

10: Regressionsdiskontinuitetsdesigns

Videregående kvantitative metoder i studiet af politisk adfærd

Frederik Hjorth

fh@ifs.ku.dk

fghjorth.github.io

@fghjorth

Institut for Statskundskab

Københavns Universitet

14. november 2016

- 1 Formalia
- 2 Opsamling fra sidst
- 3 Regressionsdiskontinuitetsdesigns
- 4 Eksempler på RD-designs
- 5 Case: Eggers & Hainmueller
- 6 Kig fremad

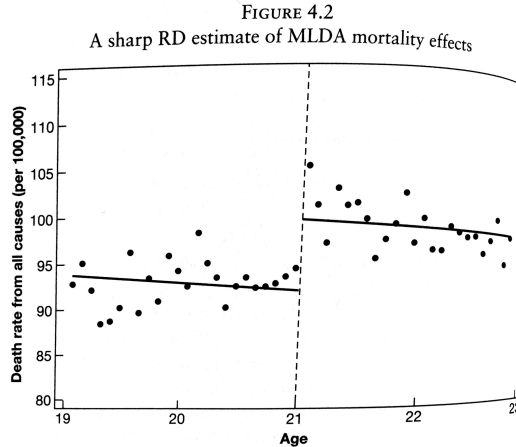
Uge	Dato	Tema	Litteratur	Case
1	5/9	Introduktion til R	Imai kap 1	
2	12/9	Regression I: OLS	GH kap 3, MM kap 2	Gilens & Page (2014)
3	26/9	Regression II: Paneldata	GH kap 11	Larsen et al. (2016)
4	29/9	Regression III: Multileveldata, interaktioner	GH kap 12	Berkman & Plutzer
5	3/10	Introduktion til kausal inferens	Hariri (2012), Samii (2016)	
6	10/10	Matching	Justesen & Klemmensen (2014)	Ladd & Lenz (2009)
	17/10	*Efterårsferie*		

Uge	Dato	Tema	Litteratur	Case
	17/10	*Efterårsferie*		
7	24/10	Eksperimenter I	MM kap 1, GG kap 1+2	Gerber et al. (2008)
8	31/10	Eksperimenter II	GG kap 3+4+5	Gerber & Green (2000)
9	14/11	Instrumentvariable	MM kap 3	Arunachalam & Watson
10	14/11	Regressionsdiskontinuitetsdesigns	MM kap 4	Eggers & Hainmueller
11	21/11	Difference-in-difference designs	MM kap 5	Enos (2016)
12	28/11	'Big data' og maskinlæring	Grimmer (2015), Varian (2014)	
13	5/12	Scraping af data fra online-kilder	MRMN kap 9	
14	12/12	Tekst som data	Grimmer & Stewart (2013), Imai kap 5	

- kriterier for instrumentvaliditet
- first stage, second stage, reduced form
- LATE ctr. TOT
- case: Arunachalam & Watson

Spørgsmål?

Motiverende eksempel: mortalitet og alkoholaldersgrænse



Formelt:

$$D_a = \begin{cases} 1 & \text{if } a \geq 21 \\ 0 & \text{if } a < 21 \end{cases}$$

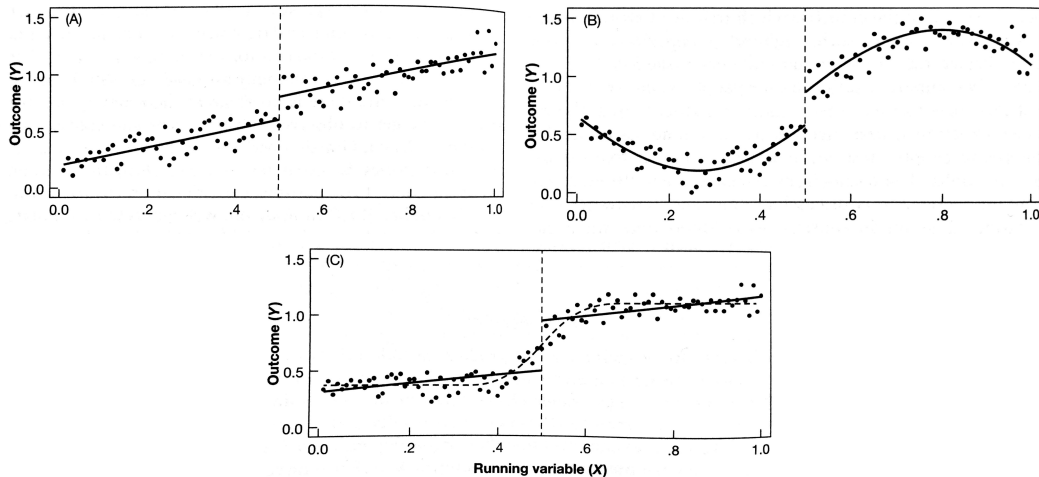
Outcome varierer med treatment D_a , men kan også variere med running/forcing variable a

RD-model:

$$\bar{M}_a = \alpha + \rho D_a + \gamma a + e_a \quad (1)$$

→ hvornår kan vi stole på at ρ fanger effekten af D_a ?

