Homework 3: Industrial Organisation

Dhananjay Ghei

November 27, 2018

N.B. The code for this exercise was written in `R` *and is available on my Github account.* `www.github.com/dhananjayghei/io_estimation`.

Question 1

There are $J_t + 1$ goods in each year $t = 1971, \dots, 1990$ including the outside good. Assume utility is given by

$$u_{ij} = \alpha(y_i - p_j) + x_j\beta + \epsilon_{ij} \qquad j = 0, \dots, J_t$$

with ϵ_{ij} type 1 extreme value, y_i income, x_j, p_j observed characteristics, and parameters β, α. Suppose we observe aggregate data - market shares - as opposed to the micro-level data of zeros and ones (no purchase, purchase). Derive the appropriate likelihood function for the aggregated data by starting with the micro-level specification and regrouping.

First, normalize the utility with respect to u_{i0}.

$$\begin{aligned}\widetilde{u_{ij}} = u_{ij} - u_{i0} &= -\alpha p_j + x_j\beta + \epsilon_{ij} - \epsilon_{i0} \\ &= \delta_j + \epsilon_{ij} - \epsilon_{i0}\end{aligned}$$

Denote $u_{ij} := \widetilde{u_{ij}}$. Now, note that the error term is type 1 extreme value and is i.i.d. we can write:

$$\begin{aligned}Pr(u_{ij} > u_{ik}\ \forall k \neq j|\epsilon_{ij}) = \prod_{k\neq j} Pr(u_{ij} > u_{ik}|\epsilon_{ij}) &= \prod_{k\neq j} Pr(\epsilon_{ik} > \epsilon_{ij} + \delta_j - \delta_k|\epsilon_{ij}) \\ &= \prod_{k\neq j} e^{-e^{-(\epsilon_{ij}+\delta_j-\delta_k)}}\end{aligned}$$

which gives us:

$$\begin{aligned}Pr(u_{ij} > u_{ik}\ \forall k \neq j) &= \int_{-\infty}^{\infty} \prod_{k\neq j} e^{-e^{-(\epsilon_{ij}+\delta_j-\delta_k)}} e^{-\epsilon_{ij}} + e^{-e^{-\epsilon_{ij}}} d\epsilon_{ij} \\ &= \int_{-\infty}^{\infty} \prod_{k} e^{-e^{-(\epsilon_{ij}+\delta_j-\delta_k)}} e^{-\epsilon_{ij}} d\epsilon_{ij}\end{aligned}$$

Define $t \equiv e^{-\epsilon_{ij}}$.

$$\begin{aligned}Pr(u_{ij} > u_{ik}\ \forall k \neq j) &= \int_{0}^{\infty} e^{-t(\sum_k e^{\delta_j-\delta_k})} dt \\ &= \frac{e^{\delta_j}}{1+\sum_k e^{\delta_k}} \\ &\equiv s_j(\alpha, \beta)\end{aligned}$$

Finally, note that the above expression is for every year and does not depend on individual i. We can now add the time subscript to each product j. Then, the maximum-likelihood function $\mathscr{L}$ is given by:

$$\mathscr{L} = \prod_{t=1}^{T} \prod_{j=1}^{J_t} s_{jt}(\alpha, \beta)^{s_{jt}^d}$$

, where s_{jt}^d denotes the market share of good j at time t in the data.

Question 2

Using the automobile data, estimate the logit demand specification using maximum likelihood and assuming prices are exogenous. What is the implied own-price elasticity of the 1990 Honda Accord (HDACCO)? What is the implied cross-price elasticity of Honda Accord with respect to the 1990 Ford Escort (FDESCO)? Pick two additional cars and report the same numbers.

I use the maximum likelihood function defined above, convert it into log and run a minimisation routine on the negative of log likelihood function to get the estimates. Table 2 gives the results from maximum likelihood estimation of the logit demand specification.
Table 1 shows the own and cross-price elasticities of four cars - namely, BMW 735i, Ford Escort, Honda Accord and Volkswagen Jetta in 1990. To calculate the elasticities, I use the formulas we derived in class for the logit demand specification. As expected, the logit model generates aggregate substitution patterns, and elasticities that cannot possess many of the features that one might expect to have (Berry et al. [1995]). For example, in Table 1, the cross-price elasticity of BMW 735i (typically, a luxury car) is the same for Mercedes Benz (another, luxury car) and Honda Accord (comparatively, cheaper car). However, one would expect that people who buy BMW 735i have a preference for luxury cars and they would be substituting more towards luxury cars than towards the other cars. This is also a drawback of such models (as pointed out by Berry et al. [1995] in their paper.)

