

# PLSC 503 – Spring 2017

## Simultaneity and Endogeneity

March 22, 2017

Consider:

$$Y_1 = \mathbf{X}_1\beta_1 + \gamma_1 Y_2 + \mathbf{u}_1$$

$$Y_2 = \mathbf{X}_2\beta_2 + \gamma_2 Y_1 + \mathbf{u}_2$$

Rewrite:

$$\begin{aligned} Y_1 &= \mathbf{X}_1\beta_1 + \gamma_1[\mathbf{X}_2\beta_2 + \gamma_2 Y_1 + \mathbf{u}_2] + \mathbf{u}_1 \\ &= \mathbf{X}_1\beta_1 + \gamma_1(\mathbf{X}_2\beta_2) + \gamma_1\gamma_2 Y_1 + \gamma_1\mathbf{u}_2 + \mathbf{u}_1 \\ Y_1 - \gamma_1\gamma_2 Y_1 &= \mathbf{X}_1\beta_1 + \gamma_1(\mathbf{X}_2\beta_2) + \gamma_1\mathbf{u}_2 + \mathbf{u}_1 \\ (1 - \gamma_1\gamma_2)Y_1 &= \mathbf{X}_1\beta_1 + \gamma_1(\mathbf{X}_2\beta_2) + \gamma_1\mathbf{u}_2 + \mathbf{u}_1 \\ Y_1 &= \mathbf{X}_1 \left( \frac{1}{1 - \gamma_1\gamma_2} \beta_1 \right) + \mathbf{X}_2 \left( \frac{\gamma_1}{1 - \gamma_1\gamma_2} \beta_2 \right) + \left( \frac{\gamma_1\mathbf{u}_2 + \mathbf{u}_1}{1 - \gamma_1\gamma_2} \right) \\ &= \Delta_1\mathbf{X}_1 + \Delta_2\mathbf{X}_2 + \mathbf{e} \end{aligned}$$

$$Y_1 = \mathbf{x}_1 \left( \frac{1}{1 - \gamma_1 \gamma_2} \beta_1 \right) + \mathbf{x}_2 \left( \frac{\gamma_1}{1 - \gamma_1 \gamma_2} \beta_2 \right) + \left( \frac{\gamma_1 \mathbf{u}_2 + \mathbf{u}_1}{1 - \gamma_1 \gamma_2} \right)$$

means

$$\frac{\partial Y_1}{\partial X_\ell} = \frac{\beta_\ell}{1 - \gamma_1 \gamma_2}.$$

But

$$\hat{\Delta}_1 \neq \hat{\beta}_1.$$

For (e.g.)

$$Y_2 = \mathbf{X}_2\beta_2 + \gamma_2 Y_1 + \mathbf{u}_2$$

we have:

$$E(Y_2, \mathbf{u}) = \frac{\gamma_2}{1 - \gamma_1\gamma_2} \sigma_{\mathbf{u}}^2$$

- OLS
- Lagged Variables
- Two-Stage Least Squares (2SLS)
- Systems of Equations / 3SLS / etc.

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

has

$$\hat{\boldsymbol{\beta}}_{OLS} = \boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}.$$

Suppose  $\text{Cov}(\mathbf{X}, \mathbf{u}) \neq \mathbf{0}$ , but we have  $\mathbf{Z}$  with

- $\text{Cov}(\mathbf{Z}, \mathbf{X}) \neq \mathbf{0}$  and
- $\text{Cov}(\mathbf{Z}, \mathbf{u}) = \mathbf{0}$ .

Then

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{IV} &= (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Y} \\ &= (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'(\mathbf{X}\boldsymbol{\beta} + \mathbf{u}) \\ &= \boldsymbol{\beta} + (\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{u}\end{aligned}$$

is consistent.

- Regress endogenous  $\mathbf{X}$ s variables on  $\{\mathbf{Z}, \mathbf{X}\}$
- Generate  $\hat{\mathbf{X}}$ s
- Regress  $Y$  on  $\hat{\mathbf{X}}$  to get  $\beta_{2SLS}$ .
- Adjust standard error estimates

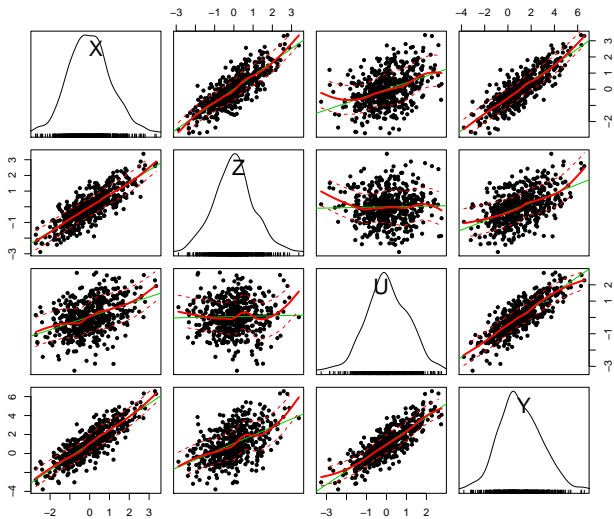
```
library(MASS)
library(sem)
library(car)

seed<-1337
set.seed(seed)

mu<-c(0,0,0) # <== X, Z, U
Sigma<-matrix(c(1,0.8,0.4,0.8,1,0,0.4,0,1),nrow=3,byrow=TRUE) #
Vars<- mvrnorm(500,mu,Sigma)
colnames(Vars)<-c("X","Z","U")
Vars<-data.frame(Vars)

Vars$Y<- 1 + Vars$X + Vars$U
```





```
> OLS<- lm(Y~X,data=Vars)
> summary(OLS)
```

