Advanced Econometrics

University of Warsaw

Jerzy Mycielski

Doctoral School of Social Sciences, 2026

Regression Discontinuity Design (RDD)

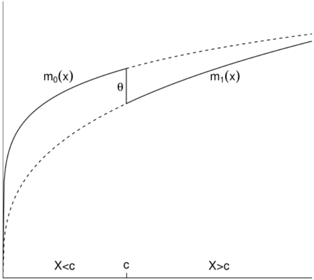
- We say that estimation of causal effect is based on RRD when treatment is determined by a threshold crossing rule.
- For example: college students admitted based on an admission exam.
- We can suppose that students which are close to the trashold (below and above) are similar to each other (as if randomly assigned)
- The covariate X which determine treatment is called running variable.
- The rule which is determining the treatment: $D = 1\{X \geq c\}$
- Let $m_0(x) = E(Y_0 | X = x)$, $m_1(x) = E(Y_1 | X = x)$
- If $m_0(x)$, $m_1(x)$ are continous at c, the causal effect at c:

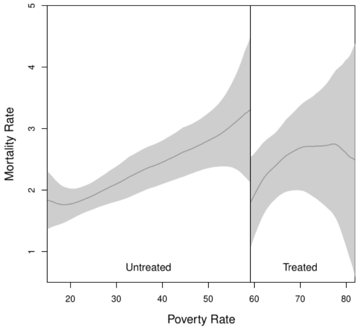
$$\overline{\theta} = E(Y_1 | X = c) - E(Y_0 | X = c) = m(c+) - m(c-)$$

where $m(c-) = lim_{x \uparrow c} m(x), m(c+) = lim_{x \downarrow c} m(x)$

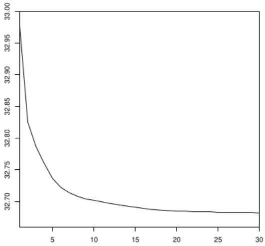
Sharp RDD

- Measuring effects of Head Start program, Ludwig and Miller (2007):
  - provided to help poor children age 3-5 and their families (data for 1965).
  - 300 poorest counties were provided assistance when applying for grants (cut-off c is 59.1984% poverty rate).
  - Outcome variables: mortality (deaths related to poverty, e.g. tuberculosis) and education.

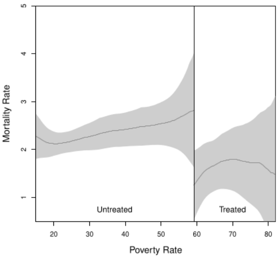




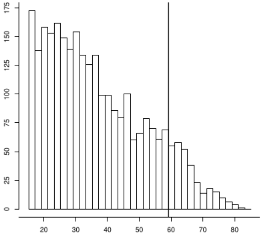
Head start program, RDD



(a) Cross-Validation Function



(b) RDD with Covariates



(c) Histogram of Poverty Rate

Source: Hansen (2022)

Estimation of causal effect, RDD

- Estimation of $m_0(x)$ and $m_1(x)$ should be based on nonparametric methods.
- Otherwise we may confuse the causal effect with the nonlinearity of the model.
- It is recomended to use the Local Linear estimator rather than Nadaray-Watson estimator as the NW estimates are inprecise on the bounds.
- The bandwith selection is based on ROT or CV criterions in a model with level shift added.
- We can also generalise this model to partially linear model of Robinson (1988).
- We can test the continuity of $m(x)$ at c by testing the continuity of $f(x)$ at c (density dicontinuity test). If $f(x)$ is discontinous than it can be treated as the evidence of running variable manipulation.

Head start program, results

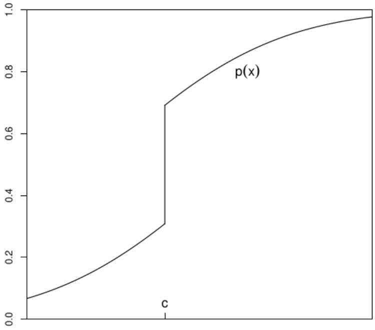
	Baseline	Covariates
$\widehat{\theta}$	-1.51	-1.56
$s(\widehat{\theta})$	(0.71)	(0.71)
% Black		0.027
$s(\widehat{\beta}_1)$		(0.007)
% Urban		-0.0094
$s(\widehat{\beta}_2)$		(0.0046)

Source: Hansen (2022)

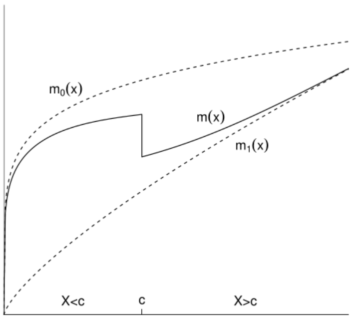
Fuzzy RDD (FRD)

- $Y_0,\ Y_1$ outcomes without and with treatment
- $\theta = Y_1 - Y_2$ treatment effect
- X running variable
- Conditional treatment effect at cutoff $\overline{\theta} = \mathbb{E}\left[\theta \middle| X = c\right]$
- D treatment indicator
- Confitional probability of treatment $p\left(x\right) = P\left[D = 1 \middle| X = x\right]$
- Left and right limits at cutoff $p\left(c+\right)$, $p\left(c-\right)$
- For FRD we assume that $p\left(c+\right) \neq p\left(c-\right)$