Table 1 Elasticities from Maximum Likelihood Estimation

The table reports the own-price and cross-price elasticities of different products using the maximum likelihood estimation.

	BMW 735i	Ford Escort	Honda Accord	Mercedez Benz S420	VW Jetta
BMW 735i	-4.48632	0.00050	0.00050	0.00050	0.00050
Ford Escort	0.00207	-0.67571	0.00207	0.00207	0.00207
Honda Accord	0.00492	0.00492	-1.10717	0.00492	0.00492
Mercedez Benz S420	0.00041	0.00041	0.00041	-5.72258	0.00041
VW Jetta	0.00057	0.00057	0.00057	0.00057	-0.91465

Table 2 Results with Maximum Likelihood Estimation

The table reports the coefficients and standard errors from maximum likelihood estimation.

	Coefficient	StdError
(Intercept)	-6.3968	0.1354
HP/Wt	1.6288	0.1432
Air	-0.2678	0.4137
MPD	0.2547	1.1290
Size	2.1963	0.4378
Price	-0.1197	5.1132

Question 3

Estimate the logit demand specification using the linearised version of this model from BLP. What is the implied own-price elasticity of the 1990 Honda Accord (HDACCO)? What is the implied cross-price elasticity of Honda Accord with respect to the 1990 Ford Escort (FDESCO)? Pick two additional cars and report the same numbers.

First, I try to replicate the descriptive statistics of Table 1 from BLP. In particular, for all the variables in the dataset, I calculate the sales weighted mean and report them. Table 3 reports the descriptive statistics and it matches the Table 1 from the BLP paper.

Table 3 Descriptive Statistics

The table reports the descriptive statistics of the data similar to Table 1 of BLP. The columns are sales weighted mean of the variables in each year.

Year	Price	HP/Wt	Size	Air	MPD
1971	7.868	0.490	1.496	0.000	1.849
1972	7.979	0.391	1.510	0.014	1.875
1973	7.535	0.364	1.529	0.022	1.818
1974	7.506	0.347	1.510	0.026	1.452
1975	7.821	0.337	1.479	0.054	1.503
1976	7.787	0.338	1.508	0.059	1.696
1977	7.651	0.340	1.467	0.032	1.835
1978	7.645	0.346	1.405	0.034	1.929
1979	7.599	0.348	1.343	0.047	1.657
1980	7.718	0.350	1.296	0.078	1.466
1981	8.349	0.349	1.286	0.094	1.559
1982	8.831	0.347	1.277	0.134	1.817
1983	8.821	0.351	1.276	0.126	2.087
1984	8.870	0.361	1.293	0.129	2.117
1985	8.939	0.372	1.265	0.140	2.024
1986	9.382	0.379	1.249	0.176	2.856
1987	9.965	0.395	1.246	0.229	2.789
1988	10.070	0.396	1.251	0.237	2.919
1989	10.321	0.406	1.259	0.289	2.806
1990	10.337	0.419	1.270	0.308	2.852

Next, I convert the logit demand specification into a linearised version by taking logarithms. Following this transformation, we can run OLS to estimate the logit demand. The results for the regression are reported in Table 6. The results from OLS match Column 1 from Table III of Berry et al. [1995] (BLP, hereafter).
Subsequently, I calculate the elasticities for four cars - namely, BMW 735i, Ford Escort, Honda Accord and Volkswagen Jetta in 1990. Table 4 shows the own and cross-price elasticities for these products. As one would expect, the own price elasticities are negative and of the right sign. The cross-price elasticities are the same for each car w.r.t the other cars. This is because of the logit specification. Under the logit specification, the cross-price elasticity of product j w.r.t. product k is given by:

$$\varepsilon_{jk} = \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = \alpha p_k s_k$$

Table 4 Elasticities from Ordinary Least Squares

The table shows the own-price and cross-price elasticities of different products using the ordinary least squares regression.

	BMW 735i	Ford Escort	Honda Accord	Mercedez Benz S420	VW Jetta
BMW 735i	-3.32269	0.00037	0.00037	0.00037	0.00037
Ford Escort	0.00154	-0.50045	0.00154	0.00154	0.00154
Honda Accord	0.00364	0.00364	-0.82000	0.00364	0.00364
Mercedez Benz S420	0.00030	0.00030	0.00030	-4.23830	0.00030
VW Jetta	0.00042	0.00042	0.00042	0.00042	-0.67742

Question 4

Use the instruments used in Berry-Levinsohn-Pakes. You will need the firmids and the year variables to calculate these instruments (they are product firm-year specific). Estimate the logit model using 2SLS and instrumenting for price. What is the implied own-price elasticity of the 1990 Honda Accord (HDACCO)? What is the implied cross-price elasticity of Honda Accord with respect to the 1990 Ford Escort (FDESCO)? Pick two additional cars and report the same numbers.

