AAEC6311 – Applied Econometrics II

Spring 2008

Instructor: Eric Belasco

Final Exam

Name: _____

You have two hours and 30 minutes to complete the exam. There are three questions. Point values will sum to 100, with points associated with each question indicated accordingly. **USE YOUR TIME WISELY**. Also, I will address clarifying questions, but not substantive questions. You are allowed to use two 3x5 note sheets (front and back), a calculator, and a pencil/pen. Use the back of each sheet of paper if you need additional room, making such a move clear to me. Good luck!

Question 1 (26 points). You are interested in evaluating factors that influence the adoption of a new technology aimed at soil conservation. The following formulation is used to evaluate this topic:

$$U_{ki} = X_i\beta_k + \varepsilon_{ki} \text{ , where k = 1,2 ; i = 1,2,…,n}$$

where U_{1i} is the utility for the ith individual under the existing technology while U_{2i} is the utility under the new technology. $D_i = 1$ if the soil conservation measure is adopted and zero otherwise, or if we assume utility maximization, $U_{2i} > U_{1i}$. Variables contained in X include soil quality, cropping system, farm size, age, education, and farming experience.

(a) (5 points) Derive the linear probability model (LPM) estimator for the above model.

(b) (5 points) Derive the Logit model and explain how you would estimate via a Logit framework. What fundamental advantages might the Logit model have over the LPM?

(c) (3 points) Discuss how you would assess the model fit between these two models.

When you present this idea, your advisor points out that there are a few different new technologies that attempt to conserve soil quality. More specifically, there are three types, which we will define as type A, B, and C. (*Note: four alternatives, with the final one being 'technology not adopted'.*) The alternative technologies can all be used on the same commodities and regions.

(d) (5 points) Set up the conditional Logit model to estimate this regression and describe the estimation procedure.

(e) (3 points) Discuss the limitations and advantages of the conditional Logit model.

(f) (5 points) Using the sampling distribution above and assuming an uninformative uniform (0,1) prior distribution, which has probability distribution $\pi(\beta) = 1/(1-0)$, you move forward with your Bayesian model. Given that the posterior distribution will be proportional to the sampling distribution, what difference is there in using Bayesian methods when compared to classical methods (maximum likelihood estimation)? Be as specific as possible.

Question 2 (46 points). Consider the following Cobb-Douglas production function relationship:

$$Y_i = \alpha + \beta_1 PUBK_i + \beta_2 PRIK_i + \beta_3 LAB_i + \beta_4 UNEM_i + \epsilon_i,$$

for i=1,2,…,48 (48 US states, excluding Alaska and Hawaii), where Y is the log of gross state product, PUBK is the log of public capital (highways, streets, water, etc.), PRIK is the log of private capital, LAB is the log of labor inputs, and UNEM is the state unemployment rate. All data is for 1986.

(a) (5 points) A colleague suggests that your model is underspecified and your estimates will be biased, since you did not include a lagged term for PUBK and PRIK. He suspects that current production should also be a function of past investments in production. Is your colleague correct in saying that your results will be biased? Explain.

(b) (5 points) Specifically, describe how you would estimate the given model if you were unable to obtain the recommended variables.

(c) (5 points) Under what conditions would your new estimation method improve on estimation from the previous specification? In what ways is your estimation improved?

You go back to your data source and identify historic data that transforms your current cross-sectional data set into a panel series. Consider the revised Cobb-Douglas production function relationship:

$$Y_{it} = \alpha + \beta_1 PUBK_{it} + \beta_2 PRIK_{it} + \beta_3 LAB_{it} + \beta_4 UNEM_{it} + \epsilon_{it},$$

where t = 1970, 1971,…,1986. The results on the regressions run are shown below.

	β_1	β_2	β_3	β_4
POLS	0.155 (0.017)	0.309 (0.010)	0.594 (0.014)	-0.007 (0.001)
Between	0.179 (0.072)	0.302 (0.042)	0.576 (0.056)	-0.004 (0.010)
Within (FE)	-0.026 (0.029)	0.292 (0.025)	0.768 (0.030)	-0.005 (0.001)
RE	0.004 (0.023)	0.311 (0.020)	0.730 (0.025)	-0.006 (0.001)

(d) (3 points) Explain the major assumptions around using POLS and its limitations and potential biases.

(e) (5 points) Interpret the parameter estimate for β_1 using the POLS and Fixed Effects (within) specifications. What can you infer from these results after controlling for state-specific constants?

(f) (3 points) Is this panel balanced or unbalanced? Explain the difference.

(g) (5 points) Explain how you control for state-specific heterogeneity. What specific impacts might you be controlling for with state-specific constants in this model?

(h) (5 points) Discuss how to choose between using the random effects or fixed effects model and the major assumptions made on each.

(i) (5 points) You suspect that the RE and FE estimators are not statistically different, which is also confirmed by your Hausman test statistics. What can you infer about model selection?

(j) (5 points) If the random effects model is the true model, is the fixed effects model still consistent? Efficient? Explain.

Question 3 (28 points). The following questions consider the difficulties associated with modeling commodity prices.

(a) (5 points) Commodity prices have been argued to follow either a random walk or a stationary process. Discuss the difference.

(b) (5 points) Explain specifically how you would test which process commodity prices follow.

(c) (5 points) Assume you wanted to fit an ARMA(p,q) model to characterize corn prices. Write out a generalized ARMA model and describe its components. How would you choose an appropriate p and q?

(d) (3 points) Why might an ARCH model be more appropriate when modeling commodity prices?

(e) (5 points) Rather than modeling corn prices independently, you decide to characterize it as part of a system, along with soybeans and cattle prices. Under which circumstances would you expect the proposed VAR model to improve upon the previously mentioned AR model? Explain.

(f) (5 points) You are interested in testing the impact of corn prices on future live cattle prices. Specifically explain any tests or tools you might use to evaluate this question and what you might be able to infer.