ECO 82800

Panel Econometrics

Midterm Exam

23 March 2017, 9:30am – 11:30pm

This exam is a closed-book, closed-notes exam. Calculators without matrix functions are allowed; apps on cell phones are not allowed. The exam consists of six questions, most of them with parts. Points per question are as indicated; parts are weighted equally. The total is 80 points. Budget your time. You may answer the questions in any order, but *keep parts of a question together and label everything clearly*.

1. (15 points) Comment on the following statements:
   a. Panel data increase the number of data points in the sample. Observing N households over T periods raises the volume of information by a factor of T relative to a simple cross-sectional sample of N households.
   b. Generating a panel data of households means asking them repeatedly for new information. This is likely to lead to attrition.
   c. Generating a panel data of households means asking them repeatedly for new information. This is likely to lead to lower quality information and greater measurement error.

2. (5 points) Under what circumstances is the following statement correct?
$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$
   where $\otimes$ denotes a Kronecker product.

3. (15 points) Consider a regression model $y_i = X_i'\beta + u_i$ for $i = 1, \ldots, n$, estimated with a sample of households drawn from different cities. For simplicity, let us assume that this regression model may be estimated with OLS. A common recommendation is that the standard errors of $\hat{\beta}$ ought to be clustered by city because of the common location that households share.
   a. Why is this common location a concern for estimating this regression model?
   b. How would you test whether this concern is valid?
   c. If this concern is valid, would you recommend to estimate this regression model with OLS? If yes, why. If no, why not and what alternative approach would you recommend instead?

4. (15 points) Consider a two-factor error components model:
$$y_{it} = \alpha + X_{it}'\beta + \mu_i + \lambda_t + \nu_{it}$$
$$i = 1, \ldots, N \qquad t = 1, \ldots, T$$
   a. In the context of a random effects model, list all the assumptions related to μ, λ and ν.
   b. Write this model in full matrix notation. Be sure to define every term that you introduce.
   c. Define $\kappa_2 = T\sigma_\mu^2 + \sigma_\nu^2$ and $Q_2 = E_N \otimes \bar{J}_T$ where $E_N = I_N - \bar{J}_N$, $\bar{J}_N = \iota_N(\iota_N'\iota_N)^{-1}\iota_N'$, and $\bar{J}_T = \iota_T(\iota_T'\iota_T)^{-1}\iota_T'$. Define and motivate the Wallace & Hussain version of the Best Quadratic Unbiased Estimator of κ_2.

5. (20 points) Baltagi and Griffin estimated a gasoline demand model with one- and two-factor error components. The sample consists of data from 18 OECD countries over the period 1960-1978. The regression model, the formulation of the two-factor model, and the MLE results are given below (with standard errors in parentheses).

$$\ln \frac{Q_{Gas}}{\#Cars} = \alpha + \beta_1 \ln \frac{Inc}{Pop} + \beta_2 \ln \frac{P_{Gas}}{P_{GDP}} + \beta_3 \ln \frac{\#Cars}{Pop} + u$$

$$u_{it} = \mu_i + \lambda_t + \nu_{it}$$

	β_1	β_2	β_3	σ_μ	σ_λ	σ_ν
one-factor	0.588	-0.378	-0.616	0.292	n.a.	0.092
	(0.066)	(0.044)	(0.029)			
two-factor	0.231	-0.254	-0.606	0.361	0.095	0.082
	(0.091)	(0.045)	(0.026)			

a. What is your best guess of the income elasticity of the demand for gasoline? Why do you choose that number and not the other?
b. According to the two-factor estimates, what do these estimates tell us about the demand for gasoline when the relative price of gasoline in the US rises by 10 percent from one year to the next?
c. Based on conceptual arguments, would you recommend this regression model to be estimated with random effects or with fixed effects?
d. If the model is estimated by fixed effects, how would you test for the existence of the time factor?

6. (10 points) The forward orthogonal difference estimator uses variables that are transformed with the following equation, where we use the dependent variable as an example:

$$y_{it}^* = \left(\frac{T-t}{T-t+1}\right)^{1/2} \left(y_{it} - \frac{1}{T-t} \sum_{s=t+1}^{T} y_{is}\right)$$

where $t = 1, \dots, T-1$. This equation in matrix form is given as

$$y_i^* = (DD')^{-1/2} D y$$

where D is the $(T-1) \times T$ matrix that generates first differences. Prove that the forward orthogonal difference estimator is the same as the one-factor (individual) fixed effects estimator.