KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and
Applied Economics

Department of Economics

Nested logit or random coefficients logit? A comparison of alternative discrete choice models of product differentiation

by

Laura GRIGOLON & Frank VERBOVEN

Econometrics

Center for Economic Studies
Discussions Paper Series (DPS) 11.24
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

September 2011

DISCUSSION
PAPER

Nested logit or random coefficients logit? A comparison of alternative discrete choice models of product differentiation

Laura Grigolon and Frank Verboven*

September 2011

Abstract

We start from an aggregate random coefficients nested logit (RCNL) model to provide a systematic comparison between the tractable logit and nested logit (NL) models with the computationally more complex random coefficients logit (RC) model. We first use simulated data to assess possible parameter biases when the true model is a RCNL model. We then use data on the automobile market to estimate the different models, and as an illustration assess what they imply for competition policy analysis. As expected, the simple logit model is rejected against the NL and RC model, but both of these models are in turn rejected against the more general RCNL model. While the NL and RC models result in quite different substitution patterns, they give robust policy conclusions on the predicted price effects from mergers. In contrast, the conclusions for market definition are not robust across different demand models. In general, our findings suggest that it is important to account for sources of market segmentation that are not captured by continuous characteristics in the RC model.

*_Laura Grigolon:_ University of Leuven. Email: Laura.Grigolon@econ.kuleuven.be. _Frank Verboven_: University of Leuven and C.E.P.R. (London). Email: Frank.Verboven@econ.kuleuven.be. We are grateful to Geert Dhaene and Øyvind Thomassen for useful comments and discussions. We also thank seminar participants at TSE, and conference participants at EARIE 2010 and ZEW 2010.

1 Introduction

Discrete choice models of product differentiation have gained considerable importance in empirical work. Because they treat products as bundles of characteristics, they offer the possibility to uncover rich substitution patterns with a limited number of parameters. Berry (1994) developed a framework to estimate a class of discrete choice models with aggregate sales data. His framework includes the logit and nested logit models, and the full random coefficients logit model of Berry, Levinsohn and Pakes (1995) (hereafter BLP).

The logit and nested logit models have been popular because of their computational simplicity, since they can be transformed to simple linear regressions of market shares on product characteristics. At the same time, they have long been criticized because they yield too restrictive substitution patterns. The logit model assumes that consumer preferences are uncorrelated across all products, implying symmetric cross-price elasticities. The nested logit model allows preferences to be correlated across products within the same group or "nest". It thus entails a special kind of random coefficients on group dummy variables (Cardell, 1997). It allows products of the same group to be closer substitutes than products of different groups, but the aggregate substitution patterns remain restrictive: cross-price elasticities within the same group are still symmetric, and substitution outside a group is symmetric to all other groups. In contrast, BLP's full random coefficients logit model incorporates random coefficients for continuously measured product characteristics (and not just for the group dummy variables in the nested logit model). This creates potentially more flexible substitution patterns, where products tend to be closer substitutes as they have more similar continuous characteristics. However, the random coefficients model is computationally more demanding, and several recent papers have studied a variety of problems relating to its numerical performance; see Knittel and Metaxoglou (2008), Dubé, Fox and Su (2011) and Judd and Skrainka (2011).

Against this background it is a particularly timely question whether and when the popular logit and nested logit models can be used as reasonable alternatives to the computationally more demanding full random coefficients logit model. In this paper we provide a systematic comparison between these demand models, and as an illustration assess how they perform in competition policy analysis. To accomplish this, we start from a general random coefficients nested logit model (RCNL) that covers the logit, nested logit (NL) and full random coefficients logit (RC) as special cases. The RCNL model thus includes both the random coefficients for continuously measured characteristics as in the RC model, and the random coefficients or "nesting parameters" for the group-specific dummy variables of the NL model. The RCNL model serves as a benchmark to assess the relative performance of the RC and

NL models.

To motivate our analysis, we begin with a simulation experiment for two data generating processes behind a RCNL model: one in which the groups or nests are good proxies for the continuous characteristics, and one in which they are not. We use the simulated datasets to compare the RCNL model with the misspecified logit, NL and RC models. We find that the NL model overestimates the nesting parameter when the groups are good proxies for the continuous characteristics. Furthermore, the RC model overestimates the random coefficient for the continuous variable.

We then turn to our main empirical analysis. We collected a unique dataset on the automobile market for nine European countries covering around 90% of the car sales in the European Union during 1998–2006. The market is commonly classified in various different segments (subcompact, compact, intermediate, standard, luxury, SUV and sports) and car manufacturers typically promote their models as belonging to one of these segments. Hence, the segments may proxy for observed product characteristics such as the size, engine performance and fuel efficiency. But it is also possible that they capture intrinsically unobserved features shared by different car models. Our dataset is therefore particularly interesting to compare the performance of the logit, NL, RC and RCNL models. Consistent with earlier findings, the logit model is rejected against both the NL and RC models. More importantly, in the general RCNL model the nesting parameters become quantitatively smaller (consistent with the results of our simulation experiment), but they remain highly significant and economically important. Furthermore, the random coefficients relating to car size become insignificant, while the random coefficients relating to engine power and fuel efficiency remain significant. These various findings suggest that the nesting parameters may proxy for random coefficients of some of the observed continuous characteristics, but also capture other unobserved dimensions of consumer preferences.

To illustrate the implications of our findings, we present own- and cross-price elasticities for the different models, and we perform policy counterfactuals common in competition policy: market definition and merger simulation. In terms of substitution patterns, the NL and RC model yield quite different results. In particular, there is much stronger substitution within segments in the NL model and much larger substitution to other (especially neighboring) segments in the RC model. Despite these different substitution patterns, merger simulations of two domestic mergers yields fairly robust conclusions across different demand models: while the simple logit clearly appears inappropriate, the NL, RC and RCNL all tend to give robust conclusions. In sharp contrast, the conclusions for market definition are not robust: the RC suggests a wide market definition at the level of all cars (similar to the logit), whereas the NL and RCNL suggest a more narrow definition at the level of the segments.

We draw two implications for competition policy. First, the lack of robustness in market definition should not be attributed to the RC model per se, but rather to the arbitrariness in selecting candidate markets (as segments) in the market definition approach. Second, the robustness in merger simulation suggests the simple NL model can be sufficient to obtain reliable policy conclusions, despite the different substitution patterns from the RC model.

More generally, one can draw two implications for the choice of demand model in applied work. First, the choice between the tractable NL model and the computationally more complex RC model may depend on the application. In our analysis of hypothetical domestic mergers consumer heterogeneity regarding the cars domestic/foreign origin is particularly relevant, and the NL model captures this reasonably well. In other applications, the most relevant aspects of consumer heterogeneity may not be captured well by nesting parameters for groups or subgroups. In these cases, it is appropriate to estimate RC models with random coefficients for the most relevant continuous characteristics.

Second, our results imply that it can be important to account for sources of market segmentation that are not captured by continuously measured product characteristics. Adding a nested logit structure to BLP's random coefficients model is a tractable way to accomplish this, since it gives closed-form expressions for the integrals in the choice probabilities. But one may also consider other tractable models from McFadden's (1978) generalized extreme value model (GEV). Examples are Small's (1987) model of ordered alternatives and Bresnahan, Stern and Trajtenberg's (1997) "principles of differentiation model", which allows for segmentation in more than one dimension without imposing a hierarchical structure. In principle, BLP's framework can of course also incorporate random coefficients on group dummies. But this is more complicated because it increases the dimensionality of the integrals that need to be simulated, and in practice it often proves difficult to estimate the coefficients as precisely as in the closed form GEV models. For example, Nevo (2001) estimates a rich demand model for the U.S. cereals market. His model includes three random coefficients for the segments (all-family, kids and adult), but two of these are estimated rather imprecisely.

Our comparison of alternative discrete choice models is timely for several related reasons. First, a few recent papers have thoroughly studied several (often commonly known) numerical difficulties with the aggregate random coefficients model. Knittel and Metaxoglou (2008) mainly focus on global convergence problems, in particular the role of starting values and different optimization algorithms. Dubé, Fox and Su (2011) focus on the properties of BLP's "inner loop" contraction mapping algorithm for inverting the market share system. They stress the importance of a tight convergence criterion for the contraction mapping, and suggest a mathematical program with equilibrium constraints (MPEC) as an alternative approach. Reynaerts, Varadhan and Nash (2010) explore alternative algorithms to the contracting mapping to invert the market share system. Judd and Skrainka (2011) focus on problems of pseudo-Monte Carlo integration to compute the aggregate market share system, in particular without variance reduction methods. They consider a variety of alternative integration methods. We draw from these findings in our own empirical analysis, by cautiously considering multiple starting values, using a tight inner loop contraction mapping and taking a large number of Halton draws for approximating the integrals.[1]

Second, there is a large and rapidly growing empirical literature estimating aggregate discrete choice models of product differentiation, with applications in industrial organization, international trade, environmental and public economics, marketing, finance, etc. A complete review of the applied aggregate discrete choice literature is beyond the scope of this introduction, so we limit attention here to early work. Much of this work has actually also looked at automobiles. Bresnahan (1981) and Feenstra and Levinsohn (1995) are important contributions preceding the seminal work of Berry (1994) and BLP. Verboven (1996) and Fershtman and Gandal (1998) are early applications of Berry's (1994) aggregate nested logit model. Nevo (2001), Petrin (2002) and Sudhir (2001) are early applications with interesting extensions of BLP's full random coefficients model. In recent years, academic work appears to focus more exclusively on the random coefficients models, whereas competition policy practitioners often use the logit and nested logit models. Our findings on the automobile market suggest that the nested logit model may not only be a reasonable approximation in competition policy, but also in other applications where the market segments are the most relevant differentiating dimensions, for example an analysis of trade liberalization. In contrast, applications on quality discrimination or environmental policy would warrant estimating BLP's random coefficients logit model, since the relevant random coefficients (engine power and fuel efficiency) are not well-captured by the nesting parameters.[2]

The rest of this paper is organized as follows. Section 2 presents the model and conducts Monte Carlo experiments. Section 3 uses the dataset for the European car market to estimate the logit, NL, RC and RCNL models and the implied price elasticities. Section 4 draws implications for competition policy analysis, applying market definition and merger simulation. Conclusions follow in section 5.

[1] We do not however consider Dubé, Fox and Su's (2011) alternative MPEC approach, because we have a large number of products/markets, implying a large number of nonlinear constraints in their constrained optimization algorithm. Nor do we pursue Judd and Skrainka's (2011) alternative integration methods here.

