

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

Vitor Kamada

July 2018

## Endogeneity: Instrumental Variables

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

$$\text{Cov}(x_j, u) = 0, \quad j = 1, 2, \dots, K - 1$$

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

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

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

# Instrumental Variables Estimation

$$y = x\beta + u$$

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

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

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

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

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

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

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

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

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

## Two-Stage Least Squares (2SLS)

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

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

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

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

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

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

2) Run OLS

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

# Mroz (1987)

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

use mroz.dta

summarize wage exper educ motheduc fatheduc husedu if  
inlf==1

Variable	Obs	Mean	Std. Dev.	Min	Max
wage	428	4.177682	3.310282	.1282	25
exper	428	13.03738	8.055923	0	38
educ	428	12.65888	2.285376	5	17
motheduc	428	9.516355	3.3081	0	17
fatheduc	428	8.988318	3.523405	0	17
huseduc	428	12.61215	3.035163	4	17



# reg lwage exper expersq educ if inlf==1

lwage	Coef.	Std. Err.	t	P> t
exper	.0415665	.0131752	3.15	0.002
expersq	-.0008112	.0003932	-2.06	0.040
educ	.1074896	.0141465	7.60	0.000
_cons	-.5220406	.1986321	-2.63	0.009

reg educ exper expersq motheduc fatheduc huseduc if inlf==1

educ	Coef.	Std. Err.	t	P> t
exper	.0374977	.0343102	1.09	0.275
expersq	-.0006002	.0010261	-0.58	0.559
notheduc	.1141532	.0307835	3.71	0.000
fatheduc	.1060801	.0295153	3.59	0.000
huseduc	.3752548	.0296347	12.66	0.000
_cons	5.538311	.4597824	12.05	0.000

test motheduc fatheduc huseduc

( 1) **motheduc** = 0

( 2) **fatheduc** = 0

( 3) **huseduc** = 0

F( 3, 422) = 104.29

Prob > F = 0.0000

## predict educhat, xb

reg lwage exper expersq educhat if inlf==1

lwage	Coef.	Std. Err.	t	P> t
exper	.0430973	.013876	3.11	0.002
expersq	-.0008628	.0004144	-2.08	0.038
educhat	.0803918	.0227772	3.53	0.000

ivregress 2sls lwage exper expersq (educ = motheduc  
fatheduc husedu) if inlf==1

lwage	Coef.	Std. Err.	z	P> z
educ	.0803918	.021672	3.71	0.000
exper	.0430973	.0132027	3.26	0.001
expersq	-.0008628	.0003943	-2.19	0.029