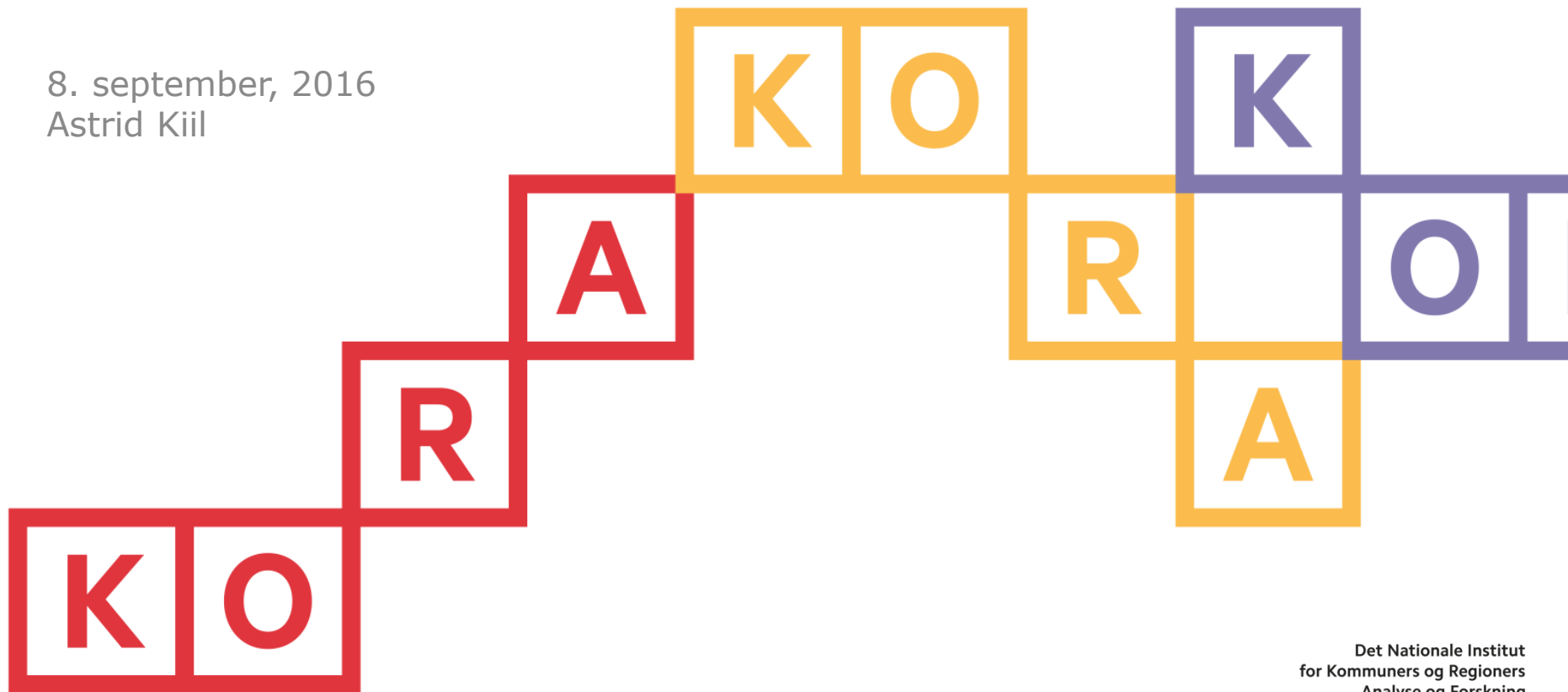


Dansk Evalueringselskabs Konference 2016 – læringsseminar 9:

Brug af matching metoder til identifikation af kausale effekter

8. september, 2016
Astrid Kiil





Kort om mig

- Cand.oecon. fra SDU i 2008
- Ph.d. i Sundhedsøkonomi fra SDU i 2012
- Ansat som forsker ved AKF i 2011 -> KORA
 - Specialist i kvantitativ metode, herunder forskellige former for regressioner og matching, benchmarking og økonomisk evaluering
 - Har udført opgaver for bl.a. kommuner, styrelser og fonde
 - Er pt. tilknyttet Trygfondens Børneforskningscenter på AU som postdoc

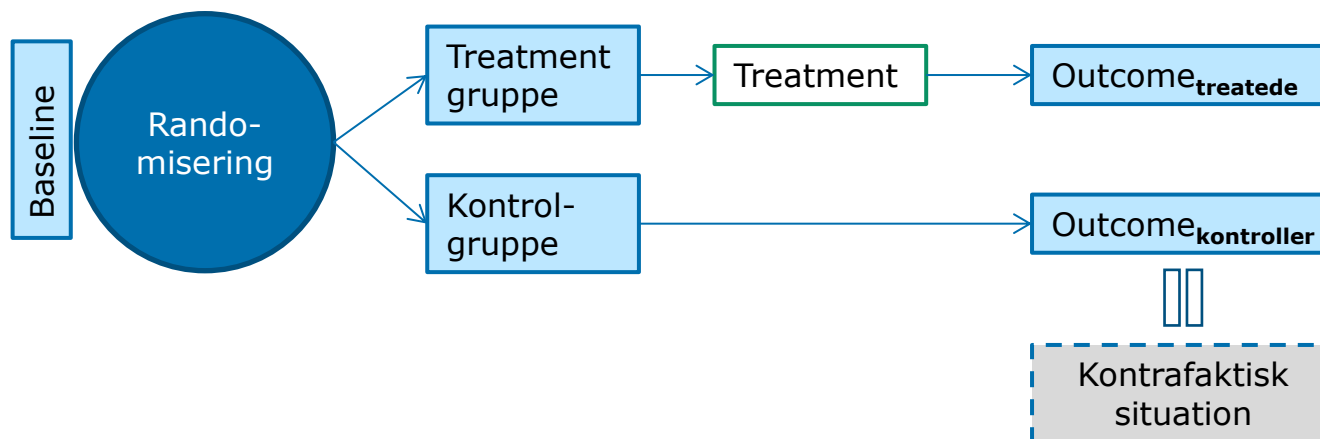
“Køreplan” for de næste 3 timer

- Modul 1: Introduktion og teori (45 min.)
 - Baggrund og intuitiv introduktion til propensity score matching
 - Estimering af propensity score og matching på propensity score
 - Metodens antagelser -> styrker og begrænsninger
 - Gruppearbejde: Diskuter eksempler fra jeres hverdag, hvor det kunne give mening at bruge propensity score matching
- Modul 2: Øvelser i Stata (60 min.)
- Modul 3: Eksempler på anvendelse (45 min.)
 - Does employment-based private health insurance increase the use of covered health care services? A matching estimator approach
 - Effekter af løntilskud på kommunale arbejdspladser
 - Evaluering af indsats for forløbskoordination: Erfaringer med fremskudt visitation i fire kommuner

Kausalitet og *the missing counterfactual problem*

- Kan illustreres med et lille konstrueret eksempel:
 - Hvad er effekten af at deltage i et kommunalt jobsøgningskursus for ledige?
 - En sammenligning af ledige kursusdeltagere og andre ledige viser, at kursusdeltagerne er dobbelt så tilbøjelig til at være i arbejde 1 måned efter kurset er afsluttet
 - Men hvad nu hvis kursusdeltagerne er yngre, bedre uddannede, har et lavere medicinforbrug og færre diagnoser, og har mere arbejdsmarkedserfaring end de ledige, der ikke deltog?
- Det væsentlige spørgsmål er, hvordan det var gået for kursusdeltagerne, hvis de ikke havde været på kursus?

RCT er *golden standard*



Men...

- Praktiske og etiske barrierer for kontrollerede randomiserede forsøg inden for samfundsvidenskaben
- I nogle situationer er der et tradeoff mellem intern og ekstern validitet

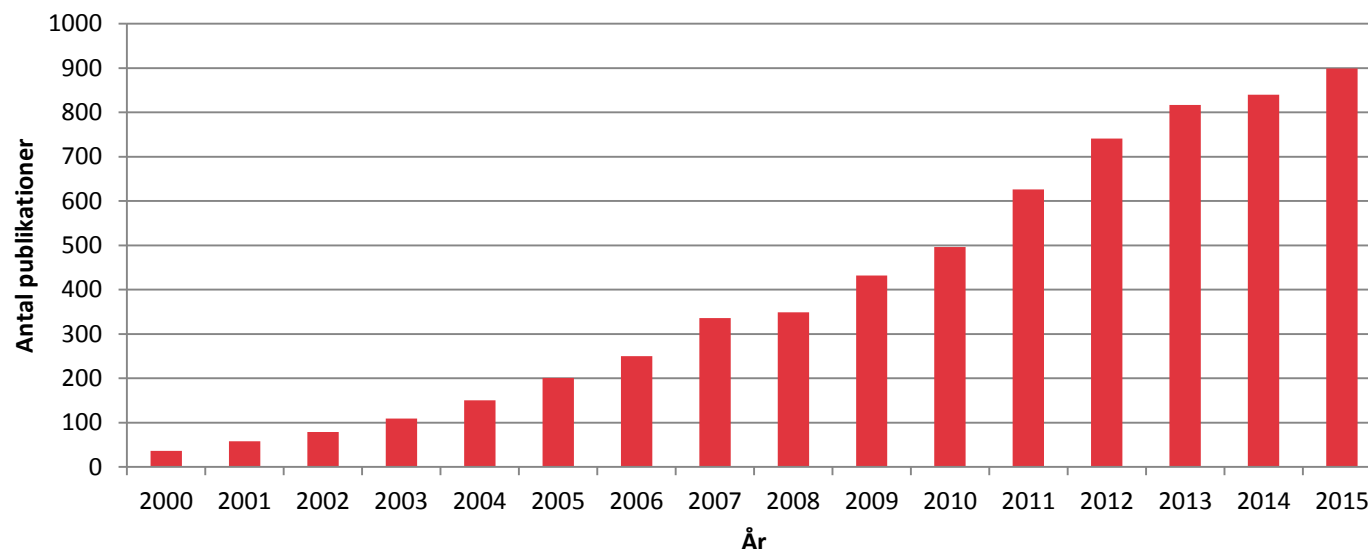


Observationelle metoder – alternativ til RCT

- Forskellige statistiske metoder til at håndtere den ærgerlige kendsgerning, at vi ikke kan observere i data, **hvad der ville være sket, hvis...**
 - Naturlige eksperimenter og diff-in-diff
 - Forskellige former for regression
 - Instrument-variabel estimation
 - **Propensity score matching**
 - Ikke-parametriske bounds

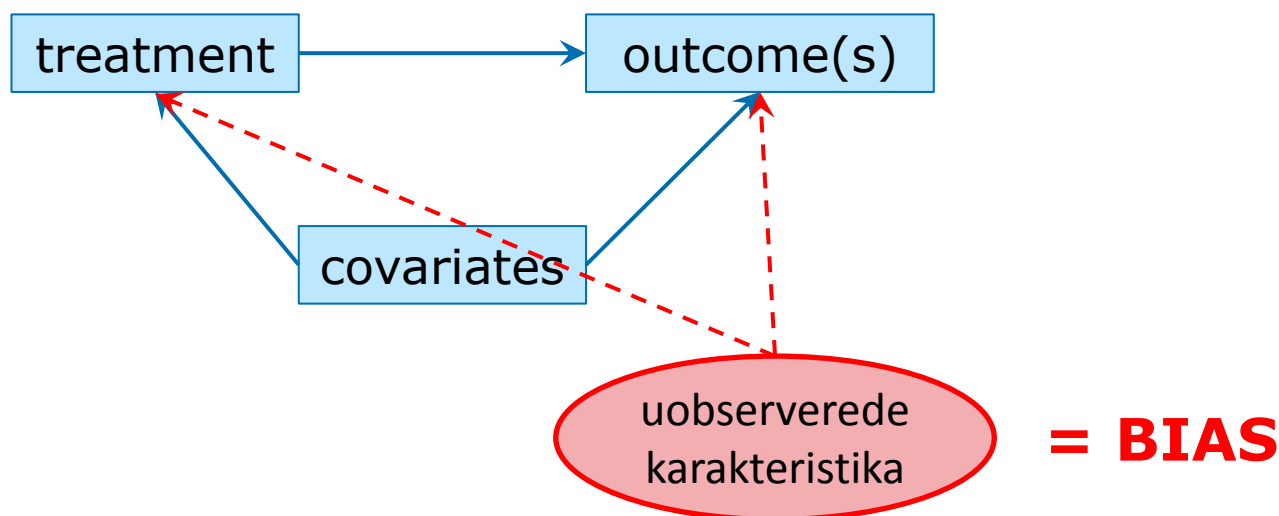
Metodens udbredelse inden for den samfundsvidenskabelige forskning

Publikationer i databaserne EconLit, Sociological Abstracts og ProQuest Dissertations & Theses, der indeholder sætningen "propensity score"



Matching: Den væsentlige identificerende antagelse

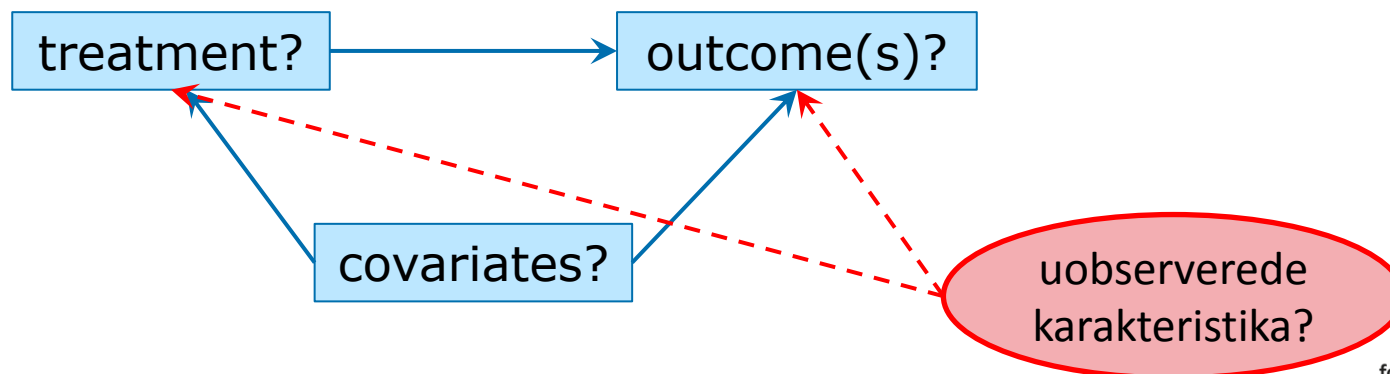
- Udelukkende selektion på observerede karakteristika
-> ingen selektion på uobserverede karakteristika



- Der er også andre antagelser, men dem venter vi lige lidt med...