Call:

```
lm(formula = Y ~ X, data = Vars)
```

Residuals:

Min	1Q	Median	3Q	Max
-3.3809	-0.6058	-0.0102	0.6320	2.9470

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.04770	0.04209	24.89	<2e-16 ***
X	1.40254	0.04005	35.02	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9413 on 498 degrees of freedom

Multiple R-squared: 0.7112, Adjusted R-squared: 0.7106

F-statistic: 1226 on 1 and 498 DF, p-value: < 2.2e-16

```
> TSLS<-tsls(Y~I(X),data=Vars,instruments=~Z)
> summary(TSLS)
```

2SLS Estimates

Model Formula:  $Y \sim I(X)$

Instruments:  $\sim Z$

Residuals:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-3.29300	-0.68210	-0.06139	0.00000	0.76270	2.70300

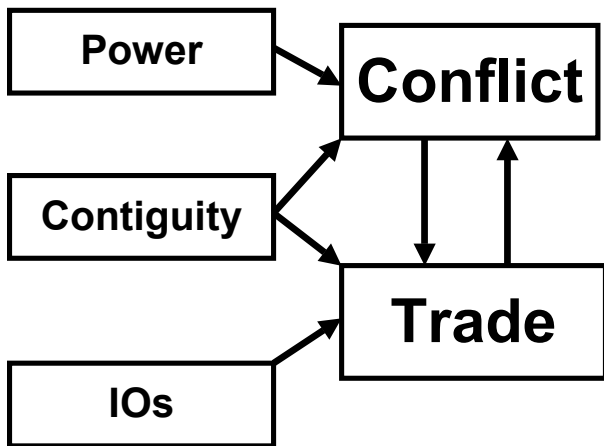
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.0491828	0.0456017	23.00754	< 2.22e-16 ***
I(X)	1.0302012	0.0536909	19.18763	< 2.22e-16 ***

---

Signif. codes: 0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 1

Residual standard error: 1.0196738 on 498 degrees of freedom

## IV: A (Toy) Example



```
> summary(IRData)
```

dyadid	logdisputes	logtrade	I0s
Min. : 2020	Min. : -0.6931	Min. : -0.6931	Min. : 4.579
1st Qu.: 135155	1st Qu.: -0.6931	1st Qu.: 2.4079	1st Qu.: 19.500
Median : 220484	Median : -0.6931	Median : 5.5786	Median : 27.704
Mean : 275526	Mean : -0.2627	Mean : 4.6518	Mean : 30.891
3rd Qu.: 385710	3rd Qu.: 0.0000	3rd Qu.: 7.1248	3rd Qu.: 39.289
Max. : 900920	Max. : 3.4965	Max. : 11.5037	Max. : 93.700

contiguity	capratio	GDPgrowth
Min. : 0.0000	Min. : 1.081	Min. : -9.0800
1st Qu.: 0.0000	1st Qu.: 4.849	1st Qu.: -0.2923
Median : 0.0000	Median : 26.577	Median : 0.8363
Mean : 0.3207	Mean : 196.310	Mean : 0.5097
3rd Qu.: 1.0000	3rd Qu.: 144.035	3rd Qu.: 1.7106
Max. : 1.0000	Max. : 7451.982	Max. : 7.0460

```
> OLSWar<-lm(logdisputes~logtrade+contiguity+capratio)
> summary(OLSWar)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.82840	-0.32644	-0.26860	-0.08972	3.45504

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-4.253e-01	6.020e-02	-7.065	3.46e-12	***
logtrade	8.558e-03	1.057e-02	0.809	0.4185	
contiguity	4.623e-01	7.124e-02	6.489	1.50e-10	***
capratio	-1.296e-04	6.467e-05	-2.003	0.0455	*
---					

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.853 on 813 degrees of freedom  
Multiple R-squared: 0.08301, Adjusted R-squared: 0.07962  
F-statistic: 24.53 on 3 and 813 DF, p-value: 3.345e-15

