14.75, Recitation 1

Outline of the recitation

Linear Regression Model: Ordinary Least Squares

- Causality
- Bias
- Estimation
- Standard error

The Wald estimator - Two Stage Least Squares

Linear Regression Model

$$y_i = \beta_0 x_i + \epsilon_i$$

where

- y_i is the “outcome”, the “independent” variable. Example: country’s income
- x_i is a “dependent” variable - either the “variable of interest” or a “control” variable. Example: a dummy indicating democracy.
- ϵ_i is the error term

What is β_0?

Causality

β_0 is the impact of democracy on income.

Indeed, consider any country, for instance France.

- If France is a democracy, its income is $y_{France}(1) = \beta_0 + \epsilon_{France}$
- If it is not, its income is $y_{France}(0) = \epsilon_{France}$
- Thus, $\beta_0 = y_{France}(1) - y_{France}(0)$

Similarly, considering all n countries in our sample,

$$\mathrm{E}\left[y_i(1) - y_i(0)\right] = \beta_0 + \mathrm{E}\left[\epsilon_i\right] - \mathrm{E}\left[\epsilon_i\right] = \beta_0$$

Do we directly observe $\mathrm{E}\left[y_i(1) - y_i(0)\right]$?

NO! For any country, we either observe $y_i(1)$ OR $y_i(0)$ but not the two of them:

- if France is actually a democracy, we observe $y_{France}(1)$
- if it is not, we observe $y_{France}(0)$
- but in no state of the world do we observe both: indeed, we observe $y_{France}(1)$ and $y_{NorthKorea}(0)$.

What can we get from the data?

$$\mathrm{E}\left[y_i(1)|x_i = 1\right] - \mathrm{E}\left[y_i(0)|x_i = 0\right]$$

How does this relate to $\mathrm{E}\left[y_i(1) - y_i(0)\right]$?

Potential bias

$$\underbrace{\mathrm{E}\left[y_i(1)|x_i=1\right]-\mathrm{E}\left[y_i(0)|x_i=0\right]}_{Observed} =$$
$$\underbrace{\mathrm{E}\left[y_i(1)|x_i=1\right]-\mathrm{E}\left[y_i(0)|x_i=1\right]}_{Goal} +$$
$$\underbrace{\mathrm{E}\left[y_i(0)|x_i=1\right]-\mathrm{E}\left[y_i(0)|x_i=0\right]}_{Bias}$$

Goal:

$$\begin{aligned}\mathrm{E}\left[y_i(1)|x_i=1\right]-\mathrm{E}\left[y_i(0)|x_i=1\right] &= \beta_0+\mathrm{E}[\epsilon_i|x_i=1]-\mathrm{E}[\epsilon_i|x_i=1] \\ &= \beta_0.\end{aligned}$$

Bias:

$$\mathrm{E}\left[y_i(0)|x_i=1\right]-\mathrm{E}\left[y_i(0)|x_i=0\right]=\mathrm{E}[\epsilon_i|x_i=1]-\mathrm{E}[\epsilon_i|x_i=0].$$

Key assumption

What do we need to assume to derive the goal from the observed?

$$cov(\epsilon_i, x_i) = 0$$

which implies
$\mathrm{E}[\epsilon_i | x_i = 1] = \mathrm{E}[\epsilon_i | x_i = 0] = 0.$

Why might we think $cov(\epsilon_i, x_i) \neq 0$?

What if we fear that $cov(\epsilon_i, x_i) \neq 0$?

$$y_i = \beta_0 x_i + \epsilon_i$$

1. Control for things $\epsilon_i = w_i\gamma + \eta_i$ and regress

$$y_i = x_i\beta_0 + w_i\gamma + \eta_i$$

   where η is uncorrelated with x and w.
2. Find an Instrument and run Wald estimator (later today)
3. Difference-in-Difference Estimation, RDD (later this semester))

Back to our linear estimation

$y_i = \beta_0 x_i + \epsilon_i$: How do we estimate this?

