

IIA-3 Econometrics: Supervision 5

Emre Usenmez

Christmas Break 2024

Topics Covered

Faculty Qs:

Supplementary Qs: Endogeneity, measurement errors; simultaneous equations;

Very grateful to Dr Oleg Kitov and Dr Clive Lawson for the very informative stylized answers to previous iterations of the supervision questions.

FACULTY QUESTIONS

QUESTION 1

SUPPLEMENTARY QUESTIONS

QUESTION 1

(a) Explain what is meant when it is said that the explanatory variables and the disturbance term in a regression equation are not independent. What can be said about the properties of the OLS estimates in this case?

Answer: If the disturbance term and the explanatory variables are not independent then they are correlated. Those explanatory variables that are correlated with the error term are called *endogenous variables*. Since unbiasedness depends on $Cov(\varepsilon_i, X_i) = 0$, this dependency between the error term and the explanatory variables would yield biased estimates.

(b) Suppose that $Y_i = \alpha + \beta X_i + \lambda W_i + \varepsilon_i$ where there also exists a relationship between X_i and W_i of the form $W_i = \rho + \phi X_i + v_i$. Show that if Y_i is estimated using only the X_i variable then the estimate of β obtained is biased. Under what circumstances would this estimate of β be biased downwards?

Answer: Let's start by substituting in the latter expression into the former:

$$\begin{aligned} Y_i &= \alpha + \beta X_i + \lambda W_i + \varepsilon_i \\ &= \alpha + \beta X_i + \lambda(\rho + \phi X_i + v_i) + \varepsilon_i \\ &= (\alpha + \lambda\rho) + (\beta + \lambda\phi)X_i + (\lambda v_i + \varepsilon_i) \\ &= \gamma_0 + \gamma_1 X_i + u_i \end{aligned}$$

Now notice that both W_i and u_i depend on v_i . This means the assumption of exogeneity, i.e. independence between the explanatory variable and the disturbance term, would be violated when Y is regressed on X . As a result, $\hat{\gamma}_1$ would be *inconsistent* and *biased*.

To see this, start by looking at the expression for the regression coefficient γ_1

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2} = \gamma_1 + \frac{\sum_{i=1}^n (X_i - \bar{X})(u_i - \bar{u})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Since X and u are not distributed independently of each other, we can't summarize the distribution of the error term, or obtain an expression for its expected value. The most we can do is to determine how the error term would behave if the sample were very large.

However, neither the numerator nor the denominator tends to a particular limit as n increases. To get around this, we can divide both the numerator and the denominator by n . Then the probability limit of $\hat{\gamma}_1$ as n tends to infinity becomes

$$\begin{aligned}
 plim(\hat{\gamma}_1) &= \gamma_1 + \frac{plim\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(u_i - \bar{u})\right)}{plim\left(\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2\right)} \\
 &= \gamma_1 + \frac{Cov(X, u)}{Var(X)} \\
 &= \gamma_1 + \frac{Cov\left(\left(\frac{W - \rho - v}{\phi}\right), (\lambda v + \varepsilon)\right)}{Var\left(\frac{W - \rho - v}{\phi}\right)} \\
 &= \gamma_1 + \frac{Cov\left(\frac{W - \rho}{\phi}, \lambda v\right) + Cov\left(\frac{W - \rho}{\phi}, \varepsilon\right) + Cov\left(\frac{-v}{\phi}, \lambda v\right) + Cov\left(\frac{-v}{\phi}, \varepsilon\right)}{Var\left(\frac{W - \rho - v}{\phi}\right)}
 \end{aligned}$$

If we then assume that the error term in the original model, ε , is distributed independently of W , and the error term in the second model, v , is distributed independently of W and ε , then the first, second and fourth terms of the numerator are zero. Then

$$\begin{aligned}
 plim(\hat{\gamma}_1) &= \gamma_1 + \frac{0 + 0 + Cov\left(\frac{-v}{\phi}, \lambda v\right) + 0}{Var\left(\frac{W - \rho - v}{\phi}\right)} \\
 &= \gamma_1 + \frac{-\frac{\lambda}{\phi} Var(v)}{Var\left(\frac{W - \rho}{\phi}\right) + Var\left(\frac{-v}{\phi}\right) + 2Cov\left(\frac{W - \rho}{\phi}, \frac{-v}{\phi}\right)} \\
 &= \gamma_1 + \frac{-\frac{\lambda}{\phi} Var(v)}{\frac{1}{\phi^2} Var(W) + \frac{1}{\phi^2} Var(v) + 0} \\
 &= \gamma_1 - \lambda \phi \frac{Var(v)}{Var(W) + Var(v)}
 \end{aligned}$$

Thus $\hat{\gamma}_1$ is subject to bias whereby the bias is negative if $\lambda\phi$ is positive, and the bias is positive if $\lambda\phi$ is negative.