Online Quiz 4

Total: 20 Marks

Half (or 0.5) marks are the minimum unit and the correction in two digits is required in the answers. No mark if the answer cannot be identified.

IV Regression

We want to estimate a supply equation for young construction workers in Australia: $Hour_i = \beta_0 + \beta_1 Wage_i + \beta_2 Educ_i + \beta_3 Age_i + u_i$, in which $Hour_i$ is the supply of labour, $Wage_i$ is hourly wage, and $Educ_i$ is years of education. Suppose large outliers are unlikely. All variables are i.i.d. draws from their joint distribution and have nonzero finite fourth moments.

(1, 1 mark) Explain why this supply equation cannot be consistently estimated by OLS regression.

(2, 1 mark) One of your friends argued that $Exper_i$ and its square, $Exper_i^2$, to be instruments for $Wage_i$. Explain how these variables satisfy the condition of instruments.

(3, 1 mark) How to check weak instruments in this study?

(4, 1 mark) Is the supply equation identified? Explain.

(5, 1 mark) Can we statistically test the assumption that the instruments are exogenous in this study?

(6, 1 mark) Describe the steps (or STATA command) you would take to obtain IV regression.

(7, 1 mark) If the two conditions for a valid instrument hold, what is the assumption for your IV regression.

(8, 1 mark) Suppose $y_i = \beta_1 + \beta_2 x_i + u_i$ and z_i is a valid instrumental variable for the endogenous variable x_i. u_i is the error term. Show that $E(y_i | z_i) = \beta_1 + \beta_2 E(x_i | z_i)$ if $E(u_i | z_i) = 0$.

(9, 1 mark) Show β_2 could be expressed as a fraction of the conditional expectation in (8) for the two cases with $z_i = 1$ and $z_i = 0$.

(10, 1 mark) Explain how $E(x_i|z_i = 1) - E(x_i|z_i = 0)$ could be viewed as a measure of the strength of the instrumental variable z_i.

The condition $z_i = 1$ defines a subset of the population, and the sample average of the subset of observations (i.e. $\bar{y}_1$) is a consistent estimator for the population average, $E(y_i|z_i = 1)$. Replacing each element by its consistent estimator gives us the Wald Estimator ($\hat{\beta}_{WALD} = \frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0}$), in honour of Abraham Wald.

Time Series

Suppose we have a stationary process: $y_t = \beta_0 + \beta_1 y_{t-1} + u_t$ and u_t follows the standard normal distribution.

(11, 1 mark) Explain what is the meaning of stationarity.

(12, 2 mark) Show the expected value and variance of y_t.

(13, 1 mark) R^2 is always increased whenever we include the lags and can we include the lags as much as possible?

(14, 1 mark) How to choose the number of lags p in an $AR(p)$?

(15, 1 mark) Are the forecasts from the time series model the OLS predicted values? Why?

(16, 2 marks) Compute the 1st and 2nd autocovariance of y_t.

(17, 1 marks) Compute the 1st and 2nd autocorrelation of y_t.

(18, 1 mark) What is the difference between BIC and AIC?