*Addressing Few Econometric Problems

Marcelo Neri

Daniel Duque

Selectivity Bias

$$Y_i = \alpha + \beta S_i + X'_i\gamma + \varepsilon_i$$

- Selection bias is certainly one of the major problems incurred in estimating through OLS the impact of a particular policy.
- That is because almost always those who are subject to a policy are different from those who are not.
- For example, people who go to a hospital are probably sicker than those who are not, and people who study most are probably more disciplined than average.
- Fortunately, there are ways to specifically address this problem: Heckit, and Matching. Differences in Differences, the so-called "individual fixed-effects" approach, for panel data and Regression Discontinuity Design allows to approach causality.

Heckman adjustment (**Heckit**)

- Considering $Y_i = \alpha + \beta S_i + X'_i\gamma + \varepsilon_i$, but where not all Y_i is observed, but rather there is a selection bias, where:
- $P(Labor)_i = \delta_0 + \sum_k \delta_k z_{k,i} + \epsilon_i$ Where cor(ε, ϵ) = $\rho > 0$
- That is, an individual's labor supply generates a selection bias on mincerian equation estimates. That is particularly true for women, affecting the estimated gender gap.
- Heckman procedure adjust the bias of this correlation,adding to the mincerian equation the cdf of predicted P(Labor) .

Matching

- Propensity-score matching (PSM) uses an average of the outcomes of similar subjects who get the other treatment level to impute the missing potential outcome for each subject. The average treatment effect (ATE) is computed by taking the average of the difference between the observed and potential outcomes for each subject. PSM generates artificial distributions of the control group that matches the treatment group profile.

Again Difference in Difference Methodology

g3 = (Treatment After Intervention – Treatment Before Intervention) –
(Control After Intervention – Control Before Intervention);

Difference in Difference

Y = g0 + g1*dT + g2*dA + (D-D)*dT*dA + other controls

Follows key reference: https://www.povertyactionlab.org/full-search?search_api_views_fulltext=handbook

Individual Fixed-effects (for panel data)

$$\boldsymbol{Y_{i,t} = \beta_0 + \beta_1 S_{i,t} + \rho X_{i,t} + \sum_{\tau=1}^{T} \delta_\tau \mathbf{1}(t=\tau) + \sum_{\varphi=1}^{I} \gamma_\varphi \mathbf{1}(i=\varphi) + \varepsilon_{i.t}}$$

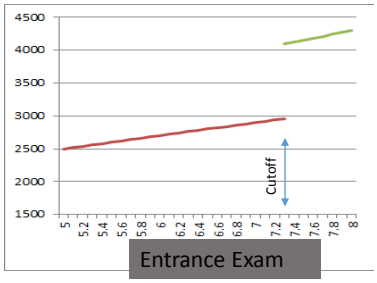
Time-Fixed-Effect **Individual-Fixed-Effect**

- Panel data are those in which you have the same set of observations in more than one point in time. So you can use the "fixed-effects" strategy.
- By including fixed effects, you control all fixed mean differences between people in any observable or unobservable predictors, such as genetic and environmental differences between hospitalized and non-hospitalized. Fixed time effects also control mean changes that have occurred for all groups.
- However, changes that happened differently for each group are not controlled.

Regression Discontinuity Design (RDD)

- RDD estimates the average treatment effect by exploring a selection rule based on a discontinuity in a particular observable variable.
- For example, let's think about the college entrance examination, in which students enter for college if they take a certain cutoff.
- In principle, students with higher grades in the college entrance examination also tend to perform better in the job market, which may bias the estimate of the impact. But in this case, we can compare students who took notes very close to x0, up and down to isolate the college entrance impact.

- We can compare students who took notes very close to x, above and below, because, in the neighborhood of x, students above and below will tend to be very similar in observable and unobservable characteristics.
- However, those just to the right of discontinuity x passed in college; the ones on the left do not.



RDD $D_i = D(x_i) = f[x_i > x_m]$ $Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i$

Where:

- f [.] is the function, x is the continuous variable
- xm is the treatment threshold separating the units (individuals) into two groups: those receiving treatment (D = 1) and those who did not receive (D = 0)
- Dit is a dummy indicating the treatment
- Yit is the outcome of interest
- Usually, the sample is subset into a window of x around xm.

$$X_i = \theta_0 + \theta_1 D_i + \varepsilon_i \qquad Y_i = \beta_0 + \beta_1 \hat{X}_i + \varepsilon_i$$

- In this first example, we have the case of a "sharp RDD", in which the cutoff 100% determinant for the treatment. However, in some cases it may not be so.
- Sometimes passing from the left to the right side of the cutoff does not guarantee treatment, but significantly increases its likelihood what is called "fuzzy RDD", in which the methods of "sharp RDD" and instrumental variables are combined. More specifically, the discontinuity will instrument the treatment.