Comprehensive Exam in Econometrics
Suffolk University
June 2006

Answer all questions. You have four hours to take this test.

A1. Consider $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, where $\mathbf{X}$ represents a non-stochastic T x k data matrix, including a constant.

a) State the assumptions of the classical linear regression model. Use these assumptions to show that $\hat{\beta} = (X'X)^{-1}X'y \; and \; \hat{\sigma}^2 = e'e/T-k$ are unbiased estimators for the regression coefficients (β) and the error term variance (σ^2).

b) Briefly explain the principle of maximum likelihood (ML) estimation. Given the above assumptions along with the normality of the error term ε, derive the maximum likelihood estimators of β_{MLE} and σ^2_{MLE}.

c) Let $\mathbf{X}\beta = \mathbf{X}_1\beta_1 + \mathbf{X}_2\beta_2$ where $\mathbf{X}_1$ and $\mathbf{X}_2$ are non-stochastic T x k_1 and T x k_2 matrices respectively. Derive the OLS estimators for β_1 and β_2. Briefly discuss the consequences of erroneously (i) excluding X_2 when it should have been included, (ii) including X_2 when it should have been left out.

A2. Consider the classical linear regression model, $\mathbf{y} = \mathbf{X}\beta + \varepsilon$, where $\mathbf{X}$ is non-stochastic. Let $E(\varepsilon) = 0$ and $E(\varepsilon\varepsilon')=\sigma^2\Omega$ where Ω is symmetric positive definite matrix.

a) If Ω is known, derive the GLS estimates and briefly explain why they are more efficient than the OLS estimates

b) What is the problem of serial correlation? Consider $y_t = x_t'\beta + \varepsilon_t$. Further, $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$ where $-1 < \rho < 1$ and v is i.i.d. Explain the Durbin-Watson and the LM tests for the 1st order serial correlation and the estimated GLS method for correcting it.

c) Consider $y_t = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \gamma y_{t-1} + \varepsilon_t$ where $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$. How would you test for serial correlation in this dynamic model? How would you estimate the parameters efficiently if the 1st order serial correlation is detected?

(B1) Consider the following simple work-trip mode (automobile vs. public transit) choice model

$$y_i^* = \beta_0 + \beta_{1i} x_i + u_i, \quad i = 1, \ldots, n$$

where x_i is transit fare minus automobile travel cost, $u_i \sim \mathrm{N}(0,1)$, and y_i^* is a latent variable that measures the difference between the utility individual i gets from driving to work and using the public transit system. A new variable y_i is created such that $y_i = 1$ if automobile is chosen, i.e., $y_i^* \geq 0$ and $y_i = 0$ otherwise. The coefficient on the x variable could vary across individuals according to $\beta_{1i} = \beta_1 + \tau_i$ where $\tau_i \sim \mathrm{N}(0, \sigma_\tau^2)$. Assume that x is a fixed regressor and u and τ are independent.

(a) Derive the log-likelihood function for this model.

(b) Discuss in detail how you would test for random coefficient using likelihood ratio, wald, and lagrange multiplier tests. Specify the null and the alternative hypothesis, write down (but *do not* derive) the general expressions of the test statistics, etc.

(B2) Consider the following wage equation for n individuals observed over T periods

$$Y_{it} = X_{it}\beta + \varepsilon_{it} \quad i = 1, \ldots, n \text{ and } t = 1, \ldots, T \tag{1}$$

where $Y_{it} = \log(wage_{it})$, X_{it} is a $1 \times k$ vector of observable characteristics that affect the wage of individual i at time period t, and β is a $k \times 1$ vector of coefficients.

(a) Suppose that $E(\varepsilon_i \varepsilon_j^\top) = \sigma_{ij} \mathrm{I}_T$ where $\varepsilon_i = (\varepsilon_{i1}\ \varepsilon_{i2}\ \ldots\ \varepsilon_{iT})^\top$ is $T \times 1$.

(*i*) Describe in words what this covariance structure is and what it implies about the data.

(*ii*) Describe in detail an *approximate* maximum likelihood procedure to estimate β.

(*iii*) Describe how you would test the hypothesis

$$\mathrm{H}_0 : \Sigma \text{ is diagonal} \quad \text{v.s.} \quad \mathrm{H}_1 : \sim \mathrm{H}_0$$

where Σ is $n \times n$ with $\Sigma_{(ij)} = \sigma_{ij}$. Also describe in words what this null and alternative hypothesis implies about the data.

(b) Suppose you believe that unobserved heterogeneity (in the form of individual productivity) should be controlled for in the wage equation. What changes would you make in equation (1). Make all the necessary assumptions and describe in detail the resulting model, including the covariance structure. Also discuss OLS estimation of the resulting model.

(B3) Capital Asset Pricing Model (CAPM) quantifies the tradeoff between risk and return. CAPM implies the following linear relation between the return of an asset and the return of a market index

$$R_{it} = \alpha_i + \beta_i R_{mt} + u_{it}, \quad t = 1, \ldots, T$$

where R_{it} is the (excess, over the riskfree rate) return of asset i (e.g., individual stock or a portfolio of stocks) at time t and R_{mt} is the (excess) return of a market index (e.g., S&P500) at time t. The intercept term α_i measures the mispricing of the asset relative to the market and the slope coefficient β_i measures the volatility (risk) of asset i and is usually referred to as the beta of the asset.

(a) What is the beta of the market, i.e., $\hat{\beta}_m$?

(b) Suppose you invest in a portfolio consisting of n stocks and the return of this portfolio at time t is $R_{pt} = \sum_{i=1}^{n} w_i R_{it}$ where w_i is the weight (e.g., based on market capitalization) of stock i in the portfolio with $\sum_{i=1}^{n} w_i = 1$. You first calculate the beta of each stock, i.e., $\hat{\beta}_i$, $i = 1, \ldots, n$, from OLS estimation of the CAPM model n times separately for each stock. How can you calculate (derive) the beta of the portfolio, i.e., $\hat{\beta}_p$ from the betas of the stocks?

(c) Comment on the estimation of the CAPM model for the n stocks jointly by the seemingly unrelated regressions (SUR) technique.