Economics 125 Fall 2009
Linus Yamane Pitzer College

Problem Set #4

I. Chickens

Lets examine the demand for chickens in the United States from 1960 to 1982. Data on per capita consumption of chickens (in pounds) (PCCON), real disposable per capita income (DISP), real retail price of chickens (PRICE), and the composite real price of chicken substitutes (SUBPR) are available on the course web page under "**chicken.dta**."

Consider the following three regression specifications:

$$PCCON_t = \gamma_1 + \gamma_2\, DISP_t + \gamma_3\, PRICE_t + \varepsilon_t \quad (A)$$

$$\ln PCCON_t = \alpha_1 + \alpha_2 \ln DISP_t + \alpha_3 \ln PRICE_t + v_t \quad (B)$$

$$\ln PCCON_t = \beta_1 + \beta_2 \ln DISP_t + \beta_3 \ln PRICE_t + \beta_4 \ln SUBPR_t + u_t \quad (C)$$

a. Estimate equations (A) and (B), and use the Ramsey RESET test to determine which is a better regression specification. **estat ovtest**

b. Suppose (C) is the true model, but you have estimated (B). What are the theoretical consequences? What kind of specification tests would you do?

c. Suppose you estimate (C), the true model, and β_4 turns out to be statistically insignificant. Should we drop the price of substitute products as an explanatory variable in the demand function? Does this mean there is no specification error if we fit (B) to the data?

d. Now assume that (B) is the true demand function. What type of specification error is committed in this instance from estimating (C)? What are the theoretical consequences of this specification error? Illustrate your point with the regression results.

II. Simultaneous Equations

Consider the following simultaneous equations model for wheat:

$$Q_t^d = a + bP_t + cY_t + e_t$$
$$Q_t^s = d + fP_t + gW_t + u_t$$

where:

$Q_t^d = Q_t^s$ = Quantity Demanded and Quantity Supplied in tons
P_t = Price per ton
Y_t = Income
W_t = Weather (index of rainfall)

The annual data from 1986-2000 are provided in file "**simult.dta**". Use the following Stata commands:

reg3 (quant price income) (quant price weather), ols endog(price)
reg3 (quant price income) (quant price weather), 2sls endog(price)
reg3 (quant price income) (quant price weather), 3sls endog(price)

1. Identification
a. Identify the variables which are endogenous. Identify which are exogenous.
b. Check the identifiability of the equations in the model by applying the order condition.

2. Ordinary Least Squares
a. Estimate both equations using OLS and report the summary statistics.
b. What are the statistical properties of the estimated coefficients? Provide a sentence or two of explanation.

3. Indirect Least Squares (ILS)
a. Solve the structural equations for the reduced form.
b. Estimate the reduced form equations and report the summary statistics.
c. Calculate all of the structural parameters. Append a worksheet. This requires some algebra.
d. What are the statistical properties of the estimated coefficients? Provide a sentence or two of explanation.

4. Two Stage Least Squares (2SLS)
a. Estimate the demand equation and the supply equation using two stage least squares and report the summary statistics.
b. What are the statistical properties of the estimated coefficients? Provide a sentence or two of explanation.
c. Compare the coefficient estimates from the ILS and 2SLS estimates.

5. Three Stage Least Squares (3SLS)
a. Estimate the demand equation and the supply equation using three stage least squares and report the summary statistics.
b. Compare the coefficient estimates and the standard errors from the 2SLS and 3SLS estimates.

III. Fertility Decisions

We can examine the issue of fertility by looking at the decision of women in Peru to have children. The file **fertility.dta** contains data on 864 Peruvian women from 1985. We have information on the following variables: children (number of children the women has had); haschildren (a dummy equal to 1 if she has had children and 0 if not); education (in years); age; married (a dummy variable equal to 1 if married and 0 if not); fathereduc and mothereduc (education of the woman's father and mother); members (of the woman's household); adults (in the woman's household); householdexp (annual household expenditure); land (owned by the household, in hectares); rural (a dummy equal to 1 if she lives in a rural area and 0 if she lives in an urban area); and employee (a dummy variable equal to 1 if she works as an employee of a firm and 0 if she does not).

Suppose the decision to have children depends on the following

haschildren = f(age, education, married, rural, adults, fathereduc, mothereduc, land)

1. Using OLS, estimate the parameters of a linear probability model of the fertility decision. Do the signs of the OLS estimated parameters make sense? Why or why not? Comment on the appropriateness of using the OLS estimated standard errors to conduct tests of statistical

significance. Next examine the fitted values from this estimated linear probability model. For how many observations are the fitted values negative? For how many are they greater than 1? What is the probability that a woman with average characteristics will have a child?

2. Now assume that the cumulative distribution of the stochastic disturbances is the logistic. Estimate the above relationship using the logit model. Do the signs of these estimated logit parameters make sense? Why or why not? Which of the estimated logit parameters is significantly different from zero? Interpret. What is the probability that a woman with average characteristics will have a child?

3. Now assume that the cumulative distribution of the disturbances is normal. Estimate the above relationship using the probit model. Do the signs of these estimated probit parameters make sense? Why? Which of the estimated probit parameters is significantly different from zero? Interpret. What is the probability that a woman with average characteristics will have a child?

Due Thursday 19 November