Discrete Choice Modeling
William Greene
Stern School of Business, New York University

Lab Session 11

This exercise will suggest some open ended investigations for you to carry out on your own, to design and use your own discrete choice model.
.
Discrete Choice Models for Health Care and Moral Hazard

This set of exercises uses the health care data contained in healthcare.lpj.

A. We are interested in the variable HSAT, which contains the answer to a general question about the individuals satisfaction with their health. This is an ordered outcome coded 0, 1, …, 10.

1. Using the entire sample, formulate an appropriate model for health satisfaction, and estimate the parameters of your model. Include ; MARGINAL EFFECTS in your model command so you can obtain a listing of the partial effects of the variables in your model. You might want a more elaborate analysis of the partial effects, such as given by

       PARTIALS ; effects: ..the variable & other variable = low(change)high
                    ; Outcome = one of the specific values that HSAT takes $
   or
       PARTIALS ; effects: .. the variable | other variable = v1,v2,… ; plot(ci)
                    ; Outcome = one of the specific values that HSAT takes $

   (The ;Plot, which can be used in either command, will produce a plot of the partial effect with confidence limits.) Now, using what we know to be an inappropriate estimator, fit your model by OLS and compare the coefficients you get (a) to the ordered choice model coefficients and (b) to the marginal effects.

2. The sample contains roughly equal numbers of men and women. The variable FEMALE is a dummy variable which equals 1 for women and 0 for men. You can use this variable to deal with the samples of men and women separately. Test the hypothesis that the same model that you specified in part 1 applies to both men and women, versus the alternative hypothesis that different equations should be fit for the two sexes. Use a likelihood ratio test to carry out the test.

3. Your professor is uncertain whether income and education are significant determinants of health satisfaction. Test the hypothesis, separately for men and women. What do you find?

B. We are going to analyze the individual's choice of whether to obtain public insurance (PUBLIC). (We will motivate this in the next exercise.) This is a binary choice.

   1. Formulate an appropriate model and analyze this choice. What variables should appear in the model. Use both probit and logit models and compare your results. Does it matter?

   2. As in part A. we are interested in whether the model differs for men and women. Fit the model separately for men and women and test whether the two groups can be described by the same model. Use a likelihood ratio test.

   3. A middle ground in part 2 might be to fit the same model for both men and women, but allow certain variables to impact the outcome differently. Examine this approach using the education variable. Instead of having a single constant in your model, include the two dummy variables FEMALE and MALE = 1 – FEMALE. (You must CREATE MALE.) Now, instead of including EDUC in the equation, include the two variables MALE*EDUC and FEMALE*EDUC. Now, fit your binary choice model using the entire sample and with these four variables in the model. Request the partial effects. Do you find a noticeable difference between the marginal effects for the two education variables?

C. The study from which this data set is taken was concerned with the variables DocVis (number of visits to the doctor) and HospVis (number of hospital visits). During our discussion, we have also used created variables Doctor = 1(DocVis > 0) and Hospital = 1(HospVis > 0). The authors of the paper were interested in the question of moral hazard, that is, whether the presence of insurance significantly determined usage of the health care system. They focused on the variable ADDON, which is an indicator of whether the person has private insurance. For purpose of this exercise, we will examine variable PUBLIC, which is an indicator of whether the individual has public insurance. (One must have public insurance to obtain the Addon insurance.) For this exercise, you are to examine the issue of moral hazard in the context of a count model (Poisson or negative binomial) for DocVis and a binary choice model (probit or logit) for Doctor. The question is whether the insurance variable is a significant determinant of the dependent variable. Formulate a model to test the hypothesis. What else should be in the equation? Carry out the test using both count data models and the binary choice models. Carry out the analysis for men and women separately. What do you find?

D. As we discussed in class, these are panel data. (The group count variable is called TI.) The hypothesis tests you carried out in part B are influenced by the effect of the clustering of the observations in the panel. Devise an appropriate panel data treatment (cluster, robust inference), fixed effects, or random effects, and repeat the analysis. Do your results change?