Linear Models Lecture 15: IV

Robert Gulotty

University of Chicago

May 7, 2024

2SLS and IV

- iv_robust(Y $\sim D + X|Z + X$, data = dat)
- IV formula:

$$\hat{\beta}_{IV} = (Z'X)^{-1}Z'y$$

- Two stage least squares:
  - Suppose in the first stage we regress

$$X = Z\gamma + v$$

  - In the second stage, we use $\hat{X} = Z\hat{\gamma} = Z(Z'Z)^{-1}Z'X = P_Z X$,

$$\hat{\beta}_{2SLS} = (\hat{X}'\hat{X})^{-1}\hat{X}'y$$

Equivalence Between 2SLS and IV

- 2SLS is exactly identical to IV

$$
\begin{aligned}
\hat{\beta}_{2SLS} &= (\hat{X}'\hat{X})^{-1}\hat{X}'y \\
&= (X'Z(Z'Z)^{-1}Z'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y \\
&= (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'y \\
&= (Z'X)^{-1}(Z'Z)(X'Z)^{-1}X'Z(Z'Z)^{-1}Z'y \qquad ((ABC)^{-1} = C^{-1}B^{-1}A^{-1}) \\
&= (Z'X)^{-1}(Z'Z)(Z'Z)^{-1}Z'y \\
&= (Z'X)^{-1}Z'y = \hat{\beta}_{IV}
\end{aligned}
$$

Challenges with IV

- The IV estimator is among the most common tools of econometrics.
- However, it has several weaknesses.
  - Imprecision
  - Small sample Bias
  - Sensitivity to Weak Instruments

Problems with IV estimator: Imprecision

- Suppose Z and X are mean 0, $y = X\beta + e$,

$$Z'X = X'Z = \sum z_i x_i = n * cov(z, x)$$
$$Z'Z = \sum z_i^2 = n * var(z)$$
$$X'X = \sum x_i^2 = n * var(x)$$
$$\hat{\beta}_{IV} = (Z'X)^{-1} Z'y$$
$$\hat{\beta}_{OLS} = (X'X)^{-1} X'y$$
$$Avar(\hat{\beta}_{OLS}) = \sigma_e^2 (X'X)^{-1}$$
$$Avar(\hat{\beta}_{IV}) = \sigma_e^2 (Z'X)^{-1} Z'Z (X'Z)^{-1}$$

Problems with IV estimator: Imprecision

$$
\begin{aligned}
Avar(\hat{\beta}_{OLS}) &= \sigma_e^2 (X'X)^{-1} = \frac{\sigma_e^2}{n} \frac{1}{var(x)} \\
Avar(\hat{\beta}_{IV}) &= \sigma_e^2 (Z'X)^{-1} Z'Z (X'Z)^{-1} = \frac{\sigma_e^2}{n^2} \frac{n * var(z)}{cov(x,z)^2} \\
&= \frac{\sigma_e^2}{n} \frac{1}{var(x)} \frac{var(x) var(z)}{cov(x,z)^2} \\
&= \frac{\sigma_e^2}{n} \frac{1}{var(x)} \frac{1}{\rho_{xz}^2} \\
&= Avar(\hat{\beta}_{OLS}) \frac{1}{\rho_{xz}^2}
\end{aligned}
$$

- As $\rho_{xz}^2 \rightarrow 0$, $Avar(\hat{\beta}_{IV}) \rightarrow \infty$

Problems with IV estimator: Bias

- IV is often is neither biased nor unbiased because it does not even have an expectation.
- Kiviet has shown that the IV estimator has M moments, the number of overidentifying restrictions. If $q = 0$, IV has no expectation.