[2] Wojcik (2000) also compares the NL and RC model. She claims the NL model is likely to be superior, but Berry and Pakes (2001) raise serious methdological problems with her comparison. Our approach is rather different from Wojcik since we start from a more general model that covers the NL and RC models as special cases. Furthermore, we follow prediction excercises in the spirit of those advocated by Berry and Pakes (2001). Our conclusions are much more nuanced since we focus on identifying circumstances where the NL may, or may not, be a reasonable alternative.

2 The model

2.1 Demand

We consider a random coefficients nested logit model (RCNL) that contains the logit, nested logit (NL) and random coefficients logit (RC) as special cases. There are T markets, $t = 1, \ldots, T$. In each market t there are L_t potential consumers. Each consumer i may either choose the outside good 0 or one of the J differentiated products, $j = 0, \ldots, J$. Consumer i's conditional indirect utility for the outside good is $u_{i0t} = \overline{\varepsilon}_{i0t}$. For products $j = 1, \ldots, J$ it is

$$u_{ijt} = x_{jt}\beta_i + \xi_{jt} + \overline{\varepsilon}_{ijt}, \tag{1}$$

where x_{jt} is a $1 \times K$ vector of observed product characteristics (including price), β_i is a $K \times 1$ vector of random coefficients capturing the individual-specific valuations for the product characteristics, ξ_{jt} refers to unobserved product characteristics (to the econometrician), and $\overline{\varepsilon}_{ijt}$ is a remaining individual-specific valuation for product j.

The random coefficients vector, β_i, can be specified as follows. Let β be a $K \times 1$ vector of mean valuations of the characteristics, σ be a $K \times 1$ vector with standard deviations of the valuations, and ν_i be a $K \times 1$ vector with standard normal random variables. We then specify

$$\beta_i = \beta + \Sigma\nu_i, \tag{2}$$

where Σ is a $K \times K$ diagonal matrix with the standard deviations σ on the diagonal.[3] The individual valuations for the products j, $\overline{\varepsilon}_{ijt}$, may be modeled as iid random variables with an extreme value or "logit" distribution, as in BLP. Here, we suppose that the $\overline{\varepsilon}_{ijt}$ follow a more general "nested logit" distribution, which allows preferences to be correlated across products in the same group or segment. More specifically, following Berry's (1994) discussion of Cardell (1997), suppose we can assign each product j to a group g, where the groups $g = 0, \ldots, G$ are collectively exhaustive and mutually exclusive and group 0 is reserved for the outside good 0. Write

$$\overline{\varepsilon}_{ijt} = \zeta_{igt} + (1 - \rho)\varepsilon_{ijt}, \tag{3}$$

where ε_{ijt} is iid extreme value and ζ_{igt} has the (unique) distribution such that $\overline{\varepsilon}_{ijt}$ is extreme value. The parameter ρ is a "nesting" parameter, $0 \leq \rho \leq 1$, and can be interpreted as a random coefficient proxying for the degree of preference correlation between products of

[3] In principle, one may also specify non-zero off-diagonal elements in Σ to allow consumer valuations to be correlated across characteristics.

the same group.[4] As ρ goes to one, the within-group correlation of utilities goes to one, and consumers perceive products of the same group as perfect substitutes relative to other products. As ρ goes to zero, the within-group correlation goes to zero, and the model reduces to the simple logit.

Using (2) and (3) and defining the mean utility for product j, $\delta_{jt} \equiv x_{jt}\beta + \xi_{jt}$, we can write consumer i's conditional indirect utility (1) as

$$u_{ijt} = \delta_{jt} + x_{jt}\Sigma\nu_i + \zeta_{igt} + (1-\rho)\varepsilon_{ijt}.$$

Indirect utility can thus be decomposed as the sum of three terms: a mean utility term δ_{jt} common to all consumers; an individual-specific term $x_{jt}\Sigma\nu_i$ relating to continuous product characteristics x_{jt}; and an individual-specific term $\zeta_{igt} + (1-\rho)\varepsilon_{ijt}$ relating to the products' discrete characteristics, the groups or nests. If $\sigma_k = 0$ for all elements in σ (or in Σ), we obtain the standard nested logit model. If $\rho = 0$, we obtain BLP's random coefficients logit model. And if all $\sigma_k = \rho = 0$, the simple logit model results.

Each consumer i in market t chooses the product j that maximizes her utility. The aggregate market share for product j in market t is the probability that product j yields the highest utility across all products (including the outside good 0). The predicted market share of product $j = 1, \ldots, J$ in market t, as a function of the mean utility vector δ_t and the parameter vector $\theta = (\beta, \sigma, \rho)$, is the integral of the nested logit expression over the standard normal random variable vector ν_i:

$$s_{jt}(\delta_t, \theta) = \int_\nu \frac{\exp\left(\left(\delta_{jt} + x_{jt}\Sigma\nu\right) / \left(1-\rho\right)\right)}{\exp\left(I_g / \left(1-\rho\right)\right)} \frac{\exp I_g}{\exp I} \phi(\nu) d\nu, \tag{4}$$

where I_g and I are McFadden's (1978) "inclusive values" defined by

$$\begin{aligned} I_g &= (1-\rho) \ln \sum\nolimits_{k=1}^{J_g} \exp\left(\left(\delta_{kt} + x_{kt}\Sigma\nu\right) / \left(1-\rho\right)\right), \\ I &= \ln\left(1 + \sum\nolimits_{g=1}^{G} \exp I_g\right), \end{aligned}$$

and J_g is the number of products in segment g (such that $\sum_{g=1}^{G} J_g = J$). If $\rho = 0$, we obtain BLP's random coefficients logit model:

$$s_{jt}(\delta_t, \theta) = \int_\nu \frac{\exp\left(\delta_{jt} + x_{jt}\Sigma\nu\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_{kt} + x_{kt}\Sigma\nu\right)} \phi(\nu) d\nu.$$

[4] One can extend the nested logit model to group-specific nesting parameters ρ_g, $g = 1, \ldots, G$.

We approximate the integral over ν_i in (4) by simulating R draws over the density of ν:

$$s_{jt}(\delta_t, \theta) = \frac{1}{R} \sum\nolimits_{i=1}^{R} \frac{\exp\left(\left(\delta_{jt} + x_{jt}\Sigma\nu_i\right) / \left(1-\rho\right)\right)}{\exp\left(I_g / \left(1-\rho\right)\right)} \frac{\exp I_g}{\exp I}. \tag{5}$$

To estimate the demand parameters θ, we follow Berry (1994), BLP and the subsequent literature. We equate the observed market share vector (i.e. unit sales per product divided by the number of potential consumers L_t) to the predicted market share vector, $s_t = s_t(\delta_t, \theta)$. We solve this system for δ_t in each market t, using a slight modification of BLP's contraction mapping for the nested logit model; see Brenkers and Verboven (2006). Since the error term enters additively in δ_t, this gives a solution for the error term ξ_{jt} for each product $j = 1, \ldots, J$ in market t. We can then interact this with a set of instruments providing the moment conditions to proceed with GMM, as we discuss in more detail in section 3.

2.2 Monte Carlo experiments

Set-up To compare the different demand models, we begin with a Monte Carlo experiment. We assume a data generating process according to the most general RCNL model and will estimate the logit, NL, RC and RCNL model with the generated data sets. We mainly focus on the consequences from estimating misspecified models, and do this by comparing two data generating processes: one where a product's group is informative about an omitted continuous characteristic, and one where it is not. We also take the opportunity to comment on the numerical performance of the different models, in light of the above recent literature on these issues.

We generate 500 datasets, each consisting of $T = 50$ independent markets and $J = 25$ products per market. Each product j in each market t has one continuous characteristic, x_{jt}^1 and one discrete characteristic, d_{jt}, a dummy variable referring to the product's group or nest (either group 0 or group 1). So the observed product characteristics vector (including a constant) is $x_{jt} = (1, x_{jt}^1, d_{jt})$. Furthermore, each product has an unobserved characteristic ξ_{jt}.

To generate the data, we assume that ξ_{jt} is normally distributed, $\xi_{jt} \sim N(0,1)$, and uncorrelated with x_{jt}. Hence, the observed product characteristics are exogenous. It will be convenient to treat the group dummy variable d_{jt} as the realization of a latent continuous variable d_{jt}^*: the correlation between d_{jt}^* and x_{jt}^1 measures the extent to which the product's group is informative about the continuous characteristic, for which the NL model omits the random coefficient. More specifically, assume that

$$\begin{pmatrix} x_{jt}^1 \\ d_{jt}^* \end{pmatrix} \sim N\begin{pmatrix} 0 & 1 & \varsigma_{xd} \\ 0 & \varsigma_{xd} & 1 \end{pmatrix},$$

and $d_{jt} = \mathbf{1}_{\{d_{jt}^*>0\}}$. To consider the implications of omitting a random coefficient for x_{jt}^1 in the NL model, we consider $\varsigma_{xd} = 0$ and $\varsigma_{xd} = 0.9$, i.e. no or strong correlation between x_{jt}^1 and d_{jt}^*.

We specify consumer preferences for the product characteristics $x_{jt} = (1, x_{jt}^1, d_{jt})$ as follows. We set the mean valuations to $\beta = (-5, -1, -1)$ and their standard deviations to $\sigma = (0, \sigma_1, 0)$, with either $\sigma_1 = 0.25$ or $\sigma_1 = .5$.[5] Furthermore, we set the nesting parameter associated with the product group d_{jt} equal to $\rho = 0.3$. The true model is thus a RCNL model, where consumers are heterogeneous for the continuous characteristic x_{jt}^1 (through the random coefficient σ_1) and for the discrete characteristic d_{jt} (through the nesting parameter ρ, and not through a "BLP-type" random coefficient). Consumers have homogeneous preferences for the constant.

The market shares are computed from the market share equation (5), using the generated observed and unobserved product characteristics (x_{jt} and ξ_{jt}) and the assumed parameters $\theta = (\beta, \sigma, \rho)$. To approximate the integral in (5), we take $R = 500$ independent standard normal draws per market (and we use the same draws to estimate the different demand models).

For each of the 500 generated datasets, we use GMM to estimate the correctly specified RCNL model and the three other misspecified models. We generate the set of instruments from within the model, following Chamberlain's (1987) approach as applied in Berry, Levinsohn and Pakes (1999). Given the demand parameters $\theta = (\beta, \sigma, \rho)$, this instrument vector is the expected value of $\partial\xi_{jt}(\theta)/\partial\theta'$. This includes the characteristics vector itself (x_{jt}^1) and nonlinear functions of the characteristics and the parameters.

