# Quant II

## RDD and Mediation

Ye Wang

4/08/2019

# Outline

- RD: various perspectives
  - RD as an approximation of conditional means
  - RD as a local experiment
- Mediation
  - The history of mediation analysis
  - Non-parametric mediation analysis

# RD as an approximation of conditional means

- It is a model-based approach.

# RD as an approximation of conditional means

- It is a model-based approach.
- We assume that there is a hyperthetical population and the sample is drawn from it.

# RD as an approximation of conditional means

- It is a model-based approach.
- We assume that there is a hyperthetical population and the sample is drawn from it.
- We are interested in the difference between the two population means across the threshold.

# RD as an approximation of conditional means

- It is a model-based approach.
- We assume that there is a hyperthetical population and the sample is drawn from it.
- We are interested in the difference between the two population means across the threshold.
- The difference may occur at any order (RD, Kink, etc.).

# RD as an approximation of conditional means

- It is a model-based approach.
- We assume that there is a hyperthetical population and the sample is drawn from it.
- We are interested in the difference between the two population means across the threshold.
- The difference may occur at any order (RD, Kink, etc.).
- We fit an outcome model to extract the difference.

# Some caveats

- Do not fit a global polynomial.
- Do not use two linear regressions either.
- Just use the package!
- It requires strong assumptions to aggregate estimates from different thresholds.
- You may control for covariates (Calonico et al. 2019) or use discrete running variable (Kolesár and Rothe, 2018).
- The equivalence test may work better when testing the balance of covariates (Hartman and Hidalgo, 2020).

# RD as a local experiment

- Another perspective is to treat RD as a local experiment (Cattaneo, Frandsen and Titiunik; 2015).
- Within the chosen bandwidth, the treatment is randomly assigned.

# RD as a local experiment

- Another perspective is to treat RD as a local experiment (Cattaneo, Frandsen and Titiunik; 2015).
- Within the chosen bandwidth, the treatment is randomly assigned.
- Then all the old techniques could be applied.

# RD as a local experiment

- Another perspective is to treat RD as a local experiment (Cattaneo, Frandsen and Titiunik; 2015).
- Within the chosen bandwidth, the treatment is randomly assigned.
- Then all the old techniques could be applied.
- We choose the bandwidth to minimize the imbalance of covariates.

# RD as a local experiment

- Another perspective is to treat RD as a local experiment (Cattaneo, Frandsen and Titiunik; 2015).
- Within the chosen bandwidth, the treatment is randomly assigned.
- Then all the old techniques could be applied.
- We choose the bandwidth to minimize the imbalance of covariates.
- A true design-based approach.
- More works should be done on this topic.

# An example

- We will work with the Meyersson (2014) paper: "Islamic Rule and the Empowerment of the Poor and the Pious"
- The paper shows a (local) result: the victory of Islamic parties in Turkey resulted in better outcomes for women.
- Running variable: the difference in vote share between the largest Islamic party and the largest secular party (not two party)
- Outcome that we'll look at: high school education

# Set up the data

. . .

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
## -1.0000 -0.4600 -0.3102 -0.2786 -0.1061  0.9905     544
```

# Estimation

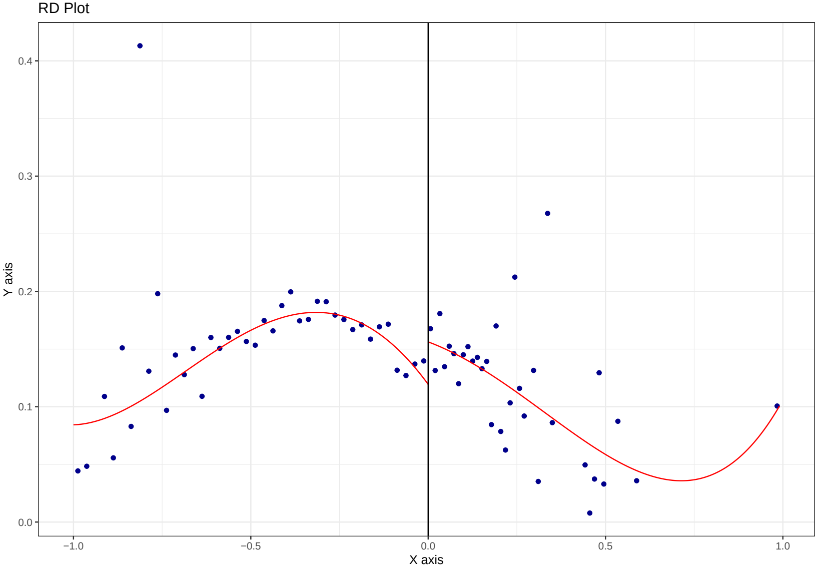
. . .

# Results

