ECMA31100 Introduction to Empirical Analysis II

Winter 2022, Week 2: Discussion Session

Conroy Lau

The University of Chicago

January 19, 2022

Welcome to ECMA31100!

Your TA

- Name: Conroy Lau
- Current research interests: Econometric theory and applied microeconomics
- E-mail: `ccplau@uchicago.edu`
  - Put "ECMA31100" on the subject line
  - Please let me know if you have any suggestions or if you find any typos in the slides

TA office hours

- Wednesdays, 5:30pm–7:30pm, Zoom
- Feel free to send me an e-mail if you plan to come, so I can prepare ahead

TA sessions

Tentative plans of the TA sessions

- Review materials that are based on the lectures
- Discuss research papers to complement the lectures
- Introduce research papers that might be useful for your project or MA thesis

Today's outline

- Logistics and general advice
- Discussion on propensity score matching
- Overview of a few microeconometric methods

Self-introduction

Please introduce yourself according to the alphabetical order on Zoom

- Name
- Program and major
- What courses are you taking in this quarter?
- Research interests
- Anything else!

Contents

Some general advice

Problem sets

- Start early!
- Do not copy answers

Typesetting

- I suggest using $\LaTeX$ to type your answers — it's a good investment!
- Build your own templates!
- Learn how to import figures and tables generated from your code
- UChicago offers **free** professional Overleaf accounts for students

Some general advice

Coding

- Consider using version control (I use Git and GitHub)
- Do not exceed 80 characters per line
- Feel free to choose any programming language (I mainly use `Julia` and `R` recently)
- Follow style guides (e.g., see the tidyverse style guide by Wickham (2021) for `R`)
- Include informative but not excessive comments in your code

Some excellent sources of readings for your project

- Annual Review of Economics
- Journal of Economic Literature
- Various Handbooks in Economics

Any questions?

Contents

Matching estimators

Neyman-Rubin potential outcome model

- Suppose we have a sample of N units indexed by $i = 1, \ldots, N$
- Let D_i be the binary variable that equals 1 if individual i receives treatment
- Let X_i be the vector of covariates for individual i
- Denote $Y_i(1)$ and $Y_i(0)$ as the potential outcomes with and without treatment resp.
- We can write $Y_i = Y_i(1)D_i + Y_i(0)(1 - D_i)$
- For each individual i, we cannot observe both $Y_i(1)$ and $Y_i(0)$

Two assumptions

- Unconfoundedness: $(Y_i(0), Y_i(1)) \perp\!\!\!\perp D_i | X_i$
- Overlap condition: $0 < \mathbb{P}[D_i = 1 | X_i = x] < 1$

Matching estimators

Some target parameters

- Average treatment effect $\tau_{\text{ATE}} \equiv \mathbb{E}[Y_i(1) - Y_i(0)]$
- Average treatment effect on the treated $\tau_{\text{ATT}} \equiv \mathbb{E}[Y_i(1) - Y_i(0)|D_i = 1]$
- Average treatment effect on the untreated $\tau_{\text{ATU}} \equiv \mathbb{E}[Y_i(1) - Y_i(0)|D_i = 0]$

Matching

- Matching on covariates
- Matching on the propensity scores

Matching estimators

Matching on covariates (see Imbens (2015) for more details)

- Let M be the number of matches per unit
- In the absence of ties, let $\mathcal{J}_M(i)$ be the set of matches for unit i:

$$\mathcal{J}_M(i) \equiv \left\{ j = 1, \ldots, N : D_j = 1 - D_i \quad \text{and} \quad \sum_{k : D_k \neq D_i} \mathbb{1}\left[||X_i - X_k|| \leq ||X_i - X_j|| \right] = M \right\},$$

- Then, we can impute the potential outcomes as follows:

$$\widehat{Y}_i(0) = \begin{cases} Y_i & , D_i = 0 \\ \dfrac{1}{M} \displaystyle\sum_{j \in \mathcal{J}_M(i)} Y_i & , D_i = 1 \end{cases} \quad \text{and} \quad \widehat{Y}_i(1) = \begin{cases} \dfrac{1}{M} \displaystyle\sum_{j \in \mathcal{J}_M(i)} Y_i & , D_i = 0 \\ Y_i & , D_i = 1 \end{cases}$$

