Question 1: Instrumental Variables

Consider the model:

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$$

where x_{1i} is endogenous (i.e. $\mathbb{E}[x_{1i}\epsilon_i] \neq 0$) and where x_{2i} is exogenous (i.e. $\mathbb{E}[x_{2i}\epsilon_i] = 0$). Let $x_i = [x_{1i}, x_{2i}]'$ and let $\hat{\beta} = [\hat{\beta}_1, \hat{\beta}_2]' = (\frac{1}{n}\sum_{i=1}^n x_i x_i')^{-1}(\frac{1}{n}\sum_{i=1}^n x_i y_i)$ be the OLS estimator of β.

(a) Show that plim$(\hat{\beta}_1) \neq \beta_1$ and plim$(\hat{\beta}_2) \neq \beta_2$. (Hint: write out each of the individual elements of each matrix.)

(b) Suppose that there is an instrument z_{1i} that satisfies $\mathbb{E}[z_{1i}\epsilon_i] = 0$ and $\mathbb{E}[z_{1i}x_{1i}] \neq 0$. Carefully write out the first-stage equation for TSLS specifying exactly which variables are included on the right-hand side.

(c) Let z_i represent the column vector of covariates from the first-stage equation. Let $\hat{\pi}$ represent the OLS estimator of the coefficients from the first-stage equation. An economist estimates the following 2nd-stage equation by OLS:

$$y_i = \gamma_1 \widehat{x_{1i}} + \gamma_2 x_{2i} + u_i$$

where $\widehat{x_{1i}} = z_i'\hat{\pi}$. Show that this estimator is consistent for β_1 and β_2.

Question 2: Measurement Error

Consider the following population model with one covariate:

$$y_i = \beta x_i^* + u_i$$

in which x_i^* is latent (unobserved). The observed regressor is:

$$x_i = x_i^* + v_i$$

in which v_i is independent of u_i. Further, we know that v_i and u_i are symmetrically distributed around 0.

(a) Detail why, in the presence of measurement error, β is not identified.

(b) Derive the plim of $\hat{\beta}_{OLS}$ and an expression for the attenuation bias in terms of $Var(v_i)$ and $Var(x_i^*)$. Explain how this bias would affect your interpretation of the estimate of $\hat{\beta}_{OLS}$.

(c) Suppose that there is an instrument z_i that satisfies $\mathbb{E}[z_i u_i] = 0$ and $\mathbb{E}[z_i x_i] \neq 0$. You estimate the following first-stage regression:

$$x_i = \pi z_i + \epsilon_{i1},$$

and the reduced-form equation:

$$y_i = \gamma z_i + \epsilon_{i2}.$$

Show how your estimates of $\hat{\pi}_{OLS}$ and $\hat{\gamma}_{OLS}$ can be used to construct a consistent estimator for β.

Question 3: Seemingly Unrelated Regressions

Consider the system of equations:

$$y_{i1} = x_{i1}'\beta_1 + u_{i1}$$

$$y_{i2} = x_{i2}'\beta_2 + u_{i2}$$

for which $\beta = (\beta_1, \beta_2)'$ and $U_i = (u_{i1}, u_{i2})'$ with $\mathbb{E}[u_{i1}] = \mathbb{E}[u_{i2}] = 0$ and

$$\mathbb{E}[U_i U_i'] = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

.

(a) What is the seemingly unrelated regressions (SUR) estimator for β? How does the SUR estimator compare with the equation-by-equation OLS estimator of β?

(b) Let $\sigma_{12} = 0$. Show in detail how the SUR estimator compares to the OLS estimator.

(c) Let $n = 2$ so that there is only 1 observation for each equation. (For this case, x_{i1}, x_{i2}, β_1, and β_2 are all scalars.) Show in detail how the SUR estimator compares to the OLS estimator.