Indledende overvejelser ved propensity score matching (PSM)

- Hvad vil jeg undersøge effekten af? **treatment**
- Hvad vil jeg undersøge effekten på? **outcome(s)**
- Har jeg en potentiel kontrolgruppe?
- Hvilke variable, som påvirker både treatment og outcome(s), har jeg data på både for treatede og kontroller? **covariates**
- Relevante uobserverede karakteristika? **risiko for bias**

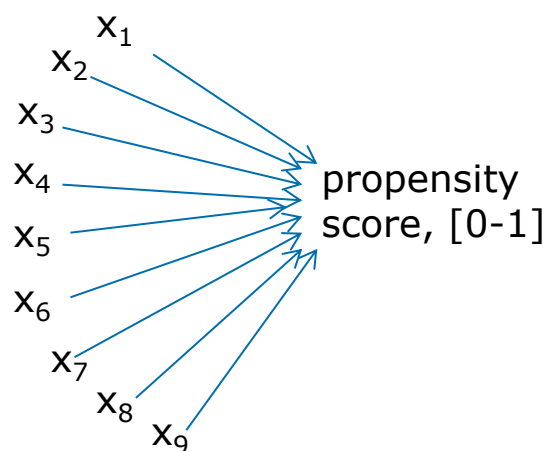


Implementering af PSM estimatoren

1. Identificere treatede, potentielle kontroller og covariates i data
2. Estimere en propensity score for alle observationer
3. Match hver treatede med en eller flere kontroller på propensity scoren
4. Estimere effekt ved at sammenligne gns. outcome for treatmentgruppen og den matchede kontrolgruppe
5. Beregne standardfejl for de estimerede effekter
6. Diskutere og teste hvor sandsynligt det er, at metodens antagelser holder

2. Estimering af propensity score

- Propensity scoren viser for hver observation sandsynligheden for at have modtaget treatment



Formålet er at koge informationen i de (typisk mange) covariates ned til et enkelt tal, som vi kan matche treatede og kontroller på

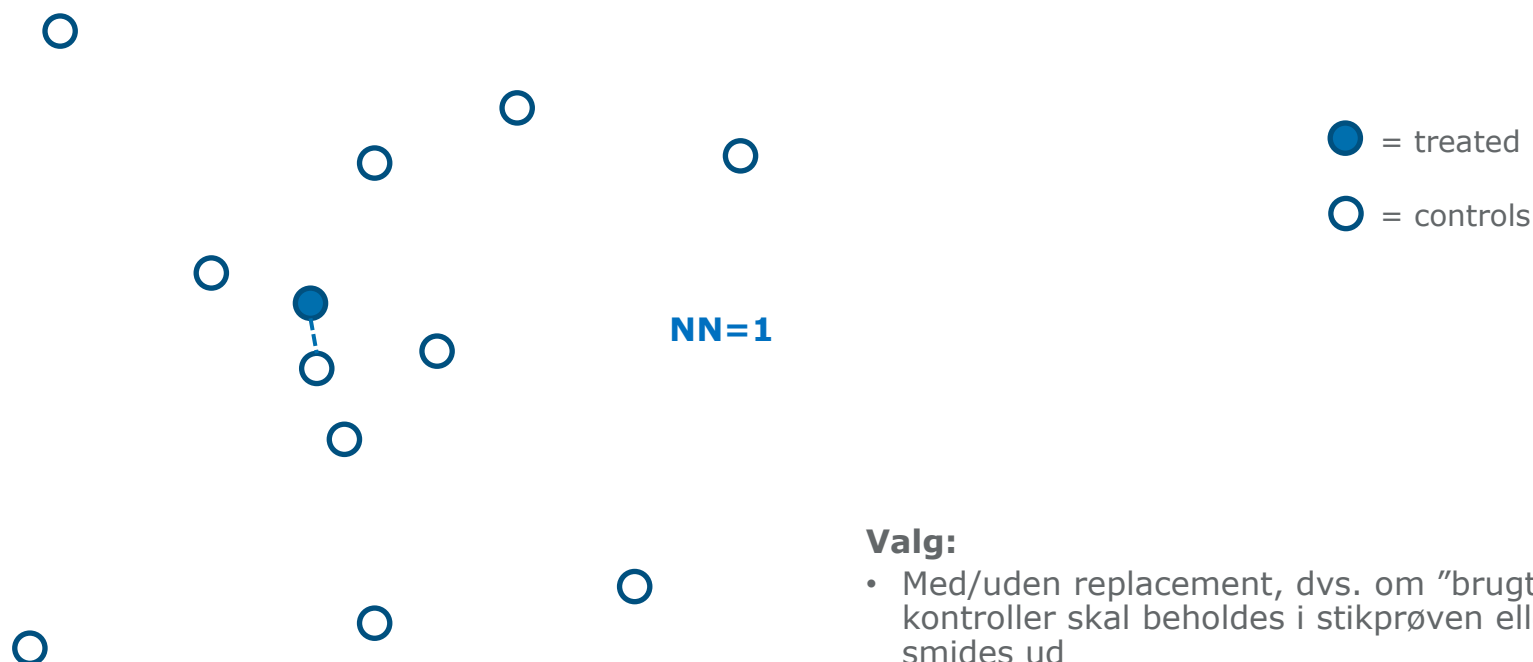
- I praksis estimeres propensity scoren ofte med logistisk eller probit regression (og er således parametrisk)

3. Matching på propensity scoren

- Treatede observationer matches med én eller flere kontroller med lignende propensity score, dvs. sandsynlighed for treatment
- Der findes forskellige algoritmer til formålet
 - Nearest Neighbour (NN)
 - Caliper & Radius
 - Stratification & Intervals
 - Kernel & Local Linear
 - Weighting

} illustreres
grafisk

Illustration af Nearest Neighbour (NN) matching algoritme

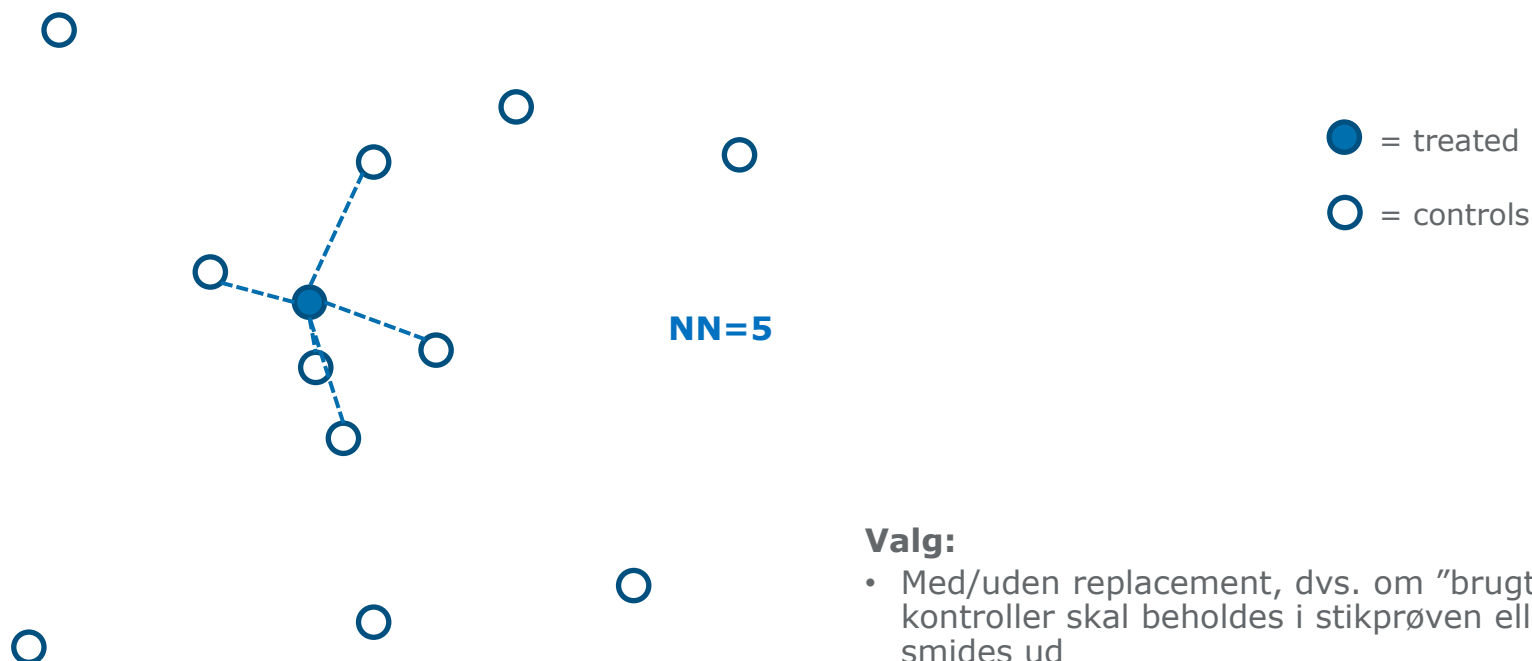


Valg:

- Med/uden replacement, dvs. om "brugte" kontroller skal beholdes i stikprøven eller smides ud
- Oversampling, dvs. hvor mange kontroller hver treated skal matches med (2-NN, 5-NN osv.)
- Vægtning af kontroller ved oversampling

Hvad nu hvis den nærmeste kontrol ligger helt her oppe?

