Post-Instrument Bias in Linear Models

Online Appendix

(Not Intended for Publication)

January 17, 2023

Derivations

Proof of Proposition 1. Assume that all variables have mean zero.[1] We start by writing the expression of the probability limit of $\widehat{\beta}^{OLS}$ in matrix notation,

$$plim\ \widehat{\beta}^{OLS} = \begin{bmatrix} \beta_x \\ \beta_w \end{bmatrix} + E \begin{bmatrix} x^2 & xw \\ xw & w^2 \end{bmatrix}^{-1} E \begin{bmatrix} x\epsilon \\ w\epsilon \end{bmatrix}.$$

After computing the inverse of the variance-covariance matrix, we have

$$\begin{aligned} plim\ \widehat{\beta}^{OLS}_x &= \beta_x + \frac{E[w^2]E[x\epsilon] - E[xw]E[w\epsilon]}{E[x^2]E[w^2] - E[xw]^2} \\ &= \beta_x + \frac{E[w^2]E[x\epsilon] - E[xw]E[w\epsilon]}{\sigma_x^2\sigma_w^2(1-\rho_{xw}^2)} \\ &= \beta_x + \frac{\sigma_w^2\sigma_x\sigma_\epsilon(\rho_{x\epsilon} - \rho_{xw}\rho_{w\epsilon})}{\sigma_x^2\sigma_w^2(1-\rho_{xw}^2)} \\ &= \beta_x + \frac{\sigma_\epsilon(\rho_{x\epsilon} - \rho_{xw}\rho_{w\epsilon})}{\sigma_x(1-\rho_{xw}^2)}. \end{aligned}$$

Similarly for the $plim\ \widehat{\beta}^{IV}$,

[1] This is without loss of generality, as shown by Giles (1984).

$$
\begin{aligned}
plim\ \widehat{\beta}^{IV} &= \begin{bmatrix} \beta_x \\ \beta_w \end{bmatrix} + E\begin{bmatrix} zx & zw \\ xw & w^2 \end{bmatrix}^{-1} E\begin{bmatrix} z\epsilon \\ w\epsilon \end{bmatrix} \\
plim\ \widehat{\beta}^{IV}_x &= \beta_x + \frac{E[w^2]E[z\epsilon] - E[zw]E[w\epsilon]}{E[zx]E[w^2] - E[zw]E[xw]} \\
&= \beta_x + \frac{-E[zw]E[w\epsilon]}{\sigma_x\sigma_w^2\sigma_z(\rho_{zx} - \rho_{xw}\rho_{zw})} \\
&= \beta_x + \frac{-\sigma_w^2\sigma_z\sigma_\epsilon\rho_{zw}\rho_{w\epsilon}}{\sigma_x\sigma_w^2\sigma_z(\rho_{zx} - \rho_{xw}\rho_{zw})} \\
&= \beta_x + \frac{-\sigma_\epsilon\rho_{zw}\rho_{w\epsilon}}{\sigma_x(\rho_{zx} - \rho_{xw}\rho_{zw})},
\end{aligned}
$$

where we used the fact that $E[z\epsilon] = 0$. Finally for the $plim\ \widehat{\beta}^{IV no\ w}$,

$$
\begin{aligned}
plim\ \widehat{\beta}^{IV\ no\ w} &= \begin{bmatrix} \beta_x \end{bmatrix} + E\begin{bmatrix} zx \end{bmatrix}^{-1} E\begin{bmatrix} z\epsilon_0 \end{bmatrix} \\
plim\ \widehat{\beta}^{IV\ no\ w}_x &= \beta_x + \frac{E[z\epsilon_0]}{E[zx]} \\
&= \beta_x + \frac{E[z(\beta_w w + \epsilon]}{E[zx]} \\
&= \beta_x + \frac{E[z\beta_w w]}{E[zx]} \\
&= \beta_x + \frac{\beta_w\sigma_z\sigma_w\rho_{zw}}{\sigma_z\sigma_x\rho_{zx}} \\
&= \beta_x + \frac{\beta_w\sigma_w\rho_{zw}}{\sigma_x\rho_{zx}}
\end{aligned}
$$

where we used the fact that $E[z\epsilon] = 0$.