```
## Call: rdrobust
##
## Number of Obs.                 2630
## BW type                       mserd
## Kernel                   Triangular
## VCE method                       NN
##
## Number of Obs.                2315          315
## Eff. Number of Obs.            529          266
## Order est. (p)                   1            1
## Order bias  (q)                  2            2
## BW est. (h)                  0.172        0.172
## BW bias (b)                  0.286        0.286
## rho (h/b)                    0.603        0.603
## Unique Obs.                   2313          315
##
## ================================================================================
##         Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
## ================================================================================
##   Conventional     0.030     0.014     2.116     0.034     [0.002 , 0.058]
##         Robust         -         -     1.776     0.076    [-0.003 , 0.063]
## ================================================================================
```

# Plot it

```r
rdplot(d$hischshr1520f, d$iwm94, p = 4)
```



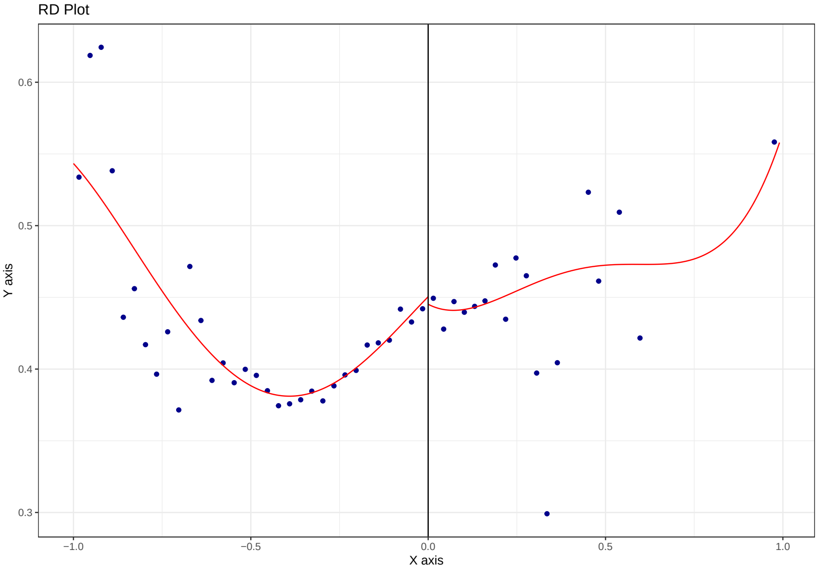
# Placebo tests

- Do placebo tests on other covariates and other outcomes.

. . .

```
## $coef
##                        Coeff
## Conventional   0.004097863
## Bias-Corrected 0.008070629
## Robust         0.008070629
##
## $se
##                   Std. Err.
## Conventional   0.01227104
## Bias-Corrected 0.01227104
## Robust         0.01408919
```

# Placebo plot



# More Placebos

```
## $coef
##                      Coeff
## Conventional   0.01285314
## Bias-Corrected 0.01263466
## Robust         0.01263466
##
## $se
##                  Std. Err.
## Conventional   0.01403235
## Bias-Corrected 0.01403235
## Robust         0.01691543
```

```
## $coef
##                        Coeff
## Conventional   0.0039695120
## Bias-Corrected 0.0008440962
## Robust         0.0008440962
##
## $se
##                  Std. Err.
## Conventional   0.01537558
## Bias-Corrected 0.01537558
```

# Sorting

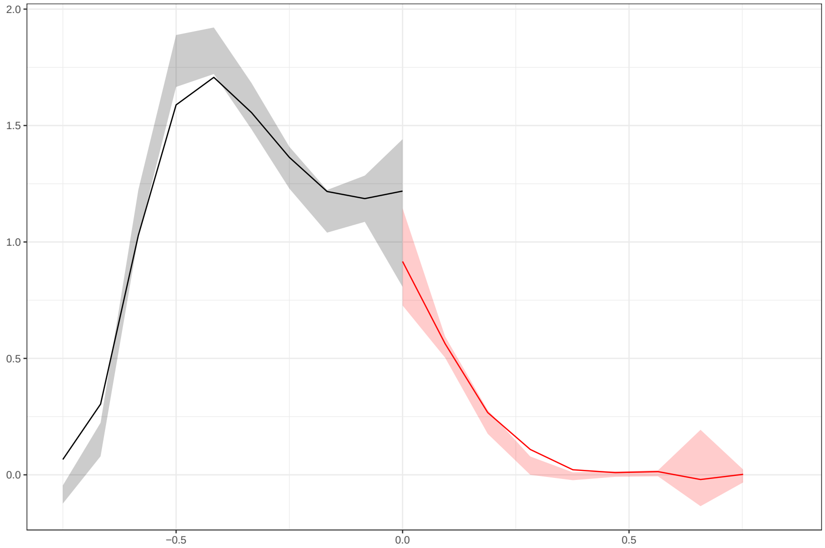
- Density tests are also a good way to examine the possibility of sorting.

. . .

