6) Control Function Approach

Vitor Kamada

July 2018

Endogeneity

$$y_1 = z_1\delta_1 + \alpha_1 y_2 + u_1$$

$$E(z'u_1) = 0$$

$$y_2 = z\pi_2 + v_2$$

$$E(z'v_2) = 0$$

Endogeneity: $Cov(u_1, v_2) \neq 0$

$$u_1 = \rho_1 v_2 + e_1$$

Control Function Approach

$$y_1 = z_1\delta_1 + \alpha_1 y_2 + u_1$$

$$y_1 = z_1\delta_1 + \alpha_1 y_2 + \rho_1 v_2 + e_1$$

$$y_2 = z\pi_2 + v_2$$

$$y_1 = z_1\delta_1 + \alpha_1 y_2 + \rho_1 \hat{v}_2 + error$$

$$error_i = e_{i1} + \rho_1 z_i(\hat{\pi}_2 - \pi_2)$$

Mroz (1987): CF with Bootstrap

```
reg educ exper expersq motheduc fatheduc huseduc if ///
inlf==1, vce(boot, reps(400) seed(7) nodots)
predict vhat, residuals
reg lwage educ exper expersq vhat if inlf==1, ///
vce(boot, reps(400) seed(7) nodots)
```

lwage	Observed Coef.	Bootstrap Std. Err.	z	P>|z|
educ	.0803918	.0225192	3.57	0.000
exper	.0430973	.0156212	2.76	0.006
expersq	-.0008628	.0004348	-1.98	0.047
vhat	.047189	.0268704	1.76	0.079
_cons	-.1868572	.3170912	-0.59	0.556

reg educ exper expersq motheduc fatheduc huseduc if /// inlf==1, vce(boot, reps(400) seed(7) nodots)

educ	Observed Coef.	Bootstrap Std. Err.	z	P>|z|
exper	.0374977	.0331465	1.13	0.258
expersq	-.0006002	.0010203	-0.59	0.556
motheduc	.1141532	.0312135	3.66	0.000
fatheduc	.1060801	.0298867	3.55	0.000
huseduc	.3752548	.0343345	10.93	0.000
_cons	5.538311	.4530623	12.22	0.000

test (motheduc=0) (fatheduc=0) (huseduc=0)

```
 ( 1)  motheduc = 0
 ( 2)  fatheduc = 0
 ( 3)  huseduc = 0

       chi2(  3) =  334.18
     Prob > chi2 =    0.0000
```

Card (1995): OLS

$$Log(wage) = \alpha_1 educ + \alpha_2 black \cdot educ + z_1\delta_1 + u_1$$

reg lwage educ blackeduc black exper expersq south ///smsa reg662-reg669 smsa66

lwage	Coef.	Std. Err.	t	P>|t|
educ	.0707788	.0037548	18.85	0.000
blackeduc	.0178595	.006271	2.85	0.004
black	-.4191076	.0794021	-5.28	0.000
exper	.0821556	.0066828	12.29	0.000
expersq	-.0021349	.0003207	-6.66	0.000
south	-.1441927	.0259827	-5.55	0.000
smsa	.1340694	.0200931	6.67	0.000
reg662	.0988865	.0358663	2.76	0.006

2SLS (nearc4: grew up near a four-year college)

ivregress 2sls lwage black exper expersq south smsa ///
reg662-reg669 smsa66 (educ blackeduc = nearc4 ///
blacknearc4), vce(robust)

lwage	Coef.	Robust Std. Err.	z	P>|z|
educ	.1273557	.0560034	2.27	0.023
blackeduc	.0109036	.0398149	0.27	0.784
black	-.282765	.4997958	-0.57	0.572
exper	.1059116	.0248758	4.26	0.000
expersq	-.0022406	.0004888	-4.58	0.000
south	-.1424762	.0298096	-4.78	0.000
smsa	.1111555	.0309714	3.59	0.000
reg662	.1021697	.0365977	2.79	0.005

CF for Two Endogeneus Variables

```
quietly reg educ black exper expersq south smsa ///
reg662-reg669 smsa66 nearc4 blacknearc4, vce(boot,
reps(400) seed(7) nodots)

predict vhat1, residuals

quietly reg blackeduc black exper expersq south smsa ///
reg662-reg669 smsa66 nearc4 blacknearc4, vce(boot, ///
reps(400) seed(7) nodots)

predict vhat2, residuals

reg lwage educ blackeduc black vhat1 vhat2 exper ///
expersq south smsa reg662-reg669 smsa66, vce(boot, ///
reps(400) seed(7) nodots)
```

Endogeneity Test

lwage	Observed Coef.	Bootstrap Std. Err.	z	P>|z|
educ	.1273557	.0523323	2.43	0.015
blackeduc	.0109036	.0370742	0.29	0.769
black	-.282765	.4640906	-0.61	0.542
vhat1	-.0568274	.0525634	-1.08	0.280
vhat2	.0070106	.0379315	0.18	0.853

test (vhat1=0) (vhat2=0)

```
 ( 1)  vhat1 = 0
 ( 2)  vhat2 = 0

           chi2(  2) =    1.17
         Prob > chi2 =    0.5563
```

First Stage

reg educ nearc4 blacknearc4 black exper expersq south smsa ///
reg662-reg669 smsa66 , vce(boot, reps(400) seed(7) nodots)

educ	Observed Coef.	Bootstrap Std. Err.	z	P>|z|
nearc4	.3191761	.0982935	3.25	0.001
blacknearc4	.0029741	.1655514	0.02	0.986

reg blackeduc nearc4 blacknearc4 black exper expersq south smsa ///
reg662-reg669 smsa66, vce(boot, reps(400) seed(7) nodots)

blackeduc	Observed Coef.	Bootstrap Std. Err.	z	P>|z|
nearc4	-.0908895	.0321739	-2.82	0.005
blacknearc4	.874705	.1763268	4.96	0.000