$\square$

Proof of Corollary 3. First we examine the case $\rho_{x\epsilon} = \rho_{w\epsilon}$ and $\rho_{zx} = \rho_{zw}$. The OLS and IV biases are

$$\left|Bias(\widehat{\beta}_x^{OLS})\right| = \left|\frac{\sigma_\epsilon \rho_{w\epsilon}}{\sigma_x(1+\rho_{xw})}\right|,$$
$$\left|Bias(\widehat{\beta}_x^{IV})\right| = \left|\frac{\sigma_\epsilon \rho_{w\epsilon}}{\sigma_x(1-\rho_{xw})}\right|.$$

Note that $\rho_{xw} = \rho_{zx} \cdot \rho_{zw} > 0$, and so $\left|Bias(\widehat{\beta}_x^{OLS})\right| < \left|Bias(\widehat{\beta}_x^{IV})\right|$.

Consider the case $\rho_{x\epsilon} = \rho_{w\epsilon}$ and $\rho_{zx} = -\rho_{zw}$. The OLS and IV biases are the same

$$\left|Bias(\widehat{\beta}_x^{OLS})\right| = \left|Bias(\widehat{\beta}_x^{IV})\right| = \left|\frac{\sigma_\epsilon \rho_{w\epsilon}}{\sigma_x(1+\rho_{xw})}\right|.$$

Similarly, if $\rho_{x\epsilon} = -\rho_{w\epsilon}$ and $\rho_{zx} = \rho_{zw}$, we have

$$\left|Bias(\widehat{\beta}_x^{OLS})\right| = \left|Bias(\widehat{\beta}_x^{IV})\right| = \left|\frac{\sigma_\epsilon \rho_{w\epsilon}}{\sigma_x(1-\rho_{xw})}\right|.$$

Finally, if $\rho_{x\epsilon} = -\rho_{w\epsilon}$ and $\rho_{zx} = -\rho_{zw}$, $\rho_{xw} = \rho_{zx} \cdot \rho_{zw} < 0$ and

$$\left|Bias(\widehat{\beta}_x^{OLS})\right| = \left|\frac{\sigma_\epsilon \rho_{w\epsilon}}{\sigma_x(1-\rho_{xw})}\right|,$$
$$\left|Bias(\widehat{\beta}_x^{IV})\right| = \left|\frac{\sigma_\epsilon \rho_{w\epsilon}}{\sigma_x(1+\rho_{xw})}\right|,$$

giving us the same strict inequality of the first case. □

Proof of Proposition 2.

For the case where $\tilde{w}$ is not included, we have

$$
\begin{aligned}
plim\ \widehat{\beta}^{IV\ no\ w} &= \begin{bmatrix} \beta_x \end{bmatrix} + E\begin{bmatrix} z\tilde{x} \end{bmatrix}^{-1} E\begin{bmatrix} z\left(\tilde{w}\beta_w - u_x\beta_x - u_w\beta_w + \epsilon\right) \end{bmatrix} \\
&= \beta_x + \frac{E[z\beta_w\tilde{w}]}{E[zx]} \\
&= \beta_x + \frac{\beta_w\sigma_z\sigma_{\tilde{w}}\rho_{z\tilde{w}}}{\sigma_z\sigma_{\tilde{x}}\rho_{z\tilde{x}}} \\
&= \beta_x + \frac{\beta_w\sigma_{\tilde{w}}\rho_{z\tilde{w}}}{\sigma_{\tilde{x}}\rho_{z\tilde{x}}}
\end{aligned}
$$

where we used the fact that $y = (\tilde{x} - u_x)\beta_x + (\tilde{w} - u_w)\beta_w + \epsilon$.

For the cases where $\tilde{w}$ is included, we start by writing the expression of the probability limit of $\widehat{\beta}^{OLS}$ in matrix notation as before,

$$
plim\ \widehat{\beta}^{OLS} = \begin{bmatrix} \beta_x \\ \beta_w \end{bmatrix} + E\begin{bmatrix} \tilde{x}^2 & \tilde{x}\tilde{w} \\ \tilde{x}\tilde{w} & \tilde{w}^2 \end{bmatrix}^{-1} E\begin{bmatrix} \tilde{x}\left(-u_x\beta_x - u_w\beta_w + \epsilon\right) \\ \tilde{w}\left(-u_x\beta_x - u_w\beta_w + \epsilon\right) \end{bmatrix},
$$

