jean-François Houde. 2016. "Measuring Substitution Patterns in Differentiated Products Industries." Cornell Working Paper.

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

Grigolon, Laura, and Frank Verboven. 2014. "Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation." *The Review of Economics and Statistics*, 96(5): 916–935.

Grigolon, Laura, Nina Leheyda, and Frank Verboven. 2016. "Scrapping subsidies during the financial crisis: Evidence from Europe." *International Journal of Industrial Organization*, 44(C): 41–59.

Heiss, Florian, and Viktor Winschel. 2008. "Likelihood approximation by numerical integration on sparse grids." *Journal of Econometrics*, 144(1): 62–80.

Kalouptsidi, Myrto. 2012. "From market shares to consumer types: Duality in differentiated product demand estimation." *Journal of Applied Econometrics*, 27(2): 333–342.

Knittel, Christopher R., and Konstantinos Metaxoglou. 2008. "Estimation of Random Coefficient Demand Models: Challenges, Difficulties and Warnings." National Bureau of Economic Research Working Paper 14080.

Lee, Jinhyuk, and Kyoungwon Seo. 2015. "A computationally fast estimator for random coefficients logit demand models using aggregate data." *The RAND Journal of Economics*, 46(1): 86–102.

McFadden, Daniel L. 1978. "Modelling of Choice of Residential Location." In *Spatial Interaction Theory and Residential Location.* , ed. F. Snickers A. Karlquist, L. Lundquist and J. Weibull. Amsterdam: North Holland.

Reynaert, Mathias, and Frank Verboven. 2014. "Improving the performance of random coefficients demand models: The role of optimal instruments." *Journal of Econometrics*, 179(1): 83 – 98.

Schiraldi, Pasquale. 2011. "Automobile replacement: a dynamic structural approach." *The RAND Journal of Economics*, 42(2): 266–291.

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

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

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

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

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

A Appendix A

Proof GEV I show that under the assumptions that (i) M is a positive integer; (ii) σ_n and ρ_r are constants satisfying $1 > \rho_r \geq \sigma_g \geq 0$; (iii) $w_m \geq 0$ and $\sum_{m=0}^{M} w_m = 1$, the generating function G in (2) verifies the four properties of GEV generating functions. To simplify the notation, let $e^{\delta_j} = Y_j$.

1. G is non-negative since $Y_j \in \mathbb{R}_+ \forall j$ and the weigths are positive

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

$$\begin{aligned}
G(\lambda Y_0, ..., \lambda Y_J) &= \sum_{r=0}^{N+M} \left(\sum_{n \in B_r} w_{r-n} \left(\sum_{j \in S_n} \exp\left(\lambda^{\frac{1}{1-\sigma_n}} Y_j^{\frac{1}{1-\sigma_n}} \right) \right)^{\frac{1-\sigma_n}{1-\rho_r}} \right)^{1-\rho_r}, \\
& \sum_{r=0}^{N+M} \left(\sum_{n \in B_r} w_{r-n} \left(\lambda^{\frac{1}{1-\sigma_n}} \sum_{j \in S_n} \exp\left(Y_j^{\frac{1}{1-\sigma_n}} \right) \right)^{\frac{1-\sigma_n}{1-\rho_r}} \right)^{1-\rho_r}, \\
& \sum_{r=0}^{N+M} \left(\sum_{n \in B_r} w_{r-n} \lambda^{\frac{1}{1-\rho_r}} \left(\sum_{j \in S_n} \exp\left(Y_j^{\frac{1}{1-\sigma_n}} \right) \right)^{\frac{1-\sigma_n}{1-\rho_r}} \right)^{1-\rho_r}, \\
&= \lambda \sum_{r=0}^{N+M} \left(\sum_{n \in B_r} w_{r-n} \left(\sum_{j \in S_n} \exp\left(Y_j^{\frac{1}{1-\sigma_n}} \right) \right)^{\frac{1-\sigma_n}{1-\rho_r}} \right)^{1-\rho_r}, \\
&= \lambda G(Y_0, ..., Y_J).
\end{aligned}$$

3. The limit property holds since weights are non-negative and at least one is strictly positive (condition iii)

4. The property of the sign of the derivatives holds because $0 \leq \rho_r \leq \sigma_n < 1$ (condition ii). The first cross-derivative G_j is given by:

$$G_j = \sum_{r=n(j)}^{n(j)+M} \underbrace{Y_j^{\frac{\sigma_n}{1-\sigma_n}}}_{\geq 0} \cdot \underbrace{A_n^{\frac{\rho_r - \sigma_n}{1-\rho_r}}}_{\geq 0} \cdot \underbrace{B_r^{-\rho_r}}_{\geq 0},$$

where A_n and B_r are defined as follows:

$A_n = w_{r-n} \sum_{j \in S_n} Y_j^{\frac{1}{1-\sigma_n}}$,
$B_r = \sum_{n \in B_r} w_{r-n} \left(\sum_{j \in S_n} Y_j^{\frac{1}{1-\sigma_n}} \right)^{\frac{1-\sigma_n}{1-\rho_r}}$.
for $j \in B_r$. Since $Y_j > 0 \forall j$, $G_j \geq 0$ as required.