```

> library(sem)
> TwoSLSWar<-tsls(logdisputes~contiguity+capratio+I(logtrade),
  instruments=~contiguity+capratio+IOs)
> summary(TwoSLSWar)

```

## 2SLS Estimates

Model Formula: logdisputes ~ contiguity + capratio + I(logtrade)

Instruments: ~contiguity + capratio + IOs

## Residuals:

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
	-1.21e+00	-5.24e-01	-2.26e-01	-7.44e-17	-2.10e-02	3.65e+00

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-0.1515180	8.562e-02	-1.770	7.717e-02
contiguity	0.6263774	8.111e-02	7.722	3.353e-14
capratio	-0.0002664	7.252e-05	-3.674	2.543e-04
I(logtrade)	-0.0558374	1.769e-02	-3.157	1.652e-03

Residual standard error: 0.8723 on 813 degrees of freedom

```
> ITrade<-lm(logtrade~contiguity+IOs+capratio)
> summary(ITrade)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-6.0385	-1.7666	0.4139	1.6154	7.6029

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	0.7319793	0.1912570	3.827	0.000140	***
contiguity	1.3386037	0.1816041	7.371	4.17e-13	***
IOs	0.1218373	0.0055313	22.027	< 2e-16	***
capratio	-0.0013913	0.0001626	-8.555	< 2e-16	***
---					

Signif. codes: 0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 1

Residual standard error: 2.239 on 813 degrees of freedom  
 Multiple R-squared: 0.5535, Adjusted R-squared: 0.5519  
 F-statistic: 335.9 on 3 and 813 DF, p-value: < 2.2e-16



```
> IVWarByHand<-lm(logdisputes~capratio+contiguity
  +(ITrade$fitted.values))
> summary(IVWarByHand)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.0055	-0.3618	-0.2782	-0.0492	3.5301

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-1.515e-01	8.323e-02	-1.821	0.069050	.
capratio	-2.664e-04	7.049e-05	-3.780	0.000168	***
contiguity	6.264e-01	7.884e-02	7.944	6.49e-15	***
ITrade\$fitted.values	-5.584e-02	1.719e-02	-3.248	0.001210	**
---					

Signif. codes: 0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 1

Residual standard error: 0.8479 on 813 degrees of freedom

Multiple R-squared: 0.09402, Adjusted R-squared: 0.09068

F-statistic: 28.12 on 3 and 813 DF, p-value: < 2.2e-16

## Weak Instruments

```
> OLSTrade<-lm(logtrade~logdisputes+contiguity+IOs)
> summary(OLSTrade)
```

Residuals:

Min	1Q	Median	3Q	Max
-6.2467	-2.2067	0.4275	1.6659	6.1264

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.191111	0.182875	1.045	0.296
logdisputes	0.408116	0.095067	4.293	1.98e-05 ***
contiguity	1.357557	0.193109	7.030	4.38e-12 ***
IOs	0.133778	0.005614	23.831	< 2e-16 ***

---

Signif. codes: 0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 1

Residual standard error: 2.312 on 813 degrees of freedom

Multiple R-squared: 0.5241, Adjusted R-squared: 0.5223

F-statistic: 298.4 on 3 and 813 DF, p-value: < 2.2e-16

```
> TwoSLSTrade<-tsls(logtrade~contiguity+IOs+I(logdisputes),
  instruments=~contiguity+capratio+IOs)
> summary(TwoSLSTrade)
```

## 2SLS Estimates

Model Formula: logtrade ~ contiguity + IOs + I(logdisputes)

Instruments: ~contiguity + capratio + IOs

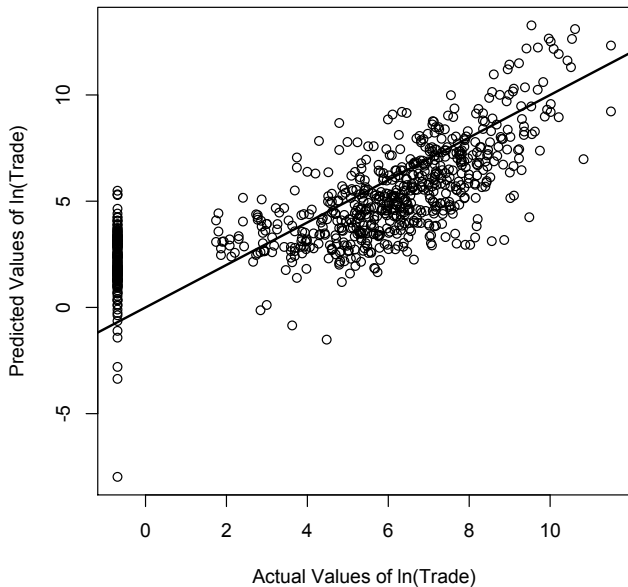
## Residuals:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-2.57e+01	-1.46e+00	1.36e+00	2.84e-14	4.00e+00	1.09e+01

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	2.150	0.85122	2.526	1.173e-02
contiguity	-2.728	1.52615	-1.787	7.427e-02
IOs	0.172	0.02045	8.408	2.220e-16
I(logdisputes)	7.371	2.45198	3.006	2.727e-03

Residual standard error: 6.3721 on 813 degrees of freedom

# Pretty Good Instrument (Trade)



# Crappy Instrument (War)

