Econ 8206: Problem Set 1

Samuel S. Kortum
Department of Economics
University of Minnesota

November 1, 2002. Due November 8, 2002

Consider the demand for automobiles as envisioned by Goldberg.
Decision Tree. To keep things simple, the decision tree is

(1) *Buy* or *Not Buy*;

(2) If *Buy*: *Van*, *Compact*, or *Luxury*

(3) If *Van*: *Van1*, *Van2*, or *Van3* (i.e. 3 models of vans);
If *Compact*: *Com1* or *Com2* (i.e. 2 models of compacts);
If *Luxury*: *Lux1*, *Lux2*, *Lux3*, or *Lux4* (i.e. 4 models of luxury car).

Thus, the total number of choices, including not to buy, is 10.

Family Characteristics. Suppose you have the characteristics and purchase decisions of N families. Rather than using real data, I want you to simulate data.

The family characteristics include family income, and family size.

- Income (I) is distributed log normal with a mean of 40 and std. dev. of 20.
- Family size (F) follows the following discrete distribution:

Family size (F)	1	2	3	4	5
Probability $\Pr(F)$	0.2	0.3	0.2	0.2	0.1

Assume family size is independent of family income.

Vehicle Characteristics.
Vehicle data include price, and vehicle size.

Vehicle	*Van1*	*Van2*	*Van3*	*Com1*	*Com2*	*Lux1*	*Lux2*	*Lux3*	*Lux4*
Price (p)	14	18	23	8	10	22	30	40	25
Size (S)	100	110	120	50	57	80	90	90	85

If the choice is *Not Buy*, the price and size are both set to zero.

Purchasing Decisions. In simulating purchasing decisions of families, assume that the choice probabilities are those derived for the nested logit model. The specification is

$$v_{ij} = -\alpha_0 p_j + \alpha_1 p_j I_i + \beta_0 S_j + \beta_1 S_j F_i - \gamma_0 D_j / F_i$$

where $D_j = 1$ for a van and 0 otherwise.

Choice parameters are $\alpha_0 = 10$, $\alpha_1 = .1$, $\beta_0 = .2$, $\beta_1 = .5$, $\gamma_0 = 400$. There is a separate Nested Logit parameter ρ for choices within each class of cars as well as for the choice between classes. The values are: $\rho_v = .8$ for choices among vans, $\rho_c = .5$ for choices among compacts, $\rho_l = .9$ for choices among luxury cars, and $\rho_0 = .7$ for choices among the different classes.

Questions:

1. Write out the equation for the the choice probability *Com1*.

2. Simulate data for 1000 families. Provide relevant summary statistics such as the fraction of families that purchase a car, etc. Put together a set of cross tabulations showing which type of families tend to like which type of car.

3. Estimate the choice among the models in a particular class using multinomial logit. Thus you will estimate a separate model for each of the three classes of cars. Which of the parameters are you able to identify this way?

4. Estimate the whole model by maximum likelihood. Assess your success at recovering the true parameters.

5. Redo everything but with data for only 250 families. How do the results change?