To summarize, we generate 500 datasets of 1,250 observations ($T = 50$ and $J = 25$) under four scenario's, where (i) $\varsigma_{xd} = 0$ or $\varsigma_{xd} = 0.9$ and (ii) $\sigma_1 = 0.25$ or $\sigma_1 = 0.5$. (i) If $\varsigma_{xd} = 0$, the product's group d_{jt} is not informative about x_{jt}^1: a probit regression of d_{jt} on x_{jt}^1 implies 51.6% correct classifications, which is only slightly above a random classification rule. If $\varsigma_{xd} = 0.9$, d_{jt} is quite informative about x_{jt}^1, implying 85.6% correct classifications. (ii) If $\sigma_1 = 0.25$, consumers are relatively homogeneous regarding x_{jt}^1, so that omitting the

[5] We set the constant to a low value of $\beta_0 = -5$ to obtain a relatively large share of the outside good, as in most empirical studies. For the data generating process where $\varsigma_{xd} = 0$, we obtain an average share of the outside good equal to 0.82, and for $\varsigma_{xd} = 0.9$, we obtain an average share of the outside good equal to 0.79 (with standard deviations of 0.1).

random coefficient for x^1_{jt} in the NL model may not be consequential. In contrast, if $\sigma_1 = 0.5$, consumers are relatively heterogeneous regarding x^1_{jt}, so that omitting the random coefficient for x^1_{jt} may have stronger effects on the parameters estimates.

Results Table 1 shows the results from estimating the correctly specified RCNL and the three other misspecified models under our four scenario's. For each demand model and scenario, we present the mean and standard deviation of the parameter estimates (as obtained from the 500 different datasets). Numbers in bold indicate that the parameter estimate is significantly different from the true value (on the left column).

We first have a look at the parameter estimates of the correctly specified RCNL model. For all four scenario's the parameter estimates are plausible: the means are very close to the true parameters, the standard deviations are quite small and the distribution (not shown) is approximately normal. This confirms that our estimation procedure, with analytical derivatives and a tight contraction mapping convergence criterion, works well in practice.

The parameter estimates for the logit, NL and RC logit give interesting results on the effects of estimating misspecified models. In the logit and RC models there are parameter biases in each of the four scenario's. Most interestingly, consider the two bottom panel scenario's with $\sigma_1 = 0.5$. The RC (which imposes $\rho = 0$ and thus ignores consumer heterogeneity for the groups) underestimates the mean valuation of x^1_{jt} ($\widehat{\beta}_1 \approx -1.3 < -1$) and overestimates the standard deviation of the valuation of x^1_{jt} ($\widehat{\sigma}_1 \approx 0.65 > 0.5$). The mean valuation parameter for the group dummy is not biased when $\varsigma_{xd} = 0$ (left part, $\widehat{\beta}_d = -.99 \approx -1$), but it is upward biased when $\varsigma_{xd} = 0.9$ (right part, $\widehat{\beta}_d = -.48 > -1$).

In contrast with the logit and RC models, the NL model does not result in notable biases if either $\varsigma_{xd} = 0$ or $\sigma_1 = 0.25$ (top and bottom left panels). The NL model only results in biases if both $\varsigma_{xd} = 0.9$ and $\sigma_1 = 0.5$ (bottom right panel). In this scenario the NL model underestimates the mean valuation for the group ($\widehat{\beta}_d = -1.43$) and overestimates the nesting parameter $\widehat{\rho} = 0.48$. Intuitively, when the group is quite informative about x^1_{jt}, the nesting parameter captures part of the omitted random coefficient for the continuous characteristic x^1_{jt}.

3 Empirical analysis

3.1 Dataset for the European car market

We make use of 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

Table 1: Monte Carlo Results: Different Demand Models under Different Scenario's

Parameter	True Value	Logit	NL	RC	RCNL	Logit	NL	RC	RCNL
		$\varsigma_{xd} = 0$				$\varsigma_{xd} = 0.9$			
β_0	-5	**-6.18**	-5.01	**-6.23**	-5.00	**-6.37**	-4.87	**-6.42**	-5.00
		(0.05)	(0.24)	(0.06)	(0.25)	(0.06)	(0.37)	(0.07)	(0.38)
β_d	-1	-1.00	-1.00	-1.00	-1.00	**-0.44**	-1.06	**-0.44**	-1.00
		(0.10)	(0.10)	(0.11)	(0.10)	(0.11)	(0.18)	(0.11)	(0.17)
β_1	-1	**-1.34**	-0.98	**-1.37**	-1.00	**-1.35**	-0.95	**-1.38**	-1.00
		(0.04)	(0.08)	(0.04)	(0.08)	(0.07)	(0.11)	(0.08)	(0.11)
ρ	0.3		0.28		0.30		0.32		0.30
			(0.06)		(0.06)		(0.08)		(0.08)
σ_1	0.25			0.30	0.25			0.31	0.24
				(0.12)	(0.07)			(0.11)	(0.09)
β_0	-5	**-6.02**	-4.56	**-6.22**	-5.01	**-6.22**	**-3.98**	**-6.41**	-5.02
		(0.05)	(0.25)	(0.06)	(0.25)	(0.06)	(0.40)	(0.07)	(0.40)
β_d	-1	-1.00	-0.99	-1.00	-1.00	**-0.51**	**-1.43**	**-0.48**	-0.99
		(0.10)	(0.10)	(0.10)	(0.10)	(0.12)	(0.20)	(0.11)	(0.17)
β_1	-1	**-1.20**	**-0.80**	**-1.31**	-1.00	**-1.22**	**-0.68**	**-1.34**	-1.00
		(0.04)	(0.08)	(0.04)	(0.07)	(0.07)	(0.11)	(0.08)	(0.11)
ρ	0.3		0.36		0.30		**0.48**		0.30
			(0.06)		(0.06)		(0.08)		(0.08)
σ_1	0.5			**0.67**	0.50			**0.63**	0.50
				(0.06)	(0.05)			(0.06)	(0.06)
% correctly classified		51.55				85.63			
		(1.07)				(0.94)			

The table reports the empirical means and standard deviations (in parentheses) of the estimated parameters. Biased parameter estimates (significantly different from the true value) appear in bold. The estimates are based on 500 random samples of 50 markets and 25 products, generated using the true values of the RCNL model.

sold during nine years (1998–2006) in nine West-European countries. This covers around 90% of the sales in the European Union. The countries are Belgium, France, Great Britain, Germany, Greece, Italy, Portugal, Spain, and the Netherlands. For each model/country/year we have information on sales, defined as total new registrations. For models introduced or eliminated within a given year, we know the number of months with positive sales in the given year. We exclude the models with extremely small market shares, e.g. Bentley Arnage or Kia Clarus. This results in a dataset of 18,643 model/country/year observations or on average about 230 models per country/year.

We combine the sales data with information on the list prices and various characteristics referring to the base model: vehicle size (curb weight, width and height), engine attributes (horsepower and displacement) and fuel consumption (liter/100km or €/100 km). We start from JATO's classification to assign each model to one of seven possible marketing segments: subcompact, compact, intermediate, standard, luxury, SUV and sports. Furthermore, we assign the models to their brands' perceived country of origin. For example, the Volkswagen Golf is perceived of German origin even if produced in Spain. We construct a dummy variable for whether a model is of foreign or domestic origin in each country. Our two-level nested logit model will use the marketing segments and foreign origin dummy to define the groups (e.g. subcompact) and subgroups (e.g. domestic subcompact, foreign subcompact). Table 2 provides summary statistics for sales, price and the product characteristics used in our empirical demand model. We show the summary statistics for all countries and for France and Germany separately (since we will focus on these countries when we present our counterfactuals).

Since our empirical analysis will focus on comparing the nested logit and random coefficients logit models, it is informative to provide background on how the continuous characteristics relate to the marketing segments. Table 3 (top panel) shows summary statistics for our four characteristics by marketing segment. Cars belonging to the same marketing segment tend to have similar horsepower, fuel consumption, width, and height. Horsepower and fuel consumption show a higher dispersion within a segment than width and height, but their segment averages also vary more widely. For example, average horsepower varies from 48.7kW in the subcompact to 134kW in the luxury segment, whereas average width varies from 162.5cm in the subcompact to 182.3 in the luxury segment. Table 3 (bottom panel) summarizes how well the four characteristics predict to which segment each model belongs. For each segment pair (e.g. subcompact–compact) we estimate a probit explaining segment assignment as a function of the four characteristics, and we ask how often the probit correctly classifies the different car models. The table shows that the continuous variables predict the SUV extremely well, with over 95% correct classifications with respect to any other segment.

Table 2: Summary Statistics

	All countries		France		Germany	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Sales (units)	5,785	14,694	8,440	19,931	11,432	21,074
Price/Income	1.19	0.94	0.90	0.53	0.95	0.63
Horsepower (in kW)	88.8	40.9	87.7	37.4	92.8	44.6
Fuel efficiency (€/100 km)	8.4	2.1	8.5	2.3	8.8	2.6
Width (cm)	173.0	8.5	173.1	8.5	173.4	8.6
Height (cm)	148.3	13.8	149.2	14.2	148.2	14.1
Foreign (0-1)	0.92	0.28	0.86	0.35	0.71	0.45
Months present (1-12)	9.89	2.55	9.70	2.65	9.77	2.56

The table reports means and standard deviations of the main variables. The total number of observations (models/markets) is 18,643, where markets refer to the 9 countries and 9 years.

Classification is also quite accurate for most other segments, for example for the luxury segment there are over 89% correct classifications with respect to any other segment. The lowest number of correct classifications occurs for a few "neighboring segments" (on the diagonal), e.g. 76.6% correct classifications between compact and intermediate, 77.9% between intermediate and standard. But even in these instances the characteristics predict the segments quite well.

In sum, this preliminary evidence indicates that a limited number of characteristics (horsepower, fuel consumption, width and height) have quite good, but not perfect predictive power for the classification in marketing segments. We will bear this in mind when comparing the NL and RC models.

3.2 Specification

To estimate the logit, NL, RC and RCNL demand models we slightly modify the model discussed in section 2: (i) we treat price separately since it is an endogenous characteristic and since we allow its random coefficient to follow the empirical distribution of income; (ii) we consider a two-level instead of one-level nested logit; and (iii) we allow the error term to include fixed effects for the car models and markets.

