LARF in Casual Inference

Theoretical Analysis and Empirical Application

Yi Chai   Chen Fang

Advanced Research Seminar

May 18, 2023

Table of Contents

Table of Contents

Abadie (2003)

- Assumption
  - Independence; Exclusion; First stage; Monotonicity
- LARF

$$E[Y|X, D, D_1 > D_0]$$

- Identify LATE

$$E[Y|X, D = 1, D_1 > D_0] - E[Y|X, D = 0, D_1 > D_0] = E[Y_1 - Y_0|X, D_1 > D_0]$$

- Abadie's Pseudo-Weight

$$\kappa = 1 - \frac{D(1 - Z)}{P(Z = 0|X)} - \frac{Z(1 - D)}{P(Z = 1|X)}$$

- Theorem 3.1

$$E[g(Y, D, X)|D_1 > D_0] = \frac{1}{P(D_1 > D_0)} E[\kappa * g(Y, D, X)]$$

Estimation LARF by Linear Regression

- Parameters

$$(\alpha_0, \beta_0) = \arg\min_{\alpha,\beta} E[\{Y - (\alpha D + X^{'}\beta)\}^2 | D_1 > D_0]$$

- By Theorem 3.1

$$(\alpha_0, \beta_0) = \arg\min_{\alpha,\beta} E[\kappa\{Y - (\alpha D + X^{'}\beta)\}^2]$$

- Two-Step Estimation
  - Construct $\hat{\kappa}$ by estimating $P(Z = 1|X)$
  - Estimate LARF using:

$$(\hat{\alpha}, \hat{\beta}) = \arg\min_{\alpha,\beta} E[\hat{\kappa}\{Y - (\alpha D + X^{'}\beta)\}^2]$$

Comparison with 2SLS

- Parameters of linear specification for LARF

$$\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} = \left(E\left[\begin{pmatrix} D \\ X \end{pmatrix} \kappa \begin{pmatrix} D \\ X \end{pmatrix}' \right] \right)^{-1} E\left[\begin{pmatrix} D \\ X \end{pmatrix} \kappa \quad Y \right]$$

- Probablity limit of 2SLS

$$\begin{pmatrix} \alpha_{2SLS} \\ \beta_{2SLS} \end{pmatrix} = \left(E\left[\begin{pmatrix} Z \\ X \end{pmatrix} \begin{pmatrix} D \\ X \end{pmatrix}' \right] \right)^{-1} E\left[\begin{pmatrix} Z \\ X \end{pmatrix} \quad Y \right]$$

- Without covariates, $\alpha_{2SLS} = \alpha_0 = LATE$
- With covariates, 2SLS estimands do not only respond to the distribution of (Y, D, X) for compliers

Estimation LARF by Linear Regression

- Abadie's Pseudo-Weight

$$\kappa = 1 - \frac{D(1-Z)}{P(Z=0|X)} - \frac{Z(1-D)}{P(Z=1|X)}$$

- Theorem 3.1

$$E[g(Y,D,X)|D_1 > D_0] = \frac{1}{P(D_1 > D_0)} E[\kappa * g(Y,D,X)]$$

- Parameters

$$(\alpha_0, \beta_0) = \arg\min_{\alpha,\beta} E[\{Y - (\alpha D + X^{'}\beta)\}^2 | D_1 > D_0]$$

- By Theorem 3.1

$$(\alpha_0, \beta_0) = \arg\min_{\alpha,\beta} E[\kappa\{Y - (\alpha D + X^{'}\beta)\}^2]$$

- Two-Step Estimation
  - Construct $\hat{\kappa}$ by estimating $P(Z=1|X)$
  - Estimate LARF using:

$$(\hat{\alpha}, \hat{\beta}) = \arg\min_{\alpha,\beta} E[\hat{\kappa}\{Y - (\alpha D + X^{'}\beta)\}^2]$$

