

# Random coefficient car demand models

Rossi ABI-RAFEH

November 1, 2016

I try to replicate the results of Berry, Levinshon and Pakes paper on car demand estimation, using a subsample of their data (years from 1977 until 1981). Instead of estimating an aggregate system of demand equations where each car model's market share depends on all the product prices, the authors reduce the number of estimated parameters (dimensionality problem) by positing an individual utility function, linear in car characteristics and car price, then aggregate over consumers to obtain moments that can be matched to the data. In the estimations, some car characteristics are assumed to be observed by both manufacturers and consumers, but not the econometrician, and enter into the decision of the buyer. However, these unobserved demand attributes are likely to be correlated with the price of the car, resulting in the usual simultaneity bias. The solution that BLP propose is to instrument the car price with other cars' characteristics. When allowing for heterogeneity in taste across consumers, the instruments enter the aggregate moment conditions in a non-linear way, a problem that BLP solve by using a contraction mapping to solve for the inverse function of the market shares, and then non-linear GMM for the parameter estimation. I assume all along the replication that car buyers are price takers.

Section 1 presents the three models that I use, as well as the estimation procedure of the full model allowing for taste heterogeneity, and the instruments used for each model. Section 2 presents the data and the results, and compares them to the results of BLP.

## 1 Demand models

The primitives of the demand models are an assumption of utility maximizing buyers, and a functional form for their utilities. In each of the following models, I assume an individual (indirect) utility  $u_{jt,i}$  from buying a car that is linear in a car's own characteristics  $x_{jt}$  and own price  $p_{jt}$ , following the discrete choice literature.

$$u_{jt,i} = x'_{jt}\beta_i - p_{jt}\alpha_i + \xi_{jt} + \epsilon_{jt,i} \quad (1)$$

The utility is decomposed into a mean utility part  $\delta_{jt} = x'_{jt}\bar{\beta}_i - p_{jt}\bar{\alpha}_i + \xi_{jt}$  and an individual specific part  $\mu_{jt,i} = \sum_{k \in K} x_j^k \sigma^k \zeta_i^k + \epsilon_{jt,i}$ .  $\xi_{jt}$  includes the car characteristics relevant to the consumers' choice, and observed both by the consumer and the manufacturer but not econometrician. Thus the model allows for explicit observed characteristic differentiation between cars, as well as unobserved product differentiation.

There exists an outside option for each buyer, in each market, that provides utility  $u_{0t}$  that is normalized to 0. I interpret the outside option share as the number of households in the market that did not purchase a car. The models use aggregate sales data, and assume that individual choices of purchases are not observed, thus ought to link the observed aggregate market shares to the individual level utilities. This is done by integrating over the values of the individual specific taste shocks that result into a consumer choosing product  $j$  over any other product. Call the set of these values  $A_{jt} = \{\epsilon_{jt,i}, s.t. u_{jt,i} \geq u_{qt,i} \forall q \neq j\}$  for each product  $j$ .

$$s_{jt} = s(\delta_{jt}, \alpha, \beta, \sigma) \quad (2)$$

This equation is then inverted (analytically in the logit and nested logit model, or numerically, in the full random-coefficients model), and the resulting  $\delta_{jt}$  are used in the estimation.

## 1.1 Logit Model

The simple logit model assumes there is no heterogeneity in taste across consumers. Thus  $\delta_{jt} = x'_{jt}\beta - p_{jt}\alpha$ , and  $\mu_{jt,i} = \epsilon_{jt,i}$ . This implies that the observed car characteristics are additively separable from the individual taste shocks. In addition, the individual taste shocks  $\epsilon_{jt,i}$  are assumed to be independent across products and individuals, and follow a Type I extreme value distribution.

Under these assumptions, the market share  $s_{jt}$  is a logistic closed form function of the mean utilities, and the inversion is done analytically :

$$\log(s_{jt}) = \delta_{jt} - \log\left(\sum_{q \in J} \exp(\delta_q)\right) \quad (3)$$

Thus, by subtracting the market share of the outside option, we obtain an estimable equation, that we estimate using both OLS and IV in section 2.

$$\log(s_{jt}) - \log(s_{0t}) = \delta_{jt} - \delta_{0t} = x'_{jt}\beta - p_{jt}\alpha + \xi_{jt} \quad (4)$$

The need for instrumental variables comes from the endogeneity in prices. The price of a product  $j$  is likely to be set by the manufacturer simultaneously with the unobserved car characteristics, leading to simultaneity bias, the differentiated product version of the simultaneity problem common in homogeneous product demand estimation. BLP motivates the use of three sets of instruments for the price of car  $jt$  :

1. For each product  $\{j,t\}$ , the observed car characteristics  $x_{j,t}$
2. The sum of observed car characteristics for cars from the same brand  $\sum_{q \neq j, q \in J_r} x_{q,t}$
3. The sum of observed car characteristics for cars from other manufacturers  $\sum_{q \neq j, q \in J_r} x_{q,t}$

The identification assumption is an exclusion restriction, where we require the observed characteristics used to compute the instruments to be correlated to the price, but mean independent from the unobserved characteristics  $E(x_{qt}\xi_{jt}) = 0$ .

This simple logit models implies unrealistic substitution patterns on the demand system: the cross and own price elasticity of any product depends exclusively on its market share. Thus two products that have the same market share will have the same cross-price elasticities with any third product, even if they had very different characteristics. However, we expect substitution between products to follow taste along other characteristics than price, for instance, if the price of a Toyota Corolla were to increase, we expect demand to shift more towards cars of similar size category than bigger cars.