Matching estimators

Average treatment effect

- Recall that ATE is given by $\tau_{\text{ATE}} \equiv \mathbb{E}[Y_i(1) - Y_i(0)]$
- By the imputation from last slide, we can estimate it by $\widehat{\tau}_{\text{ATE}} = \frac{1}{N}\sum_{i=1}^{N}[\widehat{Y}_i(1) - \widehat{Y}_i(0)]$

Average treatment effect on the treated

- For ATT, we have

$$\begin{aligned}
\tau_{\text{ATT}} &= \mathbb{E}[Y_i(1)|D_i = 1] - \mathbb{E}[Y_i(0)|D_i = 1] \\
&= \mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[\mathbb{E}[Y_i(0)|D_i = 1, X_i]|D_i = 1] \\
&= \mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[\mathbb{E}[Y_i(0)|X_i]|D_i = 1] \\
&= \mathbb{E}[Y_i|D_i = 1] - \mathbb{E}[\mathbb{E}[Y_i|D_i = 0, X_i]|D_i = 1]
\end{aligned}$$

- Similar as before, we have $\widehat{\tau}_{\text{ATT}} = \frac{1}{N_1}\sum_{i=1}^{N} D_i[\widehat{Y}_i(1) - \widehat{Y}_i(0)]$ where $N_1 = \sum_{i=1}^{N} D_i$

Matching estimators

Matching on propensity scores

- $p(x) \equiv \mathbb{P}[D_i = 1 | X_i = x]$ is the propensity score
- Under the Probit model, we have

$$\mathbb{P}[D_i = 1 | X_i = x_i] = \Phi(x_i' \beta),$$
$$\mathbb{P}[D_i = 0 | X_i = x_i] = 1 - \Phi(x_i' \beta),$$

  where $\Phi(\cdot)$ is the standard normal c.d.f. and β has the same length as X_i

- Given the data, we can estimate β by maximum likelihood:

$$\widehat{\beta} = \arg\max_{\beta} \frac{1}{N} \sum_{i=1}^{N} \left[D_i \log \Phi(X_i' \beta) + (1 - D_i) \log(1 - \Phi(X_i' \beta)) \right]$$

- Next, use a similar idea in the last slide (also see page 43 of week 1 lecture slides)

Any questions?

Blocking

Main idea (see Imbens and Rubin (2015) for more details)

- Estimate the propensity score $\widehat{p}(X)$ from the sample
- Construct J blocks by partitioning $[0, 1]$ into J intervals
- The boundary points are $\{b_0, b_1, \ldots, b_J\}$ such that $0 = b_0 < b_1 < \cdots < b_J = 1$
- Assign individual i to block j if $\widehat{p}(X_i) \in [b_{j-1}, b_j)$ and denote $B_i(j) \equiv \mathbb{1}[\widehat{p}(X_i) \in [b_{j-1}, b_j)]$
- Estimate ATE as the weighted average of each block's estimated ATE

Some questions for implementation

- How to choose the boundary points?
- How many blocks should we choose?
- How to assess unconfoundedness within blocks?

Blocking

Linearized propensity score

- Define the linearized propensity score (LPS) as $\widehat{\ell}(x) \equiv \ln \frac{\widehat{p}(x)}{1-\widehat{p}(x)}$

Checking independence (see Imbens and Rubin (2015) for more details)

- Let $N_1(j)$ and $N_0(j)$ be the number of treated and untreated units in block j resp.
- Let $\overline{\ell}_1(j)$ and $\overline{\ell}_0(j)$ be the mean LPS of treated and untreated units in block j resp.
- Let S_ℓ^2 be the sample variance of LPS in block j:

$$S_\ell^2(j) = \frac{1}{N_0(j) + N_1(j) - 2} \left\{ \sum_{i=1}^{N} B_i(j) \left[D_i(\widehat{\ell}(X_i) - \overline{\ell}_1(j))^2 + (1 - D_i)(\widehat{\ell}(X_i) - \overline{\ell}_0(j))^2 \right] \right\}$$