LARF and 2SLS

- Parameters of linear specification for LARF

$$\begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} = \left(E\left[\begin{pmatrix} D \\ X \end{pmatrix} \kappa \begin{pmatrix} D \\ X \end{pmatrix}' \right] \right)^{-1} E\left[\begin{pmatrix} D \\ X \end{pmatrix} \kappa \; Y \right]$$

- Probablity limit of 2SLS

$$\begin{pmatrix} \alpha_{2SLS} \\ \beta_{2SLS} \end{pmatrix} = \left(E\left[\begin{pmatrix} Z \\ X \end{pmatrix} \begin{pmatrix} D \\ X \end{pmatrix}' \right] \right)^{-1} E\left[\begin{pmatrix} Z \\ X \end{pmatrix} \; Y \right]$$

- Without covariates, $\alpha_{2SLS} = \alpha_0 = LATE$
- With covariates, 2SLS estimands do not only respond to the distribution of (Y, D, X) for compliers.
- When $\hat{P}(Z = 1|X)$ is linear in X or constant treatment effect, $\alpha_{2SLS} = \alpha_0$

Table of Contents

Banerjee et al. (2015)

- **Reference**
  - Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The miracle of microfinance? Evidence from a randomized evaluation. American economic journal: Applied economics, 7(1), 22-53.
- **Background**
  - Rapid Expansion of Microfinance institutions (MFIs) (7.6 million in 1997 to 137.5 million in 2010)
    - ⋆ "A type of banking service that is provided to unemployed or low-income individuals or groups who would otherwise have no other means of gaining financial services. "
  - **Support**: "Fast poverty alleviation" (e.g. Mohammed Yunus and the Grameen Bank for their contribution to the reduction in world poverty)
  - **Backlash**: "Hyperprofits off the poor" (e.g. a rash of reported suicides linked to over-indebtedness on *New York Times*, little regulation from government)

Banerjee et al. (2015)

- **Experiment Design** Timeline
  - Studies households over 3.5 years after the introduction of the system: the longest period of any study.
  - In 2005, 52 of 104 poor neighborhoods in Hyderabad, India were randomly selected for opening of an MFI branch.
  - **First endline** survey was **15-18 months** after introduction of microfinance in each neighborhood.
  - **Second endline** survey was **two years** later.
  - **Total sample**: 6,864 households, 90% maintained contact.

Banerjee et al. (2015)

- **Main Conclusions (First Endline)**
  - NO DIFFERENCE in monthly per capita consumption and monthly non-durable consumption.
  - Significant POSITIVE IMPACTS of the purchase of durables households reduced spending on what they described as "temptation goods".
  - INCREASE in the number of new businesses created, particularly by women.

Banerjee et al. (2015)

- **Simplified Empirical Strategy**
  - **Treatment**: `spandana_1` (Has outstanding loan from Spandana at endline 1, Binary)
  - **Instrument Variable**: `treatment` (whether be selected as treatment area, Binary)
  - **Outcomes**: the household expenditure and its structure.

Banerjee et al. (2015)

- **Empirical Results**
  - NO DIFFERENCE in monthly per capita consumption. (Conclusion 1)

	total expenditure	total expenditure per capita
OLS	74.4	−25.6
	(20.1)	(40.5)
2SLS	1041.0	267.8
	(812.3)	(164.2)
LARF	1105.3	275.7*
	(840.6)	(166.2)

Banerjee et al. (2015)

- **Empirical Results**
  - NO DIFFERENCE in monthly non-durable consumption. (Conclusion 1)
  - Significant POSITIVE IMPACTS of the purchase of durables households reduced spending on what they described as "temptation goods". (Conclusion 2)

	durables	nondurable	temptation
OLS	40.4	60.2	48.7**
	(84.1)	(165.6)	(21.9)
2SLS	868.6**	308.6	−189.9**
	(342.9)	(668.8)	(89.4)
LARF	880.6***	323.1	−190.1**
	(364.31)	(691.5)	(93.2)

Duflo et al. (2006)

