ECO 6375 Prof. T. Fomby
Econometrics III Fall 2008

EXERCISE 8

UNORDERED MULTINOMIAL LOGIT MODEL

Purpose: To learn how to use the **unordered multinomial logit model** and **unordered conditional logit model** to analyze dependent variables that represent three or more unordered responses.

First let us work with the unordered multinomial logit model **where the explanatory variables are individual-specific and do not change over choices**. Go to my website and download the files mlogit1.sas and UnordMultiLogit.dta. The latter file is a STATA data file. We will be working with some artificial multinomial data generated in the EVIEWS 4.0 manual to serve as an example of an unordered multinomial logit problem. For the discussion here, let us assume that we have data on two variables x1 (age) and x2 (income) on 1000 consumers who choose to shop for groceries at either store 1 (st = 1), store 2 (st = 2), or store 3 (st = 3).

(a) Use the program mlogit1.sas to generate the parameter estimates for an unordered multinomial logit model of store choice for these 1000 consumers. (SAS uses the first store as the "base" (comparison) store.) Report the parameter estimates along with their standard errors. Are the two variables statistically significant in each of the two choice equations? Explain.

(b) Now add some code to your mlogit1.sas program so that you drop the x2 (income) variable from the choice equation and re-estimate your multinomial model. Using this output and the output you produced in part (a) above, construct a likelihood ratio test of the overall significance of the x2 variable in the store choice model. Show your work.

(c) Using the output of the covariance matrix of the coefficient estimates that was produced by the covb option in Proc Logistic, test the following hypotheses using the Wald test procedure. Show your work.

$$H_0 : \beta_{1,2} - \beta_{1,3} = 0 \text{ versus } H_1 : \beta_{1,2} - \beta_{1,3} \neq 0$$

$$H_0 : \beta_{2,2} - \beta_{2,3} = 0 \text{ versus } H_1 : \beta_{2,2} - \beta_{2,3} \neq 0$$

Explain to me the implication of each of the above null hypotheses.

(d) Using the output of the covariance matrix of the coefficient estimates that was produced by the covb option in Proc Logistic, test the following hypotheses using the Wald test procedure. Show your work.

$$H_0 : \beta_{1,2} = \beta_{1,3} = \beta_{2,2} = \beta_{2,3} = 0 \text{ versus } H_1 : not H_0$$

Explain to me the implication of the above null hypothesis.

(e) Using the estimated model of part (a) above, predict which store the following 3 consumers will choose:

Consumer 1001: x1 = 0.99, x2 = 0.46, Pr(st=1) = _____, Pr(st=2) = _____, Pr(st=3) = _____, therefore the most likely store for consumer 1001 is store = _____.

Consumer 1002: x1 = 0.28, x2 = 0.18, Pr(st=1) = _____, Pr(st=2) = _____, Pr(st=3) = _____, therefore the most likely store for consumer 1002 is store = _____.

Consumer 1003: x1 = 0.11, x2 = 0.96, Pr(st=1) = _____, Pr(st=2) = _____, Pr(st=3) = _____, therefore the most likely store for consumer 1003 is store = _____.

(f) Go to my website and download the SAS program mlogit1f.sas and run it. One of my former students wrote an IML program to estimate the multinomial parameters of the above model. Are the point estimates of the parameters the same (up to minute rounding error) as those you obtained above? Print out the "Result" matrix of that program and hand it in with this exercise. (For your edification you should look at the details of the program and note its parallel construction to the previous Newton-Raphson programs we have written.)

(g) Now use STATA to run the following STATA commands on the STATA data set UnordMultiLogit.dta:

```
mlogit  s x1 x2, b(1)
mlogit s x1 x2, b(1) rr
```

Are the estimates and standard errors produced by STATA the same as what you got in SAS up to minute rounding error? Go to the RRR part of the output. Explain the meanings of the RRR numbers and how the standard errors for these number were computed. (Hint: You can go to the STATA website and use the STATA helpfile and learn a lot about the RRR numbers.)

(h) **Using STATA 9.0**, run a multinomial probit model on the above store data. Your commands will be something like

```
mprobit s x1 x2, baseoutcome(1)
```

How do the multinomial probit estimates compare with the multinomial logit estimates in terms of sign and statistical significance?

(i) The multinomial probit model is not subject to the criticism of the multinomial logit's implicit **assumption of independent irrelevant alternatives** (IRR). Explain to me what the IRR criticism of the multinomial logit model is and how the multinomial probit model manages to overcome it.

Finally we need to learn how the conditional logit model differs from the unordered multinomial model. In particular, the conditional logit model **contains explanatory variables that change not only across individuals but also across choices**.

(j) Download the SAS program Travmode.sas and run it. Briefly explain to me the economic meaning of the model and the rationale of the "predictive" part of the program.