IV Roundup

EC420 MSU

Justin Kirkpatrick

Last updated March 30, 2021

Today

Review Main Concepts of IV:

- Why?
- How?
  - Conceptual
  - Three requirements for a valid IV
- What?
  - Local Average Treatment Effect
  - Compliers, always-takers, never-takers, and defiers
- Examples
  - Card (1995)
  - Stevenson (2018)

Why?

$$Y = \beta_0 + \beta_1 D + \beta_2 X + u$$

We use an instrument when we have an endogenous variable of interest

- Something in u is correlated with D
  - D is our variable of interest
  - X is our statistical control(s)
  - $E[u|D, X] \neq 0$

This may be for one of many reasons:

Omitted variable

- E.g. $educ$ and $ability$

Selection into treatment

- $(Y_{0i}, Y_{1i}) \not\perp D_i$
- Who gets treated is determined by their potential outcomes
- E.g. $1(drop-out-of-college)$ and $1(Zuckerberg)$

Measurement error

- When D is not correctly measured
- We did not talk about this in class, but it belongs here

Conceptually, an instrument is something that causes variation in our variable of interest

But is itself exogenous, so the variation it causes can be "as good as random"

We are, essentially, borrowing the exogeneity of the instrument to estimate the effect of the endogenous variable on Y in:

$$Y = \beta_0 + \beta_1 D + \beta_2 x_1 + u$$

Three requirements for a valid IV, Z

1. *Relevant First Stage*: Z causes D
   - F-stat >15 or "weak instrument" problem (biased)
2. *Independence assumption*: Z is as good as randomly assigned
3. *Exclusion Restriction*: Z only affects Y through D, and not directly (or through $\boldsymbol{u}$)

Mechanically, we use 2SLS

Two stage least squares

$$Y = \beta_0 + \beta_1 D + \beta_2 x_1 + u$$

$$D = \gamma_0 + \gamma_1 Z + \gamma_2 x_1 + v$$

$$\hat{D} = \hat{\gamma}_0 + \hat{\gamma}_1 Z + \hat{\gamma}_2 x_1$$

$$Y = \beta_0 + \beta_1 \hat{D} + \beta_2 x_2 + \tilde{u}$$

And $\hat{\beta}_1$ is unbiased.

What are we estimating?

The estimate of $\hat{\beta}_1^{IV}$, the coefficient on D, is the **LATE**

- Local Average Treatment Effect
- The ATE for **compliers**

The four types of observations affected by the instrument

- Always-takers
- Compliers
- Never-takers
- Defiers

The LATE is the ATE for *people who would take the treatment if the instrument applies to them, but would not take the treatment otherwise.*

Whether or not this is the measure you want is application-specific!

Examples

I'm using two examples from Chapter 7 of Causal Inference: The Mixtape by Scott Cunningham. The chapter is free online and elaborates on what we've learned thus far. The book is more advanced than *Mastering Metrics* but should be mostly accessible with the knowledge from this class. Plus, it has examples.

College in the County (Card 1995)

David Card was probably the primary mover of the *credibility revolution* which sought to bring empirical testing to theory.

In this paper, Card examined the returns to schooling (as we have used as an example many times). He wanted to see how wages changed when someone attended college, but understood that people are not randomly selected into going to college (for reasons we are all familiar with by now).

- NLS Young Men Cohort of the National Longitudinal Survey
- 5,525 men 14-24 with follow-up every 5 years.
- Baseline survey asked about local labor markets
- Also asked if subject lived in the same county as a 4-year college
- **Instrument, Z, was "presence of 4-year college"**

Does this meet our 3 requirements?

Examples

```
library(AER) #<-- Has ivreg() function
library(haven)

card <- read_dta("https://raw.github.com/scunning1975/mixtape/master/card.dta")

#OLS
ols_reg <- lm(lwage ~ educ + exper + black + south + married + smsa, card)

#2SLS
iv_reg = ivreg(lwage ~ educ + exper + black + south + married + smsa | exper + black + south + married + smsa + n
```