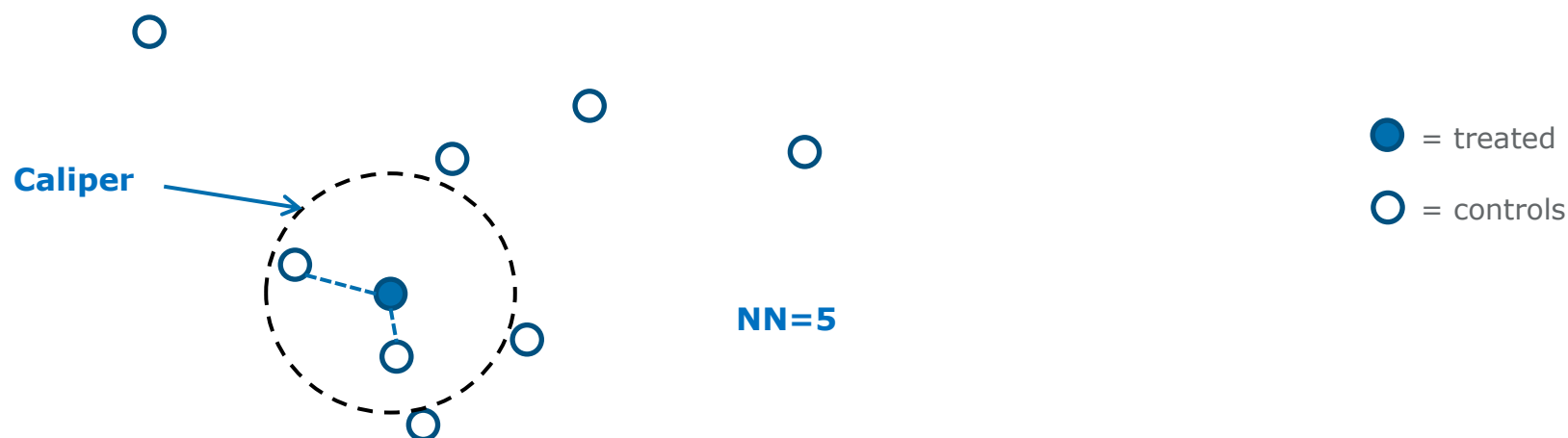
Illustration af Nearest Neighbour (NN) matching algoritme



Valg:

- Med/uden replacement, dvs. om "brugte" kontroller skal beholdes i stikprøven eller smides ud
- Oversampling, dvs. hvor mange kontroller hver treated skal matches med (2-NN, 5-NN osv.)
- Vægtning af kontroller ved oversampling

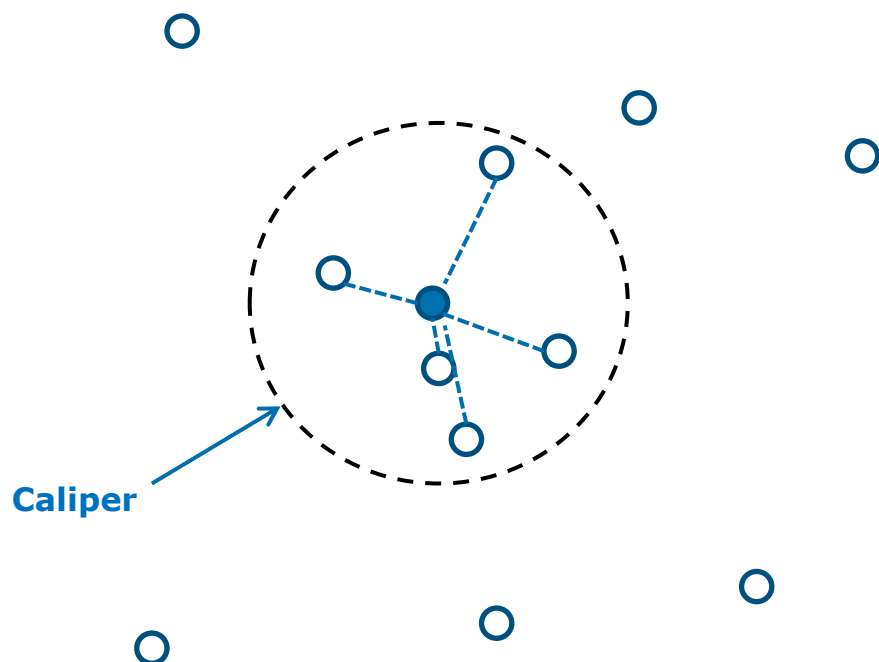
Illustration af Caliper matching algoritme



Valg:

- Oversampling, dvs. hvor mange kontroller hver treated skal matches med (2-NN, 5-NN osv.)
- Caliper, dvs. hvor langt væk man kan acceptere, at nærmeste nabo er

Illustration af Radius matching algoritme



● = treated

○ = controls

Valg:

- Caliper, dvs. inden for hvilken afstand man matcher med kontroller



Hvilken matching algoritme skal man vælge?

- Afhænger af data
- Tradeoff mellem bias og varians
- Generelt anbefales det at matche med replacement
 - God idé lige at tjekke, hvor mange gange den samme observation bruges som kontrol
- Pragmatisk tilgang: Start med at eksperimentere lidt med det og se, hvor meget det påvirker resultaterne
 - Undersøg hvad der foregår i "maskinrummet", hvis det er væsentligt for hvilken effekt vi kommer frem til

4.+5. Estimere effekt og beregne standardfejl

- Estimere effekter ved at sammenligne gns. outcome for treatment- og kontrolgruppe **efter matching**
 - ATT, average treatment effect for the treated
 - ATE, average treatment effect
- Beregne standardfejl for de estimerede effekter
 - Naive standardfejl (der ikke tager højde for at propensity score er estimeret)
 - Bootstrappede standardfejl
 - Analytiske standardfejl (Abadie-Imbens)



Ikke enighed om, hvilken af disse to metoder, der er korrekt

6. Diskutere og teste antagelser

- Den identificerende antagelse
 - Ingen selektion på uobserverede karakteristika: $(Y_1, Y_0) \perp D | X$

- Øvrige antagelser
 - Stable-unit treatment value (SUTVA)
 - Overlap/common support **-> kan testes**
 - Balance på covariates **-> kan testes**

Overlap/common support

- For enhver værdi af X , skal der være en positiv sandsynlighed for både at være både treated og kontrol, for ATT: $0 < P(D=1|X) < 1$
 - Sikrer at vi kun sammenligner de sammenlignelige observationer
- Testes ved visuel analyse af fordelingen af propensity scoren for treatment og kontrolgruppe
- Overlap/common support kan gennemtvinges på flere måder
 - Drop treatede observationer med propensity scores over maximum og under minimum for kontrolgruppens propensity scores
 - Drop en procentdel af de treatede observationer, hvor tætheden af propensity scoren for kontrollerne er lavest

Balance på covariates

- Siger noget om, hvorvidt den estimerede propensity score har succes med at eliminere forskelle i covariates mellem treatede og kontroller -> er det lykkedes at konstruere en sammenlignelig kontrolgruppe?
- Tjek for hver enkelt covariate og samlet set
- Efter matching må der ikke være nogen statistisk signifikante forskelle på covariates for treatede og kontroller
- Gns. bias skal reduceres mest muligt

Metodens styrker og begrænsninger

- Styrker
 - Særligt velegnet når man har mange covariates og en treatment, der forekommer relativt sjældent
 - Eksplicitte testbare antagelser om balance og common support
 - Det lykkedes nogle gange at formidle metode og resultater til et publikum uden de store statistisk kompetencer
 - Velegnet til matchede designs
- Begrænsninger
 - Bias hvis selektion på uobserverede karakteristika (ligesom regression) -> overvej om instrument variabel estimation er muligt
 - Fungerer bedst for større stikprøver
 - Inklusion af irrelevante covariates reducerer estimatorens efficiens

Matching vs. OLS regression

- Styrker ved matching metoden
 - Eksplicitte testbare antagelser om balance og common support sikrer, at vi danner en kontrolgruppe der reelt er sammenlignelig med vores treatmentgruppe.
 - Semi-parametrisk fordi der ikke skal specificeres funktionel form for sammenhængene mellem outcome og covariates (men dog typisk mellem treatment og covariates)
 - > men hvis OLS er korrekt specificeret, er den mere efficient

Særligt velegnet når man har mange covariates og en treatment, der forekommer relativt sjældent

- Begrænsninger ved matching metoden
 - **Bias hvis selektion på uobserverede karakteristika (ligesom OLS)**
 - Fungerer bedst for større stikprøver (rent teknisk er metoden "data hungry")
 - Inklusion af irrelevante covariates reducerer estimatorens efficiens

Gruppearbejde

- Find sammen i grupper med 3 personer
- Diskuter eksempler fra jeres hverdag, hvor det kunne give mening at bruge propensity score matching
 1. Hvad er treatment, outcome(s) og covariates?
 2. Hvem kan vi bruge som kontrolgruppe?
 3. Hvor stor er risiko for selektion på uobserverede karakteristika (som giver biased effektestimater)?
 4. Er vi mest interesserede i ATE eller ATT?



Modul 2: Øvelser

1. Find sammen i grupper med 3 personer – hver gruppe skal have en person, som har en bærbar med Stata installeret
 2. Lav opgaverne på det udleverede øvelsesark
 - OBS: Vi bruger Stata i dag, men man kan fint lave tilsvarende analyser med andet software, fx SAS, SPSS eller R
 - OBS: Øvelserne er konstrueret for at vise jer de mest basale kommandoer og output i Stata – i den virkelige verden vil de opgaver, man står overfor, typisk være mere komplekse og dermed også kræve mere omtanke og programmering
- Vi diskuterer øvelserne i fællesskab på "tavlen"

Leuven and Sianesi (2003) `psmatch2`

Mahalanobis and Propensity score Matching

```
psmatch2 depvar [indepvars] [if exp] [in range] [, outcome(varlist)
      pscore(varname) neighbor(integer) radius caliper(real)
      mahalanobis(varlist) ai(integer) kernel llr
      kerneltype(type) bwidth(real) spline nknots(integer)
      common trim(real) noreplacement descending odds index
      logit ties quietly w(matrix) ate]
```

1. Data

```

name: <unnamed>
log: U:\DES 2016 læringsseminar\DES2016.log
log type: text
opened on: 6 Sep 2016, 07:50:45

```

```

. *** 1. Data
. sysuse auto, clear
(1978 Automobile Data)

```

```

. describe

```

Contains data from C:\Program Files (x86)\Stata14\ado\base/a/auto.dta

```

obs:      74      1978 Automobile Data
vars:     12      13 Apr 2014 17:45
size:    3,182    (_dta has notes)

```

variable name	storage type	display format	value label	variable label
make	str18	%-18s		Make and Model
price	int	%8.0gc		Price
mpg	int	%8.0g		Mileage (mpg)
rep78	int	%8.0g		Repair Record 1978
headroom	float	%6.1f		Headroom (in.)
trunk	int	%8.0g		Trunk space (cu. ft.)
weight	int	%8.0gc		Weight (lbs.)
length	int	%8.0g		Length (in.)
turn	int	%8.0g		Turn circle (ft.)
displacement	int	%8.0g		Displacement (cu. in.)
gear_ratio	float	%6.2f		Gear Ratio
foreign	byte	%8.0g	origin	Car type

Sorted by: foreign

2. Beskrivende statistik for outcome

```
. *** 2. Beskrivende statistik
. ttest price, by(foreign)
```

Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
Domestic	52	6072.423	429.4911	3097.104	5210.184	6934.662
Foreign	22	6384.682	558.9942	2621.915	5222.19	7547.174
combined	74	6165.257	342.8719	2949.496	5481.914	6848.6
diff		-312.2587	754.4488		-1816.225	1191.708

diff = mean(Domestic) - mean(Foreign) t = -0.4139
 Ho: diff = 0 degrees of freedom = 72

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
 Pr(T < t) = 0.3401 Pr(|T| > |t|) = 0.6802 Pr(T > t) = 0.6599

```
. pstest price, raw treated(foreign)
```

variable	Mean		%bias	t-test	
	Treated	Control		t	p> t
price	6384.7	6072.4	10.9	0.41	0.680

2. Beskrivende statistik for covariates

```
. pstest mpg headroom weight length, raw treated(foreign)
```

variable	Mean		%bias	t-test	
	Treated	Control		t	p> t
mpg	24.773	19.827	86.0	3.63	0.001
headroom	2.6136	3.1538	-73.7	-2.61	0.011
weight	2315.9	3317.1	-172.9	-6.25	0.000
length	168.55	196.13	-160.8	-5.89	0.000

2. Beskrivende statistik – indledende OLS

```
. regress price foreign mpg headroom weight length
```

Source	SS	df	MS	Number of obs	=	74
Model	357606298	5	71521259.6	F(5, 68)	=	17.53
Residual	277459098	68	4080280.86	Prob > F	=	0.0000
				R-squared	=	0.5631
				Adj R-squared	=	0.5310
Total	635065396	73	8699525.97	Root MSE	=	2020

price	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
foreign	3545.345	649.9282	5.45	0.000	2248.433	4842.256
mpg	-14.6551	71.50353	-0.20	0.838	-157.3382	128.028
headroom	-481.9711	326.389	-1.48	0.144	-1133.27	169.328
weight	5.685127	1.007732	5.64	0.000	3.674229	7.696025
length	-82.31547	34.01132	-2.42	0.018	-150.184	-14.44694
_cons	5169.742	5202.938	0.99	0.324	-5212.559	15552.04

3. Propensity score matching

1-1 NN med replacement (beregnet både ATT, ATE og ATU)

```
. *** 3. Propensity score matching: 1-1 NN med replacement
. psmatch2 foreign mpg headroom weight length, outcome(price) neighbor(1) logit ate
```

```
Logistic regression      Number of obs      =      74
                        LR chi2(4)          =     35.72
                        Prob > chi2         =     0.0000
                        Pseudo R2          =     0.3966

Log likelihood = -27.174095
```

foreign	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
mpg	-.1687704	.0920296	-1.83	0.067	-.349145	.0116042
headroom	.0226512	.4916462	0.05	0.963	-.9409577	.98626
weight	-.0039211	.0019534	-2.01	0.045	-.0077497	-.0000925
length	.0000698	.0590076	0.00	0.999	-.115583	.1157226
_cons	13.67663	7.725704	1.77	0.077	-1.465473	28.81873

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
price	unmatched	6384.68182	6072.42308	312.258741	754.448837	0.41
	ATT	6384.68182	4318.09091	2066.59091	628.275353	3.29
	ATU	6072.42308	11122.1731	5049.75	.	.
	ATE			4162.86486	.	.

