7: Eksperimenter I

Videregående kvantitative metoder i studiet af politisk adfærd

Frederik Hjorth
fh@ifs.ku.dk
fghjorth.github.io
@fghjorth

Institut for Statskundskab
Københavns Universitet

25. oktober 2017

Fagets opbygning

Blok 1

Gang	Tema	Litteratur	Case
1	Introduktion til R	Leeper (2016)	
2	R workshop + tidy data	Wickham (2014), Zhang (2017)	
3	Regression I: OLS brush-up	AP kap 3	Newman et al. (2015), Solt et al. (2017)
4	Regression II: Paneldata	AGS kap 4	Larsen et al. (2017)

Fagets opbygning

Blok 2

5	Introduktion til kausal inferens	Hariri (2012), Samii (2016)	Eckles & Bakshy (2017)
6	Matching	Justesen & Klemmensen (2014)	Nall (2015)
Efterårsferie			
7	Eksperimenter I	AP kap 1+2, GG kap 1+2	Gerber, Green & Larimer (2008)
8	Eksperimenter II	GG kap 3+4+5	Gerber & Green (2000)
9	Instrumentvariable	AP kap 4	Lundborg et al. (2017)
10	Difference-in-differences	AP kap 5	Enos (2016)
11	Regressionsdiskontinuitetsdesigns	AP kap 6	Eggers & Hainmueller (2009)

Fagets opbygning

Blok 3

12	Tekst som data	Grimmer & Stewart (2013), Benoit & Nulty (2016)	Baturo & Mikhaylov (2013)
13	Scraping af data fra online-kilder	MRMN kap 9+14	Hjorth (2016)
14	'Big data' og maskinlæring	Varian (2014), Montgomery & Olivella (2017)	Theocharis et al. (2016)

Halvvejsevaluering:

https://fghjorth.typeform.com/to/LioGD4

- introduktion til potential outcomes framework
- problemer ved ekstrapolation og modelafhængighed
- common support
- propensity score matching
- coarsened exact matching
- balance checks
- case: Nall (2015)

»Our view is that regression can be motivated as a particular sort of weighted matching estimator, and therefore the differences between regression and matching estimates are unlikely to be of major empirical importance.« (70)

Motiverende eksempel: NHIS

Group	Sample Size	Mean Health Status	Std. Error
Hospital	7,774	3.21	.014
No hospital	90,049	3.93	.003

→ hvad indikerer dette om effekten af hospitaler på sundhed?

To studerende, **K**huzdar & **M**aria

- Y_{1i}: potentielt outcome hvis *i* tager på hospitalet
- Y_{0i}: potentielt outcome hvis *i ikke* tager på hospitalet
- i dette case: hospitalisering gavner Khuzdar, men ikke Maria

$$Y_{1K} - Y_{0K} = 4 - 3 = 1 \tag{1}$$

$$Y_{1M} - Y_{0M} = 5 - 5 = 0 \tag{2}$$

Fuldt potential outcomes schedule:

	Khuzdar	Maria
Y_{0i}	3	5
Y_{1i}	4	5
D_i	1	0
Y_i	4	5
$Y_{1i} - Y_{0i}$	1	0

Observerede outcomes:

	Khuzdar	Maria
Y_{0i}	?	5
Y_{1i}	4	?

$$\rightarrow \bar{Y}_1 - \bar{Y}_0 = 4 - 5 = -1$$

Den direkte sammenligning afspejler både ATE hos de treatede + selection bias:

$$
\begin{aligned}
Y_K - Y_M &= Y_{1K} - Y_{0M} && (3)\\
&= Y_{1K} - Y_{0K} + Y_{0K} - Y_{0M} && (4)\\
&= 1 + (-2) && (5)\\
&= -1 && (6)
\end{aligned}
$$

M. mere generel notation i AP+GG:

$$E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 0] = \\ E[Y_{1i} - Y_{0i}|D_i = 1] + E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0] \quad (7)$$

når treatment randomiseres er $Y_0 i$ uafhængig af D_i:

$$E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0] = 0 \quad (8)$$

m.a.o.: random assignment eliminerer selection bias

»This does not mean that randomized trials are problem-free, but that in principle they solve the most important problem that arises in empirical research.« (15)

Succesfuld randomisering kan efterprøves m. *balance tests*

TABLE 2.2.1
Comparison of treatment and control characteristics in the Tennessee STAR experiment

	Class Size			*P*-value for equality
Variable	Small	Regular	Regular/Aide	across groups
Free lunch	.47	.48	.50	.09
White/Asian	.68	.67	.66	.26
Age in 1985	5.44	5.43	5.42	.32
Attrition rate	.49	.52	.53	.02
Class size in kindergarten	15.10	22.40	22.80	.00
Percentile score in kindergarten	54.70	48.90	50.00	.00

Notes: Adapted from Krueger (1999), table I. The table shows means of variables by treatment status for the sample of students who entered STAR in kindergarten. The *P*-value in the last column is for the *F*-test of equality of variable means across all three groups. The free lunch variable is the fraction receiving a free lunch. The percentile score is the average percentile score on three Stanford Achievement

Regressionsanalyse af randomiseret treatment D_i:

$$Y_i = \alpha + \rho D_i + \eta_i \tag{9}$$

Med tilføjet vektor af kontrolvariable X_i':

$$Y_i = \alpha + \rho D_i + X_i'\gamma + \eta_i \tag{10}$$

Hvorfor tilføjer vi X_i'?