First, we start from the following version of the above utility specification (1):

$$u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \overline{\varepsilon}_{ijt}.$$

Table 3: Summary Statistics by Segment

Segment	Subc	Comp	Interm	Stand	Lux	SUV	Sport
Mean of the characteristics							
Sales (units)	11,155	7,450	5,009	4,632	2,889	2,205	1,517
Price/Income	0.55	0.81	1.04	1.39	2.13	1.61	1.85
Horsepower (in kW)	48.7	70.1	84.6	99.6	134.0	113.7	126.6
Fuel efficiency (€/100 km)	6.4	7.2	8.0	8.7	10.4	11.2	9.6
Width (cm)	162.5	171.4	175.3	175.1	182.3	179.4	175.1
Height (cm)	149.1	144.2	144.9	142.6	145.3	175.9	133.6
Foreign (0-1)	0.92	0.92	0.93	0.91	0.89	0.96	0.86
Months present (1-12)	9.72	9.87	9.88	9.77	9.94	10.11	10.03
Number of observations	3,788	4,095	2,656	1,711	1,764	2,521	2,108
Correct classifications into different marketing segments (in percent)							
Subcompact	-	93.7	99.4	99.9	100.0	95.5	97.6
Compact		-	76.6	91.1	97.7	99.7	92.8
Intermediate			-	77.9	91.4	99.7	91.0
Standard				-	90.0	99.9	84.4
Luxury					-	99.7	88.9
SUV						-	99.9
Sports							-

The top panel of the table reports means of the main variables by segment in the top panel. The bottom panel of the table reports the percentage of correctly classified car models, based on binary probit of a segment dummy per pair on four continuous characteristics (i.e. horsepower, fuel efficiency, width and height). Subc=subcompact, Comp=compact, Interm=intermediate, Stand=standard, Lux=Luxury, SUV=Sport Utility Vehicle.

The vector of observed product characteristics, x_{jt}, includes horsepower, fuel efficiency, width, height and a dummy variable for the product's country of origin (domestic or foreign). The corresponding random coefficients are specified as before, i.e. $\beta_{ik} = \beta_k + \sigma_k \nu_{ik}$ for characteristic k. Price p_{jt} enters slightly differently: its random coefficient is specified as $\alpha_i = \alpha / y_i$, where y_i is the income of individual i. In the RC and RCNL model we treat y_i as a random variable with a known distribution equal to the empirical distribution of income. In the NL model we treat y_i as non-random and set it equal to mean income in market t, $y_i = \overline{y}_t$. In sum, for the non-price characteristics we estimate both the mean valuations β_k and the standard deviations σ_k; for price we only estimate α so that heterogeneity in willingness to pay follows the empirical distribution of income.[6]

Second, the product-specific taste parameter $\overline{\varepsilon}_{ijt}$ follows the distributional assumptions of the two-level nested logit model (instead of the one-level nested logit of section 2). The upper level consists of the above seven different market segments (subcompact, compact, standard, intermediate, luxury, SUV and sports) and one separate segment for the outside good. The lower level divides every segment in two subsegments according to the models' country of origin (domestic or foreign). In four countries there are only foreign cars, so the subsegments of domestic cars are empty (Belgium, Greece, Portugal and the Netherlands). There are now two nesting parameters, $\rho = (\rho_1, \rho_2)$. The nesting parameter ρ_1 measures correlation of preferences across cars of the same subsegment, and ρ_2 measures correlation of preferences across subsegments of the same segment. For the model to be consistent with random utility maximization, $0 \leq \rho_2 \leq \rho_1 \leq 1$. If $\rho_1 = \rho_2$, the model reduces to a one-level nested logit where the segments are the nests; if $\rho_1 > \rho_2 = 0$, the model reduces to a one-level nested logit where the subsegments are the nests. If $\rho_1 = \rho_2 = 0$, the model reduces to a simple logit. Assuming that consumers choose the product that maximizes utility, we obtain a two-level nested logit version of the aggregate market shares (4).

Finally, we exploit the panel features of our data set to specify the error term, capturing unobserved product characteristics. More precisely, we assume that $\xi_{jt} = \xi_j + \xi_t + \Delta\xi_{jt}$, where ξ_j reflects time-invariant car model fixed effects, ξ_t captures country-specific fixed effects, interacted with a time trend and squared time trend, and $\Delta\xi_{jt}$ captures remaining unobserved characteristics. Since our data are at the annual level, we also include a set of dummy variables for the number of months each model was available in a country within a given year (for models introduced or dropped within a year).

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

3.3 Identification and estimation

To estimate the demand parameters $\theta = (\beta, \alpha, \sigma, \rho)$, we follow Berry (1994), BLP and the subsequent literature. As discussed above, we solve the system $s_t = s_t(\delta_t, \theta)$ for δ_t in each market t, to obtain a solution for the error term $\Delta\xi_{jt}$ for each product $j = 1, \ldots, J$ in market t:

$$\delta_{jt}(s_t, \alpha, \sigma, \rho) = x_{jt}\beta + \xi_j + \xi_t + \Delta\xi_{jt}. \tag{6}$$

In the (two-level) NL model the left-hand side has an analytic solution,

$$\delta_{jt}(s_t, \alpha, \sigma, \rho) = \ln s_{jt}/s_{0t} - \rho_1 \ln s_{j|hgt} - \rho_2 \ln s_{h|gt} + \alpha p_{jt}/\overline{y},$$

so that a linear estimator can be used. In the RC and RCNL model $\delta_{jt}(s_t, \alpha, \sigma, \rho)$ should be computed numerically by solving the system $s_t = s_t(\delta_t, \theta)$ for δ_t, which makes estimation considerably more complex.

For all models, we can proceed with GMM by interacting the error term with a vector of instrumental variables z_{jt} that is uncorrelated with the error term. Since there are $2K + 3$ parameters (K mean valuations β_k, K standard deviations σ_k, the price parameter α and the two nesting parameters ρ_1 and ρ_2), we need at least $2K + 3$ instruments in z_{jt}. Price p_{jt} does not qualify as an instrument since it is likely to be correlated with $\Delta\xi_{jt}$. For example, a positive demand shock for product j in market t will not only increase the demand for the product, but it may also induce the firm to raise its price. Failure to account for this endogeneity issue will lead to an estimated price coefficient (α) that is downward biased. Our identification assumption is that the observed product characteristics x_{jt} are uncorrelated with the unobserved product characteristics $\Delta\xi_{jt}$ (which is weaker than the often adopted assumption that x_{jt} is uncorrelated with ξ_{jt}). As discussed in BLP, one may use alternative functions of these characteristics as instruments to estimate the $2K + 3$ parameters. More specifically, following previous practice, our vector of instrumental variables z_{jt} includes: (i) the vector of product characteristics x_{jt}; (ii) the sum of the characteristics of other products of competing firms, (iii) the sum of the characteristics of other products of the same firm. For the NL and RCNL model we also include these sums over products belonging to the same subsegment and segment, following Verboven (1996).

The GMM objective function includes a weighting matrix to account for heteroskedasticity (obtained from the residuals using a two-step procedure). To minimize the GMM objective function with respect to the parameters $\theta = (\beta, \alpha, \sigma, \rho)$ we first concentrate out the linear parameters β (which includes a set of dummy variables for the market fixed effects ξ_t). We do not directly estimate the more than 200 car model fixed effects ξ_j, but instead we use a within transformation of the data (Baltagi, 1995). Standard errors are computed using the standard GMM formulas for asymptotic standard errors.

A few recent papers have studied several numerical difficulties with estimating the RC model (and a fortiori the RCNL model): global convergence problems and the role of starting values and different optimization algorithms (Knittel and Metaxoglou, 2008), problems with numerically solving δ_t using BLP's contracting mapping (Dubé, Fox and Su, 2011), and problems with approximating the integral over the logit probabilities using simulation (Judd and Skrainka, 2011).

We draw lessons from this recent literature and proceed as follows. First, to approximate the integral (4) using the simulator (5), we make use of Halton draws over the density $N(0, 1)$. This provides a more effective coverage of the density domain than pseudo-random draws. In particular, we take a large number of 500 Halton draws for each of the 81 markets (country/years).[7] Second, to ensure the GMM objective function is smooth, we use a tight tolerance level of $1e^{-12}$ to invert the shares using BLP's contraction mapping. This tolerance level is considerably stricter than typically used in the literature.[8] Third, we program analytic derivatives of the gradient of the objective function. While this is particularly tedious for the RCNL model, it greatly improves accuracy and computation time. Finally, even if the GMM objective function is smooth, it may not be globally convex. To minimize the function with respect to the nonlinear parameters (α, σ, ρ), we use different starting values, using a stringent convergence criterion of $1e^{-6}$ and carefully examining the gradient the solution path and the Hessian eigenvalues. We use a BFGS algorithm, which is an efficient procedure that uses information at different points to obtain a sense of the curvature of the objective function. We usually obtain the same optimum, except for very high or low starting values but in these cases the value of the objective function at convergence is always higher.[9]

3.4 Parameter estimates

Table 4 shows the parameter estimates for the four different demand models. The logit model imposes $\sigma = \rho = 0$ and $y_i = \overline{y}_t$. The NL model assumes $\sigma = 0$ and $y_i = \overline{y}_t$ and estimates ρ. The RC model assumes $\rho = 0$, estimates σ and allows y_i to follow the empirical distribution

[7]Halton draws can be very effective compared to pseudo-random draws. For example, Bhat (2001) and Train (2000) report that the simulation variance in the estimated parameters is lower with 100 Halton draws than with 1000 pseudo-random draws.

[8]For the NL and RCNL we use a slightly modified version of BLP's contraction mapping; see Brenkers and Verboven (2006).

[9]The log condition number of the Hessian matrix is, at worst, 1.9, which means that only 2 (of a total of 16) decimal places of accuracy are being lost in the calculation of the Hessian, thus suggesting accurate results.

of income. Finally, the RCNL estimates both ρ and σ, and allows y_i to follow the empirical distribution of income.

In the simple logit model both the price parameter (α) and the mean valuation parameters (β) have the expect signs and are all significantly different from zero. However, as is well-known, the model is very restrictive since it imposes symmetric cross-price elasticities. Furthermore, demand is inelastic for almost 20% of the car models across countries and years. This is inconsistent with oligopolistic profit maximizing behavior unless marginal costs would be negative.