- **Reference**
  - Duflo, E., Dupas, P., Kremer, M., & Sinei, S. (2006). Education and HIV/AIDS prevention: evidence from a randomized evaluation in Western Kenya. World Bank Policy Research Working Paper, (4024).
- **Introduction**
  - A seven-year randomized evaluation.
  - What effect the **education subsidies** and **HIV curriculum** on adolescent girls' dropout, pregnancy, marriage and sexually transmitted infection in Kenya.
  - **Education Subsidies**: providing two free school uniforms over the last three years of primary school.
  - **HIV curriculum**: three teachers in each primary school received government-provided training to help them deliver Kenya's national HIV/AIDS curriculum.
  - **Conclusion**: Both programs combined reduce STI more, but cut dropout and pregnancy less, than education subsidies alone.

Duflo et al. (2006)

- **Simplified Empirical Strategy**
  - **Treatment**: Whether dropped out of primary in 2005. (`dropout05v3`: Binary)
  - **Instrument**: If school benefit from uniform program (`Utreat`: Binary)
  - **Outcome**: Ever married, pregnant, had child (`evmar05v3`, `evpreg05v3`, `evchild05v3`: also binary)

Duflo et al. (2006)

- **Simplified Empirical Result**
  - LARF cannot get estimation if we use the "least square" method (here we use "ML" method) (Estimate is "marginal effects at the means" for LARF).
  - The coefficients are insignificant for LARF but more significant for 2SLS.

	marriage	child	pregnant
2SLS	0.67***	0.19	0.42**
	(0.19)	(0.16)	(0.17)
LARF	0.61	0.51	0.69
	(0.40)	(0.98)	(1.36)

Duflo et al. (2007)

- **Reference**
  - Duflo, E., & Pande, R. (2007). Dams. The Quarterly Journal of Economics, 122(2), 601-646.
- **Background**
  - The fast growing of construction of dams in India. (Epitome / Main form of Public Investments)
  - The productivity and distributional effects are uncertain. (Poverty, Agricultural Production & Welfare)
  - Trade-off: displacement vs. water-access.
  - Irrigation is the primary purpose of over 95% of large Indian dams.

Graph Illustration

Duflo et al. (2007)

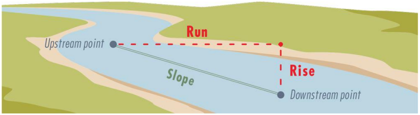
- **Data**: Balanced panel data from 1973-1999 in GIS-district level in India
- **Dependent Variables**: several social economy variables (Gross Irrigated Area, Total Production, Poverty Gap, Headcount ratio)
- **Independent Variable**: district-wise number of new dams

Duflo et al. (2007)

- **Endogeneity**: Geographic suitability, political clout and economics potentials all affect dam placement.
- **Instrument Variable**: the gradient of rivers. Gradient Distribution

$$Slope\% = \frac{Rise}{Run} \times 100\%$$

Figure: the calculation of gradient



Duflo et al. (2007)

- **Main Conclusions**
  - Dams do not affect agricultural production in the district where they are located.
  - Irrigated area and agricultural production increase in districts located downstream.
  - Poverty declines in the districts located downstream from a dam, but increases significantly in districts where dams are built.

Duflo et al. (2007)

- **Simplified Empirical Strategy**
  - Transform panel data to Cross-Section data (in 1999, within district)
  - **Treatment**: A state with more than a hundred dams by 1999 is a 'high' construction state. (calculated from `sdistrict1`: Binary)
  - **Instrument**: A district with less than 90% of river gradient below 1.5% percent is classified as a 'high' gradient district. (calculated from `damsumstate`: Binary)
  - **Outcome**: log yield of main crops (`lyield`), total value of production of water-intensive crops (`waterp`)

Duflo et al. (2007)

- **Simplified Empirical Result**
  - Conclusion 1 has been proved (Dams do not affect agricultural production in the district where they are located.)
  - The coefficients are insignificant for both outcomes using different regression method. (very different estimates)

	lyield	waterp
OLS	-0.108	-0.284^{*}
	(0.10)	(0.16)
2SLS	6.02	19.70
	(11.09)	(35.42)
LARF	-1.539	-10.69
	(4.50)	(19.26)

