School of Economics

College of Business and Economics

FINAL ASSESSMENT: NOVEMBER 2019

Course	:	Applied Econometric Techniques (AET9X01)
Moderator	:	Dr A Pholo
Date	:	22 November 2019
Time	:	180 minutes
Marks	:	100 points

Question 1 (15)

Describe succinctly, but comprehensively, the remedial measures against auto-correlation.

Question 2 (15)

Consider a simple model to estimate the effect of personal computer (PC) ownership on college grade point average for graduating seniors at a large public university:

$$\mathrm{GPA_i} = \beta_0 + \beta_1 \mathrm{PC_i} + u_i \tag{1}$$

where PC is a binary variable indicating PC ownership.

1. Why might PC ownership be correlated with u?

2. Explain why PC is likely to be related to parents' annual income. Does this mean parental income is a good IV for PC? Why or why not?

3. Suppose that, four years ago, the university gave grants to buy computers to roughly one half of the incoming students, and that the students who received grants were randomly chosen. Carefully explain how you would use this information to construct an instrumental variable for PC.

Question 3 (15)

Use the data in *LOWBIRTH.DTA* for this question

1. For 1987 and 1990, consider the state-level equation

$$\begin{aligned} \text{lowbrth}_{it} &= \theta_1 + \theta_2 \text{d90}_t + \beta_1 \text{afdcprc}_{it} + \beta_2 \log(\text{phypc}_{it}) \\ &\quad + \beta_3 \log(\text{bedspc}_{it}) + \beta_4 \log(\text{pcinc}_{it}) + \beta_5 \log(\text{popul}_{it}) + c_i + u_{it} \end{aligned}$$

   where the dependent variable is percentage of births that are classified as low birth weight and the key explanatory variable is afdcprc, the percentage of the population in the welfare program, Aid to Families with Dependent Children (AFDC). The other variables, which act as controls for quality of health care and income levels, are physicians per capita, hospitals beds per capita, per capita income, and population. Interpreting the equation causally, what sign should each β_j have? (Note: Participation in AFDC makes poor women eligible for nutritional programs and prenatal care.)

2. Estimate the preceding equation by pooled OLS, and discuss the results. You should report the usual standard errors and serial correlation-robust standard errors.

3. Difference the equation to eliminate the state fixed effects, c_i, and re-estimate the equation. Interpret the estimate of β_1 and compare it to the estimate from part 2. What do you make of $\hat{\beta}_2$?

4. Estimate the model with the within group method? Compare the results with the first differencing estimates?

Question 4 (15)

Let us consider a sample selection model where the set of regressors contains a variable y_2 that is correlated with the error term u_1.

1. Write in full details the population model, with a clear description of its components (5).

2. Apply the relevant estimation procedure to the data in MROZ Stata file. Use the constant, $exper$, $exper^2$ as elements of z_1 the vector of exogenous regressors; take $y_2 = educ$. The other elements of z should include age, $kidslt6$, $kidsge6$, $nwifeinc$, $motheduc$, $fatheduc$, and $huseduc$ (10).

Question 5 (20)

The data in the Stata file FERTIL1 are a pooled cross section on more than a thousand U.S. women for the even years between 1972 and 1984, inclusive; the data set is similar to the one used by Sander (1992). These data can be used to study the relationship between women's education and fertility.

1. Use OLS to estimate a model relating number of children ever born to a woman (*kids*) to years of education, age, region, race, and type of environment reared in. You should use a quadratic in age and should include year dummies. What is the estimated relationship between fertility and education? Holding other factors fixed, has there been any notable secular change in fertility over the time period?

2. Reestimate the model in part 1, but use *motheduc* and *fatheduc* as instruments for *educ*. First check that these instruments are sufficiently partially correlated with *educ*. Test whether *educ* is in fact exogenous in the fertility equation.

3. Now allow the effect of education to change over time by including interaction terms such as $y74 \times educ$, $y76 \times educ$, and so on in the model. Use interactions of time dummies and parents' education as instruments for the interaction terms. Test that there has been change in the relationship between fertility and education over time.

4. Test whether the instruments in part 2 and 3 are actually exogenous.

Question 6 (20)

Use information in Stata file Application 7 and estimate a Cobb-Douglas production function. After log transformation we have

$$
\begin{aligned}
y_{it} &= \beta_n n_{it} + \beta_k k_{it} + \gamma_t + (\eta_i + v_{it} + m_{it}) \\
v_{it} &= \alpha v_{i,t-1} + e_{it} \qquad |\alpha| < 1 \\
e_{it}, m_{it} &= \text{MA}\,(0)
\end{aligned}
$$

where y_{it} is log sales of firm i in year t, n_{it} is log employment, k_{it} is log capital stock and t is a year-specific intercept reflecting, for example, a common technology shock. η_i is an unobserved time-invariant firm-specific effect, v_{it} is a possibly autoregressive (productivity) shock and m_{it} reflects serially uncorrelated (measurement) errors.

Blundell and Bond (2000) consider the time series properties of these series and report estimates of this production function using a balanced panel of 509 R&D-performing US manufacturing companies observed for 8 years, 1982-89, similar to that used in Mairesse and Hall (1996). Capital stock and employment are measured at the end of the firm's accounting year, and sales is used as a proxy for output.

1. Prove that the above model has the following dynamic representation

$$y_{it} = \pi_1 n_{it} + \pi_2 n_{i,t-1} + \pi_3 k_{it} + \pi_4 k_{i,t-1} + \pi_5 y_{i,t-1} + \gamma_t^* + (\eta_i^* + w_{it})$$

   Run OLS regression and interpret your results using time trend and a correct model that account for possible heteroscedasticity and cross-section errors correlation (4).

2. Run Within-Group regression and interpret your results using time trend and a correct model that account for possible heteroscedasticity and cross-section correlation (4).

3. Compare the results obtained using OLS to those of WG (4).

4. Run a GMM model and compare the results to those of OLS and WG (4).

5. Assess the validity of instruments (4).