In the NL model the upper nest level consists of the seven marketing segments and the lower nest level consists of the segments and origin (domestic/foreign). The price parameter (α) and the mean valuation parameters (β) again have the expected sign and are significantly different from zero, with the exception of the parameter for width, which is now insignificant. The nesting parameters are estimated very precisely, $\rho_1 = 0.65$ and $\rho_2 = 0.48$. Their magnitudes are consistent with the requirements of random utility maximization ($0 \leq \rho_2 \leq \rho_1 \leq 1$) and imply that consumer preferences show the strongest correlation across cars from both the same marketing segment and origin (domestic/foreign), and show weaker but still important correlation across cars from the same segment but a different origin. This is consistent with earlier work for a more limited set of countries (Goldberg and Verboven, 2001 and Brenkers and Verboven, 2006).[10] As documented below, this implies more plausible cross-price elasticities than the simple logit model. Furthermore, the implied own-price elasticities are higher than in the simple logit: demand is now inelastic for only 3% of the car models. This may seem surprising at first, since the price coefficient α is closer to zero than in the simple logit model. However, the elasticities do not only depend on α but also on the nesting parameters ρ_1 and ρ_2.

In the RC model we estimate the price parameter (α) and the means (β) and standard deviations (σ) for the valuations of the other characteristics (including the constant). The price parameter (α) is again significantly estimated with the expected sign (negative effect). Consumers have a negative and significant mean valuation for fuel consumption, and heterogeneity is limited so that almost all consumers dislike fuel inefficient cars. Consumers have a positive and significant mean valuation for width, and the standard deviation implies that about 10% of consumers dislike large cars. Consumers have a negative mean valuation for cars from foreign origin. The standard deviation is relatively large, so that 25% of consumers actually prefer foreign cars. The mean valuation for height is insignificantly different from

[10] We also estimated a two-level NL model with the reverse nesting structure, where origin defines the upper level and origin/segment the lower level of the nests. This led to estimates of ρ_1 and ρ_2 inconsistent with random utility maximization, in line with the results of other studies on the car market.

zero, and the mean valuation for horsepower is unexpectedly negative. However,for both characteristics we find substantial and significant heterogeneity: about 50% of consumers have a positive valuation for height and about 30% have a positive valuation for horsepower. Finally, we estimate a significant standard deviation for the constant, indicating there is significant heterogeneity in the valuation of new cars relative to the outside good. Overall, the random coefficients show evidence of significant consumer heterogeneity in several dimensions, in particular height, horsepower and foreign origin. Yet it is striking that the random coefficients are estimated much less precisely than the two nesting parameters in the NL model.

In the RCNL model we combine the previous two models, so we include both the nesting parameters and the random coefficients. Both the price parameter (α) and the mean valuation parameters (β) have the expected signs and are estimated significantly with the exception of the horsepower parameter, which is insignificant. The most interesting findings relate to the estimated nesting parameters (ρ) and random coefficients (σ) in comparison with the NL and RC models.

First, compared with the NL model, the nesting parameters remain highly significant, but their magnitude becomes smaller. This is consistent with the results from our Monte Carlo study, where we found an overestimate of the nesting parameters if the random coefficients are important and the groups are correlated with the characteristics for the omitted random coefficients. Furthermore, we can no longer reject the hypothesis that $\rho_1 = \rho_2$ (P-value 0.0967) and the random coefficient for foreign origin is insignificant. So the model reduces to a one-level nested logit with no need to divide the seven segments into domestic and foreign subgroups, and it seems at first that there is no longer consumer heterogeneity for foreign origin. However, the subsegment parameter ρ_1 captures similar effects as the random coefficient for foreign origin, suggesting it is not sensible to include both. Indeed, in a one-level nested logit where we constrain $\rho_1 = \rho_2$ (so that the subgroups are no longer relevant), the random coefficient for foreign origin becomes significant again (as in the RC model). We show these results in Table A.1 in the Appendix.[11]

Second, compared with the RC model, the random coefficients for horsepower and fuel efficiency remain significant, but this is no longer the case for width, height and the constant. Intuitively, the nesting parameter for the segments captures a lot of the heterogeneity relating to the car dimensions and the outside good, but not much of the heterogeneity relating to

[11] In this case, the one-level nested logit with a random coefficient for foreign origin seems preferable to a two-level nested logit model, since it does not impose the consumer heterogeneity to enter in a hierarchical way. Nevertheless, we base our subsequent discussion on the two-level nested logit. The implied price elasticities and competition policy counterfactuals are very similar in the one-level nested logit model (not shown).

horsepower and fuel efficiency.

Table 4: Parameter Estimates for Alternative Demand Models

	Logit		Nested Logit		RC Logit		RC Nested Logit	
	Param.	St. Er.	Param.	St. Er.	Param.	St. Er.	Param.	St. Er.
	Mean valuations for the characteristics in x_{jt} (β)							
Price/income	-1.76	0.17	-1.00	0.03	-5.52	0.66	-2.75	0.18
Horsepower (kW/100)	2.30	0.24	1.34	0.08	-3.67	1.86	0.57	0.77
Fuel (€/10,000 km)	-11.48	1.43	-6.13	0.52	-20.77	3.06	-4.68	0.73
Width (cm/100)	2.51	0.55	-0.10	0.29	3.64	0.83	1.26	0.50
Height (cm/100)	3.46	0.35	1.17	0.19	0.27	1.32	2.12	0.46
Foreign (0/1)	-1.21	0.03	-0.47	0.04	-3.66	0.89	-0.57	0.14
	Standard deviations of valuations for the characteristics in x_{jt} (σ)							
Horsepower (kW/100)	n/a		n/a		4.67	0.83	0.92	0.41
Fuel (€/10,000 km)	n/a		n/a		1.15	1.69	1.66	0.57
Width (cm/100)	n/a		n/a		1.93	0.71	0.10	1.74
Height (cm/100)	n/a		n/a		4.83	0.55	0.15	1.11
Foreign (0/1)	n/a		n/a		5.46	1.05	0.22	0.84
Constant	n/a		n/a		1.18	0.43	0.21	3.00
	Nesting parameters (ρ_1 and ρ_2)							
Subsegment ρ_1	n/a		0.65	0.03	n/a		0.57	0.03
Segment ρ_2	n/a		0.48	0.03	n/a		0.47	0.07
Model fixed effects	Yes		Yes		Yes		Yes	
Market fixed effects	Yes		Yes		Yes		Yes	
Income distribution	No		No		Yes		Yes	
Random coefficients	No		No		Yes		Yes	
# inelastic demands	3,514 (19%)		556 (3%)		0		0	
χ^2 test $\rho_1 = \rho_2$	n/a		83.04		n/a		2.76	
Prob.>χ^2	n/a		(0.00)		n/a		(0.10)	

The table shows the parameter estimates and standard errors for the different demand models. The logit and NL models assume equal income ($-\alpha/\overline{y}_t$), the RC and RCNL models allow for heterogeneous income ($-\alpha/y_i$). The total number of observations (models/markets) is 18,643, where markets refer to the 9 countries and 9 years.

Since the logit, NL and RC are all restricted versions of the RCNL model, we can compare their statistical performance using likelihood ratio tests adapted to the GMM context.[12] Table 5 reports LR values and asymptotic P-values for all pairs of models, except the NL

[12]Following Hayashi (2000), we define the likelihood ratio statistic (LR) as the difference between the value of the objective function of the restricted model (re-estimated using the second-stage weighting matrix of the unrestricted model) and the value of the objective function of the unrestricted model. Under the null hypothesis, the statistic is asymptotically χ^2 distributed with degrees of freedom equal to the number of restrictions.

Table 5: Likelihood Ratio Tests for Alternative Demand Models

	Logit	Nested Logit	RC Logit
Logit	–		
Nested Logit	584.08	–	
	(0.0000)		
RC Logit	34.08	n/a	–
	(0.000)		
RC Nested Logit	534.10	30.61	423.84
	(0.0000)	(0.0002)	(0.0000)

The table reports χ^2 statistics and P-values (in parentheses) of likelihood ratio tests for different model pairs.

and RC which are not nested in each other. Each restricted model is rejected against the more general models. The logit is clearly rejected against any other model. More interestingly, both the NL and RC models are rejected against the more general RCNL model. In fact, the NL appears to provide a better fit than the RC logit relative to the RCNL, since the χ^2 statistic is lower for the NL than the RC model (30.61 versus 423.84). We already observed above that the individual random coefficients in the RC model are much less precisely estimated than the two nesting parameters in the NL model. The likelihood ratio tests thus indicate that the random coefficients of the RC model are also jointly less significant than the nesting parameters of the NL model.

Summary We can summarize our empirical results in the following four points. (i) It appears important to include the nesting parameter relating to the seven marketing segments since it remains highly significant after including the random coefficients. (ii) It does not seem appropriate to include an additional subnesting parameter relating to the origin within each segment, since the random coefficient for origin captures this well. (iii) It is relevant to include random coefficients for horsepower and fuel efficiency, but not those for the dimensions width and height since these are captured well by the marketing segments. (iv) It is striking that the nesting parameters (reflecting heterogeneity regarding segments and subsegments) are estimated much more precisely than the random coefficients (reflecting heterogeneity regarding continuous characteristics). While these findings apply to our dataset of the European car market, they can also be useful as a guide for interpretations in other applications.

3.5 Substitution patterns

We have already commented on the number of inelastic own-price elasticities implied by our estimates. We now provide a more systematic discussion on the substitution patterns. We consider own-price and cross-price elasticities at the level of the individual products and at the level of the entire segments.

Product-level price elasticities First consider the product-level own- and cross-price elasticities. We average these by segment, and distinguish between cross-price elasticities with respect to other products in the same subsegment, in a different subsegment within the same segment, and in a different segment. Table 6 shows these average product-level elasticities for Germany in 2006 (the largest country in the most recent year of our dataset). In the logit and NL model the own-price elasticities tend to increase more or less proportionally with price as one moves to higher segments, resulting in an average own-price elasticity that is almost 4 times higher in the luxury than in the subcompact segment. The near proportional relationship follows from the functional form assumption: price enters utility linearly with a homogeneous valuation across consumers ($-\alpha/\overline{y}_t$). In contrast, in the RC and RCNL models the price elasticities increase much less than proportionally, by a factor of 2.2 and 2.3 in the respective models. This follows from the less restrictive functional form: price still enters utility linearly, but consumer valuations are heterogeneous ($-\alpha/y_i$). Hence, price insensitive consumers are more likely to purchase high priced cars.

The cross-price elasticities show even more striking differences across the estimated models. In the logit model, they are extremely small even with respect to cars from the same subsegment or segment (always <0.01). In contrast, in the NL and RCNL models the cross-price elasticities are quite high with respect to products of the same subsegment (about 0.1–0.4) and they are still relevant with respect to products of other subsegments in the same segment (about 0.05). In the RC model, the cross-elasticities with respect to products of the same subsegment are still sizeable, mainly because of the magnitude and significance of the foreign ownership random coefficient. But they are negligible with respect to products of other segments within the same segment (usually <0.01). These findings illustrate the importance of accounting for consumer heterogeneity relating to the marketing segments (as done only in the NL and RCNL models) and the domestic/foreign origin (as done in all models except the simple logit).

