6

Suppose the data generation process is as

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \epsilon_i.$$

Where x_{1i} and x_{2i} are included variables, and x_{3i} and x_{4i} are omitted variables. Assume the full OLS regression leads to estimates as

$$y_i = \hat{\alpha} + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{\beta}_4 x_{4i} + \hat{\epsilon}_i,$$

and the two auxiliary regression of x_{3i} and x_{4i} on x_{1i} and x_{2i} and constant term are

$$\begin{aligned} x_{3i} &= \hat{\pi}_0 + \hat{\pi}_1 x_{1i} + \hat{\pi}_2 x_{2i} + \hat{\nu}_i, \\ x_{4i} &= \hat{\gamma}_0 + \hat{\gamma}_1 x_{1i} + \hat{\gamma}_2 x_{2i} + \hat{\mu}_i. \end{aligned}$$

Then we have

$$\begin{aligned} y_i =& \hat{\alpha} + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{\beta}_4 x_{4i} + \hat{\epsilon}_i \\ =& \hat{\alpha} + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} \\ & + \hat{\beta}_3 \left(\hat{\pi}_0 + \hat{\pi}_1 x_{1i} + \hat{\pi}_2 x_{2i} + \hat{\nu}_i\right) \\ & + \hat{\beta}_4 \left(\hat{\gamma}_0 + \hat{\gamma}_1 x_{1i} + \hat{\gamma}_2 x_{2i} + \hat{\mu}_i\right) \\ & + \hat{\epsilon}_i \\ =& \left(\hat{\alpha} + \hat{\beta}_3 \hat{\pi}_0 + \hat{\beta}_4 \hat{\gamma}_0\right) \\ & + \left(\hat{\beta}_1 + \hat{\beta}_3 \hat{\pi}_1 + \hat{\beta}_4 \hat{\gamma}_1\right) x_{1i} \\ & + \left(\hat{\beta}_2 + \hat{\beta}_3 \hat{\pi}_2 + \hat{\beta}_4 \hat{\gamma}_2\right) x_{2i} \\ & + \hat{\epsilon}_i + \hat{\beta}_3 \hat{\nu}_i + \hat{\beta}_4 \hat{\mu}_i. \end{aligned}$$

The inconsistent estimates of short regression are

$$\begin{aligned} \tilde{\alpha} &= \hat{\alpha} + \underbrace{\hat{\beta}_3 \hat{\pi}_0 + \hat{\beta}_4 \hat{\gamma}_0}_{inconsistency}, \\ \tilde{\beta}_1 &= \hat{\beta}_1 + \underbrace{\hat{\beta}_3 \hat{\pi}_1 + \hat{\beta}_4 \hat{\gamma}_1}_{inconsistency}, \\ \tilde{\beta}_2 &= \hat{\beta}_2 + \underbrace{\hat{\beta}_3 \hat{\pi}_2 + \hat{\beta}_4 \hat{\gamma}_2}_{inconsistency}. \end{aligned}$$

These are OLS estimates, because the conditions for OLS estimators are sastified,

$$\sum_{i=1}^{n} \hat{\epsilon}_i + \hat{\beta}_3 \hat{\nu}_i + \hat{\beta}_4 \hat{\mu}_i = \underbrace{\sum_{i=1}^{n} \hat{\epsilon}_i}_{=0} + \hat{\beta}_3 \underbrace{\sum_{i=1}^{n} \hat{\nu}_i}_{=0} + \hat{\beta}_4 \underbrace{\sum_{i=1}^{n} \hat{\mu}_i}_{=0} = 0.$$

$$\sum_{i=1}^{n} x_{1i}\left(\hat{\epsilon}_i + \hat{\beta}_3\hat{\nu}_i + \hat{\beta}_4\hat{\mu}_i\right) = \underbrace{\sum_{i=1}^{n} x_{1i}\hat{\epsilon}_i}_{=0} + \hat{\beta}_3\underbrace{\sum_{i=1}^{n} x_{1i}\hat{\nu}_i}_{=0} + \hat{\beta}_4\underbrace{\sum_{i=1}^{n} x_{1i}\hat{\mu}_i}_{=0} = 0.$$

$$\sum_{i=1}^{n} x_{2i}\left(\hat{\epsilon}_i + \hat{\beta}_3\hat{\nu}_i + \hat{\beta}_4\hat{\mu}_i\right) = \underbrace{\sum_{i=1}^{n} x_{2i}\hat{\epsilon}_i}_{=0} + \hat{\beta}_3\underbrace{\sum_{i=1}^{n} x_{2i}\hat{\nu}_i}_{=0} + \hat{\beta}_4\underbrace{\sum_{i=1}^{n} x_{2i}\hat{\mu}_i}_{=0} = 0.$$

These results hold because $\hat{\epsilon}_i$, $\hat{\nu}_i$, and $\hat{\mu}_i$ are OLS regression residuals.

The directions of inconsistencies in $\hat{\alpha}$, $\hat{\beta}_1$, and $\hat{\beta}_2$ depends on the signs and scales of $\hat{\beta}$s, $\hat{\pi}$s, and $\hat{\gamma}$s.