- Compute the t-statistic of block j as $t_j = \frac{\overline{\ell}_1(j) - \overline{\ell}_0(j)}{\sqrt{\left(\frac{1}{N_0(j)} + \frac{1}{N_1(j)}\right) S_\ell^2(j)}}$
- Then, compare t_j to some threshold (e.g., $t_{\max} = \max_{1 \le j \le J} t_j$)

Some remarks

Bootstrap

- Bootstrap does not work in general for matching estimators
  - This is because matching estimators are not smooth (Abadie and Imbens, 2008)
- See Abadie and Imbens (2016) for general variance formulae

Suggested readings

- Imbens (2004)
- Imbens (2015)
- Abadie and Cattaneo (2018)

Any questions?

Contents

A preview

Other topics

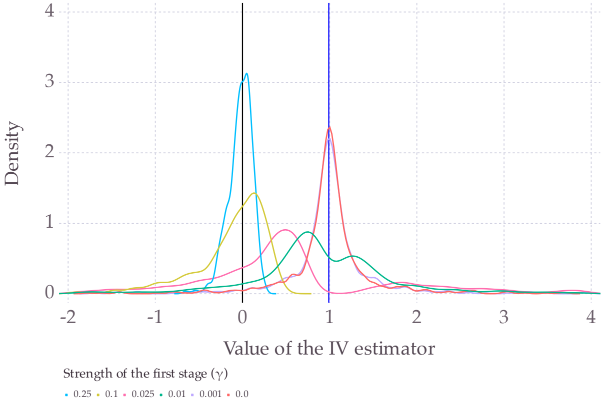
- Instrumental variables
- Difference-in-differences
- Regression discontinuity design
- Regression kink design
- Synthetic control
- ...

Next TA session

- Heterogeneous treatment effects
- Abadie (2003) (methodology, estimation and empirical example)

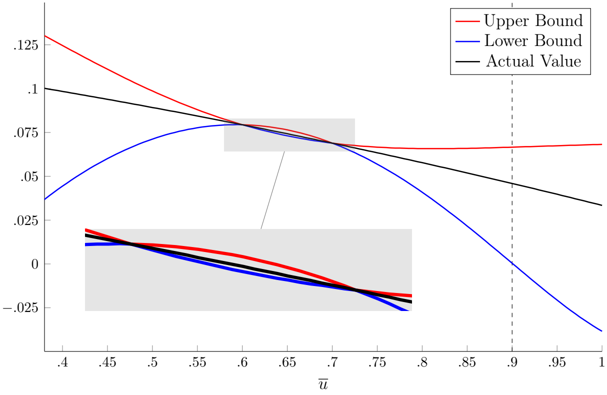
Weak instruments

Figure 1: Distribution of OLS and IV estimators (similar to that in Nelson and Startz (1990))



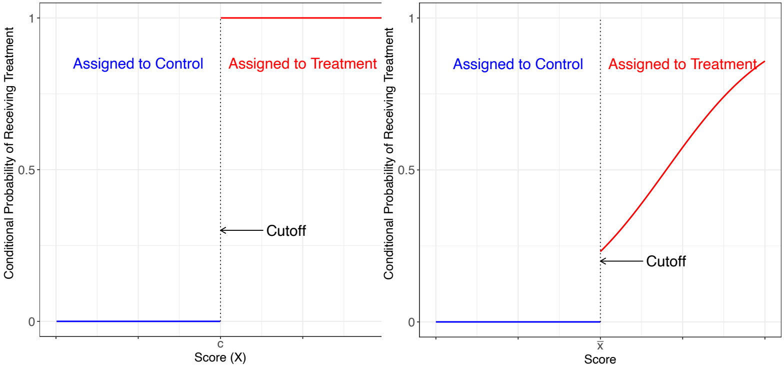
Marginal treatment effects

Figure 2: Bounds on LATE($0.35, \overline{u}$) (figure 8 of Mogstad et al. (2017))



