PROBLEM SET 4

Problem 1 (The Multinomial Logit Model)

London is served by 6 airports (http://en.wikipedia.org/wiki/Airports_of_London). Suppose the probability that consumer i chooses airport j is given by the logit model

$$s_{ij} = \frac{\exp\left(\delta_j - \gamma_1 p_{ij} - \gamma_2 d_{ij}\right)}{\sum_{k=1}^{6} \exp\left(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik}\right)}$$

where δ_j is an airport fixed effect, p_{ij} is the price of a flight from airport j to the destination of consumer i, and d_{ij} is the distance from airport j to the residence of consumer i.

(a) For consumer i, derive the elasticity of demand for airport j with respect to the price p_{ij}.

(b) Derive the elasticity of demand for airport j with respect to the price p_{ik}.

(c) Suppose all consumers pay identical price and live in the centre of London so that $d_{ij} = d_j$ for all i. Using the market shares found on (http://en.wikipedia.org/wiki/Airports_of_ what can you say about the elasticity of demand for Heathrow with respect to the price of the other airports?

(d) Do the susbstitution patterns of this model make sense? Why or why not?

(e) What data would you like to add to improve the model?

Hint: The elasticity of demand can be expressed in terms of market shares. Following extensive calculations, in (a), you should end up with

$$\frac{\partial s_{ij}}{\partial p_{ij}} \frac{p_{ij}}{s_{ij}} = -\gamma_1 p_{ij}\left(1 - s_{ij}\right).$$

In (b) you should get

$$\frac{\partial s_{ij}}{\partial p_{ik}} \frac{p_{ik}}{s_{ij}} = \gamma_1 p_{ik} s_{ik}.$$

Problem 2 (The Probit Model)

Now consider the probit model

$$P(y = 1|z, q) = \Phi(z_1\delta_1 + \gamma_1 z_2 q)$$

where q is independent of $z = [z_1, z_2]$ and distributed as $\mathcal{N}(0, 1)$; the vector z is observed but the scalar q is not.

(a) Write the model as an equivalent threshold-crossing model.

(b) Find the partial effect of z_2 on the response probability, namely,

$$\frac{\partial P(y = 1 \mid z, q)}{\partial z_2}$$

(c) Show that

$$P(y = 1 \mid z) = \Phi\left(z_1\delta_1 / \left(1 + \gamma_1^2 z_2^2\right)^{\frac{1}{2}}\right)$$

(d) Define $\rho_1 \equiv \gamma_1^2$. How do you test $H_0 : \rho_1 = 0$?

(e) If you have reasons to believe $\rho_1 > 0$, how would you estimate δ_1 along with ρ_1?

Hint for (c): Write the latent variable as $Y^* = z_1\delta_1 + \varepsilon$, where $\varepsilon = \gamma_1 z_2 q + u$. What are $E(\varepsilon|z)$ and $Var(\varepsilon|z)$? Use these expressions to derive the probability $P(y = 1|z)$.

Note that the threshold-crossing model in (a) is just the latent variable specification of the probit model. I.e., you specify Y (when is $Y = 1$, and $Y = 0$ conditional on Y^*?) and give the correct expression for Y^* (don't forget the error term).

Problem 3 (Applied Multinomial Logit and Ordered Probit)

This exercise is an extension of example 15.5 in Graduate Wooldridge (p. 507). We are interested in whether having a choice on the asset allocation in your pension plan has an impact on the share of stocks in the pension plan. (Use the 'pension.csv' data to solve the following problems)

1. Estimate the model from example 15.5 by OLS. That is, use '*pctstck*' as the dependent variable and age, years of education, gender, race (whether the person is black or not), marital status, income (dummies for different income categories), wealth and whether the plan is profit sharing as explanatory variables. What do you conclude about the effect of having a choice on share of stocks?
2. Estimate the same model using MNL. Use *pctstck=0* as the base outcome.
3. Report the relative risks of $pctstck = 50$ and $pctstck = 100$ (relative to $pctstck = 0$). What does the relative risks of the parameter *choice* imply?
4. Test whether income has any explanatory power in the MNL model (that is - test for joint significance of the *finc*-dummies) using both a Wald-test and a LR-test.
5. Examine the predictive power of the model.
    (a) Obtain the predicted probabilities of being in each of the three categories for each observation.
    (b) Find the predicted outcome of each observation as the outcome with the highest predicted probability.
    (c) Compare the predicted outcomes of the model with the actual observations.
6. Estimate an OPM on the same data. Report and interpret your results.
7. Compare the predictive power of the OPM and MNL models.
8. EXTRA: How could an ordinary probit model be estimated using the same data? Do this and compare the results with the OPM and MNL.