Cencored regression, weak IV, and quantile regression

Tutorial 1

Stanislav Avdeev

Tutorials

- 7 TA sessions
- 6 TA sessions are about lecture material
- The last session is primarily about exam and remaining questions about the course material (TBA)
- Send me any questions you want to discuss before each TA session
- Use Canvas or send me an email
- Alternately, leave your questions anonymously here: https://www.menti.com/c6uyd9qan4 (I will update the link each week on Canvas)

Assignments

- Due date: 11:59pm on Sundays (the first assignment is an exception: 11:59am on Tuesday)
- Assignments are graded within a week from the deadline
- Solutions will not be shared so if you want to discuss a specific exercise, let me know before the TA session (you submit your solutions on Sunday, thus, we can discuss any questions on the following TA session on Tuesday)

Course objective

- The key objective of the course is **applying** microeconometric techniques rather than **deriving** econometric and statistical properties of estimators.
- In other words, there's way less of this:

$$\text{plim}\hat{\beta}_{OLS} = \beta + \text{plim}(\frac{1}{N}X'X)^{-1}\text{plim}\frac{1}{N}X'\epsilon = \beta + Q^{-1} \times \text{plim}\frac{1}{N}X'\epsilon$$

- And way more of this:

```r
library(fixest)

tb <- tibble(groups = sort(rep(1:10, 600)), time = rep(sort(rep(1:6, 100)), 10)) %>%
  mutate(Treated = I(groups > 5) * I(time > 3)) %>%
  mutate(Y = groups + time + Treated*5 + rnorm(6000))
m <- feols(Y ~ Treated | groups + time, data = tb)
```

If you would like to go deeper into the former, take Advanced Econometrics I and II next year

Weak instrument problem

- Weak instrument problem means that we probably shouldn't be using IV in small samples
- This also means that it's really important that $cov(X, Z)$ is not small
- If Z has only a trivial effect on X, then it's not *relevant* - even if it's truly exogenous, it does not matter because there's no variation in X we can isolate with it
- And our small-sample bias will be big

Weak instrument problem

- There are some rules of thumb for how strong an instrument must be to be counted as "not weak"
- A t-statistic above 3, or an F statistic from a joint test of the instruments that is 10 or above
- These rules of thumb aren't great - selecting a model on the basis of significance naturally biases your results
- What you really want is to know the *population* effect of Z on X - you want the F-statistic from *that* to be 10+. Of course we don't actually know that.

Weak instrument problem: simulation

- Let's look at the output of `feols()` using a simulated dataset

```
library(fabricatr)
set.seed(777)

df <- fabricate(
  N = 200,
  Y = rpois(N, lambda = 4),
  Z = rbinom(N, 1, prob = 0.4),
  X1 = Z * rbinom(N, 1, prob = 0.8),
  X2 = rnorm(N),
  G = sample(letters[1:4], N, replace = TRUE)
)
```

Weak instrument problem: simulation

```r
iv <- feols(Y ~ X2 | X1 ~ Z, data = df, se = 'hetero')
thef <- fitstat(iv, 'ivf', verbose = FALSE)$`ivf1::X1`$stat
iv
```

```
## TSLS estimation, Dep. Var.: Y, Endo.: X1, Instr.: Z
## Second stage: Dep. Var.: Y
## Observations: 200
## Standard-errors: Heteroskedasticity-robust
##             Estimate Std. Error  t value  Pr(>|t|)
## (Intercept) 3.616743   0.164506 21.98548 < 2.2e-16 ***
## fit_X1      1.066064   0.369839  2.88251 0.0043831 **
## X2          0.326317   0.153746  2.12244 0.0350497 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## RMSE: 1.93622   Adj. R2: 0.009568
## F-test (1st stage), X1: stat = 522.5, p < 2.2e-16 , on 1 and 197 DoF.
##             Wu-Hausman: stat =  11.1, p = 0.001029, on 1 and 196 DoF.
```

- 522.47 is way above 10! We're probably fine in this particular regression

Overidentification tests

- "Overidentification" just means we have more identifying conditions (validity assumptions) than we actually need. We only need one instrument, but we have two (or more)
- So we can compare what we get using each instrument individually
- If we assume that *at least one of them is valid*, and they both produce similar results, then that's evidence that *both* are valid

Overidentification tests: simulation

- We can do this using `diagnostics = TRUE` in `iv_robust` again

```
set.seed(1000)
# Create data where Z1 is valid and Z2 is invalid
df <- tibble(Z1 = rnorm(1000), Z2 = rnorm(1000)) %>%
  mutate(X = Z1 + Z2 + rnorm(1000)) %>%
  # True effect is 1
  mutate(Y = X + Z2 + rnorm(1000))

iv <- feols(Y ~ 1 | X ~ Z1 + Z2, data = df, se = 'hetero')
fitstat(iv, 'sargan')
```

```
## Sargan: stat = 267.8, p < 2.2e-16, on 1 DoF.
```

- That's a small p-value! We can reject that the results are similar for each IV, telling us that one is endogenous (although without seeing the actual data generating process we couldn't guess if it were $Z1$ or $Z2$)

Overidentification tests: simulation

- How different are they? What did the test see that it was comparing?

```
iv1 <- feols(Y ~ 1 | X ~ Z1, data = df)
iv2 <- feols(Y ~ 1 | X ~ Z2, data = df)
export_summs(iv1, iv2, statistics = c(N = 'nobs'))
```

	Model 1	Model 2
(Intercept)	-0.01	0.00
	(0.04)	(0.05)
fit_X	1.08 ***	1.92 ***
	(0.04)	(0.05)
N	1000	1000

*** p < 0.001; ** p < 0.01; * p < 0.05.

Notice the first model gives an accurate coefficient of 1