## 1.2 Nested Logit Model

In order to obtain more realistic substitution patterns, one can incorporate heterogeneity in taste by nesting cars into groups  $r \in R$ , and allowing for correlation in the individual utility terms in choices within the same group. The nested logit specification includes an extra parameter  $\sigma$  to allow for correlation between unobservable taste factors between products of the same nest. The nests, or car groups, are imposed ex-ante however.

$$u_{jt,i} = \delta_{jt} + \zeta_r + (1 - \sigma)\epsilon_{jt,i} \quad (5)$$

The individual taste shocks  $\{\epsilon_{jt,i}\}$  are assumed to be independent (across products in the same market, and individuals) Extreme Value Type I.  $\zeta_r$  has the unique distribution (for each  $\sigma$ ) such that  $\zeta_r + (1 - \sigma)\epsilon_{jt,i}$  is itself Type I Extreme Value (proof of existence and uniqueness of this distribution in Cardell 1991). Under these assumptions, the mean value function is analytically solved, and leads to the following estimable moment.

$$\log(s_{jt}) - \log(s_{0t}) = x'_{jt}\beta - p_{jt}\alpha + \sigma \log(\bar{s}_{j/g,t}) + \xi_{jt} \quad (6)$$

When  $\sigma$  is close to 1, the only individual taste parameter is the one relative to the group of products, creating correlation between tastes within that group. When  $\sigma$  is 0, we are back in the simple logit model.

Compared to the simple logit specification, the linear regression here include an extra term as a dependant variable :  $\bar{s}_{j/g,t}$  the share of car  $j$ 's sales as a fraction of its category  $g$  total sales. This term is correlated to the intra-group unobservable term  $\zeta$ , thus is also endogenous, and needs to be instrumented for. I include for each product  $jt$  the sum of characteristics of other products in the same category  $\sum_{q \in J_g, q \neq j, j \in J_g} x_{qt}^k$  as a fourth set of instruments as suggested by Berry 1994.

### 1.3 BLP - full random coefficients model

The heterogeneity in the nested logit is defined ex-ante within set groups. The full random coefficients model used in BLP allows for heterogeneity between consumers along one or all taste attributes. The model is in the general form written above. In this case, we have only one taste parameter with a random coefficient : size, which simplifies the notation.

Now the market share function of the mean utilities is not a linear function, and cannot be computed in closed form easily (as it involves integrals). The inversion is done based on the contraction mapping result of Berry 1994, and the mean utilities found are used to compute a GMM objective function based on the moment conditions of the model and the instruments. Numerically, I follow the steps in the Appendix of Nevo 2001, and recreate his code in R and simplify it to have  $\theta_2$  of dimension 1.

After solving the model for the value of  $\sigma$  that minimizes the GMM objective function (thus minimizes the distance between the observed market shares and the ones predicted by the model), I compute standard errors using the inverse of the jacobian of the contraction mapping at the minimum. These standard errors do not take into account the simulation error. I also use the jacobian to compute the elasticities as given in the appendix of Nevo 2001.

### 1.4 Supply side and markups

In order to derive the markups over marginal cost, we assume that car manufacturers are multi-product firms in a Bertrand competition . The Nash assumption is the final primitive of the model. The profit of firm  $f$  in market  $t$  of size  $M$  is :

$$\Pi_{ft} = \sum_{jt \in J_{ft}} (p_{jt} - mc_{jt}) M s_{jt} \quad (7)$$

The first order condition of the profit maximizing problem, and re-arranging the terms lead to an expression of thee markups for product  $j$  :

$$markup_{jt,r} = \Delta_{jrt}^{-1} s_{jt} \quad (8)$$

where  $\Delta_{jrt} = -\frac{\partial s_{rt}}{\partial p_{jt}}$  if  $j \in J_r$  and 0 otherwise.

## 2 Data and Results

### 2.1 Data

The price variable is the retail price in thousands of 1993 US dollars. As in BLP, I use the Consumer Price Index to deflate, thus should have similar prices. The quantity variable is the number of sales of a given model in the US market, in 1000 units. I observe sales, prices and car characteristic from 1977 until 1981. BLP's sample includes mine but also data from 1971 until 1990. My panel has 501 product/year observations. Overall, there are 192 distinct car models. The panel is unbalanced due to some models appearing only in a subset of markets. BLP's model restrict different products to have conditional distribution-independent unobservables, but "the unobservables for different years of the same model are allowed to be freely correlated".

Instead of using dollars per mile (DpM), I use the inverse measure : miles per dollar (MpD = Miles / Gallon \* Price of a gallon gasoline adjusted for inflation = 1/DpM). The number of miles one can drive per dollar spent on gasoline is a better measure of fuel efficiency, and the obtained estimate is comparable to

BLP's. As discussed above, the supply side is assumed to follow Bertrand-Nash competition between multi-product firms. Thus, for each product, I identify the brand/manufacturer. Brands that are from the same manufacturers are pooled together : for instance, Audi and Volkswagen (7); Buick, Cadillac, Chevrolet and Pontiac (General Motors, 19). The country of origin of each product is the country of the manufacturer. I scale the variables of interest to match those in the BLP data, in order to get comparable estimate magnitudes.