```
##
## RD Manipulation Test using local polynomial density estimation.
##
## Number of obs =       2660
## Model =               unrestricted
## Kernel =              triangular
## BW method =           comb
## VCE method =          jackknife
##
## Cutoff c = 0          Left of c           Right of c
## Number of obs         2332                328
## Eff. Number of obs    769                 314
## Order est. (p)        2                   2
## Order bias (q)        3                   3
## BW est. (h)           0.25                0.282
##
## Method                T                   P > |T|
```

# Density Plot

```r
rdplotdensity(rd_density, d$iwm94)
```




```
## $Estl
```

# Kink

. . .

```
## Call: rdrobust
##
## Number of Obs.                 2630
## BW type                       mserd
## Kernel                   Triangular
## VCE method                       NN
##
## Number of Obs.                 2315         315
## Eff. Number of Obs.             428         238
## Order est. (p)                    1           1
## Order bias  (q)                   2           2
## BW est. (h)                   0.139       0.139
## BW bias (b)                   0.286       0.286
## rho (h/b)                     0.486       0.486
## Unique Obs.                    2313         315
##
## ================================================================================
##         Method     Coef. Std. Err.         z     P>|z|      [ 95% C.I. ]
## ================================================================================
##   Conventional     0.020     0.242     0.082     0.934    [-0.455 , 0.495]
##         Robust         -         -     0.205     0.838    [-0.689 , 0.849]
## ================================================================================
```

# Some potential directions

- RD in TSCS analysis
- RD under interference
- RD combined with structural models

# The history of mediation analysis

- Sometimes knowing the correct mechanism saves lives.

# The history of mediation analysis

- Sometimes knowing the correct mechanism saves lives.
- Oranges can cure blood poisoning.

# The history of mediation analysis

- Sometimes knowing the correct mechanism saves lives.
- Oranges can cure blood poisoning.
- But why?

# The history of mediation analysis

- Sometimes knowing the correct mechanism saves lives.
- Oranges can cure blood poisoning.
- But why?
- Is it due to the acid?

# The history of mediation analysis

- Sometimes knowing the correct mechanism saves lives.
- Oranges can cure blood poisoning.
- But why?
- Is it due to the acid?
- Many died because of the wrong mechanism.

# The history of mediation analysis

- Sewall Wright (1921): path analysis

# The history of mediation analysis

- Sewall Wright (1921): path analysis
- Barbara Burks (1926): IQ of parents $\rightarrow$ social status $\rightarrow$ IQ of children

# The history of mediation analysis

- Sewall Wright (1921): path analysis
- Barbara Burks (1926): IQ of parents $\rightarrow$ social status $\rightarrow$ IQ of children
- In the 1970s, path analysis was re-invented by Otis Duncan and Arthur Goldberger.

# The history of mediation analysis

- Sewall Wright (1921): path analysis
- Barbara Burks (1926): IQ of parents $\rightarrow$ social status $\rightarrow$ IQ of children
- In the 1970s, path analysis was re-invented by Otis Duncan and Arthur Goldberger.
- It was developed into the SEM framework in econometrics.

# The history of mediation analysis

- Sewall Wright (1921): path analysis
- Barbara Burks (1926): IQ of parents $\rightarrow$ social status $\rightarrow$ IQ of children
- In the 1970s, path analysis was re-invented by Otis Duncan and Arthur Goldberger.
- It was developed into the SEM framework in econometrics.
- Barro and Kenny (1986) $\rightarrow$ Imai and Yamamoto (2011).

