PROBLEM SET 4

Problem 1 (The Multinomial Logit Model)

1a. For consumer i, derive the elasticity of demand for airport j with respect to the price p_{ij}.

The elasticity is

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

Proof:

Suppose we want to calculate the elasticity of demand for airport 1 with respect to the price from that airport to the consumer's chosen destination, p_{i1}. Then, we need to calculate (using the quotient rule for differentiation):

$$\eta_{i,1,1} = \frac{\partial s_{i1}}{\partial p_{i1}}\frac{p_i}{s_i}$$

$$= \frac{p_{i1}}{s_{i1}}[\frac{-\gamma_1 \exp(\delta_1 - \gamma_1 p_{i1} - \gamma_2 d_{i2})\sum_{k=1}^{6}\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik}) + \gamma_1(\exp(\delta_1 - \gamma_1 p_{i1} - \gamma_2 d_{i1}))^2}{(\sum_{k=1}^{6}\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik}))^2}$$

$$= -\gamma_1\frac{p_{i1}}{s_{i1}}[\frac{\exp(\delta_1 - \gamma_1 p_{i1} - \gamma_2 d_{i1})\sum_{k=2}^{6}\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik}}{(\sum_{k=1}^{6}\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik}))^2} \tag{2}$$

where the last line follows from cancelling terms in the numerator.

Now note that:

$$(1 - s_{i1}) = 1 - \frac{\exp(\delta_j - \gamma_1 p_{i1} - \gamma_2 d_{i1})}{\sum_{k=1}^{6}\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{i2})} = \frac{\sum_{k=2}^{6}\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik})}{\sum_{k=1}^{6}\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik})} \tag{3}$$

Now, all we have to do is substitute (3) into (2) and we get that,

$$\eta_{i,1,1} = -\gamma_1 p_{i1}(1 - s_{i1}) \tag{4}$$

1b. The cross-price elasticity is

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

Proof:

An alternative to the above use of the quotient rule is instead to first take logs (this is entirely equivalent - which one you use will depend on which one you find easier) noting that:

$$\eta_{i,j,k} = \frac{\partial s_{ij}}{\partial p_{ik}}\frac{p_{ik}}{s_{ij}} = \frac{\mathrm{dlog} s_{ij}}{\partial p_{ik}} p_{ik}$$

Taking logs of the probability we have,

$$\begin{aligned} \log(\mathrm{s_{ij}}) &= (\delta_j - \gamma_1 p_{ij} - \gamma_2 d_{ij}) - \log(\sum_{k=1}^{6} \exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik})) \\ \implies \frac{\mathrm{dlog}(s_{ij})}{\partial p_{ik}} p_{ik} &= p_{ik}\gamma_1 \frac{\exp(\delta_k - \gamma_1 p_{ik} - \gamma_2 d_{ik})}{\sum_{l=1}^{6} \exp(\delta_l - \gamma_1 p_{il} - \gamma_2 d_{il})} = \gamma_1 p_{ik} s_{ik} \end{aligned} \tag{5}$$

1c. 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?

The elasticity of LHR w.r.t

LGW	$0.252\gamma_1 p_1$
STD	$0.135\gamma_1 p_2$
LUT	$0.07\gamma_1 p_3$
LCY	$0.022\gamma_1 p_4$
STH	$0.003\gamma_1 p_5$

The differences across airports depend only on the price and share ratios.

1d. Do the susbstitution patterns of this model make sense? Why or why not?

The substitution patters are highly restrictive because you would also expect the substitution patterns to depend on distance, the quality of the airport. Also, you would expect the airports that are near one another to be closer substitutes.

1e. What data would you like to add to improve the model?

Think about why do people choose airports: the cost of getting there, the destinations of the routes, the quality of the services. Also, you would like to ensure that you have exogenous variation in prices. With 6 airports you would need prices to vary either across routes or across time. If the variation is not exogenous, you could instrument using fuel prices for example,

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.

Solution:

$$\begin{aligned} Y^* &= Z_1\delta_1+\gamma_1 Z_2 q+\epsilon, \\ Y &= 1\left[Y^*>0\right]. \end{aligned}$$

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

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

Solution:

$$\frac{\partial P(y=1\mid z,q)}{\partial z_2}=\gamma_1 q\phi\left(z_1\delta_1+\gamma_1 z_2 q\right)$$

(c) Show that

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

Solution: This follows by integrating out q and then using the normality assumption. Then apply integration by parts.

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

Solution: see below.

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

Solution: We can answer parts (d) and (e) in one step. First, the model fully specifies $P\left(Y=1|z\right)$, so we can use ML to estimate δ_1 and ρ_1, as long as the log-likelihood has a unique maximum with respect to these parameters. This is indeed the case, as it can be shown that the Hessian of the log-likelihood is negative definite and therefore concave. Then one can use any of the trinity of LM, LR, and Wald tests for ML to test the hypothesis of part d. For example one can use the LR test to compare the likelihood of the model of part 3 to the likelihood of the restricted model that imposes the null in part d.