Table 1: Descriptive Statistics

Year	No. of Models	Quantity	Price	Japan	Euro	HP/Wt	Size	A/C	MpG	MpD
1977	94	98.232	7.653	0.107	0.051	0.340	1.469	0.032	1.946	1.912
1978	92	97.674	7.618	0.088	0.040	0.345	1.409	0.035	1.965	2.037
1979	101	81.016	7.653	0.132	0.040	0.350	1.353	0.048	2.028	1.706
1980	102	72.112	7.707	0.192	0.034	0.350	1.296	0.078	2.216	1.518
1981	112	63.846	8.319	0.214	0.043	0.349	1.285	0.094	2.367	1.609

Table 2: Descriptive Statistics - continued

Year	3 doors	4 doors	5 doors	Automatic	Power steering	Front wheel drive	Horsepower	Weight	Wheelbase
1977	0.019	0.586	0.059	0.346	0.480	0.073	1.121	3.277	1.095
1978	0.026	0.610	0.052	0.241	0.306	0.128	1.046	3.031	1.072
1979	0.057	0.501	0.086	0.192	0.265	0.167	1.004	2.863	1.052
1980	0.066	0.480	0.142	0.217	0.241	0.228	0.944	2.702	1.038
1981	0.061	0.470	0.109	0.254	0.382	0.335	0.929	2.666	1.032

Table 3: Dispersion of continuous demand characteristics

Car characteristic	Mean	Median	SD	Min	Max
Price	10.886	7.910	7.817	4.586	46.526
Sales	81.654	47.345	90.780	0.589	527.939
HP/Wt	0.362	0.348	0.083	0.170	0.786
HP	1.042	0.980	0.344	0.460	2.200
Weight	2.887	2.967	0.702	1.445	5.000
Wheelbase	1.045	1.049	0.091	0.866	1.272
Size	1.319	1.293	0.243	0.852	1.864
MpD	1.729	1.631	0.528	0.891	3.832
MpG	2.088	2.000	0.571	1.000	3.900

Tables 1 and 2 provide summary statistics for the car sales, and the mean prices, and characteristics, weighted by sales. Sales are in 1000 of cars sold, prices are in 1000 USD (1983). Horsepower per weight (HP/Wt) is in HP per 10 lbs. Size is measured as length\*width of the car, in 10,000 square inches, e.g. a size of 0.8 is then 8,000 square inches. Related measures are the wheelbase in meters, and the weight. My data is not exactly similar to the one in BLP: 1 to 4 models less in my sample each year. However, this does not substantially affect the level of mean characteristics, nor their overall trends. Car sales decreased by 35 % on average during the sample period, but prices increased, as well as fuel cost-efficiency. Only 9.4 % of new cars had air-conditioning (A/C). The variables weight and wheelbase are highly ( $\rho$  0.9) correlated with the size of the vehicle. To avoid problems of near multi-collinearity, I exclude both from the following regressions and keep only *size*. Faced with two cars of the same overall size, I can hardly think of any reason why a car buyer would choose one over the other because of weight or the distance between tyre axes! Arguably, size is also the most salient information out of the three measures, and the easiest to assess for a buyer.

Table 3 summarizes the distribution of the continuous car characteristics. Continuous car characteristics enlarge the space of possible product differentiation, relative to discrete attributes like the number of doors or whether a car has automatic transmission or not. Car size varies from 8 to 18 thousand square inches. To

give an idea, a hatchback Honda Civic is around 8,000 square inches, and a Pontiac F-Bird is around 14,000 sqin. Fig ?? shows that the distribution of sizes among new cars in 1977 and in 1981 is wide and has fat tails, an indication of a lot of product differentiation along this dimension.

## 2.2 Logit

Table 4: Market Shares

Year	Sales	Market size	$s$	$s_0$	$\log(s) - \log(s_0)$
1977	98.232	74.142	0.132	87.546	-7.217
1978	97.674	76.030	0.128	88.181	-7.251
1979	81.016	77.330	0.105	89.419	-7.404
1980	72.112	80.776	0.089	90.894	-7.679
1981	63.846	82.368	0.078	91.318	-7.871

The first set of results are based on a simple logit specification for the underlying utility function. The additive separability between observable and unobservable characteristics in the utility, the independence assumption and the type I extreme value type 1 distribution assumed on the  $\epsilon_{ijt}$  allow to analytically invert the market share function. In practice, I regress the *logit* mean utility  $\delta_{jt} - \delta_{0t} = \log(s_{jt}) - \log(s_{0t})$  on car characteristics and price, where  $s_{jt}$  is the market share of car model  $j$  in market  $t$ . The market shares are computed as the sales to market size ratio, and the market size is the number of households (in units of 1000 in Table 4) in the United States. For each year, I compute the market share of the outside option, and the corresponding mean utility  $\delta_{0t}$  is normalized to 0 for every market. The average market share in the sample is very low relative to the outside option.

I do the regression on two sets of car characteristics : the one used in BLP (HP/Wt, a dummy for air conditioning, miles per dollar MpD, size and a constant), and the full set (in addition to the previous set, this one includes, dummies for the number of doors, a dummy for front wheel drive DRV, a dummy for automatic transmission AT, horsepower HP, a dummy for power steering PS and a dummy for the nationality of the manufacturer Euro and Japan.)

The first column of Table 5 gives the results of the OLS estimation closest (if not identical) to the one in BLP. I obtain parameters of the expected sign for all characteristics, and statistically significant. The price parameter is negative, as in BLP, even before instrumenting. Increasing horsepower to weight ratio, or adding air-conditioning to a car increases the utility of the buyer, unlike the BLP results. The magnitude on the price parameter is higher (in absolute value) than in BLP, which is likely to lead to more with elastic demands than what they argue in their paper (although I am not how exactly they define the number of car models with inelastic demand there!). The magnitudes of the parameters is not easily interpretable per se, as the utility is made cardinal through a normalization of the outside option utility, but the elasticities and markups are.

