# Homework 7: Discrete Choice Models II

Due Date:

## Exercise 1: Nested Logit

Use fictional data on 300 families and their choice of restaurants. Freebirds and Mama's Pizza are fastfood restaurants, Cafe Eccell, Los Nortenos, and Wings'n'More are family restaurants, and Christophers and Mad Cows are fancy restaurants. You will model the decision of where to eat as a function of income in thousands of dollars, the number of children in the house, the rating of the restaurant, the average cose of the mean per person, and the distance between the home and the restaurant. INCOME and KIDS are attributes of the family, RATING is an attribute of the alternative restaurants, and COST and DISTANCE are attributes of the alternative as perceived by the family.

1. Get the data by typing

   ```
   use http://www.stata-press.com/data/r10/restaurant
   ```

   Look at the data.

2. Use NLOGITGEN to set up a decision tree such that the families first choose between a fast food restaurant, a family restaurant, and a fancy restaurant, and then they choose the actual restaurant.

3. Use NLOGITTREE to look at the tree structure.

4. Estimate a nested logit model where the alternative-specific attributes (COST, RATING, DISTANCE) apply to the bottom alternative set (the seven restaurants) and family-specific attributes (INCOME, KID) apply to the first decision level (the three types of restaurants); use family restaurants as the base alternative.

5. Interpret the coefficients on COST, RATING, and INCOME.

6. Is the nested logit model preferred to the standard logit model? How do you know?

7. Interpret the log-sum coefficients. Is this model consistent with random-utility maximization?

8. Get a summary of the choices that were made by using ESTAT ALTERNATIVES. Interpret some of them.

9. Calculate p2=Pr(restaurant), p1=Pr(type of restaurant), condp=Pr(restaurant—type of restaurant), and the inclusive value. Interpret the results for family 2.

10. What is the expected value of the fast food branch? What is the expected value of the fancy food branch? What is the expected value of the family food branch?

**Exercise 2: Multinomial Probit**

Use fictional data on 152 groups of people traveling for their vacation, choosing between $J = 3$ alternative modes of travel: train, bus, or car. TIME is an alternative-specific variable indicating how long a group thinks it will take them to travel using a given mode of transportation. INVC is another alternative-specific variable capturing the in-vehicle cost of the trip. hinc and psize are group-specific variables indicating household income and group size respectively.

1. Get the data by typing

   ```
   use http://www.stata-press.com/data/lf2/travel2.dta, clear
   ```

   Familiarize yourself with the data.

2. Estimate a conditional logit model with TIME INVC TRAIN and BUS as independent variables. TRAIN and BUS are just dummy variables that will give us the alternative-specific constants; the omitted category is CAR. Get the predicted probabilities.

3. Now fit the corresponding ASMPROBIT model i.e. estimate a multinomial probit model but constrain it so that the errors are uncorrelated with unit variance. Get the predicted probabilities.

4. Get the covariance and correlation matrices.

5. Compare the predicted probabilities from the CLOGIT and ASMPROBIT models.

6. Estimate a full covariance matrix version of multinomial probit i.e. STATA's default. Look at the covariance and correlation matrices. State exactly what information is given in these matrices. How substantively interpretable is this information?

7. Now estimate a restricted covariance matrix version of multinomial probit i.e. estimate STATA's default STRUCTURAL model. Look at the covariance and correlation matrices. State exactly what information is given in these matrices?

8. After the ESTAT COVARIANCE command, do the following:

   ```
   return list;
   matrix cov=r(cov);
   matrix M=(-1, 1, 0\-1, 0, 1);
   matrix cov1=M*cov*M';
   matrix list cov1;
   ```

   What did you just do? What does this imply about the relationship between the STRUCTURAL model and the default model in STATA?

9. Now estimate the model with an EXCHANGEABLE correlation matrix. Look at the covariance and correlation matrices. Interpret what you just did. Why are the results exactly the same in for this particular data with the default structural model? Would this normally be the case?

10. Now do the same with the UNSTRUCTURED correlation matrix. Look at the covariance and correlation matrices. Interpret what you just did. Why are the results from this model the same as from STATA's default model? Would this normally be the case?

11. Let's add some group-specific information to our model. We can do this with the CASEVARS() command. Use CASEVARS(HINC PSIZE) and estimate STATA's default full covariance matrix model. Interpret the coefficients on PSIZE and HINC.

12. Predict the probability of each mode of travel based on this model using STATA's command. Interpret your results for group 1.

13. Use the ASPRVALUE command to calculate the predicted probabilities for each mode of transportation at the mean for all variables under the assumption that the value of all alternative-specific variables are constant across alternatives. Interpret the results.

14. Do exactly the same as before but allow the value of the alternative-specific variables to be at their mean value for each mode of transportation. Interpret the results.

15. Now look at the change in probability associated with an increase in PSIZE from PSIZE=1 to PSIZE=2. Interpret the results.