Fuzzy RDD (FRD)



(a) Conditional Treatment Probability



(b) Fuzzy Regression Discontinuity

Source: Hansen (2022)

Fuzzy RDD

Theorem

Assume that $m_0(x)$ and $m_1(x)$ are continous at $x = c$, $p(x)$ is discontinous at $x = c$, and D is independent of θ for X close to c. Then

$$\bar{\theta} = \frac{m(c+) - m(c-)}{p(c+) - p(c-)}$$

- Indeed

$$\begin{aligned} m(x) &= (1 - p(x)) m_0(x) + p(x) m_1(x) \\ &= m_0(x) + p(x)(m_1(x) - m_0(x)) \end{aligned}$$

then

$$m(c+) = m_0(c+) + p(c+)\theta$$

$$m(c-) = m_0(c-) + p(c-)\theta$$

and solving for θ we obtain the formula.

Estimation of FRD

- Estimates of $m(c+)$ and $m(c-)$ are obtained using local linear regression of Y on X.
- Estimates of $p(c+)$ and $p(c-)$ are obtained using local linear regression of D on X.
- We use the estimator

$$\widehat{\overline{\theta}} = \frac{\widehat{m}(c+) - \widehat{m}(c-)}{\widehat{p}(c+) - \widehat{p}(c-)}$$

- Identification of $\overline{\theta}$ depends on discontinuity of $p(x)$ at $x = c$. Small discontinuity will result in weak identification.
- What we should observe is that $\widehat{m}(x)$ and $\widehat{p}(x)$ are both discontinous at $x = c$
- Standard errors can be calculated as follows

$$\widehat{se}\left(\widehat{\overline{\theta}}\right) = \frac{\widehat{se}\left(\widehat{m}(c+) - \widehat{m}(c-)\right)}{\widehat{p}(c+) - \widehat{p}(c-)}$$

- Rule based enrollment. Instrumental variable classsize based Maimonides rule:

$$f_j = \frac{r_j}{1 + \lfloor (r_j - 1)/\, 40 \rfloor}$$

where r_j is the enrollment of fifth graders at school j and $\lfloor x \rfloor$ is the largest integer less or equal to x

- First stage regression of the class size s_{ij} attended by student i on f_j and control variables:

$$s_{ij} = \pi f_j + \rho_1 r_j + \delta_1^{'} X_{ij} + \gamma_t + \varepsilon_{ij}$$

where X_{ij} is a vector of school and student characteristics.

- Second stage model

$$y_{ij} = \beta s_{ij} + \rho_2 r_j + \delta_2^{'} X_{ij} + \mu_T + \eta_{ij}$$

y_{ij} is reading score and β is the causal effect of interest.

- Functions $h_1\left(\right)$ and $h_2\left(\right)$ are polynomial functions controling for running variable.

Maimonides rule

FIGURE 1.—Fourth grade class size by initial enrollment, actual average size and as predicted by Maimonides' Rule. *Notes*: Adapted from Angrist and Lavy (1999); data from Israel in 1991.

Source: Angrist (2022)

Class size and students achievements

TABLE IV
2SLS ESTIMATES FOR 1991 (FIFTH GRADERS)

	Reading comprehension						Math					
	Full sample				+/− 5 Discontinuity sample		Full sample				+/− 5 Discontinuity sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mean score		74.4			74.5			67.3			67.0	
(s.d.)		(7.7)			(8.2)			(9.6)			(10.2)	
Regressors												
Class size	−.158	−.275	−.260	−.186	−.410	−.582	−.013	−.230	−.261	−.202	−.185	−.443
	(.040)	(.066)	(.081)	(.104)	(.113)	(.181)	(.056)	(.092)	(.113)	(.131)	(.151)	(.236)
Percent disadvantaged	−.372	−.369	−.369		−.477	−.461	−.355	−.350	−.350		−.459	−.435
	(.014)	(.014)	(.013)		(.037)	(.037)	(.019)	(.019)	(.019)		(.049)	(.049)
Enrollment		.022	.012			.053		.041	.062			.079
		(.009)	(.026)			(.028)		(.012)	(.037)			(.036)
Enrollment squared/100			.005						−.010			
			(.011)						(.016)			
Piecewise linear trend				.136						.193		
				(.032)						(.040)		
Root MSE	6.15	6.23	6.22	7.71	6.79	7.15	8.34	8.40	8.42	9.49	8.79	9.10
N		2019		1961	471			2018		1960	471	

The unit of observation is the average score in the class. Standard errors are reported in parentheses. Standard errors were corrected for within-school correlation between classes. All estimates use f_{sc} as an instrument for class size.

References I

Angrist, Joshua D (2022). "Empirical strategies in economics: Illuminating the path from cause to effect". In: *Econometrica* 90.6, pp. 2509–2539.

Angrist, Joshua D et al. (2019). "Maimonides rule redux". In: *American Economic Review: Insights* 1.3, pp. 309–324.

Hansen, B. (2022). *Econometrics*. Princeton University Press. ISBN: 9780691235899.

Ludwig, Jens and Douglas L Miller (2007). "Does Head Start improve children's life chances? Evidence from a regression discontinuity design". In: *The Quarterly journal of economics* 122.1, pp. 159–208.

Robinson, P. M. (1988). "Root-N-Consistent Semiparametric Regression". In: *Econometrica* 56.4, pp. 931–954. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1912705 (visited on 04/20/2026).