One would assert that the price in the OLS model is correlated with the ξ_j term (unobserved product characteristic) and thus, the OLS estimates will be biased. Thus, one could think of instrumenting for price. I follow BLP's approach and instrument for price in the model. BLP uses three types of instruments:

1. Observed product characteristics (z_{rk})
2. Sum of observed product characteristics of a single firm in the market (excluding the product) ($\sum_{r \neq j} z_{rk}$)
3. Sum of observed product characteristics of rival firms in the market ($\sum_{r \neq j, r \neq J_f} z_{rk}$)

I follow this approach and construct the instruments (for all four product characteristics and the constant term). It is important to note here that the instruments used in the original BLP paper are incorrect (as pointed out by Gentzkow and Shapiro in their replication file) and hence, our results are not directly comparable to the Column 2 of Table III in BLP. However, we can compare these results with Column 2 of Table I in Gandhi et al. [2011]. Table 6 shows the results from the instrumental variable regression. Table 5 shows the own and cross-price elasticities for the same 4 cars as mentioned earlier. Note that, instrumenting leads to an increase in the own-price elasticity when compared with the OLS estimates. Table 6 shows the results from the regressions.

Table 5 Elasticities from Instrument Variable Regression

The table shows the own-price and cross-price elasticities of different products using the instrumental variable regression.

	BMW 735i	Ford Escort	Honda Accord	Mercedez Benz S420	VW Jetta
BMW 735i	-5.08716	0.00056	0.00056	0.00056	0.00056
Ford Escort	0.00235	-0.76621	0.00235	0.00235	0.00235
Honda Accord	0.00558	0.00558	-1.25545	0.00558	0.00558
Mercedez Benz S420	0.00046	0.00046	0.00046	-6.48899	0.00046
VW Jetta	0.00064	0.00064	0.00064	0.00064	-1.03715

Table 6 Results with Logit Demand

The table reports the estimates of the coefficients and the standard errors from the OLS and IV regressions on the demand equations. Standard errors are in parenthesis. The data set contains 2217 observations from BLP. Significance levels are reported in the last row of the table. Column I matches the Column I of Table III from [Berry et al., 1995], whereas Column II matches the Column II of Table I from Gandhi et al. [2011]

	Dependent variable:	
	OLS	IV
	(1)	(2)
Constant	-10.072^{***}	-9.915^{***}
	(0.253)	(0.263)
HP/Weight	-0.124	1.226^{***}
	(0.277)	(0.404)
Air	-0.034	0.486^{***}
	(0.073)	(0.133)
MPD	0.265^{***}	0.172^{***}
	(0.043)	(0.049)
Size	2.342^{***}	2.292^{***}
	(0.125)	(0.129)
Price	-0.089^{***}	-0.136^{***}
	(0.004)	(0.011)
Observations	2,217	2,217
R^2	0.387	0.349
Adjusted R^2	0.386	0.348
Residual Std. Error (df = 2211)	1.083	1.116
F Statistic	279.243^{***} (df = 5; 2211)	
Note:		*p<0.1; **p<0.05; ***p<0.01

Question 5

Write

$$u_{ij} = \alpha(y_i - p_j) + x_j\beta_i + \xi_j + \epsilon_{ij} \qquad j = 0, \ldots, J$$

Now add random coefficients for each characteristic and estimate the means and variances of these normally distributed random coefficients. Estimate the demand side of the model only (unless you are ambitious and want smaller standard errors - then add the supply side too). What is the implied own- price elasticity of the 1990 Honda Accord (HDACCO)? What is the implied cross-price elasticity of Honda Accord with respect to the 1990 Ford Escort (FDESCO)? Pick two additional cars and report the same numbers.

While the estimates look better with instrumental variable regressions, the simple model still does not generate the substitution patterns as one would expect and is seen in the data. In order to correct for this, we introduce the random coefficients model. Table 9 show the estimates from the random coefficients model. The first two columns are the estimates without any demographic data. The next two columns are the estimates with the income data (Question 6).

As pointed out earlier, I constructed the instruments using the BLP approach which are different from the actual instruments used in their paper and hence, the results from Table 9 are not directly comparable to the results in Table IV of Berry et al. [1995]. However, the numbers I get are in the similar range as in the original paper albeit with larger standard errors as I do not include the supply side of the model.