After computing the inverse of the variance-covariance matrix, we have

$$
\begin{aligned}
plim\ \widehat{\beta}_x^{OLS} &= \beta_x + \frac{E[\tilde{w}^2]E[\tilde{x}\,(-u_x\beta_x - u_w\beta_w + \epsilon)] - E[\tilde{x}\tilde{w}]E[\tilde{w}\,(-u_x\beta_x - u_w\beta_w + \epsilon)]}{E[\tilde{x}^2]E[\tilde{w}^2] - E[\tilde{x}\tilde{w}]^2} \\
&= \beta_x + \frac{E[\tilde{w}^2]\,(-E[\tilde{x}u_x]\beta_x - E[\tilde{x}u_w]\beta_w + E[\tilde{x}\epsilon]) - E[\tilde{x}\tilde{w}]\,(-E[\tilde{w}u_x]\beta_x - E[\tilde{w}u_w]\beta_w + E[\tilde{w}\epsilon])}{E[\tilde{x}^2]E[\tilde{w}^2] - E[\tilde{x}\tilde{w}]^2} \\
&= \beta_x + \frac{\sigma_{\tilde{w}}^2(-\sigma_{u_x}^2\beta_x + \sigma_{\tilde{x}}\sigma_\epsilon\rho_{\tilde{x}\epsilon}) - \sigma_{\tilde{x}}\sigma_{\tilde{w}}\rho_{\tilde{x}\tilde{w}}(-\sigma_{u_w}^2\beta_w + \sigma_{\tilde{w}}\sigma_\epsilon\rho_{\tilde{w}\epsilon})}{\sigma_{\tilde{x}}^2\sigma_{\tilde{w}}^2(1-\rho_{\tilde{x}\tilde{w}}^2)} \\
&= \beta_x\left(1 - \frac{\sigma_{u_x}^2}{\sigma_{\tilde{x}}^2(1-\rho_{\tilde{x}\tilde{w}}^2)}\right) + \frac{\sigma_\epsilon(\rho_{\tilde{x}\epsilon} - \rho_{\tilde{x}\tilde{w}}\rho_{\tilde{w}\epsilon})}{\sigma_{\tilde{x}}(1-\rho_{\tilde{x}\tilde{w}}^2)} + \frac{\sigma_{u_w}^2\rho_{\tilde{x}\tilde{w}}\beta_w}{\sigma_{\tilde{x}}\sigma_{\tilde{w}}(1-\rho_{\tilde{x}\tilde{w}}^2)}.
\end{aligned}
$$

Similarly for the $plim\ \widehat{\beta}^{IV}$,

$$
\begin{aligned}
plim\ \widehat{\beta}^{IV} &= \begin{bmatrix} \beta_x \\ \beta_w \end{bmatrix} + E\begin{bmatrix} z\tilde{x} & z\tilde{w} \\ \tilde{x}\tilde{w} & \tilde{w}^2 \end{bmatrix}^{-1} E\begin{bmatrix} z\,(-u_x\beta_x - u_w\beta_w + \epsilon) \\ \tilde{w}\,(-u_x\beta_x - u_w\beta_w + \epsilon) \end{bmatrix} \\
plim\ \widehat{\beta}_x^{IV} &= \beta_x + \frac{E[\tilde{w}^2]E[z\,(-u_x\beta_x - u_w\beta_w + \epsilon)] - E[z\tilde{w}]E[\tilde{w}\,(-u_x\beta_x - u_w\beta_w + \epsilon)]}{E[z\tilde{x}]E[\tilde{w}^2] - E[z\tilde{w}]E[\tilde{x}\tilde{w}]} \\
&= \beta_x + \frac{\sigma_{\tilde{w}}\sigma_z\rho_{z\tilde{w}}\sigma_{u_w}^2\beta_w}{\sigma_{\tilde{x}}\sigma_{\tilde{w}}^2\sigma_z(\rho_{z\tilde{x}} - \rho_{x\tilde{w}}\rho_{z\tilde{w}})} + \frac{-\sigma_{\tilde{w}}^2\sigma_z\sigma_\epsilon\rho_{z\tilde{w}}\rho_{\tilde{w}\epsilon}}{\sigma_{\tilde{x}}\sigma_{\tilde{w}}^2\sigma_z(\rho_{z\tilde{x}} - \rho_{\tilde{x}\tilde{w}}\rho_{z\tilde{w}})} \\
&= \beta_x + \frac{\sigma_{u_w}^2\rho_{z\tilde{w}}\beta_w}{\sigma_{\tilde{x}}\sigma_{\tilde{w}}(\rho_{z\tilde{x}} - \rho_{\tilde{x}\tilde{w}}\rho_{z\tilde{w}})} - \frac{\sigma_\epsilon\rho_{z\tilde{w}}\rho_{\tilde{w}\epsilon}}{\sigma_{\tilde{x}}(\rho_{z\tilde{x}} - \rho_{\tilde{x}\tilde{w}}\rho_{z\tilde{w}})}.
\end{aligned}
$$

□

References

Giles, David E.A. 1984. "Instrumental Variables Regressions Involving Seasonal Data." *Economic Letters* 14:339–343.