In the second column of Table 5, I re-estimate the logit utility specification, with instruments to allow for car characteristics unobserved to the econometrician. Since I am including a constant as a car characteristic (to have mean 0 unobserved disturbances), the instruments include the number of car lines by the same manufacturer. For the second set of instruments (characteristics of other car models by the same manufacturer), the number of models, and the sum of characteristics of MpD, 4 doors, automatic transmission, power steering, horsepower to weight ratio, and horsepower are highly correlated (coefficient higher than 0.9, if not perfectly collinear). I keep only the number of car lines by the same manufacturer as instrument among them. I do a similar correlation analysis on the instruments of the set 3 (characteristics of car models by all other manufacturers), and keep from them the number of car models by other manufacturers, and the sum of characteristics 3 doors, automatic transmission, air conditioning and front wheel drive. The estimation is a two-stages least-squares estimation (using

`ivreg`

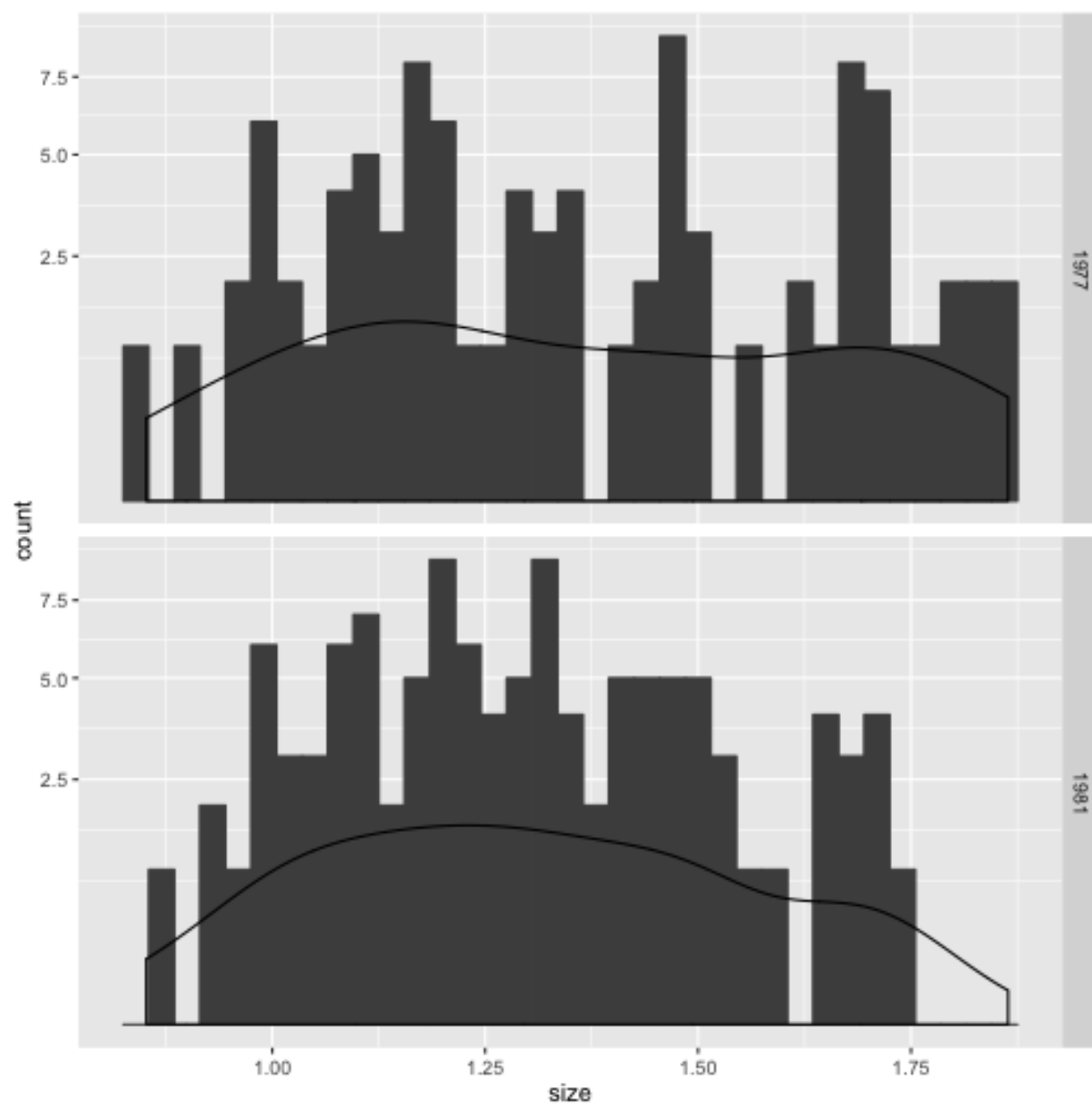


Figure 1: Car Size Distribution

Table 5: Results with Logit Demand (510 Observations)