Table 6: Product-level Price Elasticities in Germany for Alternative Demand Models

Segment	Own-	Cross-price elasticity same subseg	same seg	differ seg
		Logit		
Subcompact	-0.76	<0.01	<0.01	<0.01
Compact	-1.09	<0.01	<0.01	<0.01
Intermediate	-1.49	<0.01	<0.01	<0.01
Standard	-1.94	<0.01	<0.01	<0.01
Luxury	-2.94	<0.01	<0.01	<0.01
SUV	-2.32	<0.01	<0.01	<0.01
Sports	-2.73	<0.01	<0.01	<0.01
		Nested Logit		
Subcompact	-1.23	0.02	0.01	<0.01
Compact	-1.74	0.03	0.02	<0.01
Intermediate	-2.38	0.05	0.03	<0.01
Standard	-3.04	0.13	0.05	<0.01
Luxury	-4.64	0.17	0.07	<0.01
SUV	-3.73	0.05	0.04	<0.01
Sports	-4.40	0.08	0.03	<0.01
		RC Logit		
Subcompact	-2.85	0.03	<0.01	<0.01
Compact	-3.66	0.02	<0.01	0.01
Intermediate	-4.38	0.03	<0.01	0.01
Standard	-4.96	0.04	0.01	0.01
Luxury	-6.24	0.06	0.03	0.01
SUV	-5.67	0.04	<0.01	0.01
Sports	-6.13	0.02	<0.01	0.02
		RC Nested Logit		
Subcompact	-2.57	0.03	0.03	<0.01
Compact	-3.33	0.05	0.05	<0.01
Intermediate	-3.90	0.06	0.06	<0.01
Standard	-4.54	0.15	0.09	<0.01
Luxury	-5.75	0.17	0.11	<0.01
SUV	-5.01	0.07	0.06	<0.01
Sports	-5.42	0.10	0.05	<0.01

The table reports product-level own- and cross-price elasticities, based on the parameter estimates in Table 4. Elasticities are averages by segment for Germany in 2006. Cross-price elasticities are averaged across products from the same subsegment, from a different subsegment within the same segment, and from different segments.

Segment-level price elasticities Now consider the segment-level price elasticities, i.e. the effect of a joint 1% price increase of all cars in a given segment on demand in the various segments. Table 7 reports these segment-level own- and cross-price elasticities. We can summarize these results as follows. First, as is well-known, both the logit and NL model imply fully symmetric substitution patterns at the segment-level (i.e. identical cross-elasticities per row). For example, a price increase of all compact cars by 1% raises the demand in all other segments by 0.02% (more precisely, by 0.017%). In sharp contrast, the RC model implies more intense substitution to "neighboring segments". Taking the same example, a price increase of all compact cars by 1% has the highest effect on the demand for subcompact (+0.76%) and compact cars (+0.66%), and lowest effects on the demand for luxury (0.26%) and SUV cars (+0.39%). Finally, the RCNL model implies cross-price elasticities somewhere in between the NL and RC model, though closer to the NL model: the cross-price elasticities to other segments are fairly (but not completely) symmetric, and they are somewhat higher than in the NL model, but not nearly as high as in the RC model.

We stress that, even though the substitution patterns of the most general RCNL model appear to be better approximated by the NL model than by the RC model, this does not necessarily mean that the NL model should be preferred over the RC model. The main message is that it is important to account for consumer heterogeneity regarding the marketing segments. The NL model is one simple way to capture this, but there may be alternative ways. For example, one may consider adding random coefficients for the segments at an increased computational cost.

Summary We can summarize the differences in the estimated substitution patterns across models as follows. First, the own-price elasticities at the product level increase roughly proportionally with price in the logit and NL model, but less than proportionally in the RC and RCNL model. This is because the latter two models allow for consumer heterogeneity in the price parameter. Second, the product-level cross-price elasticities show that products of the same segment are strong substitutes in the NL and RCNL model, but not in the logit and RC models. Finally, the segment-level cross-price elasticities show that there is quite strong substitution across segments (especially the neighboring ones) in the RC model, but only weak (and symmetric) substitution in the logit, NL and RCNL models.

4 Implications for competition policy analysis

The previous section showed how the different demand models generate quite different substitution patterns. But how relevant are the found differences for applications in industrial

Table 7: Segment-level Price Elasticities in Germany for Alternative Demand Models

Segment	Subc	Comp	Interm	Stand	Lux	SUV	Sport
				Logit			
Subcompact	-0.77	0.02	0.02	0.02	0.02	0.02	0.02
Compact	0.02	-1.12	0.02	0.02	0.02	0.02	0.02
Intermediate	0.01	0.01	-1.41	0.01	0.01	0.01	0.01
Standard	0.01	0.01	0.01	-1.75	0.01	0.01	0.01
Luxury	0.01	0.01	0.01	0.01	-2.59	0.01	0.01
SUV	0.01	0.01	0.01	0.01	0.01	-2.24	0.01
Sports	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	-2.05
				Nested Logit			
Subcompact	-0.44	0.01	0.01	0.01	0.01	0.01	0.01
Compact	0.01	-0.64	0.01	0.01	0.01	0.01	0.01
Intermediate	<0.01	<0.01	-0.81	<0.01	<0.01	<0.01	<0.01
Standard	<0.01	<0.01	<0.01	-1.00	<0.01	<0.01	<0.01
Luxury	<0.01	<0.01	<0.01	<0.01	-1.48	<0.01	0.01
SUV	<0.01	<0.01	<0.01	<0.01	<0.01	-1.28	<0.01
Sports	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	-1.17
				RC Logit			
Subcompact	-1.72	0.67	0.47	0.19	0.07	0.33	0.29
Compact	0.75	-2.77	0.66	0.54	0.26	0.39	0.41
Intermediate	0.29	0.39	-3.47	0.43	0.30	0.45	0.42
Standard	0.12	0.32	0.44	-3.55	0.56	0.43	0.45
Luxury	0.05	0.16	0.32	0.61	-4.05	0.86	0.67
SUV	0.15	0.18	0.37	0.43	0.92	-4.13	0.75
Sports	0.08	0.11	0.20	0.25	0.42	0.49	-4.36
				RC Nested Logit			
Subcompact	-1.08	0.04	0.04	0.04	0.04	0.04	0.04
Compact	0.04	-1.42	0.05	0.06	0.06	0.06	0.05
Intermediate	0.03	0.03	-1.65	0.04	0.04	0.04	0.04
Standard	0.03	0.03	0.04	-1.90	0.05	0.05	0.05
Luxury	0.03	0.04	0.05	0.06	-2.37	0.08	0.07
SUV	0.03	0.04	0.05	0.06	0.08	-2.12	0.07
Sports	0.02	0.02	0.03	0.03	0.04	0.04	-2.03

The table reports the segment-level own- and cross-price elasticities (when all products in the same segment raise their price by 1%), based on the parameter estimates in Table 4. The elasticities refer to Germany in 2006. Subc=subcompact, Comp=compact, Interm=intermediate, Stand=standard, Lux=Luxury, SUV=Sport Utility Vehicle.

organization or related fields? To address this question, we consider two areas of competition policy, market definition and merger simulation, and we ask whether the different demand models yield robust conclusions.

Much of competition policy still heavily relies on market definition and an assessment of the firms' market shares within the defined market. It is simple and widely applicable to mergers and horizontal or vertical agreements because it makes few assumptions about oligopoly behavior. However, the choice of candidate relevant markets can often be quite arbitrary and artificial. Furthermore, because it is not based on a specific model of oligopoly behavior, it cannot make precise predictions about market power effects, and it cannot incorporate other considerations in an integrated framework. In merger cases, one increasingly resorts to simulation analysis to assess market power effects and incorporate efficiencies or other elements; see e.g. Werden and Froeb (1994), Hausman, Leonard and Zona (1994), Nevo (2000) and Peters (2006). While merger simulation may in principle extend to other types of competition investigations, this is difficult in practice because it requires the specification of an appropriate oligopoly model for the specific competition issue under investigation.

These relative advantages and disadvantages of market definition and merger simulation have been widely discussed. We will instead look at this from a different angle: we ask to which extent both approaches are sensitive to the adopted demand model. If one approach gives more robust conclusions across demand models, this provides a new motivation to prefer it over the other approach.

4.1 Market definition

Market definition in the European car market is not only relevant for the evaluation of mergers, but also for the implementation of the Block Exemption Regulation for the selective and exclusive distribution system. According to this Regulation, automobile manufacturers may impose selective or exclusive distribution to their dealers, provided they have market shares below 30% or 40%. Some niche manufacturers such as Mercedes or BMW may meet these thresholds if markets are defined widely to include all cars, but not if they are defined narrowly at the level of the marketing segments. Hence, it is important to know whether the segments by themselves can be considered relevant markets.

According to the SSNIP test, the relevant market is the smallest group of products for which a hypothetical monopolist could profitably impose a small, non-transitory but significant increase in price (typically 5%-10%). Since the profitability of a price increase depends on the extent of substitution to other goods, the estimated demand model is of central importance. We will apply the SSNIP test to the various estimated demand models and ask whether the seven marketing segments can be considered as separate relevant markets, or whether a broader market definition is appropriate. For each of the four estimated demand models, we first compute all products' implied marginal costs assuming multiproduct price-setting firms (following BLP, Nevo, 2000 and others). Given the estimated demand systems and the marginal costs, we then ask whether a 10% price increase by all products in a given marketing segment raises total profits in the considered segment.

Table 8 shows the SSNIP-test results for France and Germany in 2006. The logit model suggests that none of the seven marketing segments can be considered as separate relevant markets. For example, a joint 10% price increase in the compact segment in France reduces profits by 0.6%. The RC model yields a similar conclusion: only the subcompact segment can be defined as a relevant market in both France and Germany. In sharp contrast, the NL and RCNL model imply that all marketing segments constitute separate relevant markets. A joint 10% price increase in the compact segment in France would raise profits by 7.21% according to the NL model and even by 10.84% according to the RCNL model. This narrow market definition follows, of course, from the high significance of the nesting parameter for the segments in the NL and RCNL models.

