

Assignment 8

Michael Cai

April 19, 2016

1. Does putting criminals in jail reduce crime? Suppose you regressed crime rates on incarceration rates and some control variables for economic conditions, demographics, etc.

$$CRIME_i = \alpha_0 + \alpha_1 INCARC_i + CONTROLS + u_i$$

(a) There might be simultaneous causality bias because higher incarceration rates affect crime rates (since there might be fewer criminals on the street); however, higher crime rates may also affect incarceration rates as well (more crime rates means more criminals, which may mean more people that need to be put in jail).

(b) The rank/relevance condition states that $Cov(Z, X) \neq 0$ or more broadly that none of the variables $\{X..., W..., 1\}$ are perfectly collinear, where X represents all of the endogenous variables, and W represents all of the exogenous variables.

The order/exogeneity condition states that $Cov(Z, u) = 0$ or intuitively that the the instrumental variables are actually exogenous.

Lawsuits aimed at reducing prison overcrowding might be a valid instrumental variable for incarceration rates because:

(1) $Cov(Z, X) \neq 0$

Lawsuits aimed at reducing prison overcrowding is correlated with incarceration rates (Think higher incarceration rates, more prisoners, more lawsuits aimed at reducing prison overcrowding).

(2) $Cov(Z, u) = 0$

The u in this equation likely accounts for the more intangible measures of the locations in question such as the culture of the neighborhoods, attitudes toward crime, etc.

We can expect that the lawsuits aimed at reducing prison overcrowding would not be in any way related to the local proclivities for illegal activity.

Thus, the two IV conditions should hold.

(c) (1) $Cov(Z, X) \neq 0$

For a similar reason, the number of lawyers is correlated with the number of lawsuits (since the more lawsuits you have the more lawyers you would need to handle the suits) and is directly correlated with the incarceration rates for the same reason.

(2) $Cov(Z, u) = 0$

Again for a similar reason, the number of lawyers is not in any way related to the local proclivities for crime.

2. Consider the IV regression model: $Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i$ when X_i is correlated with u_i and Z_i is an instrument. Suppose that the first three assumptions for TSLS are satisfied. Which TSLS assumption is not satisfied when:

(a) If Z_i is independent of (Y_i, X_i, W_i) then we know that the $Cov(Z, X) = 0$ and thus the relevance/rank condition is not satisfied.

(b) If $Z_i = W_i$ then we know that Z_i and W_i are perfectly collinear, which means that the rank/relevance condition is not satisfied.

(c) If $W_i = 1$ for all i then W_i ceases to be a random variable and is absorbed into the constant β_0 . Thus, W_i and 1 are perfectly collinear and the rank/relevance condition is not satisfied.

(d) If $Z_i = X_i$ and we know X_i to be endogenous then we know that $Cov(X_i, u) \neq 0 \implies Cov(Z_i, u) \neq 0$, which breaks the exogeneity/order condition.

3. Consider the regression model: $Y_i = \beta_0 + \beta_1 X_i^* + u_i$ where X_i^* is exogenous (further instructions...)

(a) $Z_i = X_i^* + v_i$

$$Cov(Z_i, X_i) = Cov(X_i^* + v_i, X_i^* + \epsilon_i)$$

Because we know that X_i^* is uncorrelated with v_i and ϵ_i this $Cov = Var(X_i^*)$, which we know cannot equal 0.

(b) $Z_i = X_i^* + v_i$

$$Cov(Z_i, u_i) = Cov(X_i^* + v_i, u_i)$$

We assume that X_i^* is exogenous and thus cannot be correlated with u_i and we also assume that the measurement error v_i is uncorrelated with u_i .

Thus $Cov(Z_i, u_i) = 0$.

(c) The estimator of β_1 that I propose is the standard β^{IV} estimator which equals $\frac{S_{Z,Y}}{S_{Z,X}}$ which $\rightarrow_p \frac{Cov(Z,Y)}{Cov(Z,X)}$.

If we expand this out:

$$\frac{Cov(Z,Y)}{Cov(Z,X)} = \frac{Cov(X_i^* + v_i, \beta_0 + \beta_1 X_i^* + u_i)}{Cov(X_i^* + v_i, X_i^* + \epsilon_i)}$$

You get a bunch of terms that cancel out with the exception of:

$$= \frac{Cov(X_i^*, \beta_1 X_i^*)}{Cov(X_i^*, X_i^*)}$$

$$= \beta_1$$

Thus the instrumental variables estimator is a consistent estimator of β_1 .

4. In the following question you're going to replicate some results from Card (1995).

(a) **Sorry Ilari/Laurent. Didn't have enough time to finish this one.