

Question 1 (30 points). Suppose that the true wage equation is given by

$$\ln(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 IQ + \beta_4 \text{exper}^2 + \beta_5 \text{educ} \cdot \text{exper} + u \quad (1)$$

where *wage* is the hourly wage, *educ* is years of schooling, *exper* is years of experience and *IQ* is intelligence test score in points. The OLS estimation of the model, using a random sample of individuals, gives the following results

$$\begin{aligned} \ln(\widehat{\text{wage}}) &= \underset{(.171)}{-.233} + \underset{(.012)}{.117} \text{educ} + \underset{(.009)}{.062} \text{exper} + \underset{(0.0010)}{0.0059} IQ - \underset{(.0001)}{.0008} \text{exper}^2 - \underset{(.0005)}{.0014} \text{educ} \cdot \text{exper} \quad (2) \\ n &= 526, R^2 = 0.31 \end{aligned}$$

- (a) Interpret the coefficient on *IQ*. Why is *IQ* included in the model? If *IQ* increases by 20, what is the percentage change in the wage? Construct 95% confidence interval for the coefficient on *IQ*. Use the computed confidence interval to test the hypothesis that *IQ* has no ceteris paribus effect on $\ln(\text{wage})$ at 5% significance level.
- (b) What is the percentage return to experience of an individual with 10 years of schooling and 5 years of experience? What is the percentage return to education of an individual with 5 years of experience?
- (c) Write down the null hypothesis that the return to education of an individual with 5 years of experience is zero. State the null hypothesis that the return to education does not depend on the level of *exper*. What is the appropriate alternative hypothesis?
- (d) Now suppose that we do not know the true model and estimate the equation

$$\ln(\text{wage}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{exper} + \beta_3 IQ + u. \quad (3)$$

What are the likely consequences of estimating (3) instead of (1) for the OLS estimators. Describe briefly how you would test if (3) is the appropriate model when you do not know the exact form of the true model in (1).

- (e) Suppose that in model (1) $E(u|\text{educ}) \neq 0$. Discuss the consequences of this on the properties of the OLS estimator of β_1 . Propose a consistent estimator of β_1 when $E(u|\text{educ}) \neq 0$.

Question 2 (30 points). Consider the model

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

where $\{y_i, x_i\}_{i=1}^n$ is a random sample and $E(u_i|x_i) = 0$. Let $\widehat{\beta}_1$ denote the OLS estimator and $z_i = g(x_i)$, where $g(x)$ is some function of x (for example, $g(x) = x^2$). Define the estimator

$$\widetilde{\beta}_1 = \frac{\sum_{i=1}^n (z_i - \bar{z}) y_i}{\sum_{i=1}^n (z_i - \bar{z}) x_i},$$

where $\sum_{i=1}^n (z_i - \bar{z}) x_i \neq 0$.

- (a) Show that the estimator $\tilde{\beta}_1$ is unbiased.
- (b) In addition to the above assumptions, assume that $Var(u_i|x_i) = \sigma^2$. Show that

$$Var(\tilde{\beta}_1) = \frac{\sigma^2 \sum_{i=1}^n (z_i - \bar{z})^2}{[\sum_{i=1}^n (z_i - \bar{z})x_i]^2}.$$

Using $[\sum_{i=1}^n (z_i - \bar{z})(x_i - \bar{x})]^2 \leq [\sum_{i=1}^n (z_i - \bar{z})^2] [\sum_{i=1}^n (x_i - \bar{x})^2]$, show that $Var(\hat{\beta}_1) \leq Var(\tilde{\beta}_1)$. Discuss the result.

- (c) Now suppose that $Var(u_i|x_i) = \sigma^2 x_i^2$. Find the variance of $\tilde{\beta}_1$. Suggest an estimator that is more efficient than $\tilde{\beta}_1$ and $\hat{\beta}_1$.

Question 3 (40 points).

- (a) Suppose that the time series process y_t is generated by

$$y_t = e_t + \theta e_{t-1},$$

where θ is a finite constant and $e_t \sim iid(0, \sigma^2)$ with $\sigma^2 < \infty$. Show if the process y_t is covariance (weakly) stationary and weakly dependent.

(Hint: Recall that the covariance stationarity puts some restrictions on the mean, the variance and the covariance of the process and the weak dependence imposes some additional conditions on the correlation between two realizations of the process).

- (b) Consider the model

$$\begin{aligned} y_t &= \beta y_{t-1} + u_t \\ u_t &= \rho_1 u_{t-1} + \rho_2 u_{t-2} + e_t, \end{aligned} \tag{4}$$

where $e_t \sim iid(0, \sigma^2)$. Show that the process for y_t can be rewritten as $y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + e_t$, where α_1, α_2 and α_3 are functions of β, ρ_1 and ρ_2 . Discuss the consequences on the estimator of β if model (4) is estimated by OLS.

- (c) Consider the FDL(2) model

$$y_t = \delta_0 x_t + \delta_1 x_{t-1} + \delta_2 x_{t-2} + u_t, \tag{5}$$

where u_t is serially uncorrelated. Rewrite the model (5) in a form that would allow you to test directly the hypothesis that the long-run propensity (multiplier) is equal to 1.