

Economics 215B

(Lecture 6)

Ethan Kaplan

March 7, 2008

1 Instrumental Variables: Intro.

- Bias in OLS:
 - Consider a linear model:

$$Y = X\beta + \epsilon$$

- Suppose that

$$\text{cov}(X, \epsilon) = \rho$$

- then OLS yields:

$$\begin{aligned}\hat{\beta}_{OLS} &= (X'X)^{-1} X'Y = \\ &= (X'X)^{-1} X'(X\beta + \epsilon) \\ \implies E\hat{\beta}_{OLS} &= \beta + (X'X)^{-1} \rho\end{aligned}$$

- Two Stage Least Squares

- One solution to the problem of bias in OLS is to find variables correlated with Y only through their correlation with X and use only the variation in X correlated with these other variables (called instruments) Z . First run:

$$X = Z\gamma + \delta$$

- From this we get an estimate of γ which we call $\hat{\gamma}$ and a predicted X :

$$Z (Z'Z)^{-1} Z'X$$

- then run:

$$\left(\begin{array}{c} [X'Z (Z'Z)^{-1} Z'Z (Z'Z)^{-1} Z'X]^{-1} \\ X'Z (Z'Z)^{-1} Z'Y \end{array} \right)$$

- Three cases:

1. Under-identified: number of regressors in $Z <$ number of regressors in X
2. Just Identified: number of regressors in $Z =$ number of regressors in X
3. Over identified: number of regressors in $Z >$ number of regressors in X

- In the under-identified case, the model can not be estimated

- In the just identified case the dimension of $X'Z$ is the dimension of $Z'Z$ in which case:

$$\left(\begin{array}{c} [(Z'X)^{-1} (Z'Z) (Z'Z)^{-1} (Z'Z) (X'Z)^{-1}] \\ X'Z (Z'Z)^{-1} Z'Y \end{array} \right)$$

$$= (Z'X)^{-1} Z'Y$$

- Weak Instruments Problem

- One problem is that if $Z'X \approx 0 = Z'Y$, then the distribution $(Z'X)^{-1} Z'Y$ is the ratio of two normals with mean zero and is approximated well even in very large samples by a Cauchy Distribution, whose mean and variance do not exist. This can be very problematic.

- What is?

$$E \left[(Z'X)^{-1} Z'Y - \beta \right]$$

- In general, we don't know. However,

$$\begin{aligned} & p \lim \left[(Z'X)^{-1} Z'Y - \beta \right] \\ &= p \lim \left[(Z'X)^{-1} Z'(X\beta + \epsilon) - \beta \right] \\ &= p \lim \left[\beta + (Z'X)^{-1} Z'\epsilon - \beta \right] \\ &= p \lim [\beta - \beta] = 0 \end{aligned}$$

- Bias in OLS vs. 2SLS: $\frac{(Z'X)^{-1}Z'\epsilon}{X'\epsilon}$

– Re-expressed:

$$\frac{\sigma_{Z\epsilon}}{\sigma_{XZ}\sigma_{X\epsilon}}$$

– Another way to write the 2SLS estimator is:

$$(\hat{X}'\hat{X})^{-1}\hat{X}'\epsilon = \frac{\sigma_{\hat{X}\epsilon}}{\sigma_{\hat{X}}^2}$$

– as opposed to the OLS bias:

$$(X'X)^{-1}X'\epsilon = \frac{\sigma_{X\epsilon}}{\sigma_X^2}$$

– So the small bias of the 2SLS is in the direction of the OLS estimator.

- Wald Estimator:

- One special example is the so-called Wald Estimator:

$$\begin{aligned} Y_i &= \alpha_1 + \beta X_i + \epsilon \\ X_i &= \alpha_2 + \gamma D_i + \delta \end{aligned}$$

- where D_i is a dummy variable taking on the values of $\{0, 1\}$. Then:

$$\hat{\beta}_{WALD} = \frac{\bar{Y}_1 - \bar{Y}_0}{\bar{X}_1 - \bar{X}_0}$$

- where \bar{Y}_1, \bar{Y}_0 are the average Y when $Z = 1, 0$ respectively and \bar{X}_1, \bar{X}_0 are the average X when $Z = 1, 0$ respectively.

- Small Sample Bias of 2SLS:

- Approximate Bias Formula for small samples (derived using power series approximations):

$$\begin{aligned} & \frac{\sigma_{Z,\delta}}{\gamma'Z'Z\gamma} (K - 2) \\ &= \frac{\sigma_{Z,\delta}}{\sigma_\delta^2} \frac{\sigma_\delta^2}{\gamma'Z'Z\gamma} (K - 2) \end{aligned}$$

- where K is the number of excluded instruments.

$$\tau^2 = \frac{\sigma_\delta^2}{\gamma'Z'Z\gamma}$$

- τ^2 is called a concentration parameter and is equal to $\frac{1}{R^2}$ from the first stage regression.
- Commonly thought that bias is proportional to K . In fact, this is only true in the case where $\sigma_{XZ} = 0$ (or $\sigma_{XZ} \approx 0$). Otherwise $\gamma'Z'Z\gamma$ and thus τ^2 depend upon K .
- So is adding more instruments a good thing? Depends if they are correlated with LHS variables

conditional on the other instruments. Similar to out of sample prediction... not always a good idea.

- Can you test if instruments are too weak? You can run a joint F-test on the first stage (essentially the concentration parameter). Usually you want at least F-Statistic of 4 or 5 in the literature. Some will want at least 10.

- Limited Information Maximum Likelihood

- Can also estimate with Limited Information Maximum Likelihood. It turns out that though the asymptotic distribution of the 2SLS estimator and the LIML estimator are the same, the small sample distributions can be quite different in overidentified models. In particular, with LIML, the parameter being estimated is close to its population median rather than mean. The formula for LIML is:

$$L(\beta, \pi, \Omega) = \sum_{n=1}^N \left(\begin{array}{c} -\frac{1}{2} |\Omega| - \frac{1}{2} \begin{pmatrix} Y_i - \beta\gamma Z_i \\ X_i - \gamma Z_i \end{pmatrix}' \\ \Omega^{-1} \begin{pmatrix} Y_i - \beta\gamma Z_i \\ X_i - \gamma Z_i \end{pmatrix} \end{array} \right)$$

- Example with real and random instruments:

(Single Instr.	500 Instr. 2SLS
	Real	0.089 (0.011)	0.073 (0.008)***
	Random	-1.958 (18.116)	0.059 (0.085)
		500 Instr. LIML	
	Real	0.095 (0.017)***	
	Random	-0.330 (0.1001)***	
)			

- 2SLS Inference:

- Suppose you run 2SLS in two stages. Then you compute SEs as:

$$\left(\hat{X}'\hat{X}\right)^{-1}\hat{\sigma}^2$$

- Instead you should take the asymptotic variance of:

$$\left(\begin{array}{c} \left[(Z'X)^{-1}(Z'Z)(Z'Z)^{-1}(Z'Z)(X'Z)^{-1}\right] \\ X'Z(Z'Z)^{-1}Z'Y \end{array}\right)$$

- In which case you get:

$$\left(X'Z\right)\left(Z'Z\right)^{-1}\left(Z'X\right)^{-1}\hat{\sigma}^2$$

- You can show that the true SEs are large than the second stage OLS because they include the variation from the first stage which the second stage OLS standard errors do not.