

TA Session 5

Spring 2020 TAs

May 10, 2020

1. Using instrumental variables

Instrumental variables can help us handle omitted variable bias and measurement error. How? By running through an instrument, we parse out and evaluate only the variation our treatment is responsible for.

1.1 Omitted variable bias

With omitted variable bias (OVB), we have:

$$\hat{\tau} = \tau + \beta \frac{Cov(D_i, X_i)}{Var(D_i)} \neq \tau$$

But, using a valid instrument, the second-stage IV estimate is:

$$\hat{\tau}^{2SLS} = \tau + \beta \frac{Cov(\hat{D}_i, X_i)}{Var(\hat{D}_i)}$$

By the exclusion restriction, the second term is zero (remember the numerator *is* our exclusion restriction). What else does this equation tell us? We **must** have variation in D , or else our estimator cannot exist. Anyway, if the second term is zero:

$$\hat{\tau}^{2SLS} = \tau$$

What does this mean? Our instrument is uncorrelated with the omitted variables (by the exclusion restriction) and allows us to estimate our coefficient of interest without any bias.

1.2 Measurement error

We can use an instrumental variable to solve measurement error. Basically, what happens if we don't *perfectly* observe D or Y ? Some measurement error in our outcome variable is okay. If we run the following regression, $\tilde{Y}_i = \alpha + \tau D_i + e_i$, and obtain our estimate of $\hat{\tau}$, we can show using algebra, covariance rules, and our assumptions that:

$$\hat{\tau} = \frac{Cov(\tilde{Y}_i, D_i)}{Var(D_i)} = \tau$$

However, measurement error, both classical and non-classical, in D_i is bad. Under the classical case, if we run our regression, $Y_i = \alpha + \tau \tilde{D}_i + e_i$, we find that:

$$\hat{\tau} = \tau \frac{Var(D_i)}{Var(D_i) + Var(\gamma_i)}$$

This second term is *attenuation bias*, which biases our estimate towards zero.

Fortunately, we can instrument for \tilde{D}_i with Z_i and assume measurement error is uncorrelated with treatment, the measurement error in Z_i is uncorrelated with error in \tilde{D}_i , and that measurement error is uncorrelated with original error. These assumptions satisfy our two IV assumptions of relevance and the exclusion restriction. Again, by some algebra, variance rules, and assumptions, we find that:

$$\hat{\tau}^{IV} = \frac{Cov(Y_i, Z_i)}{Cov(\tilde{D}_i, Z_i)} = \tau$$

Non-classical measurement error, where we relax our assumption of orthogonality, also produces bias, although it does not necessarily attenuate $\hat{\tau}$. But this bias could actually change the sign of the estimate relative to τ . As before, we can show that an IV helps us deal with this bias.

2. Heterogeneous treatment effects

Under heterogeneous treatment effects, we need to ask: whose treatment effect are we uncovering with IV? The first-stage of our causal chain is the impact of the instrument on the treatment status. We can define observed treatment status as:

$$D_i = D_i(Z_i = 0) + (D_i(Z_i = 1) - D_i(Z_i = 0))Z_i = \alpha + \gamma_i Z_i + v_i$$

where

$$\gamma_i = (D_i(1) - D_i(0))$$

is the individual-specific causal effect of the instrument on treatment status. We need four new-ish assumptions for IV to identify the causal effect:

1. First stage (relevance):

$$E[D_i|Z_i = 1] \neq E[D_i|Z_i = 0]$$

2. Independence (as good as random):

$$Y_i(D_i, Z_i), D_i(1), D_i(0) \perp Z_i$$

3. Exclusion restriction:

$$Y_i(Z_i = 1, D_i) = Y_i(Z_i = 0, D_i)$$

4. Monotonicity (NEW!):

$$D_i(Z_i = 1) - D_i(Z_i = 0) \geq 0$$

for all i

Under these assumptions, the IV estimator identifies the effect of treatment on compliers, those who switched treatment status as a result of the instrument. We also know this as the LATE!

$$\hat{\tau}^{IV} = E[Y_i(1) - Y_i(0)|D_i(1) > D_i(0)] = \tau^{LATE}$$

We can only tell a compelling story for our exclusion restriction, but we CAN test for first-stage relevance. How? By running our first stage. Generally, if you find an F stat of *less* than 20, you have a weak instrument, and bad things follow:

```
# where x is your treatment variable, z is your instrument,
# and controls are included in both stages of the regression

# first_stage <- lm(x ~ z + controls, data = YOUR DATA)

# read the F-stat on this, and see if there is a significant
# correlation between your x and z. If not, no dice

# there is a canned routine to do 2SLS: ivreg() from library(AER)
# doing this canned routine will give you correct standard errors
```