Variable	OLS	IV	OLS	IV	OLS	IV
CONSTANT	-11.907 (0.621)	-11.848 (0.731)	-9.959 (0.79)	-9.899 (0.796)	-8.411 (0.748)	-9.627 (0.949)
HP/Wt	2.303 (0.664)	5.715 (0.981)	1.374 (1.269)	1.224 (1.283)	-0.747 (1.244)	0.545 (1.554)
A/C	0.296 (0.167)	1.494 (0.287)	-0.009 (0.162)	0.131 (0.2)	-0.232 (0.156)	0.767 (0.413)
MpD	0.686 (0.122)	0.495 (0.148)	0.561 (0.108)	0.563 (0.109)	0.501 (0.095)	0.57 (0.128)
Size	2.794 (0.258)	2.854 (0.304)	1.837 (0.507)	1.716 (0.52)	1.181 (0.466)	1.17 (0.678)
Price	-0.123 (0.008)	-0.234 (0.021)	-0.05 (0.01)	-0.075 (0.023)	-0.077 (0.014)	-0.187 (0.066)
-						
3 doors			-0.705 (0.171)	-0.701 (0.172)	-0.585 (0.153)	-0.682 (0.203)
4 doors			-0.033 (0.098)	-0.039 (0.098)	0.067 (0.087)	-0.063 (0.117)
5 doors			0.422 (0.202)	0.434 (0.204)	0.568 (0.176)	0.488 (0.242)
AT			-0.321 (0.168)	-0.257 (0.177)	-0.26 (0.161)	0.036 (0.261)
PS			0.138 (0.155)	0.138 (0.156)	0.084 (0.148)	0.135 (0.183)
DRV			0.195 (0.11)	0.162 (0.114)	0.26 (0.108)	0.016 (0.156)
HP			-0.367 (0.418)	-0.055 (0.494)	0.299 (0.4)	1.363 (0.955)
Euro			-1.664 (0.15)	-1.442 (0.238)	-2.902 (0.416)	-0.43 (0.608)
Japan			-0.077 (0.155)	-0.036 (0.159)	-0.25 (0.206)	0.15 (0.212)
-						
Brand FE					YES	YES
R2	0.476	n.a.	0.476	n.a.	0.731	n.a.

from the package AER), and the standard errors are corrected.

The use of instruments doubles the magnitude of the price parameter : products of higher unobserved quality have higher prices on average, and not controlling for this unobserved quality in the OLS leads to a downward bias (in absolute value) in the price parameter. This is also what BLP find when they instrument, and conclude that correcting for the endogeneity of prices matters. They also interpret their low goodness of fit of the logit OLS model as an indication of the importance of unobservable demand characteristics. However in our case, the R2 for the logit OLS is just under 50%, 10% higher than that of BLP.

The third and fourth columns in Table 5 repeats similar estimations but includes all of the available car characteristics. The most remarkable change is the great decrease (in absolute value) in the magnitude of the price coefficient, although it stays significantly different than 0. Although the price parameter using IV is higher (in absolute value) than with OLS, it is still much lower than when only the BLP attributes were included. In the BLP paper, the authors do sensitivity analyses on the included attributes and find that overall the trends in the markups are elasticities are not much different, and the values do not change dramatically. This will not be the case in my estimation as the price parameter (part of both the markup and the elasticities through  $\frac{\partial \delta - 1}{\partial p}$ ) is 6 times lower.

Buying a Ford may not hold the same intrinsic value for the buyer as buying a Cadillac, even if the two cars were identical in all other aspects. If this is the case, brand fixed effects are not explicitly accounted for in the model, my estimates will be biased because the car brand is likely to be correlated with price and car characteristics. I run additional regressions including brand dummies to control for any time-invariant

brand effect. The results are in column 5 and 6, respectively OLS and IV. Including firm dummies increases the magnitude of the estimated marginal utility of money, and even more so when price is instrumented. Brand effects seem to play an important role in the choice of buying a car as the overall goodness of fit of the model jumps to 70 % (even the adjusted R2, not reported.)

## 2.3 Nested Logit

The nests/groups I use are the car categories : Compact, Midsize and Large. The market share of compact cars is stable over the sample period, but Figure ?? shows the market share of large cars in 1981 is less than half what it was in 1977. The drop in the total car market share relative to the outside option (the number of people not buying new cars in the US) during the sample period is mostly traced back to the drop in sales of large cars. The market share of midsize cars increased (tripled!) during the period, as well as the number of new midsize models on the market.

The first column in Table 6 gives the estimates from a simple OLS estimation. The estimated parameters are of the expected signs, and all are statistically different than 0 except for air conditioning. The estimated parameter on price has a 5 times smaller magnitude than the closest logit specification. All the other magnitudes also decreased. The estimate of  $\sigma$  is 0.9, close to 1, and estimated precisely enough to be different than 0, indicating that there is substantial correlation in taste for products in the same size group. Adding all the attributes as before has the now usual effect of reducing the magnitude, and the precision of the estimates, although the price estimate stays negative and significant.

Columns 3, and 4, give the estimates for the IV 2SLS estimation of the model with the BLP attributes. The model is overidentified, as I have more instruments than endogenous variables (2). Now the additional instruments include the number of car lines within the same size category, and the characteristics of these cars. I only keep the number of car models in the same nest, and the sum of the characteristics A/C, miles per dollar and size. I use all of my instruments in column 3, however, the Sargan J test reported at the bottom of the table is rejected, indication that some of the moment conditions used are not valid in the data. The Sargan J test uses the excess information from overidentifying moment conditions to test the null hypothesis of having all the moment conditions precisely at 0. In an attempt to alleviate this potential endogeneity of my instruments, I present in column 4 the results from an IV estimation using instruments : number of car models by the same manufacturer, number of car models by all other manufacturers, number of car models within the same size category, and the fourth set of instruments. The Sargan J test does not reject the null up to a level of risk of 74%, and the weak instruments test for both endogenous variables are rejected (indication that there is substantial correlation between the instruments used and the endogenous variables). The coefficient on price is now as high in magnitude as what we find in the IV estimation of the simple logit specification. However, the IVs used do not allow to precisely estimate the coefficient  $\sigma$ . Including firm dummies in the OLS and the IV estimation reduces even further the estimate of the marginal utility of price.  $\sigma$  has a problematic negative sign, and cannot be interpreted because it is estimated very imprecisely.