# Mediation analysis and TSCS data analysis

- There are a lot of similarities between mediation analysis and TSCS data analysis under the sequential ignorability assumption.

# Mediation analysis and TSCS data analysis

- There are a lot of similarities between mediation analysis and TSCS data analysis under the sequential ignorability assumption.
- Mediation analysis is built upon similar assumptions.

# Mediation analysis and TSCS data analysis

- There are a lot of similarities between mediation analysis and TSCS data analysis under the sequential ignorability assumption.
- Mediation analysis is built upon similar assumptions.
- The same estimation techniques can be applied.

# Mediation analysis and TSCS data analysis

- There are a lot of similarities between mediation analysis and TSCS data analysis under the sequential ignorability assumption.
- Mediation analysis is built upon similar assumptions.
- The same estimation techniques can be applied.
- C4 is basically the past outcome.

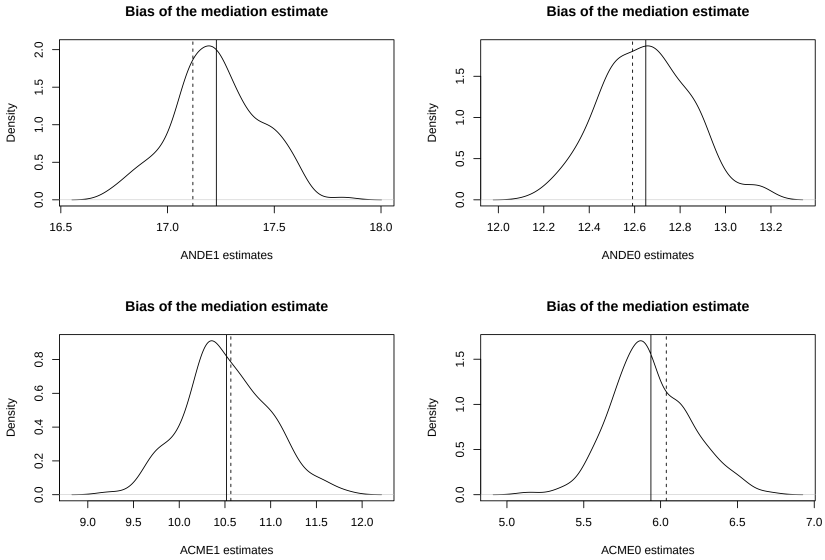
# Mediation analysis and TSCS data analysis

- There are a lot of similarities between mediation analysis and TSCS data analysis under the sequential ignorability assumption.
- Mediation analysis is built upon similar assumptions.
- The same estimation techniques can be applied.
- C4 is basically the past outcome.
- Identifying the effect of the entire history is easy.
- Identifying the effect of the past is not.

# Mediation analysis and TSCS data analysis

- There are a lot of similarities between mediation analysis and TSCS data analysis under the sequential ignorability assumption.
- Mediation analysis is built upon similar assumptions.
- The same estimation techniques can be applied.
- C4 is basically the past outcome.
- Identifying the effect of the entire history is easy.
- Identifying the effect of the past is not.
- Hard to distinguish “carryover” from heterogeneity.

# Mediation analysis in R



# Mediation analysis with multiple pathways

- Sometimes there are more than one mediators.
- One mediator may come before another.
- democracy $\rightarrow$ perceived costs and benefits $\rightarrow$ moral concerns $\rightarrow$ opposition to war

$$
\begin{aligned}
\mathbb{E}[Y(1)-Y(0)] &= \mathbb{E}[Y\big(1, L(1), M(1, L(1))\big) - Y\big(0, L(0), M(0, L(0))\big)] \\
&= \underbrace{\mathbb{E}[Y\big(1, L(0), M(0, L(0))\big) - Y\big(0, L(0), M(0, L(0))\big)]}_{A \to Y} \\
&\quad + \underbrace{\mathbb{E}[Y\big(1, L(0), M(1, L(0))\big) - Y\big(1, L(0), M(0, L(0))\big)]}_{A \to M \to Y} \\
&\quad + \underbrace{\mathbb{E}[Y\big(1, L(1), M(1, L(1))\big) - Y\big(1, L(0), M(1, L(0))\big)]}_{A \to L \to Y;\, A \to L \to M \to Y} \\
&\equiv \tau_{A \to Y}(0) + \tau_{A \to M \to Y}(1) + \tau_{A \to L \rightsquigarrow Y}(1).
\end{aligned}
$$

# Mediation analysis with multiple pathways

- Yamamoto and Zhou (2019) provide two impuation-based estimators:

$$\mathbb{E}[Y(1, L(0), M(0, L(0)))] = \mathbb{E}\Big[\mathbb{E}\big[\mathbb{E}[Y|X, A=1, L, M]|A=0, X\big]\Big]$$
$$\mathbb{E}[Y(1, L(0), M(1, L(0)))] = \mathbb{E}\Big[\mathbb{E}\big[\mathbb{E}[Y|X, A=1, L]|A=0, X\big]\Big].$$

$$\mathbb{E}[Y(1, L(0), M(0, L(0)))] = \mathbb{E}\big[\mathbb{E}[Y|X, A=1, L, M]\frac{\mathbb{P}[A=0]}{\mathbb{P}[A=0|X]}|A=0\big]$$
$$\mathbb{E}[Y(1, L(0), M(1, L(0)))] = \mathbb{E}\big[\mathbb{E}[Y|X, A=1, L]\frac{\mathbb{P}[A=0]}{\mathbb{P}[A=0|X]}|A=0\big].$$

- "pure imputation estimator" and "imputation-based weighting estimator."
- Also no $C4$, but more flexible.

# DAG vs. the Rubin model

- DAG has its comparative advantage in mechanism analysis.
- It does help us pin down ideas.
- But how can we know that the DAG we have is correct?
- Sensitivity analysis helps.
- How can you present the LATE theorem using DAG?