Should we conclude that the RC model fails to define the markets narrowly at the segment level, in contrast with the more general RCNL model against which it was rejected? The answer may be yes, since we found that the RC model omits important unobservables relating to the marketing segments that are captured in the more general RCNL. However, proper caution is warranted. First, the RCNL model is itself restrictive since it imposes largely symmetric substitution across the segments. For example, a variant of the RCNL model where consumers would be more likely to substitute to neighboring segments might lead one to conclude that market definition should include the neighboring segments. Second, the RC model may itself also give rise to "narrow" market definitions, albeit not at the "segment-level". For example, one may define relevant markets of car models that are not necessarily in the same marketing segment but that share similar horsepower, height and origin (the dimensions for which we estimated most consumer heterogeneity). Such a market definition process would however be somewhat tedious. As another simpler example, one may define two neighboring segments as the relevant market in the RC model (as suggested by the above cross-price elasticities). Our SSNIP-test results at the level of neighboring segments (not shown) confirm that neighboring segments constitute relevant markets in the RC model: a joint 10% price increase raises profits for compact+intermediate (+1.6%) but not for, e.g., compact+luxury (–1.2%).

Table 8: Relevant Market Definition in France and Germany

Segment	Logit		Nested Logit		RC Logit		RC Nested Logit	
	France	Germany	France	Germany	France	Germany	France	Germany
Subcompact	-0.1	-0.2	5.0	6.7	4.5	4.9	8.8	11.0
Compact	-0.6	-0.5	7.2	8.7	-5.1	-1.4	10.8	12.6
Intermediate	-1.0	-1.0	7.4	8.4	-8.6	-5.3	10.4	10.4
Standard	-1.6	-1.5	13.5	11.1	-7.8	-5.1	16.3	13.3
Luxury	-3.4	-3.2	16.2	15.0	-9.5	-5.9	16.6	15.2
SUV	-2.4	-2.6	16.5	15.7	2.9	-5.9	18.1	16.0
Sports	-1.4	-2.4	10.1	13.9	-11.2	-9.1	12.6	14.2

The table reports percentage profit increases implied by a joint 10% price increase of all products in the same segment, based on the parameter estimates in Table 4 and assuming marginal costs implied by multiproduct Bertrand competition. The effects refer to France and Germany in 2006. Subc=subcompact, Comp=compact, Interm=intermediate, Stand=standard, Lux=Luxury, SUV=Sport Utility Vehicle.

4.2 Merger simulation

We consider the effects of two hypothetical mergers. The first merger is between the two French manufacturers PSA (Peugeot and Citroën) and Renault, and the second merger is between the two German manufacturers BMW and Volkswagen (Volkswagen, Audi, Seat and Skoda). As shown in Table 9, PSA and Renault are strong in their home market France, with a combined market share of 56% (mainly due to the mass segments). BMW and Volkswagen are slightly less strong in their home market Germany, with a combined market share of 41%. But they have a particularly strong presence in specific segments, i.e. the standard segment (71%) and the luxury segment (58%).

We first compute the products' marginal costs assuming multiproduct price-setting firms, as we also did to implement market definition. Given the estimated demand systems and the marginal costs, we then predict the new Nash equilibrium resulting from the changed ownership structure after the merger. Intuitively, a merger will entail high price effects if the merging firms sell close substitutes with respect to each other (low cross-price elasticities) and weak substitutes with respect to outsider firms (low own-price elasticities).

Table 9 shows the predicted price effects of the two mergers in the firms' home markets. We also briefly comment on the effects in the foreign markets, and show these results in Table A.3 of the Appendix. We show the percentage price increases both for the entire market and for each of the seven marketing segments (using price indices, where postmerger market shares are the weights).

For both mergers, the logit model predicts very small domestic price effects, despite the merging firms' strong domestic market presence. In sharp contrast, the NL, RC and RCNL models give robust conclusions. The PSA–Renault merger would result in large aggregate price increases in France (between 8.3% and 20.2%).[13] The BMW–VW merger entails more modest price increases in its home country Germany, but the results are again robust across all models except the logit model (between 1.9% and 3.0%). In particular, the predicted price increases are the largest in the standard segment, where the German producers have the strongest presence (between 4.9% and 10.0%). While the NL, RC and RCNL give robust conclusions regarding the predicted merger effects, the NL model gives more precise predictions than the RC model, as shown by the smaller confidence intervals in Table A.3 of the Appendix. This follows from the fact that the nesting parameters were estimated more precisely than the random coefficients.

The predicted price effects in the foreign markets are much smaller. But there is again a notable difference between the logit model and the other three models (where the predicted effects are between 0.4% and 0.6% for the BMW–VW merger in France, and between 0.2% and 0.4% for the PSA–Renault merger in Germany).

In sum, these findings show that it is clearly inappropriate to use a simple logit model with its symmetric substitution patterns. But it does not appear important whether to generalize the model to a NL, RC or RCNL model, since they give robust conclusions.

4.3 Summary

We can summarize our findings on market definition and merger simulation as follows. Merger simulation yields fairly clear conclusions across different demand models: the simple logit model is clearly inappropriate, but a generalization to the NL, RC or RCNL gives robust conclusions. In contrast, market definition depends more heavily on the adopted demand model. In particular, the RC model suggests a too wide definition at the level of all cars (similar to the logit model), whereas the NL and RCNL models suggest a more narrow definition at the level of the segments. We discussed that this lack of robustness should not be attributed to the RC model per se, but rather to the arbitrariness in selecting candidate relevant markets in the market definition approach.

[13] The overall predicted price increases are most close for the NL and the RC model (15.5% and 20.1%). They are somewhat lower for the RCNL model (20.2%), but the bootstrapped 95% confidence intervals show a small overlap, as shown in Appendix in Table A.3.

Table 9: The Effects of Two Hypothetical Mergers in France and Germany

France	All	Subc	Comp	Interm	Stand	Lux	SUV	Sport
	PSA–Renault merger in France							
	Domestic market shares (in percent)							
PSA	33.4	35.3	38.8	46.0	-	19.1	-	37.3
Renault	22.7	29.8	20.9	17.8	-	9.5	-	13.5
	Predicted domestic price increase (in percent)							
Logit	0.9	1.6	0.9	0.75	0.0	0.2	0.0	0.5
Nested Logit	15.5	31.2	13.5	12.8	0.0	2.1	0.0	7.0
RC Logit	20.2	37.1	22.6	24.1	0.6	4.8	0.1	14.0
RC Nested Logit	8.3	15.9	8.0	8.2	-0.1	1.5	-0.1	4.5
Germany	VW-BMW merger in Germany							
	Domestic market shares (in percent)							
BMW	10.6	2.1	7.9	-	39.6	25.3	15.2	10.8
VW	30.8	23.1	36.3	53.8	31.3	32.4	12.0	21.4
	Predicted domestic price increase (in percent)							
Logit	0.3	0.3	0.4	0.3	0.6	0.3	0.2	0.2
Nested Logit	2.9	0.6	2.8	0.1	10.0	4.3	1.6	1.1
RC Logit	2.2	0.6	2.0	1.8	4.9	3.2	1.7	1.5
RC Nested Logit	1.9	0.6	1.8	0.5	5.8	3.0	1.1	0.9

The table reports percentage price increases for two hypothetical mergers, PSA–Renault and BMW–VW, in their domestic markets France and Germany, based on the parameter estimates in Table 4 and assuming multiproduct Bertrand competition. The effects refer to France and Germany in 2006. 95% confidence intervals, based on a bootstrapping procedure, are shown in Appendix in Table A.4. For example, the 95% confidence interval for the overall predicted price increase after the PSA–Renault merger is [0.7–1.8]% for the logit, [12.5–18.3]% for the NL, [14.6–27.2]% for the RC and [5.4–15.7]% for the RCNL model. Subc=subcompact, Comp=compact, Interm=intermediate, Stand=standard, Lux=Luxury, SUV=Sport Utility Vehicle.

5 Conclusion

We started from a general aggregate RCNL model to provide a systematic comparison between the simple logit and NL models and the computationally more complex RC model. We first used simulated data to document parameter biases from estimating a NL or RC model, when the true model is in fact a RCNL model. We then use data on the automobile market to estimate the different models, and as an illustration assess what they imply for competition policy analysis. Our main findings on the advantages and disadvantages of the NL and RC model can be summarized as follows.

In terms of the statistical performance, both the NL and the RC model are rejected against the more general RCNL model. The NL model appears to be less strongly rejected (much lower χ^2) than the RC model, and the nesting parameters of the NL model (ρ) drop by only a modest amount after including random coefficients on continuous variables (σ) in the RCNL model. Furthermore, the nesting parameters are estimated more precisely than the random coefficients, suggesting that the marketing segments capture a substantial part of consumer heterogeneity.

In terms of substitution patterns, the NL and RC model yield quite different results. The own-price elasticities increase nearly proportionally with price in the NL model and less than proportionally in the RC model, because the latter model allows for consumer heterogeneity in the price parameter. Furthermore, products within the same segment are much closer substitutes in the NL model, whereas there is strong substitution to other segments (especially to neighboring ones) in the RC model.

Despite the rather different substitution patterns the NL and RC model generate quite robust conclusions on the predicted price effects from mergers. In sharp contrast, the conclusions for market definition are not robust: markets are defined narrowly at the segment level in the NL model, and at the wider level of all cars in the RC model (similar to the logit). This suggests two implications for competition policy. First, the lack of robustness in market definition should not be attributed to the RC model per se, but rather to the arbitrariness in selecting candidate relevant markets. Second, the robustness in merger simulation suggests the simple NL model can be sufficient to obtain reliable policy conclusions, despite the different substitution patterns.

More generally, one can draw two implications for the choice of demand model in applied work. First, the choice between the tractable NL model and the computationally more complex RC model may depend on the application. In our merger analysis we considered two domestic mergers. A particularly relevant aspect of consumer heterogeneity is then the cars' domestic/foreign origin, which the NL model captures reasonably well. In other applications, the most relevant aspects of consumer heterogeneity may not be captured well by nesting parameters for groups or subgroups. In these cases, it is appropriate to estimate RC models with random coefficients for the most relevant continuous characteristics.

Second, our findings show that it is important to account for sources of market segmentation that are not captured by the continuously measured characteristics in the RC model. We established this by adding a nested logit structure to BLP's random coefficients model. But in future research one may also consider other tractable models.

References

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

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

_ and Ariel Pakes, "Comments on Alternative models of demand for automobiles, by Charlotte Wojcik," *Economics Letters*, December 2001, *74* (1), 43–51.

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

_ , _ , and _ , "Voluntary Export Restraints on Automobiles: Evaluating a Trade Policy," *American Economic Review*, June 1999, *89* (3), 400–430.