Udfordring for RDD: running-variablens funktionelle form



Model (1) estimeres i et smalt 'vindue' rundt om cutoff \rightarrow i fastsættelsen af vinduets 'bandwidth', et klassisk tradeoff:

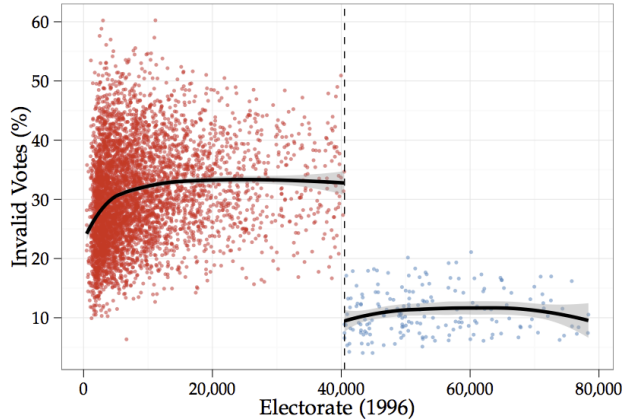
»[I]f the window is very narrow, there are few observations left, meaning the resulting estimates are likely to be too imprecise to be useful. Still, we should be able to trade the reduction in bias near the boundary against the increased variance suffered by throwing data away, generating some kind of optimal window size.« (161)

Metoder til optimal bandwidth selection, fx. Imbens, G., & Kalyanaraman, K. (2011). "Optimal bandwidth choice for the regression discontinuity estimator". *The Review of economic studies*.

Samii om RD (o.a.) -designs' 'localness':

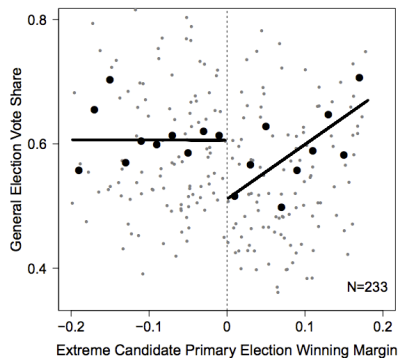
»The LATE theorem states that under a set of basic identifying conditions, an instrumental variable identifies the average causal effect for the subpopulation of units whose treatment status is in fact moved by the instrument. (...) Similarly, regression discontinuity identifies effects local to the relevant cut points, matching with calipers identifies effects local to the region of common covariate support, experiments identify effects local to the typically nonrepresentative sample of experimental subjects, and so on.« (950)

Hidalgo (2010): elektronisk stemmeafgivning og valide stemmer



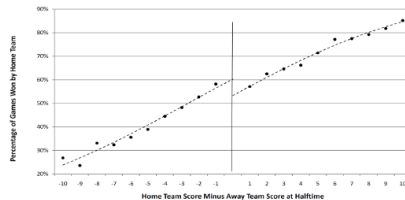
Hall (2014): ideologisk ekstreme primærvindere og valgresultater

Figure 2 – General-Election Vote Share After Close Primary Elections Between Moderates and Extremists: U.S. House, 1980–2010. The close election of the more extreme primary candidate causes a decrease in general-election vote share for the party.

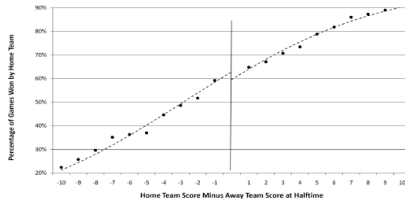


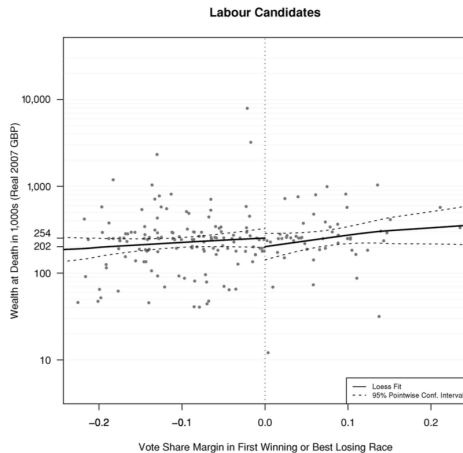
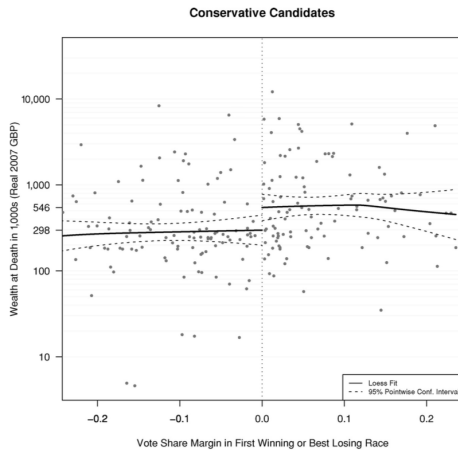
Berger & Pope (2011): er det bedre at være bagud ved pausen?

1a. NBA DATA



1b. NCAA DATA





Næste gang:

- difference-in-differences designs
- MM kap. 5
- case: Enos

Tak for i dag!