2.1 Monotonicity

In words, monotonicity means that all people who react to the instrument must change their treatment status in the same way. This rules out defiers. The instrument could have no effect on some people, but for those who are affected, they are all affected in the same direction (positive or negative, but not both). Without monotonicity, the IV estimator is the weighted average of the treatment effect for compliers and defiers, and there is no way to separate out the effect for each group. We can no longer estimate the LATE.

2.2 Non-compliance

As we've done before, we'll assume away defiers, and consider three groups:

1. Compliers:

$$D_i(1) > D_i(0)$$

2. Always-takers:

$$D_i(1) = D_i(0) = 1$$

3. Never-takers:

$$D_i(1) = D_i(0) = 0$$

The instrument does nothing for our stubborn always-takers and never-takers, so we don't know their treatment effects. This means we are estimating the LATE for compliers.

Non-compliance might manifest itself just among the treatment group, just among the controls, or both treatment and control. Regardless of the form, we can show that the IV estimator is the LATE. This is because our instrument is as good as random, but we need to scale by the proportion of compliers.

This yields the Wald estimator, or the treatment effect for compliers, which is:

$$\frac{\pi_c(E[Y_i(1)|C] - E[Y_i(0)|C])}{\pi_c} = E[Y_i(1) - Y_i(0)|C]$$

So, we have an as-good-as-random estimate. Now, we just need to figure out this complier proportion. The fraction of compliers is just:

$$\pi_c = \text{Prob}(D_i(1) > D_i(0)) = E[D_i|Z_i = 1] - E[D_i|Z_i = 0]$$

But who exactly *are* these LATE compliers? Sometimes it is worth doing a micro-evaluation of this subpopulation as well, since you'd like to know who might actually comply with your policy treatment should it go full-scale. In lecture, we talked about the probability that folks in your complier sample are men:

$$\begin{aligned} & \frac{\text{Prob}(\text{Male}_i(D_i = 1)|D_i(1) > D_i(0))}{\text{Prob}(\text{Male}_i = 1)} \\ &= \frac{E[D_i|Z_i = 1, \text{Male}_i(1) = 1] - E[D_i|Z_i = 0, \text{Male}_i(1) = 1]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]} \end{aligned}$$

In real person words, this is just that first-stage for the men divided by overall first stage for our IV! So basically, parsing out the IV for only this subsample - fun!

2.3 Multiple instruments

With heterogenous treatment effects, different instruments may affect different groups of people, meaning the LATE is not necessarily the same using one instrument versus another. Compare this to homogenous treatment effects, where all instruments should result in the same LATE. Ultimately, the IV estimator using 2SLS yields the weighted average of each instrument's IV estimator.

Remember that multiple instruments is a pipe dream most of the time. However, if you do find the opportunity where multiple instruments satisfy the exclusion and relevance restrictions, R has useful packages for you:

```
# ivreg from library(AER) again works here:  
# x being your treatment, and z's being your IVs  
  
# ivreg(y ~ x1 + x2 | z1 + z2 + z3, data=YOUR_DATA)
```

2.4 Non-binary treatment

With non-binary treatment (let's say different amounts of allocated cash endowments), there are many potential outcomes, and many causal effects. Using our assumptions from before (relevance, independence, exclusion, monotonicity), we can calculate the weighted average of the unit causal response. The unit causal response is the effect for compliers at a given treatment.

3. Summary

Homogeneous treatment effects:

$$\hat{\tau}^{IV} = \tau^{ATE}$$

Perfect compliance:

$$\hat{\tau}^{IV} = \tau^{ATE}$$

Heterogenous treatment effects:

$$\hat{\tau}^{IV} = \tau^{LATE}$$

Heterogenous treatment effects, multiple IVs:

$$\hat{\tau}^{IV} = \frac{1}{K} \sum_k w_k \tau_k^{LATE}$$

Non-binary treatment:

$$\hat{\tau}^{IV} = \sum_{s=1}^S w_s E[Y_i(s) - Y_i(s-1) | S_i(1) \geq s \geq S_i(0)]$$