Demand for Differentiated Products

Problem Set 1,
Empirical Industrial Organization

Zixuan Fu Anoosheh Mirkhushal Shuo Song

October 2024

1 Introduction

Q: Should we include a table on descriptive stats?

2 Logit

Using the following specification for the consumer's random indirect utility,

$$u_{ijt} = \boldsymbol{X_{jt}\beta} + \alpha p_{jt} + \xi_{jt} + \epsilon_{ijt}$$

we obtain the following linear estimable equation in the logit model (where ϵ_{ijt} are iid and follow a Type-I extreme value distribution:

$$log\left(\frac{s_{jt}}{s_{0t}}\right) = \boldsymbol{X_{jt}\beta} + \alpha p_{jt} + \xi_{jt}$$

The estimates obtained for the demand equation above estimated via OLS and IV are presented in Table 1. In the IV specification, following Berry, Levinsohn and Pakes (1995), we can make use of a function of rival cars' characteristics as a determinant in the price-setting decision of firms to act as instruments for the (endogenous) price. One set of instruments comes from product characteristics of the *same* firm, another from those of all *other* firms. This is a measure of the competitive environment surrounding firms, capturing the substitutability of other available cars to any particular model.
We observe that the coefficient on price, α, is larger (in absolute terms) but noisily estimated using IV than when we make use of OLS. This makes sense as, from the simple OLS regression, we obtain estimates which are downward biased (in terms of magnitude) because our failure to account for endogeneiety yields an underestimate the effect of price on market share.

Table 1: Logit

	OLS		IV	
Price (in 000s)	-0.0379**	-0.0317*	0.0115	0.0251
	(0.0148)	(0.0189)	(0.0245)	(0.0387)
Dollars per mile	-14.33***	-12.67***	-20.94***	-20.43***
	(3.558)	(3.646)	(4.337)	(5.763)
Automatic transmission	-0.390**	-0.331**	-0.443***	-0.318**
	(0.151)	(0.144)	(0.155)	(0.143)
Power steering	0.296**	0.255*	0.306**	0.284**
	(0.133)	(0.137)	(0.137)	(0.139)
Air conditioning	0.210	-0.224	0.00864	-0.443**
	(0.157)	(0.153)	(0.181)	(0.204)
Front-wheel drive	0.101	0.129	0.132	0.0967
	(0.101)	(0.0999)	(0.102)	(0.0982)
Weight	-0.00243***	-0.00135***	-0.00250***	-0.00128***
	(0.000378)	(0.000314)	(0.000367)	(0.000315)
Horsepower/weight	-9.764***	-5.222***	-9.879***	-4.399**
	(2.383)	(1.967)	(2.204)	(2.038)
Horse power	0.0329***	0.0139**	0.0314***	0.00941
	(0.00798)	(0.00680)	(0.00772)	(0.00739)
European	-1.953***	-1.879***	-2.225***	-2.702***
	(0.143)	(0.320)	(0.178)	(0.401)
Japanese	-0.102	0.608**	-0.177	-0.424*
	(0.146)	(0.293)	(0.144)	(0.230)
Size	2.939***	1.011	2.970***	0.861
	(0.694)	(0.735)	(0.665)	(0.699)
Wheelbase	0.0227	0.0440**	0.0291*	0.0521***
	(0.0183)	(0.0211)	(0.0174)	(0.0202)
Number of observations	501	501	501	501
R-squared	0.647	0.734	0.637	0.728
Brand FE		✓		✓

Notes:
Standard errors are heteroskedasticity-robust
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3 Nested Logit

In the Nested Logit model, random utility is modeled as

$$u_{ijt} = \boldsymbol{X_{jt}\beta} + \alpha p_{jt} + \xi_{jt} + \eta_{igt} + (1-\sigma)\epsilon_{ijt}$$

and from Berry (1994), the linear estimating equation for the relevant demand parameters is given by

$$log\left(\frac{s_{jt}}{s_{0t}}\right) = \boldsymbol{X_{jt}\beta} + \alpha p_{jt} + \sigma log\left(\frac{s_{jt}}{s_{gt}}\right) + \xi_{jt}$$

There is an additional endogenous variable to account for now, namely the within-group share $log\left(\frac{s_{jt}}{s_{gt}}\right)$. We continue to use the idea behind BLP instruments exploiting the competition posed by rivals with the added feature of this 'competition' coming from within the same group (i.e. size category). Our two sets of instruments are now (i) characteristics from products of the *same* firm in the *same* size category and (ii) characteristics from products of the *other* firms in the *same* size category.