Note: S.E. does not take into account that the propensity score is estimated.

psmatch2: Treatment assignment	psmatch2: Common support on support	Total
Untreated	52	52
Treated	22	22
Total	74	74

4. Tjek balance for hver enkelt covariate

```
. *** 4. Tjek balance
. pstest mpg headroom weight length, both sum
```

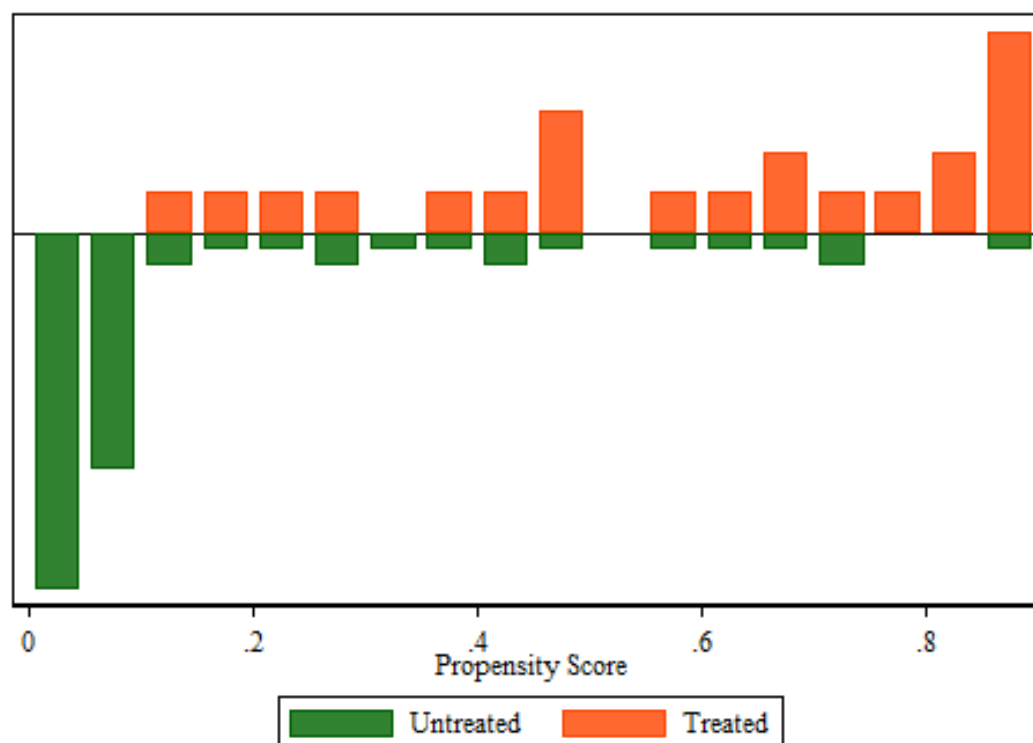
variable	Unmatched Matched	Mean		%bias	%reduct bias	t-test	
		Treated	Control			t	p> t
mpg	U	24.773	19.827	86.0		3.63	0.001
	M	24.773	25.773	-17.4	79.8	-0.62	0.537
headroom	U	2.6136	3.1538	-73.7		-2.61	0.011
	M	2.6136	2.2273	52.7	28.5	2.02	0.049
weight	U	2315.9	3317.1	-172.9		-6.25	0.000
	M	2315.9	2265.5	8.7	95.0	0.38	0.704
length	U	168.55	196.13	-160.8		-5.89	0.000
	M	168.55	165.36	18.5	88.5	0.72	0.475

4. Tjek balance samlet set

Summary of the distribution of the abs(bias)					
BEFORE MATCHING					
	Percentiles	Smallest			
1%	73.68111	73.68111			
5%	73.68111	85.96092			
10%	73.68111	160.7589	obs		4
25%	79.82102	172.8502	Sum of wgt.		4
50%	123.3599		Mean	123.3128	
		Largest	Std. Dev.	50.71041	
75%	166.8046	73.68111			
90%	172.8502	85.96092	Variance	2571.545	
95%	172.8502	160.7589	Skewness	-.0008845	
99%	172.8502	172.8502	Kurtosis	1.075513	
AFTER MATCHING					
	Percentiles	Smallest			
1%	8.710572	8.710572			
5%	8.710572	17.38058			
10%	8.710572	18.54009	obs		4
25%	13.04557	52.6975	Sum of wgt.		4
50%	17.96033		Mean	24.33219	
		Largest	Std. Dev.	19.41219	
75%	35.6188	8.710572			
90%	52.6975	17.38058	Variance	376.833	
95%	52.6975	18.54009	Skewness	.9723596	
99%	52.6975	52.6975	Kurtosis	2.223364	
Sample	Pseudo R2	LR chi2	p>chi2	MeanBias	MedBias
Raw	0.404	36.40	0.000	123.3	123.4
Matched	0.092	5.58	0.233	24.3	18.0

5. Tjek common support

psmatch2: Treatment assignment	psmatch2: Common support On suppor	Total
untreated	52	52
Treated	22	22
Total	74	74



6. Gentag øvelse 3-5 for forskellige matching algoritmer

	ATT	t-værdi (systematisk forkert)	Pseudo R ² efter matching	Gns. bias efter matching	# cov. med sig. forskelle efter matching	Treatede off common support
"Rå" sammenligning	312,26	0,41	n.a.	n.a.	n.a.	n.a.
OLS regression	3545,35	5,45	n.a.	n.a.	n.a.	n.a.
1-1 NN med replacement	2066,59	3,29	0,092	24,3	1	0
1-1 NN uden replacement	2040,86	3,53	0,157	47,4	2	0
1-5 NN med replacement	2133,21	3,60	0,029	9,4	0	0
1-5 NN med replacement og caliper på 0.15	2068,31	3,46	0,026	11,8	0	0
Radius matching inden for caliper på 0.25	1867,57	1,89	0,044	16,1	0	0
Kernel matching med epanechnikov kernel og bandwidth på 0.04	2094,43	3,16	0,057	13,8	0	2



Ekstra: Hvad sker der i "maskinrummet" ved de forskellige matching algoritmer

```
. tab _weight if foreign==1
```

psmatch2: weight of matched controls	Freq.	Percent	Cum.
1	22	100.00	100.00
Total	22	100.00	

```
. tab _weight if foreign==0
```

psmatch2: weight of matched controls	Freq.	Percent	Cum.
1	9	69.23	69.23
2	3	23.08	92.31
7	1	7.69	100.00
Total	13	100.00	

```
. list _pscore _weight if _weight>1 & _weight!=.
```

	_pscore	_weight
7.	.6512013	2
24.	.87405316	7
25.	.44998885	2
29.	.48174447	2

Modul 3: Eksempler på anvendelse

1. Does employment-based private health insurance increase the use of covered health care services? A matching estimator approach (*videnskabelig artikel*)
2. Effekter af løntilskud på kommunale arbejdspladser (*KORA-rapport*)
3. Evaluering af indsats for forløbskoordination: Erfaringer med fremskudt visitation i fire kommuner (*KORA-rapport*)

Eksempel 1

[International Journal of Health Care Finance and Economics](#)

March 2012, Volume 12, [Issue 1](#), pp 1–38

Does employment-based private health insurance increase the use of covered health care services? A matching estimator approach

Astrid Kiil 

Article

First Online: 26 February 2012

DOI: 10.1007/s10754-012-9104-3

Cite this article as:

Kiil, A. Int J Health Care Finance Econ (2012)

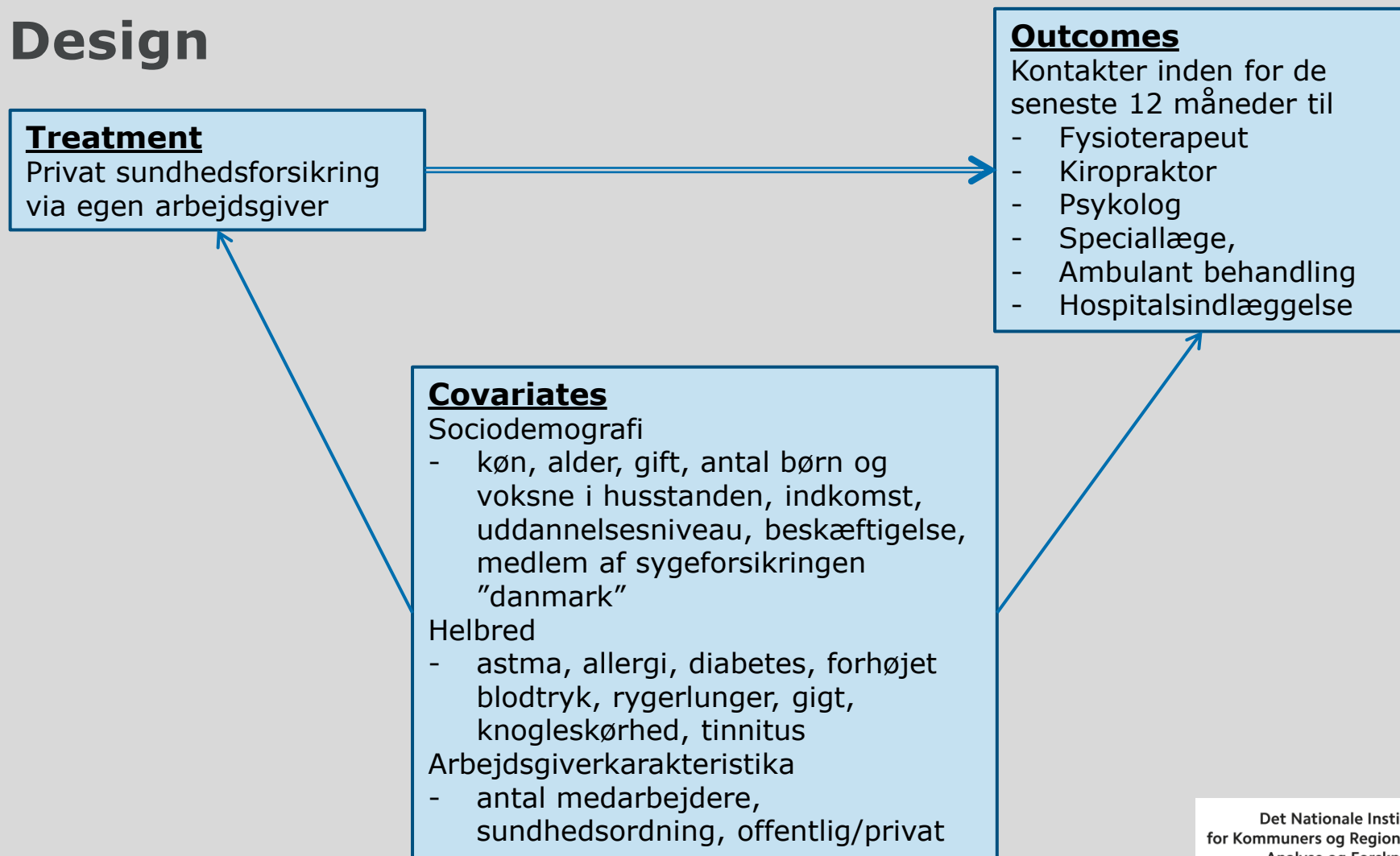
12: 1. doi:10.1007/s10754-012-9104-3



Undersøgelsesspørgsmål + data

- Hvordan påvirker det ens brug af udvalgte sundhedsydelser, at man har fået en privat sundhedsforsikring af sin arbejdsgiver?
- Spørgeskemadata på individniveau, indsamlet i 2009

Design



Resultater, alle beskæftigede

Table 3 Average treatment effect on the treated (ATT) estimates for all employed, n = 2,636

	Model I	Model II	Model III	Model IV	Model V	Model VI	Model VII	
	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	Baseline prob.
Outcomes (any contacts)								
Physiotherapy	0.045 (0.015)***	0.050 (0.015)***	0.043 (0.015)***	0.049 (0.016)***	0.048 (0.016)***	0.036 (0.018)**	0.023 (0.026)	0.178
Chiropractic care	0.031 (0.013)**	0.029 (0.013)**	0.028 (0.013)**	0.030 (0.015)**	0.029 (0.014)**	0.020 (0.015)	0.005 (0.024)	0.142
Psychological counselling	-0.003 (0.008)	-0.004 (0.009)	-0.005 (0.010)	-0.004 (0.010)	-0.001 (0.010)	-0.008 (0.012)	-0.013 (0.016)	0.068
Specialist care	-0.004 (0.019)	0.013 (0.018)	0.009 (0.018)	0.015 (0.018)	0.016 (0.019)	0.015 (0.020)	0.020 (0.028)	0.256
Ambulatory care	-0.034 (0.016)**	-0.032 (0.017)*	-0.031 (0.017)*	-0.025 (0.019)	-0.023 (0.019)	-0.027 (0.020)	0.033 (0.023)	0.193
Hospitalisation	-0.009 (0.011)	-0.007 (0.011)	-0.011 (0.012)	-0.008 (0.013)	-0.003 (0.012)	-0.001 (0.011)	0.028 (0.014)	0.055
# Of treated				1,092				
# Of controls				1,544				
Summary measures of matching quality								
Logit model pseudo R ² before matching		0.023	0.063	0.065	0.080	0.108	0.381	
Logit model pseudo R ² after matching		0.000	0.000	0.001	0.001	0.001	0.008	
Median bias before matching		13.048	8.727	6.018	6.402	6.599	6.402	
Median bias after matching		0.200	0.641	0.701	0.415	1.314	2.475	
# Of treated off common support		0	1	0	0	0	76	
Covariates	None (diff. in means)	Demographic	Demographic Socioeconomic	Demographic Socioeconomic Health conditions	Demographic Socioeconomic Health conditions Employer size	Demographic Socioeconomic Health conditions Employer size Health scheme	Demographic Socioeconomic Health conditions Employer size Health scheme Employment sector	