## 2.4 Full random coefficients model

Table 7 presents the results of a BLP model where the taste parameter on car size is a random parameter. This is a simpler version of the model estimated in BLP, where all observed characteristics are random. The results are somewhere sensitive to the choice of instruments : the first two columns give the results of the estimation of the parsimonious model with only BLP attributes. The first column (M1) uses all the IVs from the sets 1,2, 3 and 4 (nested logit), whereas column 2 (M2) uses only the instruments of the constant terms, and column 3 (M3) excludes the set number 3 sum of characteristics of cars from other manufacturer. I suspect that the low variability of this instrument may be driving the size standard deviation estimate towards 0, and decreases substantially the precision of the estimation. This is also the case in column 4 (M4) where I use the instruments including set number 3, and include all the attributes of a car in the estimation. When I remove the instruments of set 3, as I do in column 5 (M5), I recover values for  $\sigma$  that are more inline with the BLP values. The coefficient on price is high enough to match the values obtained with IV nested



logit/logit. However when we include all the attributes, this estimate drops, and we are not able to easily include firm dummies in the BLP method to recover higher values. The coefficients on other BLP attributes are of the expected sign.

Instead of interpreting the magnitudes of the direct estimates in the utility function, Table ?? presents a sample of own and cross price semi-elasticities, using the model specification that is most comparable with the one of BLP (BLP variables, instruments M2). Each semi-elasticity gives the percentage change in market share of the row car associated with a 1000 USD increase in the price of the column car model, in 1981. An increase of 1000 USD in the price of a Ford Escort decreases by more than 120 % its market share, and increases the market shares of the competitors by 0.3%. My estimates for price semi-elasticities are in the same range as those found in BLP, however they do not vary as much between car models. The standard deviation of the (random) taste parameter on should affect the elasticity, since it is the maximum utility that affects the choice of one car instead of the mean utility. The increase in price of the Ford Escort is expected to have disproportionate substitution effect on cars with similar compact size, relative to ones that are larger. However, this effect is not as clear in my estimates as it is in the BLP results. It is possible that this is due to having only one random coefficient in my model, whereas BLP estimated a model where all coefficients were random except the one of price. To further study the effect of larger standard deviation of taste parameters, I report the price semi-elasticities for the same sample of car models, using the specifications of model M4 (Table 9) and M5 (Table 10). As indicated above, M4 drives down the estimated value of the standard deviation on the size parameter to practically 0 (nearly equivalent to an IV estimation in a logit model), whereas M5 (excluding the instruments of set 3) gives a standard deviation of 1. As expected, Table 9 shows that the unrealistic substitution pattern of a logit model, but increasing the standard deviation of the taste parameter on size leads to richer substitution patterns in Table 10.

Table 11 presents the implicit markup over marginal cost under the assumption of Bertrand competition between multi-products firm. The values obtained do not match those of BLP, and cannot be valid, as they indicate markups that are higher than prices, thus negative marginal costs. This is at odds with profit maximizing behaviour.