Table of Contents

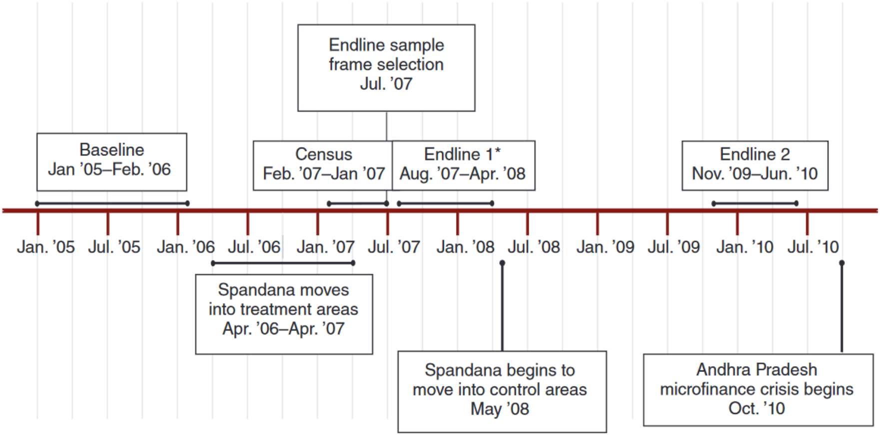
Potential Progress

- Comparation among different methods.
  - Probit, IV probit, Bivariate probit (for Binary outcome)
- Exploration in the principle of `LARF`.
  - How to explain and interpret the coefficient provided by `LARF`?
- Application and Replication using `LARF` pacakge.
  - Angrist, 1990: Veteran (Treatment) and Draft (IV)
  - Acemoglu, 2001: Modern Institute (Treatment) and Mortality (IV)
  - Binary outcome variable example...

Table of Contents

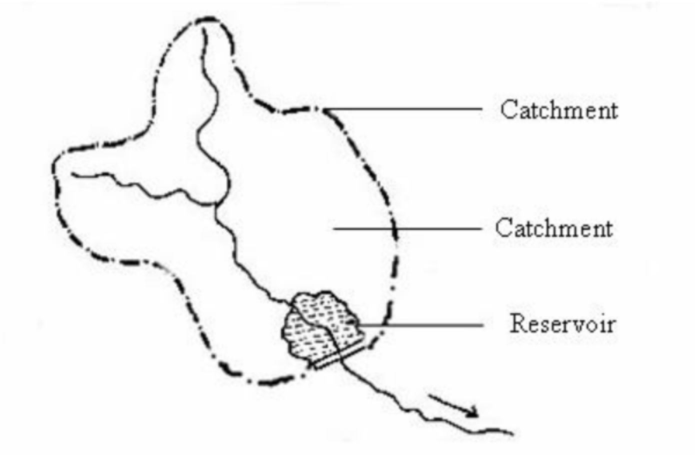
Banerjee et al. (2015) Back

Figure: Timeline of data collection and randomization



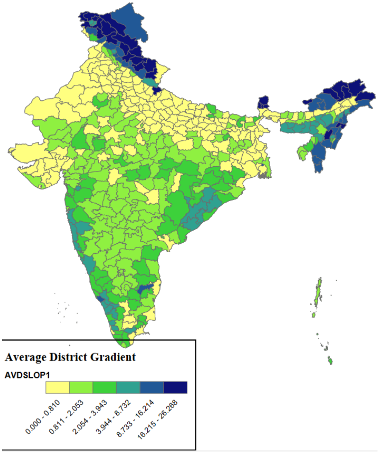
Duflo et al. (2007) Back

Figure: Illustration graph for an Irrigation Dam



Duflo et al. (2007) Back

Figure: the distribution of gradient



Reference

- Slides and Data of Banerjee et al. (2015).
- Slides and Data of Duflo et al. (2007).