* Significance at 10% level; ** significance at 5% level; *** significance at 1% level. Standard errors for the ATTs are bootstrapped with 200 replications

Common support, alle beskæftigede

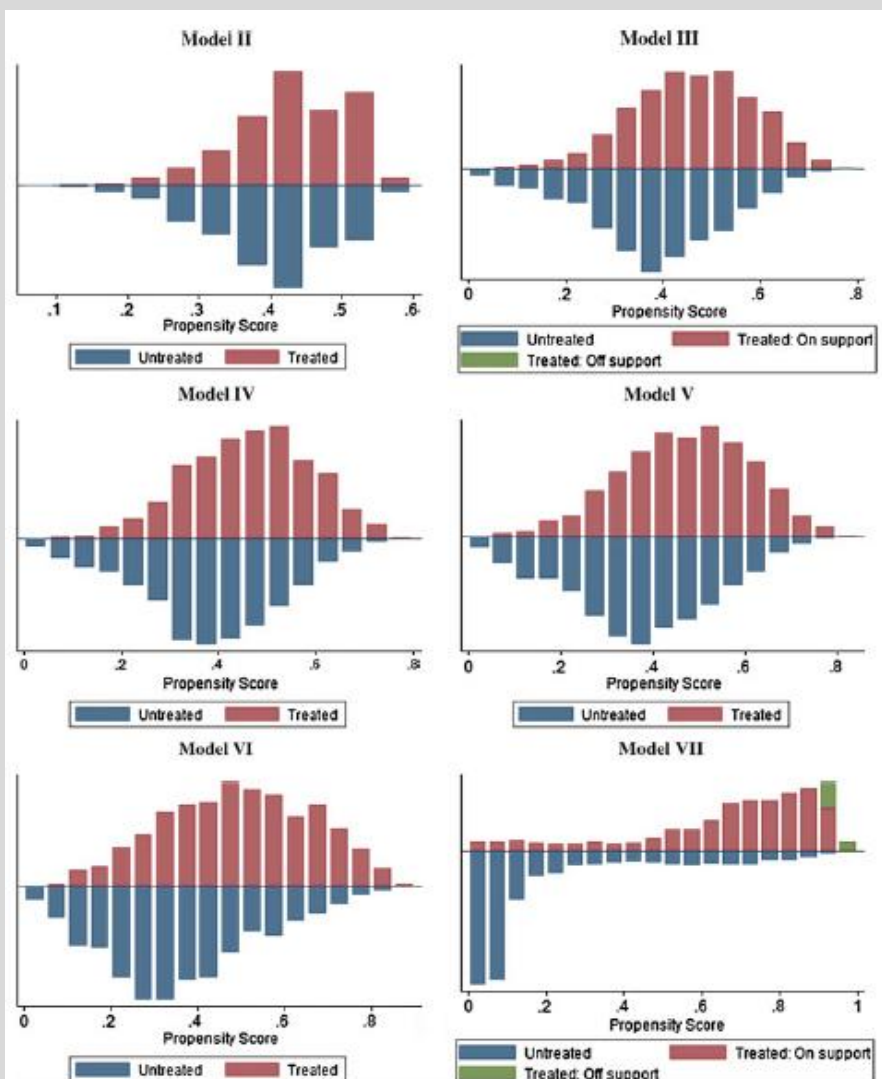


Fig. 1 Propensity scores for treated and non-treated for all employed, n = 2,636

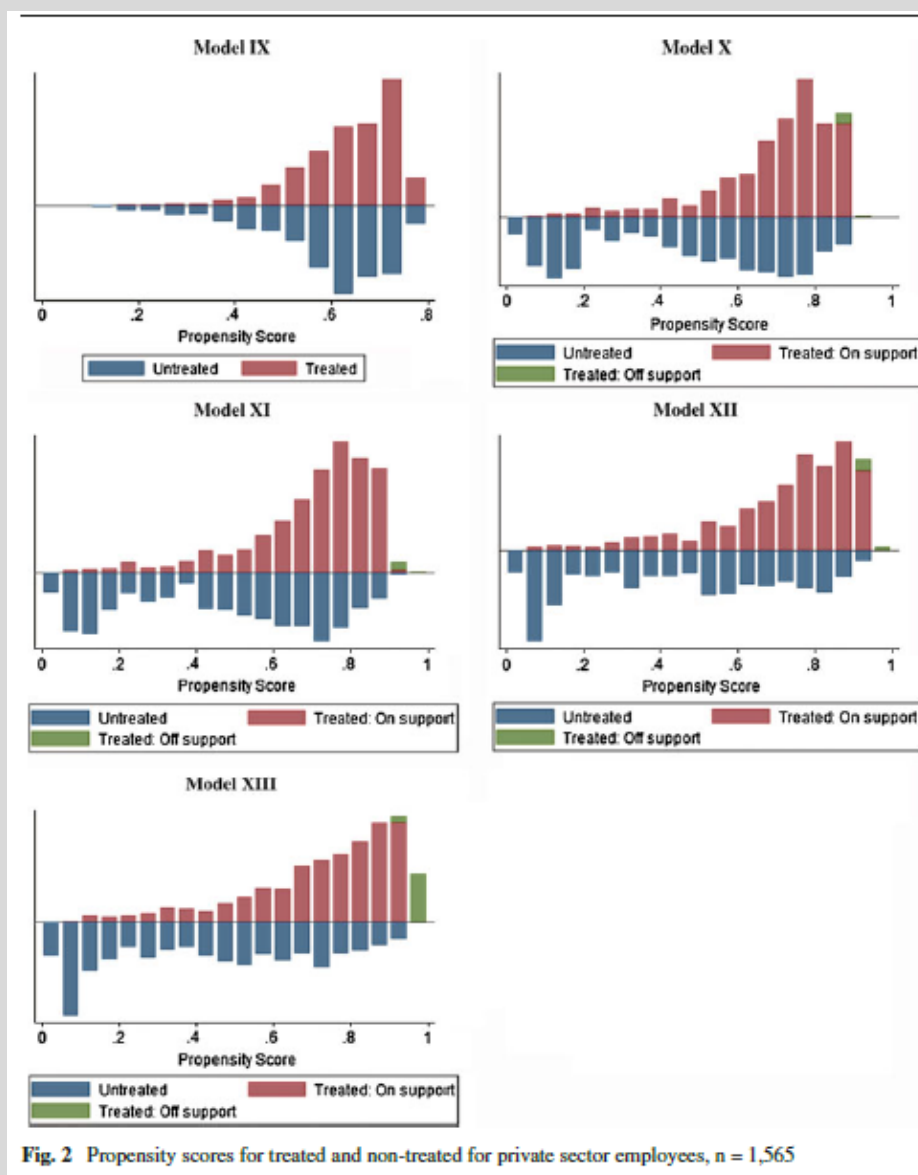
Resultater for privatansatte

Table 4 Average treatment effect on the treated (ATT) estimates for private sector employees, n = 1,565

	Model VIII	Model IX	Model X	Model XI	Model XII	Model XIII	
	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	ATT (std. err.)	Baseline prob.
Outcomes (any contacts)							
Physiotherapy	0.036 (0.022)	0.026 (0.024)	0.019 (0.026)	0.030 (0.025)	0.021 (0.031)	0.023 (0.030)	0.182
Chiropractic care	0.026 (0.018)	0.011 (0.019)	0.013 (0.022)	0.010 (0.022)	0.015 (0.026)	0.001 (0.026)	0.149
Psychological counselling	-0.004 (0.011)	-0.012 (0.013)	-0.011 (0.016)	-0.010 (0.015)	-0.007 (0.015)	-0.016 (0.020)	0.069
Specialist care	-0.009 (0.021)	-0.013 (0.027)	0.002 (0.030)	0.009 (0.028)	0.025 (0.029)	0.019 (0.032)	0.257
Ambulatory care	0.016 (0.022)	0.022 (0.023)	0.050 (0.024)**	0.053 (0.024)**	0.064 (0.026)**	0.054 (0.024)**	0.179
Hospitalisation	0.006 (0.015)	0.006 (0.015)	0.010 (0.018)	0.008 (0.017)	0.027 (0.016)*	0.034 (0.016)**	0.051
# Of treated			968				
# Of controls			597				
Summary measures of matching quality							
Logit model pseudo R ² before matching		0.043	0.179	0.185	0.239	.266	
Logit model pseudo R ² after matching		0.000	0.002	0.003	0.006	.006	
Median bias before matching		17.168	14.734	10.592	12.411	13.926	
Median bias after matching		0.785	1.824	0.890	2.132	2.494	
# Of treated off common support		0	17	12	20	76	
Covariates	None (diff. in means)	Demographic	Demographic	Demographic	Demographic	Demographic	
			Socioeconomic	Socioeconomic	Socioeconomic	Socioeconomic	
				Health conditions	Health conditions	Health conditions	
					Employer size	Employer size	
						Health scheme	

* Significance at 10% level; ** significance at 5% level. Standard errors for the ATTs are bootstrapped with 200 replications

Common support, privatanstætte





Konklusion

- Ingen statistisk signifikante effekter for alle beskæftigede
- Når man kun ser på de privatansatte, så øger sundhedsforsikring sandsynligheden for at have haft en eller flere hospitalsindlæggelser eller ambulante behandlinger inden for de seneste 12 måneder
- Den begrænsede effekt kan skyldes flere ting, bl.a.
 - Rimelig groft outcome
 - Institutionelle barrierer i form af krav om henvisning fra praktiserende læge
 - Begrænsninger i sundhedsforsikringerne dækningen
 - Substitutionseffekter, dvs. skift fra brug af offentlige sundhedsydelser til privat snarere end øget samlet brug