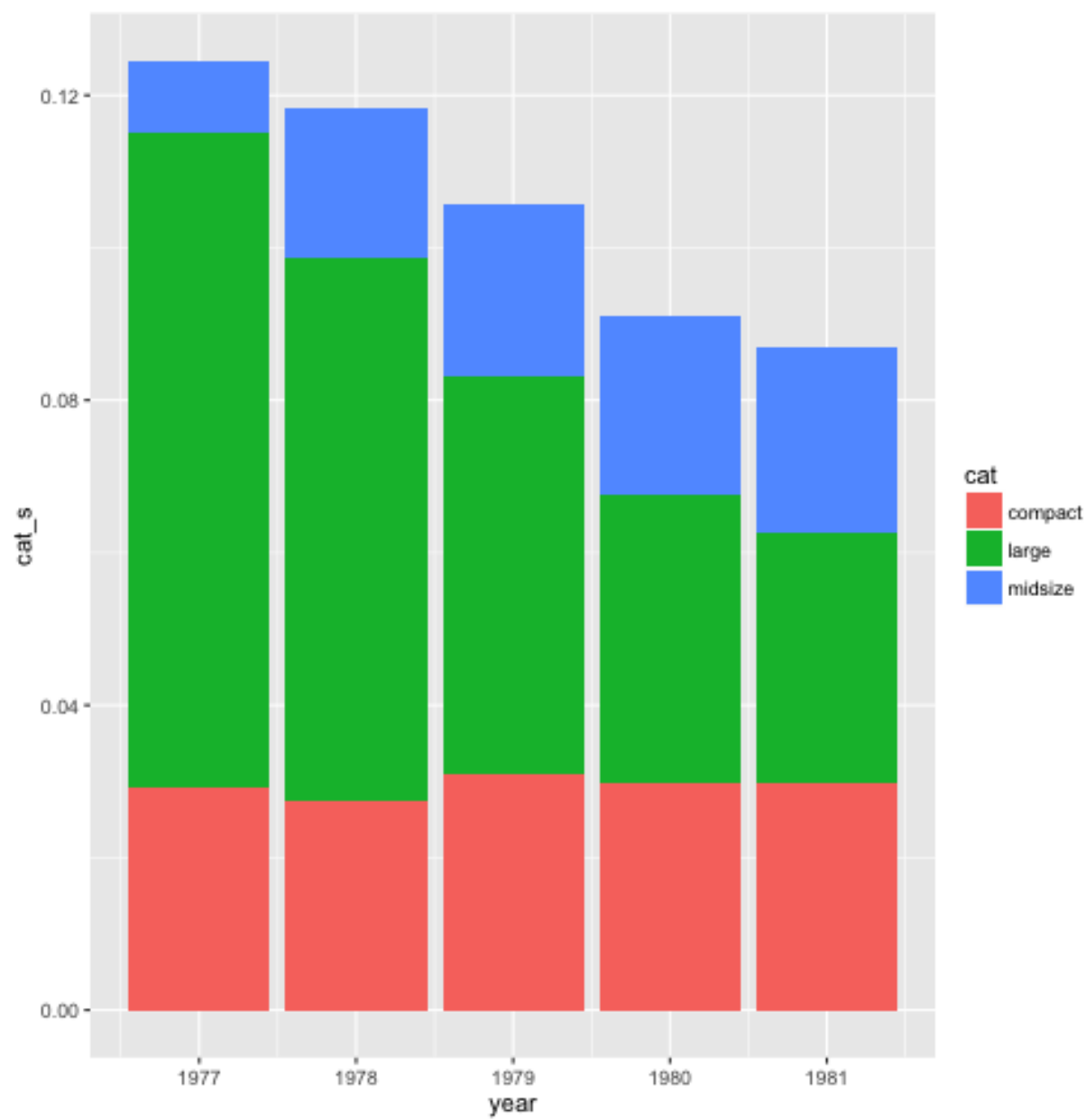


Figure 2: Market shares of car categories

Table 6: Results with Nested Logit Demand (510 Observations)

Variable	OLS	IV	IV	OLS	IV	IV	OLS	IV
CONSTANT	-6.179 (0.249)	-9.612 (0.822)	-11.135 (1.223)	-5.987 (0.335)	-7.099 (0.459)	-7.418 (1.017)	-5.777 (0.35)	-10.534 (1.973)
HP/Wt	1.037 (0.244)	4.948 (0.805)	6.681 (1.227)	-0.227 (0.524)	0.065 (0.66)	-0.587 (1.37)	-1.276 (0.572)	-0.748 (2.562)
A/C	0.042 (0.062)	1.299 (0.233)	1.856 (0.363)	-0.02 (0.067)	0.158 (0.093)	1.032 (0.268)	-0.006 (0.072)	-0.256 (0.58)
MpD	0.194 (0.046)	0.317 (0.117)	0.372 (0.163)	0.18 (0.045)	0.296 (0.059)	0.373 (0.126)	0.195 (0.044)	0.754 (0.231)
Size	1.606 (0.097)	2.384 (0.264)	2.73 (0.378)	1.457 (0.209)	1.42 (0.264)	0.738 (0.565)	1.325 (0.215)	0.864 (0.943)
Price	-0.023 (0.003)	-0.186 (0.024)	-0.258 (0.04)	-0.018 (0.004)	-0.059 (0.009)	-0.218 (0.038)	-0.02 (0.007)	-0.163 (0.089)
-								
3 doors				-0.021 (0.072)	-0.22 (0.095)	-0.314 (0.205)	0.05 (0.072)	-1.159 (0.377)
4 doors				-0.031 (0.04)	-0.038 (0.05)	-0.072 (0.102)	0.028 (0.04)	0.105 (0.153)
5 doors				0.109 (0.084)	0.217 (0.105)	0.345 (0.216)	0.202 (0.082)	0.9 (0.345)
AT				0.031 (0.069)	0.006 (0.091)	0.345 (0.209)	-0.069 (0.074)	-0.434 (0.294)
PS				-0.132 (0.064)	-0.052 (0.081)	-0.009 (0.166)	-0.052 (0.068)	0.212 (0.266)
DRV				0.13 (0.045)	0.109 (0.058)	-0.081 (0.126)	0.064 (0.05)	0.462 (0.207)
HP				0.372 (0.173)	0.54 (0.239)	2.354 (0.656)	0.601 (0.184)	0.462 (1.375)
Euro				-0.128 (0.069)	-0.31 (0.133)	0.808 (0.451)	-0.024 (0.203)	-5.278 (1.677)
Japan				0.109 (0.064)	0.104 (0.081)	0.327 (0.176)	0.094 (0.095)	-0.538 (0.402)
-								
sigma	0.93 (0.016)	0.364 (0.1)	0.113 (0.157)	0.031 (0.069)	0.006 (0.091)	0.345 (0.209)	-0.069 (0.074)	-0.434 (0.294)
-								
Brand FE							YES	YES
R2	0.93	n.a.	n.a.	0.935	n.a.	n.a.	0.943	n.a.
DIAGNOSTICS								
Weak IV - Price		17.855	20.913		11.59	6.345		6.306
p-value		0	0		0	0		0
Weak IV - sigma		25.534	33.94		7.385	11.595		3.685
p-value		0	0		0	0		0.001
Sargan J		24.743	1.977		242.325	25.571		3.081
p-value		0	0.74		0	0		0.544