The results from demand estimation are presented in Table 2. Once again, the OLS estimates of α are biased towards zero. However, σ decreases in magnitude in going from OLS to IV which suggests a lower degree of intra-group substitutability once we have accounted for the endogeneity of prices and within-group shares.

From the results of the test of overidentifying restrictions, we are able to reject the null that *all* instruments are valid. Our instruments also appear to be weak as indicated by the conditional likelihood ratio (CLR) test for weak IV-robustness (for an over-identified system). We cannot reject the joint null of $H_0 : \alpha = 0,\ \sigma = 0$.

This could point to a possible failure of the exogeneity of the instruments we use.

4 Random-Coefficients Logit Model

Next, we turn to our final specification of a random-coefficients model à la BLP (1995) where indirect utility is defined as,

$$u_{ijt} = \boldsymbol{X_{jt}\beta} + \alpha p_{jt} + \xi_{jt} + \eta_i Sizejt + \epsilon_{ijt} \quad \text{where } \eta_i \sim \mathcal{N}(0, \sigma_\eta^2)$$

The difference here is that the coefficient η_i is allowed to vary across households i. After refactoring the indirect utility, we arrive at the following expression, where the last line arises since we have learned that $\bar{\eta}$ is zero. So that, even if the ϵ_{ij}'s are still i.i.d. TIEV, the composite error is not.

$$\begin{aligned} u_{ij} &= X_j\beta + \alpha p_j + \xi_j + \eta_i Size_j + \epsilon_{ij} \\ &= \delta_j + (\eta_i - \bar{\eta})Size_j + \epsilon_{ij} \\ &= \delta_j + \eta_i Size_j + \epsilon_{ij} \end{aligned}$$

Table 2: Nested Logit

	OLS		IV	
Price (in 000s)	-0.010**	0.008	-0.040**	0.004
	(0.005)	(0.009)	(0.016)	(0.027)
Log $\left(\frac{s_{jt}}{s_{0t}}\right)$	0.888***	0.884***	0.519***	0.254***
	(0.020)	(0.022)	(0.057)	(0.093)
Dollars per mile	-6.926***	-9.527***	-7.531***	-15.009***
	(1.206)	(1.417)	(2.687)	(3.915)
Automatic transmission	-0.012	-0.084	-0.149*	-0.255**
	(0.070)	(0.079)	(0.088)	(0.115)
Power steering	-0.061	0.026	0.084	0.202*
	(0.073)	(0.080)	(0.084)	(0.112)
Air conditioning	0.041	-0.026	0.186*	-0.259*
	(0.067)	(0.081)	(0.108)	(0.143)
Front-wheel drive	0.128***	0.030	0.105*	0.087
	(0.037)	(0.037)	(0.056)	(0.074)
Weight	-0.000**	-0.000	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Horsepower/weight	-1.918*	-0.931	-5.135***	-3.647**
	(0.980)	(0.919)	(1.395)	(1.586)
Horse power	0.008**	0.003	0.019***	0.009
	(0.004)	(0.004)	(0.005)	(0.006)
European	-0.289***	0.000	-0.879***	-1.891***
	(0.076)	(.)	(0.163)	(0.429)
Japanese	0.086	0.020	0.036	-0.276
	(0.057)	(0.143)	(0.085)	(0.185)
Size	1.828***	1.222***	2.278***	1.009*
	(0.325)	(0.354)	(0.421)	(0.551)
Wheelbase	0.001	0.011	0.008	0.038**
	(0.008)	(0.009)	(0.011)	(0.016)
Number of observations	501	501	501	501
R-squared	0.938	0.947	0.886	0.838
Brand FE		✓		✓
CLR Weak IV (p-value)			0.001	0.212
Hansen J statistic (p-value)			0.000	0.001

Notes:

Standard errors are heteroskedasticity-robust

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Now, the aggregate market share becomes

$$s_j = \int \frac{\exp\left(\delta_j + \eta_i Size_j\right)}{1 + \sum_{j'}^{J} \exp\left(\delta_{j'} + \eta_i Size_{j'}\right)} dF(\eta_i).$$

BLP (1995) propose a “nested” estimation strategy, with an “inner loop” nested within an “outer loop”. In the outer loop, we iterate over σ_η and let $\hat{\sigma}_\eta$ be the value of the current round. In the inner loop, given $\hat{\sigma}_\eta$, we solve for the mean utilities (J unknowns) from J equations. For the resulting mean utilities, we estimate the rest of the coefficients via GMM. Then we return to the outer loop and search until finding $\hat{\eta}_\eta$ which minimize the quadratic norm in the sample moment conditions.

5 Markups and Elasticities

6 Merger Simulation