Eksempel 2

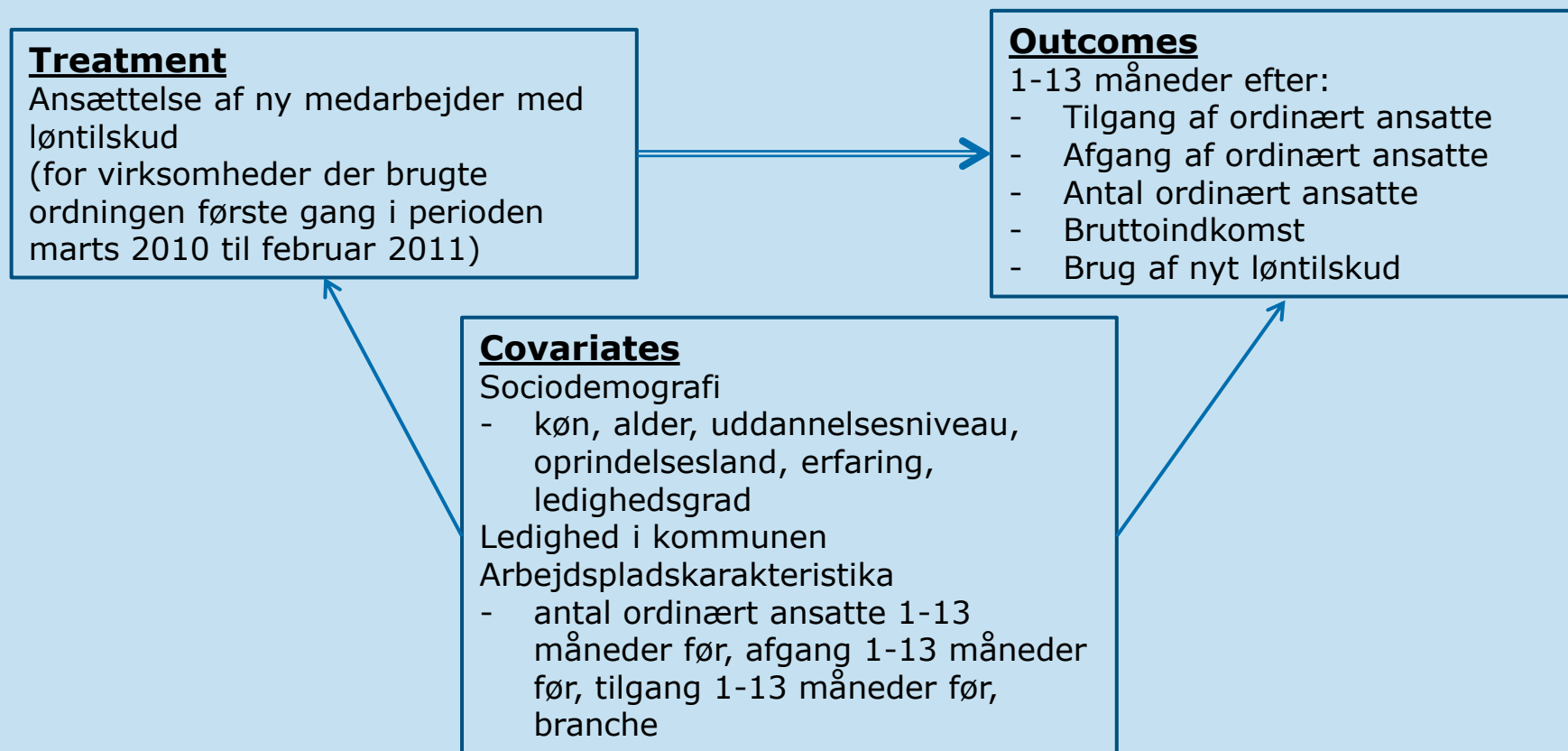




Undersøgelsesspørgsmål + data

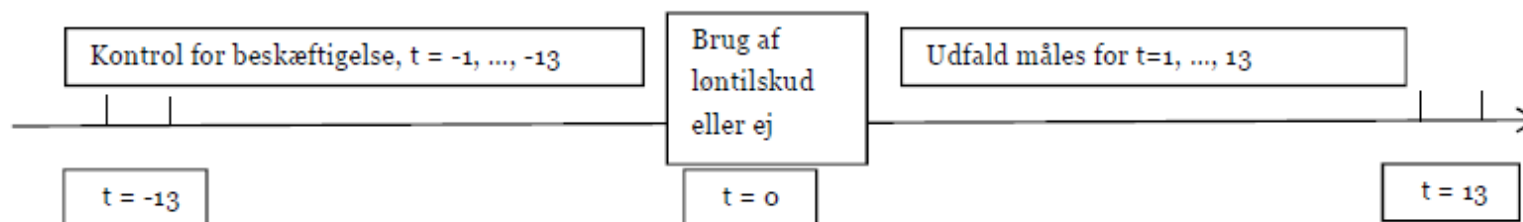
- Hvordan påvirker det beskæftigelse og lønniveau for de ordinært ansatte, når kommunale arbejdspladser ansætter en ny medarbejder med løntilskud?
- Registerdata på arbejdspladsniveau, januar 2009 – april 2012
 - Kommunernes og Regionernes Løndatakontor: løn og ansættelsesforhold (opgjort månedligt) for kommunalt ansatte
 - AMS: start- og slutdatoer for aktivering med løntilskud på kommunale arbejdspladser
 - DST: køn, alder, uddannelsesniveau, antal år som lønmodtager, arbejdsledighedsgrad

Design



Design

Figur 3.1 **Undersøgelsesdesign**



- Hver deltager matches med en relevant kontrol med en epanechnikov kernel matching-algoritme, da denne metode giver den bedste balance
- Også afprøvet: NN med hhv. 1, 5, 10 og 20 neighbours, alle med calipers på hhv. 0,0001 og 0,0005

Resultater, arbejdspladser med 5-40 medarbejdere

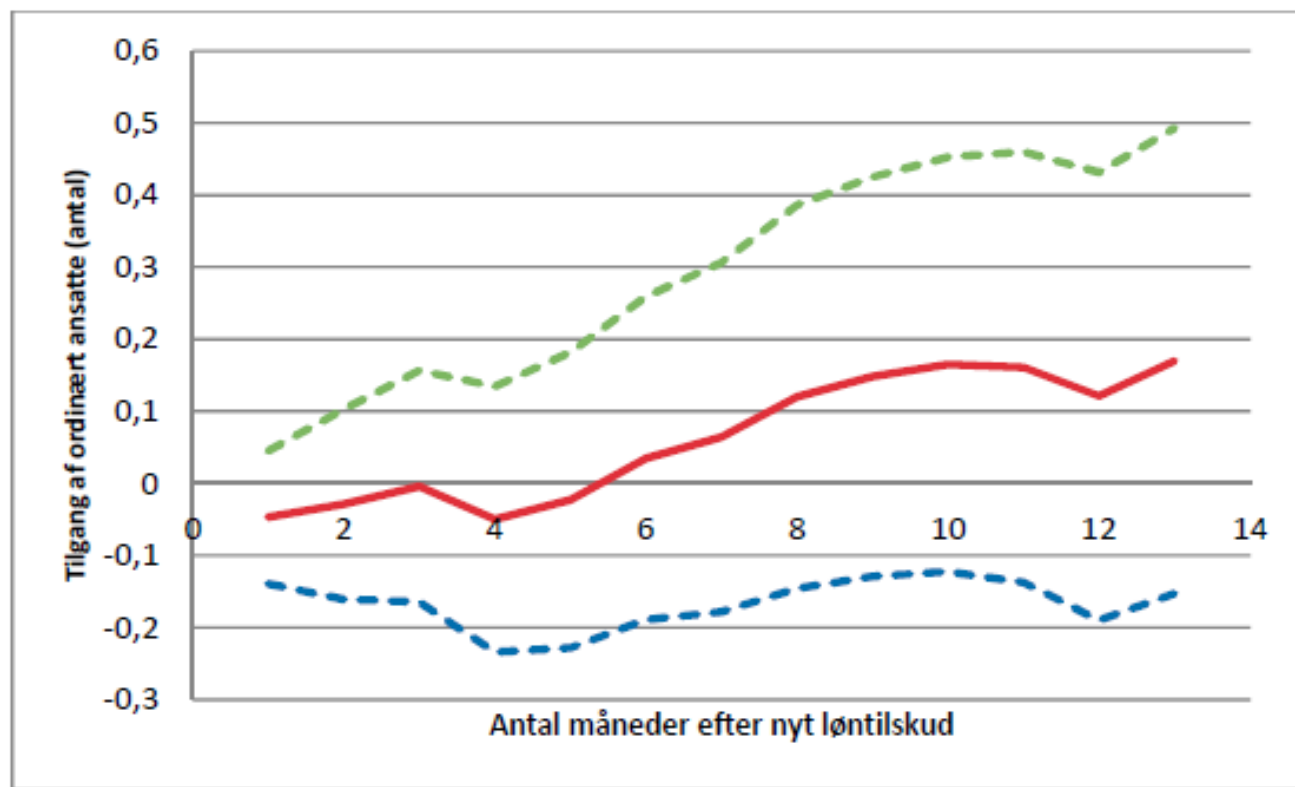
Tabel 5.2 Gennemsnitlige effekter af at ansætte en ny medarbejder med løntilskud for arbejdspladser, der har brugt ordningen i marts 2010 – februar 2011

Mdr. efter nyt løntilskud	Tilgang af ord. ansatte (std.fejl)	Afgang af ord. ansatte (std.fejl)	Ordinært ansatte (std.fejl)	Bruttoindkomst (std.fejl)	Brug af nyt løntilskud (std.fejl)
1	-0,047 (0,047)	0,049 (0,044)	0,049 (0,116)	7703 (4411)	0,036 (0,009)
2	0,018 (0,048)	-0,015 (0,046)	0,018 (0,120)	6860 (3539)	0,026 (0,008)
3	0,025 (0,047)	0,109 (0,058)	0,058 (0,118)	1095 (4137)	0,039 (0,009)
4	-0,046 (0,046)	0,043 (0,047)	-0,098 (0,124)	-3604 (3506)	0,049 (0,010)
5	0,027 (0,046)	0,082 (0,070)	-0,114 (0,126)	-5994 (3929)	0,043 (0,009)
6	0,058 (0,046)	-0,001 (0,066)	-0,138 (0,137)	-3009 (4643)	0,127 (0,014)
7	0,029 (0,047)	0,042 (0,047)	-0,109 (0,145)	-7455 (4178)	0,026 (0,009)
8	0,056 (0,056)	0,089 (0,067)	-0,095 (0,150)	-9006 (4329)	0,034 (0,009)
9	0,028 (0,040)	0,077 (0,047)	-0,156 (0,160)	-5992 (4919)	0,028 (0,009)
10	0,017 (0,039)	-0,001 (0,044)	-0,216 (0,162)	-6736 (4840)	0,038 (0,010)
11	-0,004 (0,041)	0,069 (0,044)	-0,220 (0,166)	-6854 (4681)	0,043 (0,010)
12	-0,040 (0,044)	-0,100 (0,054)	-0,328 (0,164)	-5912 (5410)	0,038 (0,010)
13	0,049 (0,045)	-0,011 (0,051)	-0,179 (0,169)	-5811 (4839)	0,049 (0,010)

Anm.: De estimerede effekter for beskæftigelse og indkomst er effekter på årlige differenser defineret som forskellen mellem en given måned og den tilsvarende måned året før. De rapporterede standardfejl er estimeret med metoden angivet i Abadie & Imbens (2006). * Signifikant på 10%-niveau; ** Signifikant på 5%-niveau; *** Signifikant på 1%-niveau.

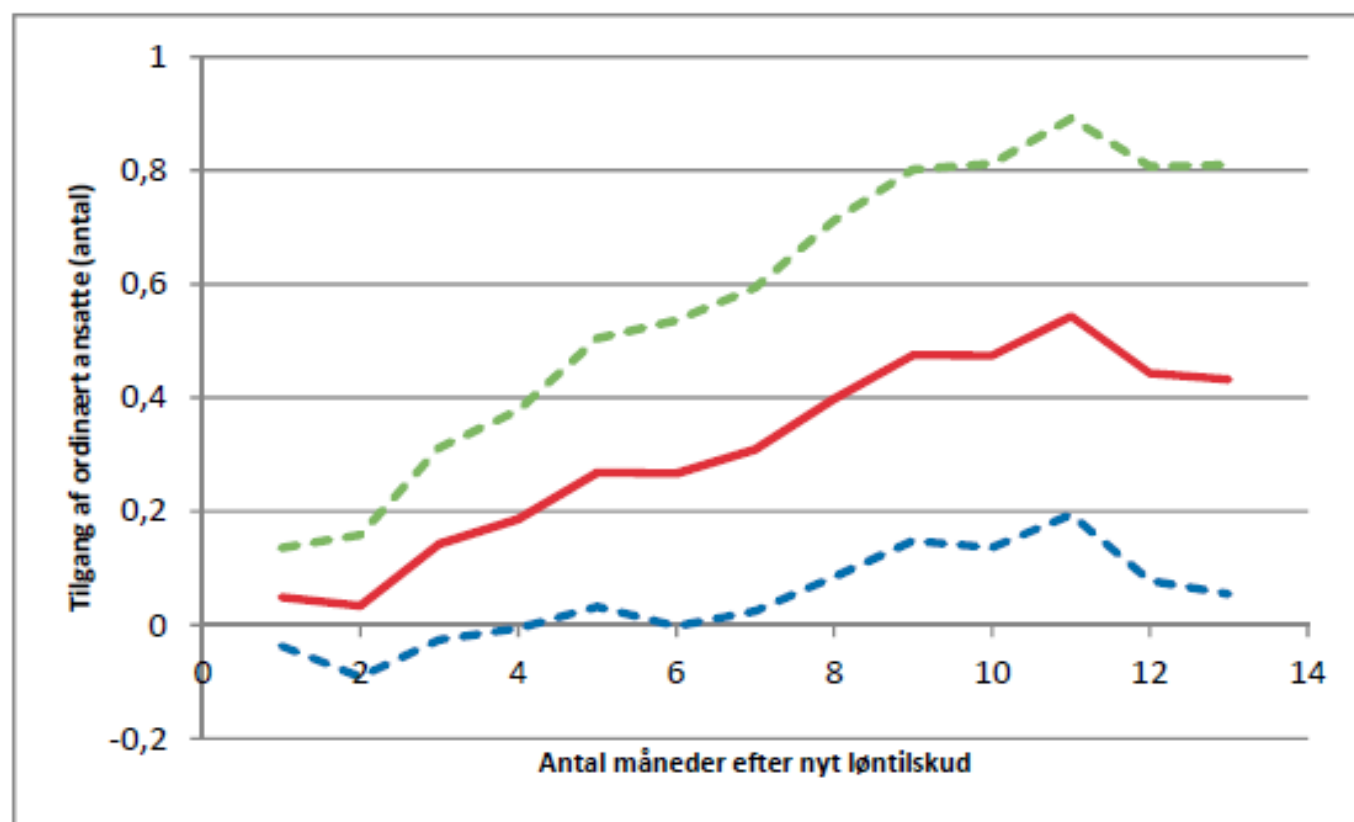
Resultater, arbejdspladser med 5-40 medarbejdere