Bhat, Chandra R., "Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model," *Transportation Research Part B: Methodological*, 2001, *35* (7), 677 – 693.

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., "Departures from marginal-cost pricing in the American automobile industry : Estimates for 1977-1978," *Journal of Econometrics*, November 1981, *17* (2), 201–227.

— , **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*, 17–44.

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

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

Dubé, Jean-Pierre H., Jeremy T. Fox, and Che-Lin Su, "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," *SSRN eLibrary*, 2011.

Feenstra, Robert C. and James A. Levinsohn, "Estimating Markups and Market Conduct with Multidimensional Product Attributes," *Review of Economic Studies*, January 1995, *62* (1), 19–52.

Fershtman, Chaim and Neil Gandal, "The Effect of the Arab Boycott on Israel: The Automobile Market," *RAND Journal of Economics*, Spring 1998, *29* (1), 193–214.

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.

Hausman, Jerry A., Gregory Leonard, and J. Douglas Zona, "Competitive Analysis with Differentiated Products," *Annales d'Economie et de Statistique*, Avril-Jui 1994, *34* (34), 159–180.

Hayashi, Fumio, *Econometrics*, Princeton University Press, December 2000.

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 Feb 2011.

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

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.

Nevo, Aviv, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, Autumn 2000, *31* (3), 395–421.

_ , "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, March 2001, *69* (2), 307–42.

Peters, Craig, "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *Journal of Law & Economics*, 2006, *49* (2), 627–49.

Petrin, Amil, "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 2002, *110*, 705–29.

Reynaerts, Jo, Ravi Varadhan, and John C. Nash, "The Convergence Properties of the BLP (1995) Contraction Mapping and Alternative Nonlinear Acceleration Algorithms," 2010. mimeo, LICOS, K.U.Leuven.

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

Sudhir, K., "Competitive Pricing Behavior in the Auto Market: A Structural Analysis," *Marketing Science*, January 2001, *20* (1), 42–60.

Train, Kenneth, "Halton Sequences for Mixed Logit," Department of Economics, Working Paper Series 1035, Department of Economics, Institute for Business and Economic Research, UC Berkeley May 2000.

Verboven, Frank, "International Price Discrimination in the European Car Market," *RAND Journal of Economics*, Summer 1996, *27* (2), 240–268.

Werden, Gregory J. and Luke M. Froeb, "The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy," *Journal of Law, Economics, & Organization*, 1994, *10* (2), pp. 407–426.

Wojcik, Charlotte, "Alternative models of demand for automobiles," *Economics Letters*, August 2000, *68* (2), 113–118.

6 Appendix

Table A.1: Parameter Estimates for Constrained One-Level RCNL model

	Constrained One-Level RCNL	
	Param.	St. Er.
	Mean valuations for the characteristics in x_{jt} (β)	
Price/income	-2.73	0.06
Horsepower (kW/100)	1.20	0.29
Fuel (€/10,000 km)	-0.45	0.03
Width (cm/100)	0.12	0.01
Height (cm/100)	0.20	0.01
Foreign (0/1)	-0.67	0.03
	Standard deviations of valuations for the characteristics in x_{jt} (σ)	
Horsepower (kW/100)	0.50	0.23
Fuel (€/10,000 km)	-1.49	0.19
Width (cm/100)	n/a	
Height (cm/100)	n/a	
Foreign (0/1)	0.55	0.05
Constant	n/a	
	Nesting parameters ($\rho_1 = \rho_2$)	
Segment ρ_1	0.56	0.01
Model fixed effects	Yes	
Market fixed effects	Yes	
Income distribution	Yes	
Random coefficients	Yes	
# inelastic demands	0	

This table shows the parameter estimates and standard errors for a constrained version of the RCNL of Table 4. We constrain $\rho_1 = \rho_2$.(so there is only one level of nesting) and the standard deviations for the valuations of width, height, and the constant are set equal to 0. The total number of observations (models/markets) is 18,643, where markets refer to the 9 countries and 9 years.

Table A.2: Product-level Price Elasticities in France for Alternative Demand Models

Segment	Own-	Cross-price elasticity same subseg	same seg	differ seg
	Logit			
Subcompact	-0.73	<0.01	<0.01	<0.01
Compact	-1.14	<0.01	<0.01	<0.01
Intermediate	-1.39	<0.01	<0.01	<0.01
Standard	-1.94	<0.01	<0.01	<0.01
Luxury	-2.97	<0.01	<0.01	<0.01
SUV	-2.22	<0.01	<0.01	<0.01
Sports	-2.15	<0.01	<0.01	<0.01
	Nested Logit			
Subcompact	-1.18	0.02	0.01	<0.01
Compact	-1.81	0.04	0.03	<0.01
Intermediate	-2.21	0.06	0.04	<0.01
Standard	-3.08	0.11	0.11	<0.01
Luxury	-4.63	0.19	0.09	<0.01
SUV	-3.59	0.06	0.06	<0.01
Sports	-3.43	0.07	0.05	<0.01
	RC Logit			
Subcompact	-2.99	0.05	<0.01	<0.01
Compact	-3.64	0.03	<0.01	0.01
Intermediate	-4.11	0.02	<0.01	0.01
Standard	-5.33	0.03	0.03	0.01
Luxury	-5.52	0.05	0.03	0.01
SUV	-4.55	0.03	0.03	0.01
Sports	-5.20	<0.01	<0.01	0.02
	RC Nested Logit			
Subcompact	-2.48	0.03	0.03	<0.01
Compact	-3.43	0.06	0.07	<0.01
Intermediate	-4.02	0.09	0.09	<0.01
Standard	-5.08	0.17	0.17	<0.01
Luxury	-6.73	0.23	0.16	<0.01
SUV	-5.61	0.09	0.09	<0.01
Sports	-5.21	0.10	0.09	<0.01

The table reports product-level own- and cross-price elasticities, based on the parameter estimates in Table 4. Elasticities are averages by segment for France in 2006, instead of for Germany as in Table 6 of the main text. Cross-price elasticities are averaged across products from the same subsegment, from a different subsegment within the same segment, and from different segments.

Table A.3: The Effects of Two Hypothetical Mergers in France and Germany - Foreign Market

France	All	Subc	Comp	Interm	Stand	Lux	SUV	Sport
	BMW–VW merger in France							
	Foreign market shares (in percent)							
BMW	3.1	0.9	2.7	-	29.2	15.8	7.9	5.6
VW	11.8	7.8	16.1	20.2	28.3	19.4	5.6	11.3
	Predicted foreign price increase (in percent)							
Logit	0.0	0.0	0.1	0.0	0.2	0.1	0.0	0.0
Nested Logit	0.6	0.2	0.6	0.0	4.7	1.5	0.2	0.6
RC Logit	0.5	0.3	0.5	0.4	1.8	1.0	0.6	0.5
RC Nested Logit	0.4	0.2	0.5	0.1	2.7	1.1	0.2	0.4
Germany	PSA-Renault merger in Germany							
	Foreign market shares (in percent)							
PSA	6.1	11.3	4.3	5.7	-	0.9	-	13.8
Renault	4.2	8.3	4.1	2.3	-	0.2	-	5.0
	Predicted foreign price increase (in percent)							
Logit	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Nested Logit	0.4	1.5	0.4	0.2	0.0	0.0	0.0	0.6
RC Logit	0.2	0.9	0.2	0.1	-0.0	0.0	-0.0	0.2
RC Nested Logit	0.2	0.6	0.2	0.1	0.0	0.0	0.0	0.3

Similar to Table 9, the table reports percentage price increases for two hypothetical mergers, BMW–VW and PSA–Renault, but now in their respective foreign markets, France and Germany, instead of the domestic markets. The results are based on the parameter estimates in Table 4 and assuming multiproduct Bertrand competition. The effects refer to France and Germany in 2006. Subc=subcompact, Comp=compact, Interm=intermediate, Stand=standard, Lux=Luxury, SUV=Sport Utility Vehicle.

Table A.4: The Effects of Two Hypothetical Mergers in France and Germany - Confidence Intervals

France	All	Subc	Comp	Interm	Stand	Lux	SUV	Sport
	PSA–Renault merger in France							
	Domestic market shares (in percent)							
PSA	33.4	35.3	38.8	46.0	-	19.1	-	37.3
Renault	22.7	29.8	20.9	17.8	-	9.5	-	13.5
	95 % Confidence Interval for predicted domestic price increase							
Logit	0.7;1.8	1.3;1.9	0.7;1.1	0.6;0.9	0.0;0.0	0.1;0.2	0.0;0.0	0.4;0.6
Nested Logit	12.5;18.3	24.9;37.2	11.1;15.9	10.2;15.4	0.0;0.0	1.5;2.7	0.0;0.0	5.3;8.6
RC Logit	14.6;27.2	28.3;48.7	15.6;31.2	14.8;35.3	0.3;0.9	2.9;7.3	0.0;0.3	9.8;19.0
RC Nested Logit	5.4;15.7	10.2;30.7	5.2;16.1	5.3;16.7	-0.3;0.0	0.9;3.04	-0.2;0.0	2.8;9.3
Germany	VW-BMW merger in Germany							
	Domestic market shares (in percent)							
BMW	10.6	2.1	7.9	-	39.6	25.3	15.2	10.8
VW	30.8	23.1	36.3	53.8	31.3	32.4	12.0	21.4
	95 % Confidence Interval for predicted domestic price increase							
Logit	0.3;0.4	0.2;0.3	0.3;0.5	0.2;0.3	0.5;0.8	0.3;0.4	0.2;0.2	0.2;0.2
Nested Logit	2.7;3.0	0.5;0.6	2.6;3.0	0.1;0.1	9.5;10.5	4.1;4.5	1.3;1.7	1.0;1.2
RC Logit	1.9;2.5	0.5;0.8	1.7;2.4	1.6;2.1	4.2;5.8	2.7;3.8	1.5;2.0	1.3;1.8
RC Nested Logit	1.6;2.4	0.4;0.8	1.4;2.3	0.2;0.8	5.0;7.0	2.6;3.6	0.8;1.6	0.7;1.1

The table reports the 95 percent confidence intervals for the percentage price increases reported in Table 9 for two hypothetical mergers, PSA–Renault and BMW–VW, in their domestic markets France and Germany, based on the parameter estimates in Table 4 and assuming multiproduct Bertrand competition. The 95% confidence intervals are based on a bootstrapping procedure. Subc=subcompact, Comp=compact, Interm=intermediate, Stand=standard, Lux=Luxury, SUV=Sport Utility Vehicle.