Examples

```r
coeftest(ols_reg, vcovHC(ols_reg, 'HC1'))
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)  5.0633165  0.0661786  76.5099 < 2.2e-16 ***
## educ         0.0711729  0.0036137  19.6951 < 2.2e-16 ***
## exper        0.0341518  0.0022668  15.0663 < 2.2e-16 ***
## black       -0.1660274  0.0173977  -9.5430 < 2.2e-16 ***
## south       -0.1315518  0.0152007  -8.6543 < 2.2e-16 ***
## married     -0.0358707  0.0035727 -10.0401 < 2.2e-16 ***
## smsa         0.1757871  0.0150994  11.6420 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Examples

```r
summary(iv_reg, diagnostics=TRUE, vcov = vcovHC, 'HC1')
```

```
##
## Call:
## ivreg(formula = lwage ~ educ + exper + black + south + married +
##     smsa | exper + black + south + married + smsa + nearc4, data = card)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.81301 -0.23805  0.01766  0.24727  1.32278
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.162476   0.837038   4.973 6.60e-07 ***
## educ         0.124164   0.049285   2.519  0.01176 *
## exper        0.055588   0.019918   2.791  0.00526 **
## black       -0.115686   0.049692  -2.328  0.01991 *
## south       -0.113165   0.023011  -4.918 8.75e-07 ***
## married     -0.031975   0.005086  -6.287 3.24e-10 ***
## smsa         0.147707   0.030398   4.859 1.18e-06 ***
##
## Diagnostic tests:
##                   df1  df2 statistic  p-value
## Weak instruments    1 2996    16.462 5.09e-05 ***
## Wu-Hausman          1 2995     1.271     0.26
## Sargan              0   NA        NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3843 on Inf degrees of freedom
## Multiple R-Squared: 0.2513,    Adjusted R-squared: 0.2498
## Wald test: 888.4 on 6 DF,  p-value: < 2.2e-16
```

Judge Fixed Effects (an IV strategy)

Researchers are very curious about the effect of prosecution for minor charges in the justice system (or the effect of lieniency on offenders).

- This requires an "incarceration" or "lieniency" treatment D that is never going to be exogenous (let's discuss).
- The "Judge Fixed Effects" literature dates back to 1933. While an offender's bail conditions are definitely endogeneous, judges have discretion and in general have different independent tendencies. **Some judges are lienient. Some are the opposite.**
- When judges are **randomly assigned** to cases, and if we believe they have some independent tendency towards lieniency, then *some* of the variation in sentence is as-good-as-randomly assigned!
    - The variation that can be explained by the judge's overall tendency for lieniency is our $\hat{D}$

Stevenson (2018) Effect of Pre-trial Detention and Cash Bail

- Looks at effect of pre-trial detention (cash bail) on the likelihood of receiving a conviction
- A little different from the question before, but same setting
- And the same problem: those who are detained pre-trial (high cash bail) are different in many ways that might also predict/affect being found guilty
- Instrument(s) $\mathbf{Z}$ are the judge fixed effects - the judge tendency to set high cash bail.

Does this meet our 3 requirements?

What exactly are we estimating?

So, how do we implement this?

- First, we generate a dummy variable for each judge (leaving one out, of course, for multicolinearity)
- Next, we regress whether or not the offender is held in jail, $\boldsymbol{jail3}$ on the judge fixed effects and other statistical controls
- Finally, using the first stage model prediction of $\widehat{\boldsymbol{jail3}}$, we estimate the effect of $\boldsymbol{jail3}$ on $\boldsymbol{guilty}$, an indicator for receiving a conviction

```r
library(haven) # to read .dta files

judge = read_dta("https://raw.github.com/scunning1975/mixtape/master/judge_fe.dta")

# There are 8 judges. We leave Judge #8 out (for multicolinearity):
ols_formula = as.formula(guilt ~ jail3 + day + day2 + bailDate +
                                  t1 + t2 + t3 + t4 + t5)

iv_formula = as.formula(guilt ~ jail3 + day + day2 + bailDate +
                                t1 + t2 + t3 + t4 + t5 |
                                day + day2 + bailDate +
                                t1 + t2 + t3 + t4 + t5 +
                                judge_pre_1 + judge_pre_2 +
                                judge_pre_3 + judge_pre_4 + judge_pre_5 +
                                judge_pre_6 + judge_pre_7)
```

Examples

```r
ols_reg_JFE = lm(ols_formula, data = judge)
coeftest(ols_reg_JFE, vcov = vcovHC(ols_reg_JFE, 'HC1'))
```