Figur 5.4 Kumulerede effekter på tilgang af ordinært ansatte



Resultater, arbejdspladser med 5-40 medarbejdere

Figur 5.5 Kumulerede effekter på afgang af ordinært ansatte



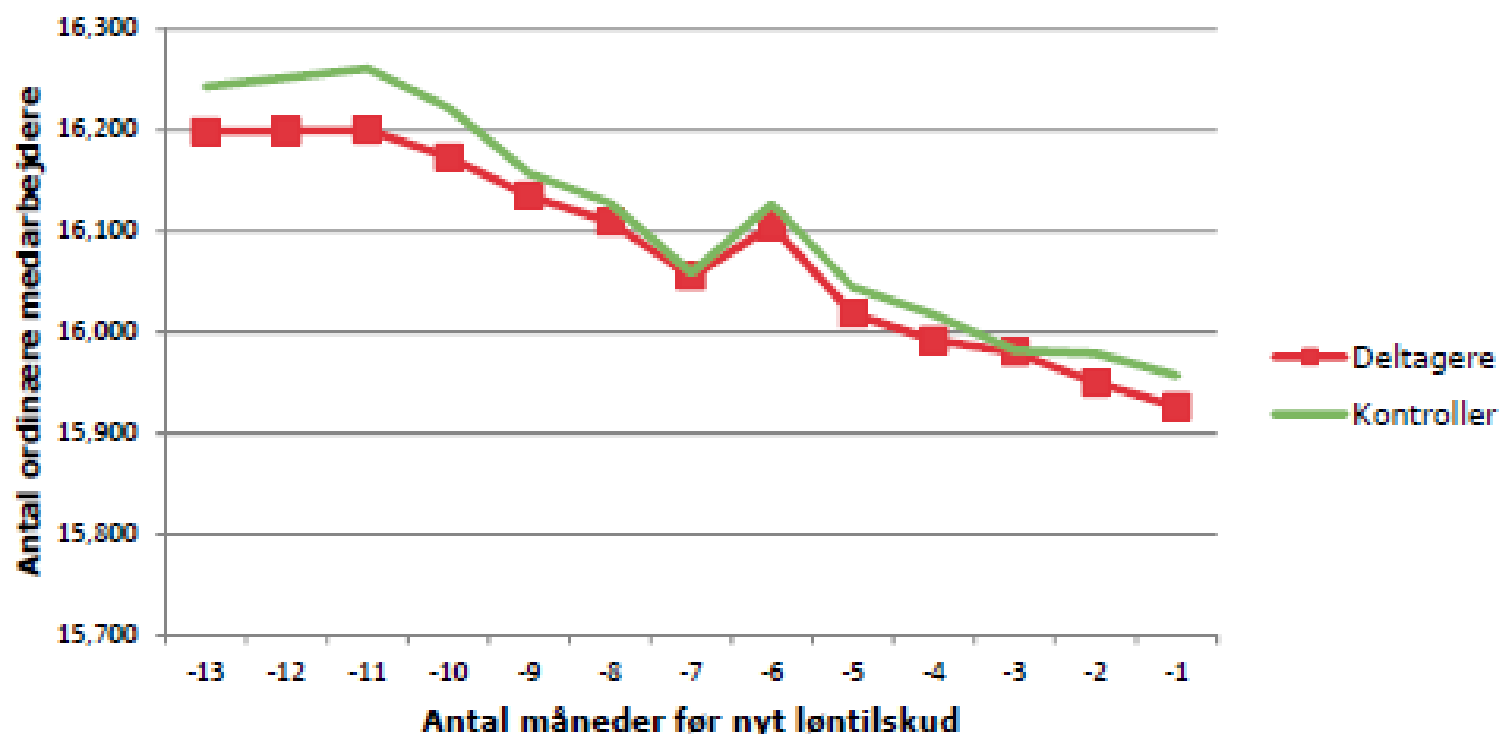
Balance, samlet set

Tabel F.2 Opsummering og fordeling af |bias| for alle covariates, marts 2010 – februar 2011

FØR MATCHING				
Percentiler		Mindste		
0,01	0,19	0,19		
0,05	0,48	0,19		
0,10	0,87	0,27	Obs	67
0,25	2,59	0,48	Sum af vægte	67
0,50	5,87		Gennemsnit	8,86
		Største	Std.afvigelse	8,56
0,75	17,34	19,44		
0,90	18,94	19,76	Varians	73,32
0,95	19,44	20,86	Skævhed	2,00
0,99	51,18	51,18	Kurtosis	9,83
EFTER MATCHING				
Percentiler		Mindste		
0,01	0,01	0,01		
0,05	0,03	0,02		
0,10	0,06	0,02	Obs	67
0,25	0,28	0,03	Sum af vægte	67
0,50	0,62		Gennemsnit	1,04
		Største	Std.afvigelse	1,41
0,75	1,31	3,89		
0,90	2,02	4,44	Varians	1,98
0,95	3,89	5,18	Skævhed	3,36
0,99	8,86	8,86	Kurtosis	16,84

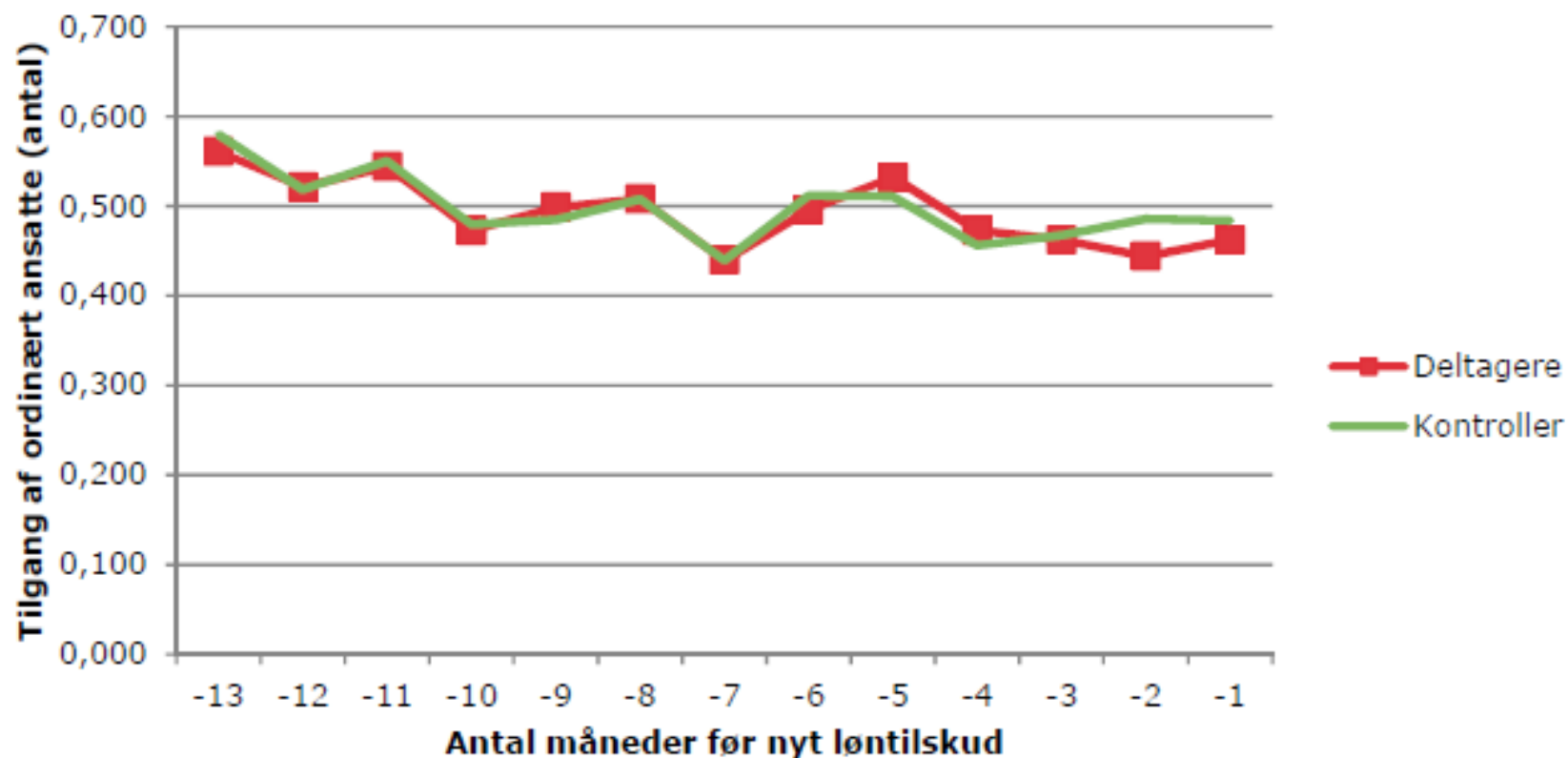
Balance, antal ordinære medarbejdere

Figur 5.1 Balance efter matching for antal ordinære medarbejdere 1-13 måneder før ansættelse af ny medarbejder med løntilskud, marts 2010 – februar 2011



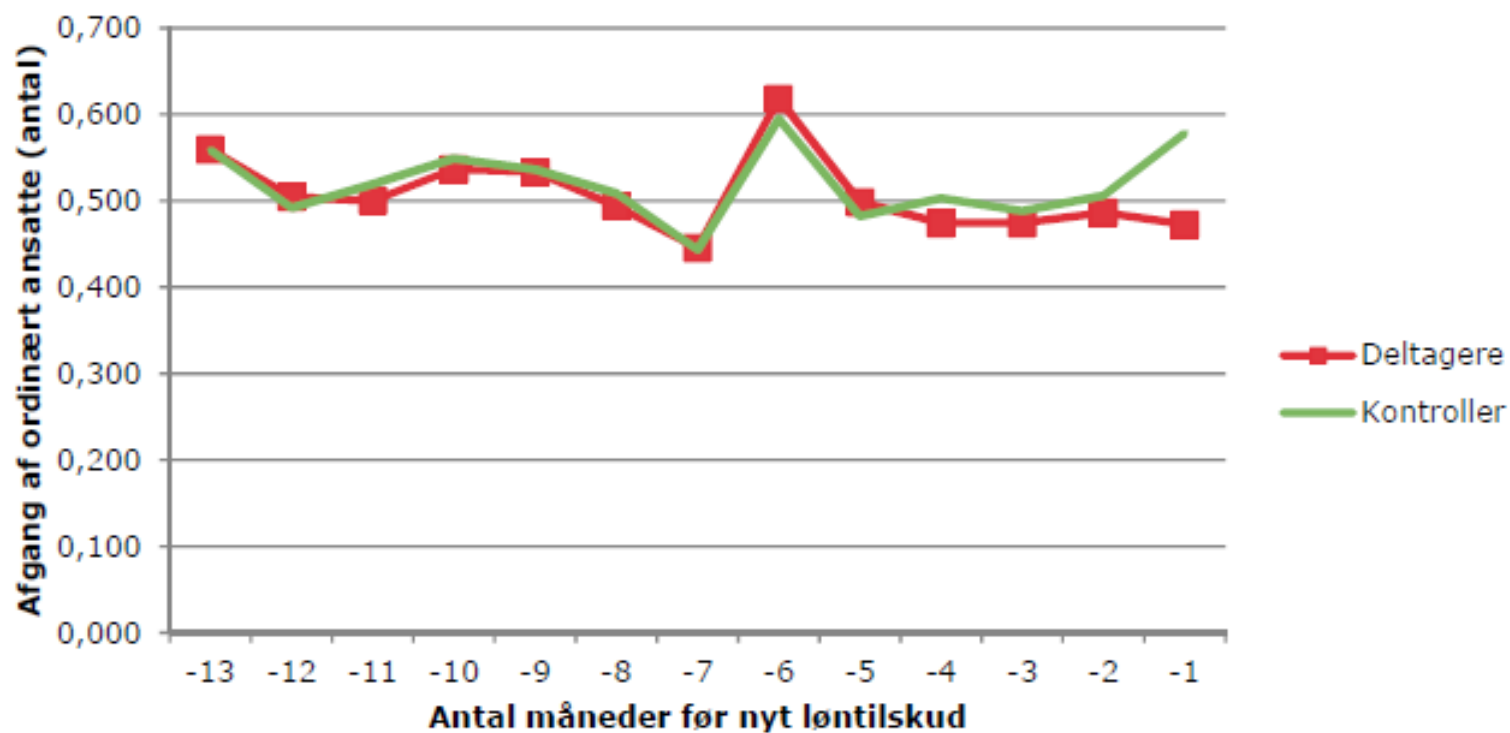
Balance, tilgang af ordinære medarbejdere

Figur 5.2 Balance efter matching for tilgang af ordinære medarbejdere 1-13 måneder før ansættelse af ny medarbejder med løntilskud, marts 2010 – februar 2011



Balance, afgang af ordinære medarbejdere

Figur 5.3 Balance efter matching for afgang af ordinære medarbejdere 1-13 måneder før ansættelse af ny medarbejder med løntilskud, marts 2010 – februar 2011



Common support

Tabel F.3 Common support, marts 2010 – februar 2011

	Common support		Total
	Antal obs. <u>uden for</u> common support	Antal obs. <u>inden for</u> common support	
Kontroller	0	44.397	44.397
Deltagere	34	660	694
Total	34	45.057	45.091