"Ordinary Least Squares" method: find $\hat{\beta}$ that minimizes

$$\sum_{i=1}^{n} (y_i - x_i\beta)^2$$

What does it mean to do this?

The answer:

$$\widehat{\beta} = \left(\sum x_i^2\right)^{-1} \sum x_i \cdot y_i.$$

Unbiasedness of $\hat{\beta}$

Why is $\widehat{\beta} = \left(\sum x_i^2\right)^{-1} \sum x_i \cdot y_i$. a good estimator of β_0?

$$
\begin{aligned}
\widehat{\beta} &= \left(\sum x_i^2\right)^{-1} \sum x_i \cdot y_i \\
&= \left(\sum x_i^2\right)^{-1} \sum x_i \cdot (x_i\beta_0 + \epsilon_i) \\
&= \left(\sum x_i^2\right)^{-1} \sum x_i^2\beta_0 + \left(\sum x_i^2\right)^{-1} \sum x_i\epsilon_i \\
&= \beta_0 + \left(\sum x_i^2\right)^{-1} \sum x_i\epsilon_i.
\end{aligned}
$$

Then

$$
\mathrm{E}\left[\widehat{\beta}|x\right] = \beta_0 + \underbrace{\left(\sum x_i^2\right)^{-1} \sum x_i \mathrm{E}\left[\epsilon_i|x\right]}_{=0} = \beta_0
$$

Standard error of $\widehat{\beta}$

Reminder from class: definition of the standard error of $\widehat{\beta}$?
It is the standard deviation of our estimate $\widehat{\beta}$ - thought experiment, run the regression on different samples, and plot the resulting distribution of $\widehat{\beta}$'s.

How do we estimate it? Writing $var\left(\epsilon_i|x\right) = \sigma^2$,

$$\begin{aligned} var\left(\widehat{\beta}|x\right) &= \left(\sum x_i^2\right)^{-2} var\left(\sum x_i\epsilon_i|x\right) \\ &= \left(\sum x_i^2\right)^{-2} \sum x_i^2 var\left(\epsilon_i|x\right) \\ &= \left(\sum x_i^2\right)^{-2} \sum x_i^2\sigma^2 \\ &= \frac{\sigma^2}{\sum x_i^2} \end{aligned}$$

We get $\hat{var}\left(\widehat{\beta}|x\right)$, an estimate of $var\left(\widehat{\beta}|x\right)$ by finding an estimate for σ^2: $\hat{\sigma^2} = \frac{\epsilon^2}{n-1}$

Valid instruments

We fear that we have omitted important variables which also affect income y and are correlated with x.
Examples?

Fix we saw in class?
Use an instrument z

Conditions for z to be a valid instrument?

1. z must affect x.
2. Exclusion restriction: z can affect y only through its effect on x:
$$\mathrm{E}\left[\epsilon_i | z_i = 1\right] - \mathrm{E}\left[\epsilon_i | z_i = 0\right] = 0$$

The Wald estimator

$\beta = \frac{\text{Reduced Form}}{\text{First Stage}}$
with

$$\begin{aligned}
\text{Reduced Form} &= \mathrm{E}\left[y_i|z_i = 1\right] - \mathrm{E}\left[y_i|z_i = 0\right] \\
&= \beta\left[\mathrm{E}\left[x_i|z_i = 1\right] - \mathrm{E}\left[x_i|z_i = 0\right]\right] \\
&- \left[\mathrm{E}\left[\epsilon_i|z_i = 1\right] - \mathrm{E}\left[\epsilon_i|z_i = 0\right]\right] \\
&= \beta\left[\mathrm{E}\left[x_i|z_i = 1\right] - \mathrm{E}\left[x_i|z_i = 0\right]\right]
\end{aligned}$$

and

$$\text{First Stage} = \mathrm{E}\left[x_i|z_i = 1\right] - \mathrm{E}\left[x_i|z_i = 0\right]$$

MIT OpenCourseWare
http://ocw.mit.edu

14.75 Political Economy and Economic Development
Fall 2012

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.