$$
\begin{aligned}
y &= X\beta + e \\
X &= Z\pi + v \\
\hat{\beta}_{IV} &= (X'P_Z X)^{-1} X'P_z y \\
&= \beta + (X'P_Z X)^{-1} X'P_z e \\
&= \beta + (X'P_Z X)^{-1} (\pi'Z' + v')P_z e \\
&= \beta + (X'P_Z X)^{-1} (\pi'Z' + v')P_z e \\
&= \beta + (X'P_Z X)^{-1} \pi'Z'P_z e + (X'P_Z X)^{-1} v'P_z e \\
&= \beta + (X'P_Z X)^{-1} \pi'Z'e + (X'P_Z X)^{-1} v'P_z e
\end{aligned}
$$

Form of small sample bias

$$
\begin{aligned}
E(\hat{\beta}_{IV}) - \beta &\approx E(X'P_ZX)^{-1}E(\pi'Z'e) + E(X'P_ZX)^{-1}E(v'P_ze) \\
&= E(X'P_ZX)^{-1}\pi'E(Z'e) + E(X'P_ZX)^{-1}E(v'P_ze) \\
&= (E(X'P_ZX))^{-1}E(v'P_ze) \\
&= (E(\pi'Z' + v')P_z(Z\pi + v)))^{-1}E(v'P_ze) \\
&= (E(\pi'Z'Z\pi + \pi'Z'v + v'Z\pi + v'P_zv))^{-1}E(v'P_ze) \\
&= (E(\pi'Z'Z\pi) + E(v'P_zv))^{-1}E(v'P_ze) \qquad \text{(b/c } E(Z'e) = E(Z'v) = 0) \\
&= (E(\pi'Z'Z\pi) + E(v'P_zv))^{-1}\sigma^2_{ev}p \\
&= (E(\pi'Z'Z\pi) + \sigma^2_v p)^{-1}\sigma^2_{ev}p \\
&= \frac{1}{\left(\frac{E(\pi'Z'Z\pi)/p}{\sigma^2_v} + 1\right)}\frac{\sigma^2_{ev}}{\sigma^2_v}
\end{aligned}
$$

F-test

$$E(\hat{\beta}_{IV}) - \beta \approx \frac{1}{\left(\frac{E(\pi' Z' Z \pi)/p}{\sigma_v^2} + 1\right)} \frac{\sigma_{ev}^2}{\sigma_v^2} \approx \frac{1}{(1 + F_{p,n-p})} \frac{\sigma_{ev}^2}{\sigma_v^2}$$

- F is the test where the null is that all instrument coefficients are 0.
- The bias of IV only goes away if $F \to \infty$
- The bias of IV is the OLS bias as $F \to 0$.
- Adding useless instruments increases p, which decreases F and increases the bias.

Weak instruments

Suppose we have a single x and a single instrument z. An instrument is weak if ρ_{zx} is small.

$$
\begin{aligned}
plim\hat{\beta}_{OLS} &= plim\frac{cov(x,y)}{var(x)} = plim\frac{cov(x,\alpha+\beta+e)}{var(x)} \\
&= \beta + plim\frac{cov(x,e)}{var(x)} = \beta + plim\frac{cov(x,e)}{\sqrt{var(x)}\sqrt{var(e)}}\frac{\sqrt{var(e)}}{\sqrt{var(x)}} \\
&= \beta + \rho_{xe}\frac{\sigma_e}{\sigma_x} \\
plim\hat{\beta}_{IV} &= plim\frac{cov(x,\alpha+\beta+e)}{cov(z,x)} \\
&= \beta + plim\frac{cov(z,e)}{cov(z,x)} = \beta + plim\frac{\frac{cov(z,e)}{\sqrt{var(x)}\sqrt{var(e)}}}{\frac{cov(z,x)}{\sqrt{var(x)}\sqrt{var(z)}}}\frac{\sqrt{var(e)}}{\sqrt{var(x)}} \\
&= \beta + \frac{\rho_{ze}}{\rho_{zx}}\frac{\sigma_e}{\sigma_x} \\
&= \beta + \frac{\rho_{ze}}{\rho_{zx}\rho_{xe}}\rho_{xe}\frac{\sigma_e}{\sigma_x} = \beta + \frac{\rho_{ze}}{\rho_{zx}\rho_{xe}}ABias(\hat{\beta}_{OLS})
\end{aligned}
$$