Følsomhedsanalyser

- NN 1, 5, 10 og 20 + calipers på 0,001 og 0,005 (i stedet for epanechnikov kernel)
- Estimation af propensity score med probit (i stedet for logit)
- Dropper 10% med højeste propensity score, fordi fordelingen er "tynd" i toppen
- Effekter estimeret på absolut antal (i stedet for differenser)
- Uden restriktion om, at løntilskud ikke må være brugt de seneste 13 måneder -> giver som ventet mindre effekter
- Arbejdspladser med 5-30 og 5-50 medarbejdere (i stedet for 5-40)



Eksempel 3



Martin Sandberg Buch, Marie Jakobsen, Christophe Kolodziejczyk og Else Ladekjær

Evaluering af indsats for forløbskoordination

– Erfaringer med fremskudt visitation i fire kommuner

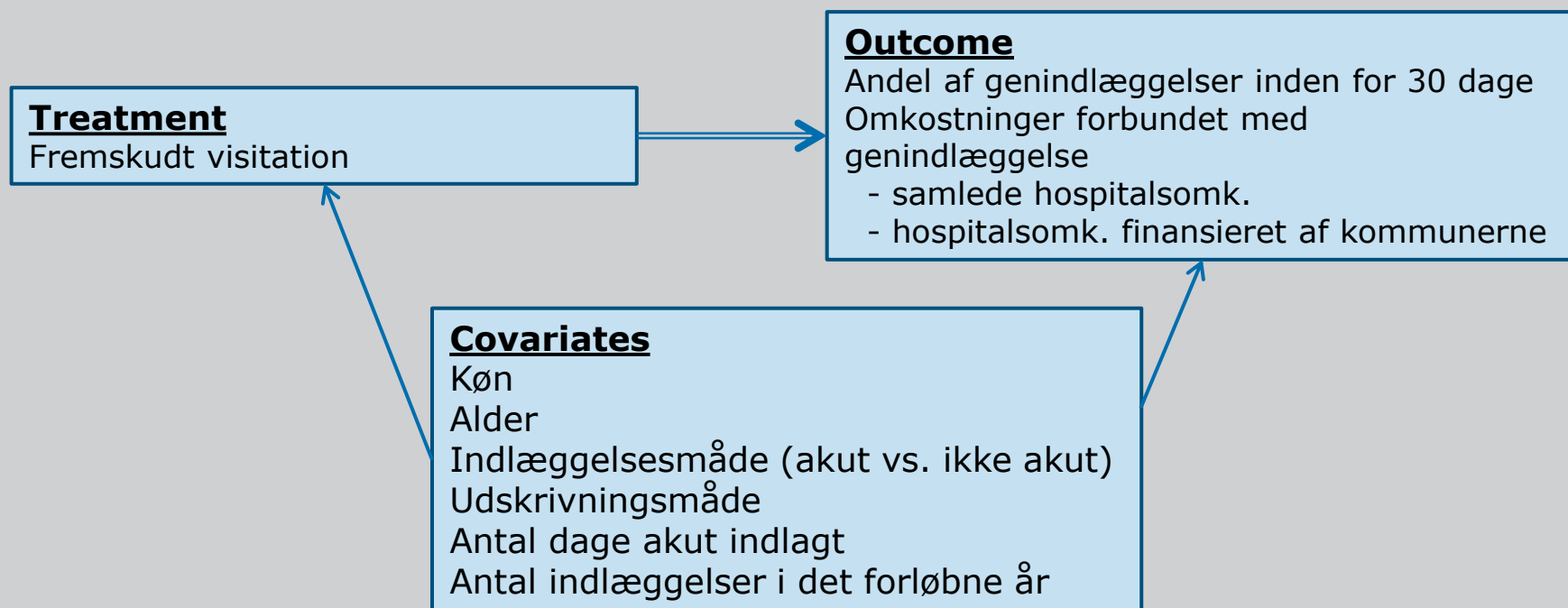




Undersøgelsesspørgsmål + data

- Overordnet er formålet at evaluere indsatsen vedrørende forløbskoordination for ældre medicinske patienter
- Registerdata på individniveau
- Population: Borgere over 65, og som er udskrevet til hjemmet efter indlæggelse på somatisk sygehusafdeling i perioden 1. januar 2014-31. maj 2015
 - Interventionsgruppe: Gribskov, Lolland, Skanderborg og Vejle Kommuner
 - Kontrolgruppe: Øvrige kommuner (hvor det antages, at der ikke er ydet fremskudt visitation)

Design



NN 1-10 med tilbagelægning

Resultater opdelt på kommuner

Tabel 6.2 Forskel i andel genindlæggelser inden for 30 dage mellem interventions- og kontrolgruppen (ATT)

	Gribskov (N=291)	Lolland (N=151)	Vejle - udvidet koordinering (N=277)	Vejle - gen- nemgang af aftaleskema (N=101)	Skanderborg (N=198)
Interventionsgruppe - kontrolgruppe (ATT)	-1,6 %	-3,6 %	-3,6 %	-12,5 %	-6,3 %
P-værdi*	0,572	0,365	0,231	0,004	0,060

Note: *) *) P-værdi er beregnet ved tosidet t-test. Hvis p-værdi > 0,05, er forskellen ikke statistisk signifikant.

Kilde: Beregnet af KORA på grundlag af data fra LPR

Resultater for genindlæggelser

Tabel 6.3 Andel genindlæggelser inden for 30 dage, de fire casekommuner under ét

	Interventi- onsgruppe	Kontrol- gruppe	Forskel (ATT)	P-værdi*
Alle indlæggelser (N=1.018)	27,2 %	31,7 %	-4,4 %	0,008
Subgruppe: Akutte indlæggelser (N=927)	27,4 %	32,0 %	-4,6 %	0,010
Subgruppe: Indlæggelser, hvor pårørende deltager i møde med kommunal udskrivningskoordinator (N=448)	27,7 %	32,5 %	-4,8 %	0,043

Note: *) P-værdi er beregnet ved tosidet t-test. Hvis p-værdi > 0,05, er forskellen ikke statistisk signifikant.

Kilde: Beregnet af KORA på grundlag af data fra LPR

Resultater for hospitalsomkostninger

Tabel 6.4 Hospitalsomkostninger ved genindlæggelser inden for 30 dage, de fire casekommuner under ét

	Interventionsgruppe	Kontrolgruppe	Forskel (ATT)	P-værdi*
Hospitalsomkostninger i alt	12.343 kr.	17.853 kr.	-5.510 kr.	<0,000
Hospitalsomkostninger finansieret af kommunerne via den kommunale medfinansiering	2.852 kr.	3.691 kr.	-839 kr.	<0,000

Note: *) P-værdi er beregnet ved tosidet t-test. Hvis p-værdi > 0,05, er forskellen ikke statistisk signifikant.

Kilde: Beregnet af KORA på grundlag af data fra LPR og DRG-grupperet LPR

Meromkostninger

Tabel 6.5 Meromkostninger til drift og udvikling af fremskudt visitation i de fire casekommuner

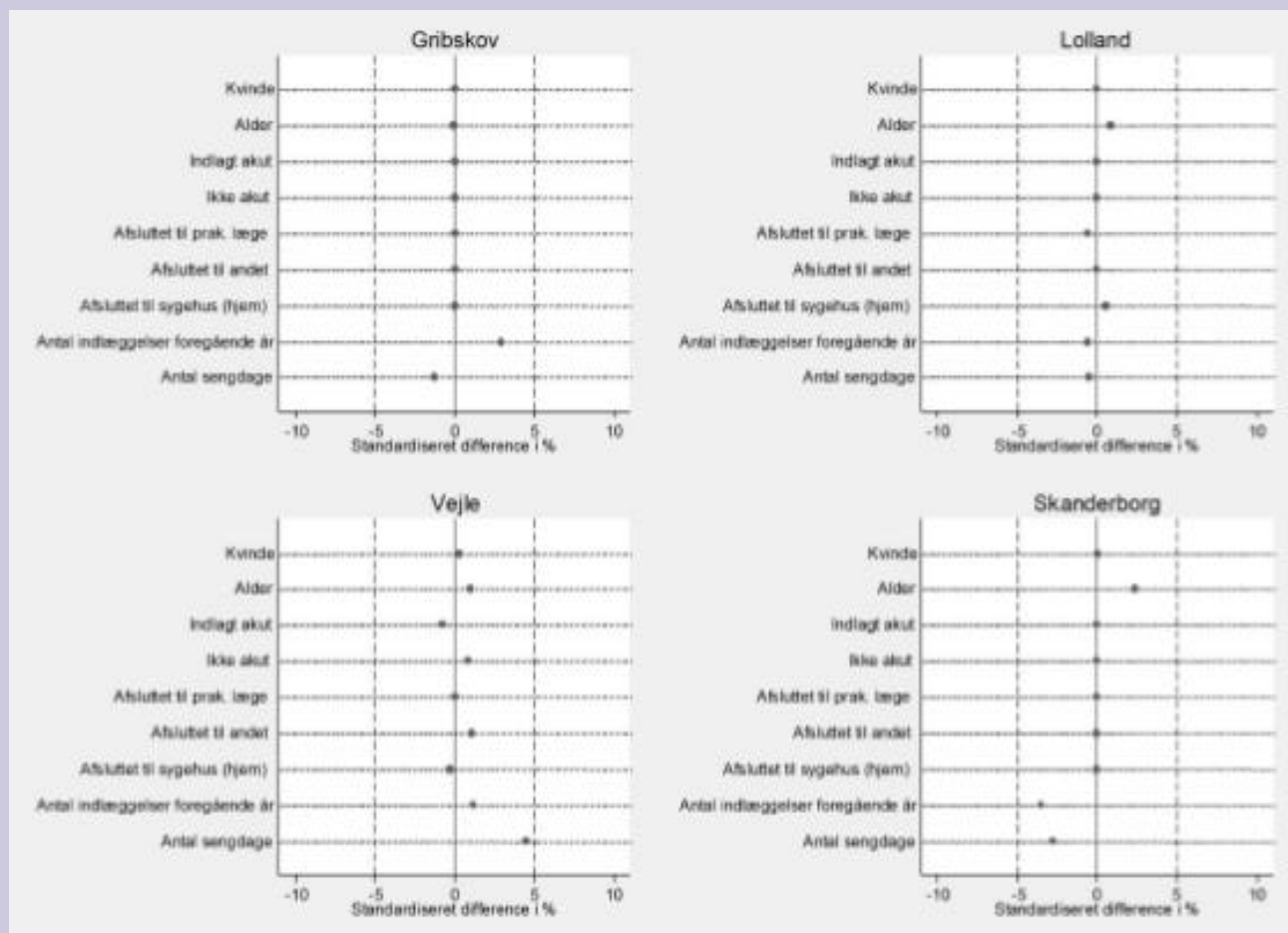
	Gribskov	Lolland	Vejle udvidet ko- ordinering	Vejle aftaleske- maer	Skander- borg
Skøn over meromkostninger til drift: Ekstra ressourceforbrug til indsatsen pr. patient/indlæggelse*, kr.	230	440	0	4.852	1.011
Skøn over meromkostninger til udvik- ling: Ekstra ressourceforbrug i opstartsfa- sen**, kr.	20.000	50.000	n.a.	35.000	250.000

Note: *) Skøn over tidsforbrug til fremskudt visitation i en typisk måned er indberettet af kommunerne sammen med oplysninger fra kommunerne om antal patienter/indlæggelser, hvor der er ydet fremskudt visitation. KORA har på grundlag af de indberettede oplysninger beregnet det gennemsnitlige ekstra tidsforbrug pr. patient/indlæggelse.

**) Skøn over tidsforbrug til udvikling af indsatsen i opstartsfasen er indberettet af kommunerne. KORA har på grundlag af de indberettede oplysninger beregnet det gennemsnitlige ekstra tidsforbrug pr. patient/indlæggelse. Opgørelsen er forbundet med stor usikkerhed, da det er vanskeligt at skønne over tidsforbrug tilbage i tid.

Kilde: Beregnet af KORA på grundlag af indberetninger fra kommunerne

Kvalitet af matching: Balance





Svagthed

- Kan ikke udelukkes, at nogle patienter i kontrolgruppen har fået fremskudt visitation eller andre lignende kommunespecifikke indsatser med henblik på at forebygge genindlæggelse
- Denne form for bias slører effekten af indsatsen og gør det mere vanskeligt at påvise statistisk signifikante effekter
- Har forsøgt at reducere dette ved at rekruttere patienter til kontrolgruppen fra kommuner i hele landet

Udvalgt litteratur

- Der er flere gode lærebogs-kapitler og sammenfatninger, se fx
 - Heckman, J.J., Ichimura, H. & Todd, P.E. (1998): *Matching as an Econometric Evaluation Estimator*, Review of Economic Studies, 65:261-294.
 - Caliendo, M. & Kopeinig, S. (2005): *Some Practical Guidance for the Implementation of Propensity Score Matching*, IZA Discussion Paper No. 1588
 - Austin, P. (2011): *An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies*, Multivariate Behavioral Research, 46:399-424.

Tak for jeres opmærksomhed!