7) Instrumental Variables (IV) and Two-Stage Least Squares (2SLS)

Vitor Kamada

December 2018

Reference

Tables, Graphics, and Figures from:

Wooldridge (2010). **Econometric Analysis of Cross Section and Panel Data.** Ch 5.

Endogeneity: Instrumental Variables

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + u$$

$$Cov(x_j, u) = 0, \quad j = 1, 2, ..., K - 1$$

$$\text{I) } Cov(z_1, u) = 0$$

$$x_k = \delta_0 + \delta_1 x_1 + ... + \delta_{k-1} x_{k-1} + \theta_1 z_1 + r_k$$

$$\text{II) } \theta_1 \neq 0$$

Instrumental Variables Estimation

$$y = x\beta + u$$

$$z'y = z'x\beta + z'u$$

$$E(z'y) = E(z'x)\beta$$

$$\hat{\beta}_{IV} = (N^{-1} \sum_{i=1}^{N} z_i' x_i)^{-1} (N^{-1} \sum_{i=1}^{N} z_i' y_i)$$

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

Weak Instrumental Variable

$$y = \beta_0 + \beta_1 x + u$$

$$plim\hat{\beta}_{1,IV} = \beta_1 + \frac{Corr(z,u)}{Corr(z,x)} \cdot \frac{\sigma_u}{\sigma_x}$$

$$plim\hat{\beta}_{1,OLS} = \beta_1 + Corr(x,u) \cdot \frac{\sigma_u}{\sigma_x}$$

Inference with the IV Estimator

$$Var(u) = E(u^2|z) = \sigma^2$$

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{n\sigma_x^2\rho_{x,z}^2}$$

Two-Stage Least Squares (2SLS)

$$\text{I) } Cov(z_h, u) = 0, \quad h = 1, 2, ..., M$$

$$\text{II) } x_k = \delta_0 + \delta_1 x_1 + ... + \delta_{k-1} x_{k-1} + \theta_1 z_1 + ... + \theta_m z_m + r_k$$

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

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

2SLS - Implementation

1) Get the fitted values $\hat{x}_k$:

x_k on $1, x_1, \ldots, x_{k-1}, z_1, \ldots, z_m$

2) Run OLS

y on $1, x_1, \ldots, x_{k-1}, \hat{x}_k$

Mroz (1987)

$$log(wage) = \beta_0 + \beta_1 exp + \beta_2 exp^2 + \beta_3 educ + u$$

```
library(AER); data(mroz, package='wooldridge')
oursample <- subset(mroz, !is.na(wage))
```

Statistic	N	Mean	St. Dev.	Min	Max
wage	428	4.178	3.310	0.128	25.000
educ	428	12.659	2.285	5	17
exper	428	13.037	8.056	0	38
motheduc	428	9.516	3.308	0	17
fatheduc	428	8.988	3.523	0	17

R Code

```
ols <- lm(log(wage)~educ+exper+I(exper^2)
+exper+I(exper^2), data=oursample)

stage1 <- lm(educ~exper+I(exper^2)+
motheduc+fatheduc, data=oursample)

man.2SLS<-lm(log(wage)~fitted(stage1)
+exper+I(exper^2), data=oursample)

aut.2SLS<-ivreg(log(wage)~educ+exper+I(exper^2) |
motheduc+fatheduc+exper+I(exper^2),
data=oursample)
```

Results (OLS, 1S, 2S, 2SLS)

	(1)	(2)	(3)	(4)
educ	0.107***			0.061*
	(0.014)			(0.031)
fitted(stage1)			0.061*	
			(0.033)	
exper	0.042***	0.045	0.044***	0.044***
	(0.013)	(0.040)	(0.014)	(0.013)
I(exper2)	-0.001**	-0.001	-0.001**	-0.001**
	(0.0004)	(0.001)	(0.0004)	(0.0004)
motheduc		0.158***		
		(0.036)		
fatheduc		0.190***		
		(0.034)		
Constant	-0.522***	9.103***	0.048	0.048
	(0.199)	(0.427)	(0.420)	(0.400)
Observations	428	428	428	428