Table 7: Results with BLP Model (510 Observations)

Parameter	Variable	M1	M2	M3	M4	M5
Means	CONSTANT	-11.622 (0.66)	-11.907 (0.69)	-11.705 (0.659)	-9.959 (0.791)	-9.778 (0.794)
	HP/Wt	2.347 (0.787)	2.303 (0.806)	2.329 (0.79)	1.374 (1.378)	1.307 (1.37)
	A/C	0.282 (0.167)	0.296 (0.15)	0.287 (0.168)	-0.009 (0.138)	-0.014 (0.151)
	MpD	0.709 (0.133)	0.686 (0.141)	0.7 (0.133)	0.561 (0.134)	0.572 (0.125)
	Size	2.354 (0.354)	2.794 (0.585)	2.504 (0.39)	1.837 (0.615)	1.611 (0.531)
	Price	-0.123 (0.009)	-0.123 (0.009)	-0.123 (0.009)	-0.05 (0.009)	-0.05 (0.01)
	-					
	3 doors				-0.705 (0.187)	-0.707 (0.189)
	4 doors				-0.033 (0.098)	-0.031 (0.102)
	5 doors				0.422 (0.226)	0.419 (0.232)
	AT				-0.321 (0.152)	-0.329 (0.159)
	PS				0.138 (0.143)	0.136 (0.148)
	DRV				0.195 (0.099)	0.191 (0.107)
	HP				-0.367 (0.41)	-0.336 (0.402)
	Euro				-1.664 (0.148)	-1.663 (0.15)
	Japan				-0.077 (0.148)	-0.083 (0.149)
	-					
	Size	1.015 (1.552)	0.001 (5.222)	0.82 (2.018)	0.001 (4.382)	0.689 (1.537)
	Std Deviations					

Table 8: Prices elasticities - random coefs model - BLP attributes model (510 Observations) – Note : Each cell entry (i, j) gives the percentage change in market share of car model i (row) , with a 1000 USD change in the price of car j (column)

CAR	TY Corolla	FD Escort	CV Chevy	HD Accord	BCK Century	LN TownCar	CD Seville	BMW733i
TY Corolla	-123.110	0.258	0.004	0.426	0.044	0.189	0.034	0.518
FD Escort	0.033	-122.886	0.004	0.426	0.044	0.189	0.034	0.518
CV Chevy	0.033	0.258	-123.140	0.426	0.044	0.189	0.034	0.518
HD Accord	0.033	0.258	0.004	-122.718	0.044	0.189	0.034	0.518
BCK Century	0.033	0.258	0.004	0.426	-123.100	0.189	0.034	0.518
LC TownCar	0.033	0.258	0.004	0.426	0.044	-122.955	0.034	0.518
CD Seville	0.033	0.258	0.004	0.426	0.044	0.189	-123.110	0.518
BMW733i	0.033	0.258	0.004	0.426	0.044	0.189	0.034	-122.626

Table 9: Prices elasticities - random coefs model - All attributes model (510 Observations)

CAR	TY Corolla	FD Escort	CV Chevy	HD Accord	BCK Century	LN TownCar	CD Seville	BMW733i
TY Corolla	-50.158	0.105	0.001	0.174	0.018	0.077	0.014	0.211
FD Escort	0.014	-50.067	0.001	0.174	0.018	0.077	0.014	0.211
CV Chevy	0.014	0.105	-50.170	0.174	0.018	0.077	0.014	0.211
HD Accord	0.014	0.105	0.001	-49.998	0.018	0.077	0.014	0.211
BCK Century	0.014	0.105	0.001	0.174	-50.154	0.077	0.014	0.211
LC TownCar	0.014	0.105	0.001	0.174	0.018	-50.095	0.014	0.211
CD Seville	0.014	0.105	0.001	0.174	0.018	0.077	-50.158	0.211
BMW733i	0.014	0.105	0.001	0.174	0.018	0.077	0.014	-49.961

Table 10: Prices semi-elasticities - random coefs model - All attributes model (510 Observations)

CAR	TY Corolla	FD Escort	CV Chevy	HD Accord	BCK Century	LN TownCar	CD Seville	BMW733i
TY Corolla	-50.126	0.131	0.002	0.216	0.025	0.103	0.019	0.258
FD Escort	0.017	-50.014	0.002	0.213	0.025	0.101	0.018	0.255
CV Chevy	0.018	0.138	-50.141	0.226	0.027	0.109	0.020	0.270
HD Accord	0.017	0.129	0.002	-49.931	0.025	0.101	0.018	0.254
BCK Century	0.019	0.146	0.002	0.240	-50.114	0.118	0.022	0.285
LC TownCar	0.018	0.138	0.002	0.227	0.027	-50.034	0.020	0.270
CD Seville	0.018	0.139	0.002	0.229	0.028	0.111	-50.123	0.272
BMW733i	0.017	0.127	0.002	0.209	0.024	0.098	0.018	-49.893

Table 11: Markups - random coefs model - BLP attributes model (510 Observations)

CAR	Price (in USD)	Markup over mc (p - mc) in USD	Variable profits $q^*(p - mc)$ in USD
Ford Escort	5158	8249.881	2350.449
Chevrolet Chevy	5399	8446.104	2924.945
Toyota Corolla	6719	8175.349	182.613
Honda Accord	7645	8157.315	1407.602
BK Century	7924	8446.104	1067.309
LN TownCar	14985	8249.881	240.715
CD Seville	23000	8446.104	191.963
BMW 733I	31980	8124.721	19.938