Barcelona School of Economics
Albrecht Glitz

Problem Set 3: Discrete Choice and Selection Models

Please submit a document with your answers, including Stata or R output and programs, to our TA Janik Deutscher via the Google Classroom by ***Wednesday, 11 February, 18:00*** *at the very latest. You can work in groups of up to four.*

Question 1. Let $y_i = 1$ if a traveller chooses air transport to travel between two cities and $y_i = 0$ if she chooses the train (and assume these are the only modes of transport). Let $\mathbf{x}_i$, a K-element vector, contain measures of characteristics of the traveller, her trip and environment. A sample of N travellers produces independent realisations of y_i and associated values of $\mathbf{x}_i$.

a) Suppose that the probability that a traveller with characteristics $\mathbf{x}_i$ takes air transport is as follows:

$$P[Y_i = 1|\mathbf{x}_i] = \frac{\exp(\mathbf{x}_i'\beta)}{1 + \exp(\mathbf{x}_i'\beta)}.$$

Explain how the maximum likelihood estimator of β is obtained in this model, show its asymptotic variance/covariance matrix, and explain how you would estimate it.

b) Table 1 (below), shows maximum likelihood estimates of three logit models of this form using responses from 121 people travelling in South East Australia who chose to travel by either air or train. Here

$$\mathbf{x}_i'\beta = \beta_0 + \beta_1 \Delta c_i + \beta_2 \Delta t_i + \beta_3 hinc_i$$

where Δc_i is the cost of travelling by air minus the cost of travelling by train (in \$100), Δt_i is the time taken to travel by air minus the time taken to travel by train (hours), and $hinc_i$ is household income (in \$1000).

i. Consider Model 1. Are the signs of the coefficients on Δc_i and Δt_i plausible? Explain. Consider an airline replacing its existing service with a new service for which travel time is one hour longer. What air fare change relative to the existing service would leave the probability of choosing air travel unchanged?

ii. Consider Model 1. Provide an approximate 95% confidence interval for β_2. Explain the sense in which this interval is "approximate".

iii. Again, consider Model 1. Test the hypothesis $H_0 : \beta_1 = 0$ against the alternative $H_1 : \beta_1 < 0$ using a Wald z-test with approximate size 5%.

iv. Compare Models 2 and 3. Conduct a likelihood ratio test of the hypothesis $H_0 : \beta_1 = 0$ against the alternative $H_1 : \beta_1 \neq 0$ using a test with approximate size 5%.

	Model 1	Model 2	Model 3
β_0	-0.661	-2.880	-2.717
	(0.494)	(0.765)	(0.640)
β_1	-0.961	0.356	
	(0.765)	(0.887)	
β_2	-0.131	-0.124	-0.128
	(0.055)	(0.062)	(0.061)
β_3		0.055	0.053
		(0.012)	(0.012)
LL	-321.75	-270.53	-270.85

Table 1: Maximum likelihood estimates of coefficients of models of transport choice. Estimated asymptotic standard errors in parentheses and LL being the value of the maximised log likelihood function.

Question 2. Consider a latent variable modeled by $y_i^* = \mathbf{x}_i'\beta + \varepsilon_i$ with $\varepsilon_i \sim N[0, \sigma^2]$. Suppose y_i^* is censored from above so that we observe $y_i = y_i^*$ if $y_i^* < U$ and $y_i = U$ if $y_i^* \geq U$, where the upper limit U is a known constant.

a) Derive the log-likelihood function for this model.

b) Obtain the expression for the truncated mean $E[y_i|\mathbf{x}_i, y_i < U]$.

c) Explain how to obtain Heckman's two-step estimator for this model.

d) Obtain the expression for the censored mean $E[y_i|\mathbf{x}_i]$.

Question 3. We analyze the decision about how many children to have using multinomial discrete choice models. For this purpose, download the SOEP practise data set soep_lebensz_en.dta, which is available on the course website.

a) Suppose, you would like to estimate the effect of education on the number of children? Would you use a Multinomial Logit or a Conditional Logit Model (if you had to use one of the two)? Why? Estimate your model. Which alternative does Stata/R normalize? Interpreting the output, how much more likely is a person with 12 years of education to have two children than to have zero?

b) Based on your estimates, compute the average marginal effect of a change in education by one year on the probability of having 0, 1, 2 and 3 children. What must be true for those four average marginal effects?

c) Would an Ordered Probit Model be better to analyze the question above? Why? Estimate an Ordered Probit Model. Interpret the meaning of each of the three thresholds that the model estimates. Without any further calculations, what can you learn from the coefficient on education? For someone with 12 years of education, how much does an additional year of education increase the probability of having two children?