ECON 421/521 Final Exam
Date: April 28, 2004; Time: 14:00-17:00, Location: H1070
Instructor: Nikolay Gospodinov

Read the questions carefully. Write clear and complete answers. Show your work to get credit. Partial answers get partial credit.

Question 1 (20 points). Suppose we are interested in studying the wage differentials between single and married individuals using the model

$$\log(wage) = \beta_0 + \beta_1 married + \beta_2 educ + \beta_3 married.educ + \beta_4 exper + \beta_5 exper^2 + u \quad (1)$$

where *married* is a dummy variable (=1 if married and =0 if single). The estimated equation is

$$\begin{aligned} \log(\widehat{wage}) &= \underset{(.039)}{.43} + \underset{(.05)}{.10}\ married + \underset{(.009)}{.075}\ educ + \underset{(.014)}{.004}\ married.educ + \underset{(.009)}{.062}\ exper - \underset{(.0001)}{.0008}\ exper^2 \quad (2) \\ n &= 526,\ R^2 = 0.35,\ SSE = 148. \end{aligned}$$

(a) What are the percentage returns to education for single and married people? Is the difference statistically significant at 5% significance level (critical values for one-sided and two-sided alternatives are 1.645 and 1.96, respectively)?

(b) What are the percentage returns to experience for single and married individuals with 10 years of experience? Describe briefly how you would test the hypothesis that the return to experience for an individual with 10 years of experience is zero.

(c) State the null hypothesis of no wage differential between single and married people who have the same level of education. How would you test this hypothesis?

(d) What are the likely consequences on the OLS estimates in (2) if the return to experience in (1) depends also on the level of *educ*. Write down the correct model that should be estimated in this case. What is the expected sign in front of the new variable? Why?

Question 2 (30 points). Let the population model be given by

$$y_i = \beta_1 x_i + u_i \quad (3)$$

i.e. the true model has no intercept ($\beta_0 = 0$). Assume that $E(u_i|x_i) = 0$.
The slope parameter is estimated using two estimators:

$$\begin{aligned} 1)\ \widetilde{\beta}_1 &= \frac{\overline{y}}{\overline{x}},\ \text{where } \overline{y} = \frac{1}{n}\sum_{i=1}^n y_i \text{ and } \overline{x} = \frac{1}{n}\sum_{i=1}^n x_i \\ 2)\ \widehat{\beta}_1 &= \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \end{aligned}$$

Suppose we know that the estimator $\widehat{\beta}_1$ is unbiased and has variance $Var(\widehat{\beta}_1) = \dfrac{\sigma^2}{\sum_{i=1}^n x_i^2}$.

(a) Show whether the estimator $\widetilde{\beta}_1 = \overline{y}/\overline{x}$ is biased or unbiased.

(b) Derive the variance of the estimator $\widetilde{\beta}_1$ and compare it to the variance of the estimator $\widehat{\beta}_1$ (Hint: Use that $\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sum_{i=1}^{n} x_i^2 - n(\overline{x})^2$).

(c) Which estimator of β_1 do you prefer, $\widetilde{\beta}_1$ or $\widehat{\beta}_1$? Explain.

(d) Now suppose that $E(u_i|x_i) \neq 0$. Explain intuitively (without any mathematical derivations) why we do not expect the above estimators to be unbiased and consistent in this case. Propose a consistent estimator of β_1 in model (3).

Question 3 (25 points). Consider the model

$$\begin{aligned} y_t &= \beta y_{t-1} + u_t \qquad (4) \\ u_t &= \rho u_{t-1} + e_t \end{aligned}$$

where e_t is $iid(0, 1)$, $0 < |\beta| < 1$, $0 < |\rho| < 1$ and $u_0 = 0$.

(a) Derive the expressions for $E(u_t)$ and $Var(u_t)$.

(b) Find the expression for $Cov(u_t, y_{t-1})$. Is the OLS estimator $\widehat{\beta}$ consistent (discuss if each assumption for consistency is satisfied by model (4))?

(c) Now rewrite model (4) as a model with uncorrelated errors and compare it to model (4). Use the structure of this model to provide some intuition for the result in part (b).

Question 4 (25 points). Consider the model

$$\begin{aligned} y_t &= \beta x_t + u_t \qquad (5) \\ u_t &= \rho u_{t-1} + e_t \end{aligned}$$

where e_t is $iid(0, 1)$ and assumptions TS.1', TS.2' and TS.3' are satisfied but it is possible that $Var(u_t|x_t)$ is not constant.

(a) Discuss how you would test the hypotheses of no serial correlation and homoskedasticity in the errors u_t (show all the steps and the sequence of tests).

(b) Suppose that the tests from part (a) indicate that there is serial correlation in the errors but no heteroskedasticity in model (5). Describe how you would proceed with the estimation and the inference (hypothesis testing) on the parameter of interest β in this case (suggest only one method that you think is the most appropriate for this problem).

(c) Now suppose that the tests from part (a) indicate that there is no serial correlation but there is heteroskedasticity in the errors. Describe how you would proceed with the estimation and the inference on the parameter of interest β in this case (suggest only one method that you think is the most appropriate for this problem).