```
##
## t test of coefficients:
##
##                Estimate  Std. Error t value Pr(>|t|)
## (Intercept) -4.1346e-01  1.2779e-01 -3.2355 0.001215 **
## jail3       -7.3042e-04  1.7573e-03 -0.4157 0.677659
## day         -6.2157e-05  2.5086e-05 -2.4777 0.013222 *
## day2         8.8970e-08  6.0051e-08  1.4816 0.138451
## bailDate     6.2899e-05  8.2088e-06  7.6624 1.83e-14 ***
## t1           5.5604e-03  1.5822e-02  0.3514 0.725264
## t2          -1.6678e-02  1.2410e-02 -1.3438 0.178998
## t3           4.9496e-04  1.0527e-02  0.0470 0.962499
## t4           1.2338e-02  6.8075e-03  1.8124 0.069932 .
## t5          -5.0521e-03  4.4226e-03 -1.1423 0.253322
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Examples

```r
iv_reg_JFE = ivreg(iv_formula, data = judge)
summary(iv_reg_JFE, diagnostics = T) # not HC robust errors
```

```
##
## Call:
## ivreg(formula = iv_formula, data = judge)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6564 -0.4798 -0.3752  0.5158  0.6248
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.162e-01  1.293e-01  -3.220  0.00128 **
## jail3        1.510e-01  6.517e-02   2.317  0.02049 *
## day         -6.573e-05  2.540e-05  -2.588  0.00965 **
## day2         8.138e-08  6.072e-08   1.340  0.18021
## bailDate     5.868e-05  8.498e-06   6.905 5.01e-12 ***
## t1           5.049e-03  1.600e-02   0.315  0.75238
## t2          -1.512e-02  1.257e-02  -1.203  0.22891
## t3           6.938e-03  1.100e-02   0.631  0.52818
## t4           1.749e-02  7.233e-03   2.418  0.01560 *
## t5           4.831e-04  5.067e-03   0.095  0.92405
##
## Diagnostic tests:
##                     df1    df2 statistic p-value
## Weak instruments      7 331955    35.312  <2e-16 ***
## Wu-Hausman            1 331960     5.548  0.0185 *
## Sargan                6     NA     7.510  0.2762
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5036 on 331961 degrees of freedom
## Multiple R-Squared: -0.01483,	Adjusted R-squared: -0.01486
## Wald test:   271 on 9 and 331961 DF,  p-value: < 2.2e-16
```

MICHIGAN STATE UNIVERSITY

A small problem with the Judge IV instruments

If we look at the first stage, we see that some of our instruments are, on their own, weak:

```
summary(lm(jail3 ~ judge_pre_1 + judge_pre_2 +
                  judge_pre_3 + judge_pre_4 + judge_pre_5 +
                  judge_pre_6 + judge_pre_7, data = judge))
```

```
##
## Call:
## lm(formula = jail3 ~ judge_pre_1 + judge_pre_2 + judge_pre_3 +
##     judge_pre_4 + judge_pre_5 + judge_pre_6 + judge_pre_7, data = judge)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4320 -0.4134 -0.3954  0.5866  0.6046
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.398146   0.002073 192.035  < 2e-16 ***
## judge_pre_1  0.004261   0.003942   1.081    0.280
## judge_pre_2  0.033810   0.004774   7.082 1.42e-12 ***
## judge_pre_3  0.019860   0.002959   6.711 1.94e-11 ***
## judge_pre_4 -0.002741   0.002928  -0.936    0.349
## judge_pre_5  0.015093   0.003389   4.454 8.43e-06 ***
## judge_pre_6  0.033792   0.002949  11.459  < 2e-16 ***
## judge_pre_7  0.015260   0.003183   4.794 1.64e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4919 on 331963 degrees of freedom
## Multiple R-squared:  0.0007151,    Adjusted R-squared:  0.000694
## F-statistic: 33.94 on 7 and 331963 DF,  p-value: < 2.2e-16
```

Questions

Any questions on the topic of IV?

Your homework will require you to use ivreg. How you tell R which variables are endogenous and which are exogenous and which are the instrument is a little tricky. Write down your endogenous, exogenous, and instruments first, then use `?ivreg` to see where you put them.