Regression discontinuity design

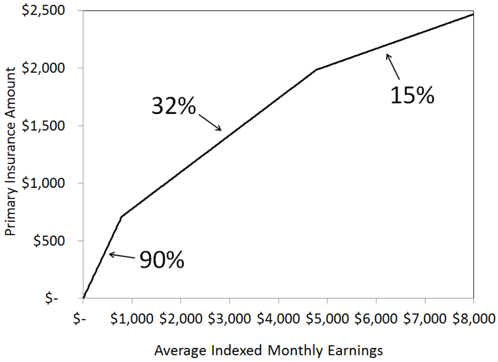
Figure 3: Sharp vs fuzzy RD (figure 4.1 of Cattaneo et al. (2018))



(a) Sharp RD (b) Fuzzy RD (One-Sided)

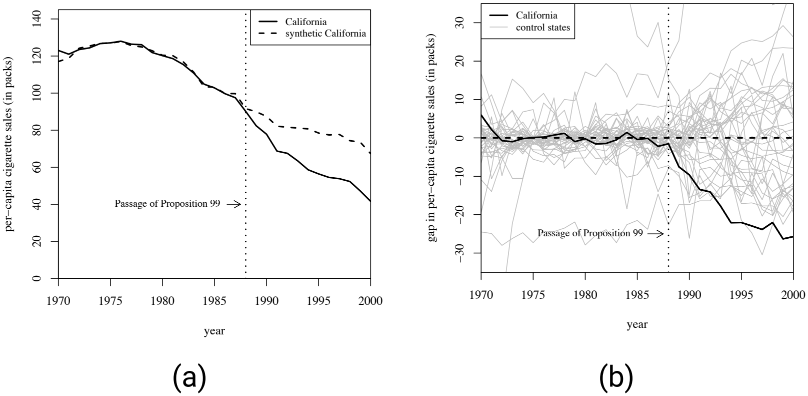
Regression kink design

Figure 4: PIA as a function of AIME (figure 1 of Gelber et al. (2016))



Synthetic control

Figure 5: Figures 2 and 4 of Abadie et al. (2007)



References 1

ABADIE, A. (2003): "Semiparametric instrumental variable estimation for causal effects," *Journal of Econometrics*, 113, 231–263.

ABADIE, A. AND M. D. CATTANEO (2018): "Econometric Methods for Program Evaluation," *Annual Review of Economics*, 10, 465–503.

ABADIE, A., A. DIAMOND, AND J. HAINMUELLER (2007): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Working Paper*.

ABADIE, A. AND G. IMBENS (2016): "Matching on the Estimated Propensity Score," *Econometrica*, 84, 781–807.

References 2

Abadie, A. and G. W. Imbens (2008): "On the Failure of the Bootstrap for Matching Estimators," *Econometrica*, 76, 1537–1557.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2018): *A Practical Introduction to Regression Discontinuity Designs: Volume II*, Working paper.

Gelber, A., T. Moore, and A. Strand (2016): "The Effect of Disability Insurance Payments on Beneficiaries' Earnings," *Working Paper*.

Imbens, G. W. (2004): "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *The Review of Economics and Statistics*, 86, 4–29.

——— (2015): "Matching Methods in Practice: Three Examples," *Journal of Human Resources*, 50, 373–419.

References 3

IMBENS, G. W. AND D. B. RUBIN (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, Cambridge University Press.

MOGSTAD, M., A. SANTOS, AND A. TORGOVITSKY (2017): "Using Instrumental Variables for Inference about Policy Relevant Treatment Effects," *Working Paper*.

NELSON, C. R. AND R. STARTZ (1990): "Some Further Results on the Exact Small Sample Properties of the Instrumental Variable Estimator," *Econometrica*, 58, 967–976.

WICKHAM, H. (2021): "The tidyverse style guide," `https://style.tidyverse.org/index.html`, accessed: 2022-01-11.