1. For at rette op på kendte ubalancer i treatment
2. For at reducere residual varians i outcome $\rightarrow \downarrow$ standardfejl

$\rightarrow$ mere om dette i næste uge!

Resultater fra STAR eksperimentet:

TABLE 2.2.2
Experimental estimates of the effect of class size on test scores

Explanatory Variable	(1)	(2)	(3)	(4)
Small class	4.82	5.37	5.36	5.37
	(2.19)	(1.26)	(1.21)	(1.19)
Regular/aide class	.12	.29	.53	.31
	(2.23)	(1.13)	(1.09)	(1.07)
White/Asian	—	—	8.35	8.44
			(1.35)	(1.36)
Girl	—	—	4.48	4.39
			(.63)	(.63)
Free lunch	—	—	−13.15	−13.07
			(.77)	(.77)
White teacher	—	—	—	−.57
				(2.10)
Teacher experience	—	—	—	.26
				(.10)
Teacher Master's degree	—	—	—	−0.51
				(1.06)
School fixed effects	No	Yes	Yes	Yes

Gerber & Greens procedure:

»First, determine N, the number of subjects in your experiment, and m, the number of subjects who will be allocated to the treatment group. Second, set a random number 'seed' using a statistics package, so that your random numbers may be reproduced by anyone who cares to replicate your work. Third, generate a random number for each subject. Fourth, sort the subjects by the random numbers in ascending order. Finally, classify the first m observations as the treatment group.« (37)

To kritiske antagelser om potential outcomes:

1. excludability
2. non-interferens (SUTVA)

Ad (1):

Lad $Y_i(z, d)$ være potential outcome for treatment assignment $z_i = z$ og faktisk treatment status $d_i = d$

Eksklusionsrestriktionsantagelsen: $Y_i(1, d) = Y_i(0, d)$

Ad (2):

Lad $Y_i(\mathbf{z}, \mathbf{d})$ være PO for Y_i for den fulde mængde af assignments og treatments

Under non-interferens: $Y_i(\mathbf{z}, \mathbf{d}) = Y_i(z, d)$

TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election

	Experimental Group				
	Control	Civic Duty	Hawthorne	Self	Neighbors
Percentage Voting	29.7%	31.5%	32.2%	34.5%	37.8%
N of Individuals	191,243	38,218	38,204	38,218	38,201

Neighbors mailing

30423-3 ||| || || | |||

For more information: (517) 351-1975
email: etov@grebner.com
Practical Political Consulting
P. O. Box 6249
East Lansing, MI 48826

PRSRT STD
U.S. Postage
PAID
Lansing, MI
Permit # 444

ECRLOT **C050
THE JACKSON FAMILY
9999 MAPLE DR
FLINT MI 48507

Dear Registered Voter:

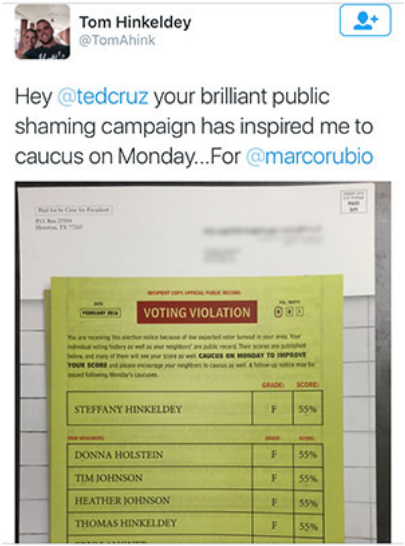
WHAT IF YOUR NEIGHBORS KNEW WHETHER YOU VOTED?

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse. This year, we're taking a new approach. We're sending this mailing to you and your neighbors to publicize who does and does not vote.

The chart shows the names of some of your neighbors, showing which have voted in the past. After the August 8 election, we intend to mail an updated chart. You and your neighbors will all know who voted and who did not.

DO YOUR CIVIC DUTY — VOTE!

MAPLE DR	Aug 04	Nov 04	Aug 06
9995 JOSEPH JAMES SMITH	Voted	Voted	_____
9995 JENNIFER KAY SMITH		Voted	_____
9997 RICHARD B JACKSON		Voted	_____
9999 KATHY MARIE JACKSON		Voted	_____


Tom Hinkeldey
@TomAhink
Hey @tedcruz your brilliant public shaming campaign has inspired me to caucus on Monday...For @marcorubio
VOTING VIOLATION
You are receiving this election notice because of low expected voter turnout in your area. Your individual voting history as well as your neighbors' are public record. Their scores are published below, and many of them will see your score as well. CAUCUS ON MONDAY TO IMPROVE YOUR SCORE and please encourage your neighbors to caucus as well. A follow-up notice may be issued following Monday's caucuses.
GRADE: SCORE:
STEFFANY HINKELDEY F 55%
DONNA HOLSTEIN F 55%
TIM JOHNSON F 55%
HEATHER JOHNSON F 55%
THOMAS HINKELDEY F 55%

Næste gang:

- eksperimenter II
- fokus: cluster random assignment, covariate adjustment, noncompliance
- case: Gerber & Green (2000)
- husk at udfylde halvvejsevaluering! `https://fghjorth.typeform.com/to/LioGD4`

Tak for i dag!