Table 7 Elasticities from Random Coefficients Model

The table shows the own-price and cross-price elasticities of different products using the random coefficients model.

	BMW 735i	Ford Escort	Honda Accord	Mercedez Benz S420	VW Jetta
BMW 735i	6.07989	-0.00025	-0.00560	-0.02271	-0.00036
Ford Escort	-0.00006	-2.99291	0.15401	-0.00008	0.02378
Honda Accord	-0.00057	0.06496	-3.73096	-0.00052	0.01755
Mercedez Benz S420	-0.02773	-0.00041	-0.00628	7.92202	-0.00043
VW Jetta	-0.00031	0.08719	0.15256	-0.00031	-3.55410

Question 6

Now add income effects as another random coefficient to the model so the utility specification is given by

$$u_{ij} = \alpha \log(y_i - p_j) + x_j\beta_i + \xi_j + \epsilon_{ij} \qquad j = 0, \ldots, J$$

Draw y_i from a log-normal with a mean that varies by year for 1971-1990 given by (2.01156, 2.06526, 2.07843, 2.05775, 2.02915, 2.05346, 2.06745, 2.09805, 2.10404, 2.07208, 2.06019, 2.06561, 2.07672, 2.10437, 2.12608, 2.16426, 2.18071, 2.18856, 2.21250, 2.18377) and a fixed variance given by $\sigma_y = 1.72$. What is the implied own-price elasticity of the 1990 Honda Accord (HDACCO)? What is the implied cross-price elasticity of Honda Accord with respect to the 1990 Ford Escort (FDESCO)? Pick two additional cars and report the same numbers.

Table 8 reports the own and cross-price elasticities of different cars using the random coefficients model and the income draws. One can see that the elasticity numbers go up and the substitution patterns as seen in the earlier logit model specification change as well. The elasticities start to look more reasonable as we introduce unobserved heterogeneity into the model. Consider for example, Honda Accord, the cross price elasticity for Honda Accord is higher w.r.t. Ford Escort than w.r.t. BMW 735i which makes more sense than the logit demand specification. Moreover, the substitution patterns are not symmetric anymore as we account for both unobserved heterogeneity and the interaction terms.

Table 8 Elasticities from Random Coefficients Model

The table shows the own-price and cross-price elasticities of different products using the random coefficients model and the income draws.

	BMW 735i	Ford Escort	Honda Accord	Mercedez Benz S420	VW Jetta
BMW 735i	-8.90218	0.00566	0.01330	0.00612	0.00164
Ford Escort	0.00136	-1.32382	0.05660	0.00116	0.00638
Honda Accord	0.00135	0.02387	-2.15531	0.00114	0.00625
Mercedez Benz S420	0.00747	0.00589	0.01380	-11.36084	0.00170
VW Jetta	0.00145	0.02338	0.05437	0.00122	-1.81453

Question 7

Put together a table with the point estimates and standard errors for all of your estimates. In the case of 5 and 6, you do not need to compute the standard errors for the parameters outside of δ, but do report the standard errors on the parameters inside of δ, as you do in the exercise 2-4.

Table 2 reports the maximum likelihood estimation results. Table 6 reports the results from the demand side for the OLS and IV models. Table 9 reports the coefficients estimated from the random coefficients model.

Table 9 Estimated parameters of the demand equation - Random Coefficients Model

The table shows the estimates of the parameters using random coefficients model using the BLP specification on 2172 distinct observations. The first two columns of estimates are from Question 5 and the last two columns of estimates are from Question 6.

Demand Side	Variable	No Demographics		With Demographics	
		Parameter	Std Error	Parameter	Std Error
Mean ($\bar{\beta}$)s	Constant	-9.261	1.444	-10.587	1.186
	HP/Weight	4.269	1.358	5.232	1.090
	Air	0.140	0.918	-1.612	3.039
	MPD	-0.314	0.343	0.004	0.404
	Size	0.453	0.825	0.525	0.862
Std. Deviations (σ_β)s	Constant	3.644	0.883	-1.782	0.816
	HP/Weight	-0.535	2.895	1.832	1.093
	Air	1.910	1.011	-3.752	2.419
	MPD	0.305	0.330	0.436	0.420
	Size	2.840	0.436	-2.154	0.560
Term on Price	$\log(y-p)$			-0.238	0.031

References

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

Amit Gandhi, Kyoo il Kim, and Amil Petrin. Identification and estimation in discrete choice demand models when endogenous variables interact with the error. Working Paper 16894, National Bureau of Economic Research, March 2011. URL `http://www.nber.org/papers/w16894`.