The second cross-derivative is given by:

$$G_{ji} = \sum_{r=n(j)}^{n(j)+M} \underbrace{Y_i^{-\frac{\sigma_{n(i)}}{1-\sigma_{n(i)}}} Y_j^{-\frac{\sigma_{n(j)}}{1-\sigma_{n(j)}}} \cdot A_{n(j)}^{\frac{\rho_r - \sigma_{n(j)}}{1-\rho_r}} A_{n(i)}^{\frac{\rho_r - \sigma_{n(i)}}{1-\rho_r}}}_{\geq 0} \cdot \left(\underbrace{-\frac{\rho_r}{1-\rho_r} B_r^{\rho_r - 2}}_{\leq 0} + \underbrace{\frac{\rho_r - \sigma_n}{(1-\rho_r) \cdot (1-\sigma_n)} B_r^{-\rho_r}}_{\leq 0} \right),$$

if $i, j \in S_n, i \neq j$. $G_{ji} \leq 0$ as required.

$$G_{ji} = \sum_{r=n(j)}^{n+M} \underbrace{Y_i^{-\frac{\sigma_{n(i)}}{1-\sigma_{n(i)}}} Y_j^{-\frac{\sigma_{n(j)}}{1-\sigma_{n(j)}}} \cdot A_{n(j)}^{\frac{\rho_r - \sigma_{n(j)}}{1-\rho_r}} A_{n(i)}^{\frac{\rho_r - \sigma_{n(i)}}{1-\rho_r}}}_{\geq 0} \cdot \left(\underbrace{-\frac{\rho_r}{1-\rho_r} B_r^{\rho_r - 2}}_{\leq 0} \right),$$

if $i, j \notin S_n$, and $i, j \in B_r, i \neq j$. $G_{ji} \leq 0$ as required.

For $i, j \notin S_n$ and $i, j \notin B_r$, $G_{ij} = 0$, which also meets the property. Higher cross-partial derivatives exhibits a similar path: the property holds if $0 \leq \rho_r \leq \sigma_n < 1$.

B Appendix B. Additional Tables

Table B.1: Monte Carlo Results: Parameter Estimates

Parameter	True	Logit	Nested Logit	Ordered Nested Logit	Random Coefficient Logit
x_{jt}^S	1.00	2.07	0.69	0.23	0.97
		(0.08)	(0.06)	(0.03)	(0.04)
x_{jt}^P	-2.00	-1.93	-0.63	-0.21	-1.99
		(0.04)	(0.04)	(0.03)	(0.04)
$\sigma_{x_{jt}^S}$	2.00				2.01
					(0.11)
σ_d	-	-	0.70	0.90	-
			(0.02)	(0.02)	
ρ	-	-	-	0.89	-
				(0.02)	

The table reports the empirical means and standard deviations (in parentheses) of the relevant model parameters: the continous characteristics x_{jt}^S and x_{jt}^P; the standard deviation of the characteristic x_{jt}^S ($\sigma_{x_{jt}^S}$), the nesting parameter (σ_d) and the neighboring nest parameter (ρ). The estimates are based on 100 random samples of 50 markets and 100 products. The true model is the Random Coefficient Logit model.

Table B.2: Summary Statistics Premium Subcompact vs Subcompact and Compact

	Premium Sub	Subcompact	p-value	Premium Sub	Compact	p-value
Price	19,038	13,039	0.000	19,038	19,468	0.771
Power (in kW)	87.75	53.62	0.000	87.75	80.18	0.138
Fuel efficiency (€/100 km)	5.68	5.23	0.131	5.68	6.22	0.054
Size (m^2)	6.44	6.24	0.612	6.44	7.87	0.000

The table reports the summary statistics of Premium Subcompact cars vs. Subcompact cars and Premium Subcompact vs Compact cars. It reports the means of four characteristic and the p-value of the difference of the means.

Table B.3: The Effect of Removing the French Feebate and Scrapping Subsidy

	2007 Observed	Nested Logit	Ordered Nested Logit I	Ordered Nested Logit II
Subcompact	57.32	58.69	58.70	58.50
Compact	25.30	25.26	25.33	25.50
Intermediate	10.50	10.00	9.99	10.04
Standard	4.13	4.09	4.03	4.05
Luxury	2.75	1.96	1.94	1.92

The table reports: (i) the 2007 observed market shares by segment (first column); (ii) the simulated market shares obtained from the 2008 market shares after setting the French feebate program and the scrapping subsidy to zero and using the fuel price of 2007. The simulations are based on the parameter estimates in Table 4.