Econometrics I

Fall 2006

Assignment 5

Today's Date: October 3
Due Date: October 10

Please show all of your work and clearly indicate your final response to each question. You are to write your own GAUSS (or MATLAB, Fortran, etc.) program in answering each of the computational questions.

Attached you will find a codebook for a data set which comes from *Consumer Demand and Labor Supply* by William Barnett (North-Holland, 1981). The ascii file containing the data set is *aggdem.out*.

You will be investigating demand relations for the FOUR goods: 1 (perishables), 2 (semi-durables), 3 (durables), and 4 (services). (You will ignore the demand for leisure). You are to use only data from the years 1890-1941 in carrying out the empirical analysis [so as not to have to deal with the missing data problems which occur during the WWII years].

We won't explicitly worry about the time series nature of the data in this assignment. You are to use the data to estimate a representative agent version of the random coefficient Cobb-Douglas demand system model we discussed in class. Think of the representative agent as having preferences that vary from period to period, with each period's preferences being determined by an independent draw from a distribution with mean $\overline{\alpha}$ and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_{44} \end{bmatrix}.$$

Begin by creating per capita consumption measures by dividing total expenditures on the four items in each year by the population size in that year. Due to the adding up constraint, we know that we can drop one equation when we estimate the model (let's all agree to drop the services equation when performing the estimation). Then the goal is to estimate the three equation model

$$x_{gt} = \overline{\alpha}_g \frac{y_t}{p_{gt}} + \varepsilon_{gt}, \; g = 1, 2, 3$$

where

x_{gt} is per capita expenditure on g in year t

y_t are total per capita expenditures in year t

p_{gt} is the price of g in year t

$$\varepsilon_{gt} = (\alpha_{gt} - \overline{\alpha}_g)\frac{y_t}{p_{gt}}.$$

All of the estimation questions relate to the parameters characterizing the three equation system only.

1. First estimate $\overline{\alpha}_g$ one equation at a time. Begin by obtaining the OLS estimate of this parameter. Is it consistent? Is it BLUE?

2. Remain in the single equation estimation environment. Define a GLS estimator of the $\overline{\alpha}_g$. Is GLS BLUE under the assumptions of the model?

3. Using your single equation estimators from (2), obtain consistent estimates of the parameters in the relevant submatrix of Σ.

4. Using your consistent estimator $\hat{\Sigma}$, compute the Seemingly Unrelated Regression (SUR) estimator of $\overline{\alpha}$.

5. Compare the single equation OLS and GLS estimates of $\overline{\alpha}$ with the SUR estimate of it. Is the SUR estimate always "better"?

6. Using the SUR estimates of the three equation subsystem, can you obtain consistent estimates of all of the remaining parameters? If no, why not? If so, derive consistent estimates of the remaining parameters.

7. Obviously this random coefficient Cobb-Douglas specification places many restrictions on the form of the conditional mean and conditional variance functions. Which restrictions do you find particularly suspect? Describe how you might go about investigating whether this specification is an adequate representation of the demand process reflected in the data.

Codebook for AGGDEM

Column	Variable
1	Year
2	Population Size (1000s)
	Quantities
3	Perishables
4	Semi-Durables
5	Durables
6	Services
7	Leisure
	Prices
8	Perishables
9	Semi-Durables
10	Durables
11	Services
12	Leisure
	Other Information
13	Leisure Consumption
14	Wage Rate
15	Unemployment Rate