Regression Discontinuity 1

Eduardo Fe

01/07/2019

Introduction.

During the class, we have discussed Lee's paper on the effect of incumbency on re-election. The data from this study is avaiable in R, as part of the package **rddtools** (written by Guido Imbens, from Harvard University). We will replicate the estimates in Lee's work in this lab. However, rather than usin Imbens' **rddtools**, we are going to use **rdrobust** a much more powerful package written by Matias Cattaneo (Princeton University).

Start by installing both **rddtools** and Cattaneo's

```
install.packages("rddtools")
install.packages("rdrobust")
library(rddtools)
library(rdrobust)
```

To load Lee's data, write

```
data(house)
```

The package **rdrobust** has a number of very advance facilities. First, the command **rdplot** can produce *optimised* graphs of the outcome (voteshare in electoin t+1) on the running variable; these graphs are very useful to get a first idea of the extent of any potential discontinuity at the 50% threshold. Among the

```
rdplot(y = house$y, x = house$x, title = "RD Plot - Senate Elections Data",
       y.label = "Vote Share in Election at time t+1",
       x.label = "Vote Share in Election at time t")
```

RD Plot – Senate Elections Data
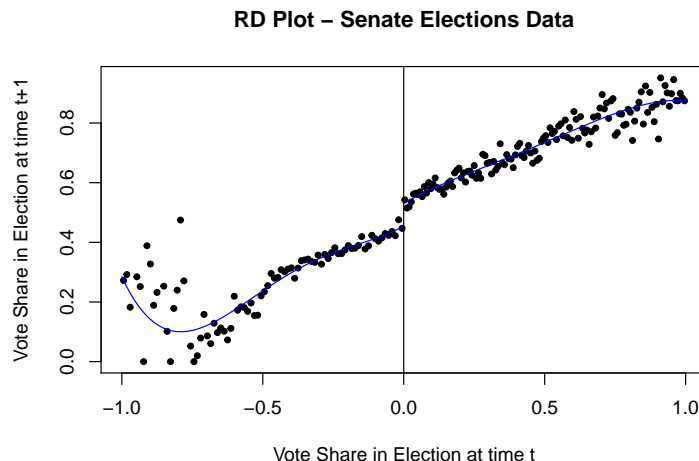
As you can see, the data already suggests that incumbency might have an effect on subsequent elections, at least in areas where the election was closely won. To estimate the LATE at the cutoff, the command to be executed is **rdrobust(y = ..., x = ...)**,

```
ate <- rdrobust(y = house$y, x = house$x, all=TRUE)
summary(ate)
```

```
## Call: rdrobust
```

```
## 
## Number of Obs.                    6558
## BW type                          mserd
## Kernel               Triangular
## VCE method                          NN
## 
## Number of Obs.                    2740          3818
## Eff. Number of Obs.                780           803
## Order est. (p)                       1             1
## Order bias  (p)                      2             2
## BW est. (h)                      0.134         0.134
## BW bias (b)                      0.238         0.238
## rho (h/b)                        0.561         0.561
## 
## =================================================================
##         Method    Coef. Std. Err.        z     P>|z|    [ 95% C.I. ]
## =================================================================
##   Conventional    0.063      0.011    5.743    0.000   [0.042 , 0.085]
## Bias-Corrected    0.059      0.011    5.350    0.000   [0.037 , 0.081]
##         Robust    0.059      0.013    4.686    0.000   [0.034 , 0.084]
## =================================================================
```

The command first returns the number of observations, the type of band-width selector and kernel. The second panel includes details the number of observations falling on each side of the cutoff, the "effective" number of observations (used in estimation at the cutoff on each side), the order ot the polynomial used in the nonpamatric regression (`Order est. (p)`, which in this case is 1), as well as the bandwidth used (`BW est. (h)`, which in his case equals 0.134). Finally, the command produces estimates of the treatment effect. In this example, the effect of incumbency at the cutoff seems to be 6.3 percentage points (significant).

In this example we don't have any further variables, so if we wanted to do a robustness check, the only viable option is to do a placebo analysis at a point away from the cutoff. To this end, you need to, first, select a point away from the cutoff where you would like to undertake this test; then you will have to modify the command above in order to tell `R` where the location of the cutoff is now (`R` automatically assumes the threshold is located at x=0). Finally, you could look at the sample sizes used at each side of the cutoff for estimation (in the original estimation, not in the placebo test). These sample sizes should be balanced. What would be your conclusion regarding this point in view of the results?

The poor and the pious.

The design in Lee's paper has been extended widely to explore other consequences of the democratic system. In a recent paper, Meyersson (2014) uses a RD design to study the effect of political Islam on women's empowerment. The idea here is that data reveal a persistent negative correlation between Islamic rule and women's rights. However, this correlation does not have a causal interpretation due to the problem of selection. Meyersson provides the first piece of evidence using a new and unique data set of Turkish municipalities. In 1994, an Islamic party won multiple municipal mayor seats across the country. Using a regression discontinuity (RD) design, Meyersson compares municipalities where this Islamic party barely won or lost elections. You task now is to use Meyersson's data and replicate his study using Cattaneo's `rdrobust`. The outcome variables of interest are the share of women and men with high school graduation in 2000 (`hischshr1520f` and `hischshr1520m` respectively). The running variable is the Islamic party's winning share (`iw94`). In addition to this, you will want to replicate the graphs in Meyerson's paper, Figure 2.a. (which gives you a check of whether or not there might be any discontinuity in the running variable at the threshold -this would invalidate the RD design) and Figure 4.