Weak/Bad instruments are worse than OLS

$$\frac{ABias(\hat{\beta}_{OLS})}{ABias(\hat{\beta}_{IV})} > 1 \rightarrow \frac{\rho_{ze}}{\rho_{zx}\rho_{xe}} > 1$$

If $\frac{\rho_{ze}}{\rho_{zx}\rho_{xe}} \geq 1$, then IV is more biased than OLS.
Suppose $\rho_{xu} = .5$, so X is super endogenous, Z is barely endogenous: $\rho_{zu} = 0.01$.
Small $\rho_{zx} = 0.019$ gives $\frac{ABias(\hat{\beta}_{OLS})}{ABias(\hat{\beta}_{IV})} = 1.052$.

Testing power of instruments

$$\frac{ABias(\hat{\beta}_{OLS})}{ABias(\hat{\beta}_{IV})} \approx \frac{1}{F}$$

F statistic of 100 means IV is 1% as biased as OLS.

Testing endogeneity via Durbin-Hausman-Wu test

- If X is exogenous, then both OLS and IV are consistent, but OLS is BLUE.
- Asymptotically, the difference between OLS and IV should converge to zero.

$$H = (\hat{\beta}_{IV} - \hat{\beta}_{OLS})'[Avar(\hat{\beta}_{IV}) - Avar(\hat{\beta}_{OLS})]^{-1}(\hat{\beta}_{IV} - \hat{\beta}_{OLS}) \sim \chi^2_{dim(\beta)}$$

- Rejecting null says that OLS and IV are not close to one another, so either X is endogneous or Z is an invalid instrument.

Control Function Regression

- Assume that X_2 is endogenous:

$$Y = X_1'\beta_1 + X_2'\beta_2 + e$$

$$X_2 = \Gamma_{12}'Z_1 + \Gamma_{22}'Z_2 + u_2$$

- The control function approach first directly models the error:

$$e = u_2'\alpha + v$$

$$\alpha = (E[u_2u_2'])^{-1}E[u_2e]$$

$$E[u_2v] = 0$$

Control Function Regression

- We then plug this in to the original structural form equation, controlling for the error.

$$
\begin{aligned}
Y &= X_1'\beta_1 + X_2'\beta_2 + e \\
Y &= X_1'\beta_1 + X_2'\beta_2 + u_2'\alpha + v \\
E[X_1 v] &= 0 \\
E[X_2 v] &= 0 \\
E[u_2 v] &= 0
\end{aligned}
$$

- After we control for u_2, the error is uncorrelated with X.
- We do so with the reduced form residual

$$\hat{u}_{2i} = X_{2i} - \hat{\Gamma}'_{12}Z_1 + \hat{\Gamma}'_{22}Z_2$$

- It is like subtracting off the endogenous part.

$$\boldsymbol{Y} = \boldsymbol{X}\hat{\beta} + \hat{\boldsymbol{U}}_e\hat{\alpha} + \hat{v}$$

Application: Heterogenous Returns to Education

- Consider the canonical returns to education model:

$$lwage_i = \mathbf{z}_{i1}\boldsymbol{\delta}_1 + g_{i1}educ_i + u_{i1}$$

- The returns to schooling for the population is $\gamma_i = E[g_{i1}]$

$$g_{i1} = \gamma_1 + v_{i1}$$

- Plugging in:

$$lwage_i = \mathbf{z}_{i1}\boldsymbol{\delta}_1 + \gamma_1 educ_i + v_{i1}educ_i + u_{i1}$$

$$educ_i = \mathbf{z}_i \pi_2 + v_{i2}$$

  Control function approach assumes that unobservables are linearly related to v_{i2}

- We then proceed estimation by controlling for $\hat{v}_{i2}$ and the interaction between $\widehat{educ}_i$ and the estimated $\hat{v}_{i2}$.