

# **Politiske elitors indflydelse på befolkningens politiske opmærksomhed**

Eksamensopgave i seminaret

”Videregående kvantitative metoder i studiet af politisk adfærd”

Institut for statskundskab

Københavns universitet

**Malte Glerup Hockenhull**

**45.070 typeenheder inkl. mellemrum**

(18,8 normalsider af 2400 typeenheder inkl. mellemrum)

## Indhold

1. Indledning, teoretisk motivation og problemstilling .....	3
2. Metoder og data .....	4
2.1. Data og centrale variable .....	5
2.2. Det fundamentale problem med kausal inferens .....	6
2.3. Coarsened Exact Matching .....	7
2.3.1. Baggrund for matching .....	7
2.3.2. Coarsening af baggrundsvariable .....	9
2.4. Instrumentvariable .....	10
3. Analyse .....	14
3.1. Analyse med CEM .....	15
3.2. Analyse med IV .....	18
4. Diskussion af resultater, kausalitet og andre designs .....	20
4.1. Analysens resultater og kausalitet .....	20
4.2. Alternative designs i undersøgelser af politisk opmærksomhed .....	21
5. Konklusion .....	23
6. Kilder .....	25
6.1. Data .....	25
6.2. Litteraturliste .....	25
7. Bilag .....	27
7.1. Centrale variable .....	27
7.2. Imbalance output før matching .....	27
7.3. Imbalance output efter matching .....	28
8. R-script .....	29

## **1. Indledning, teoretisk motivation og problemstilling**

Den nuværende amerikanske præsident er ofte på dagsordenen i de medier, han elsker at hade. Trumps kritiske syn på mediernes nyhedsdækning, som han ynder at kalde ”fake news”, er velkendt og blev et element i hans kampagne i 2016, såvel som det er blevet et element i hans politiske dagsorden i 2017. I kraft af sit embede som USA’s præsident er Trump indflydelsesrig, og hans udsagn bliver derfor hurtigt til cues, som befolkningen retter sig efter.

Det er en ukontroversiel påstand, at politikere som Trump spiller en afgørende rolle i det moderne demokrati. De er i spidsen for det politiske arbejde og alle dets facetter; det er dem, der med folkets mandat i ryggen besidder politiske embeder, udarbejder og diskuterer lovgivning og sætter retningen for samfundets udvikling.

Følger man John Zallers argument, er det samtidig også politiske eliter, der informerer befolkningen om det, de ellers ikke normalt har indblik i, f.eks. udviklingen i det politiske arbejde (Zaller 1992: 6). Zaller argumenterer for, at borgere i massesamfund er afhængige af særlige personer for at få information om verden, og at denne information typisk kommer fra personer, der beskæftiger sig med politik eller public affairs; hvad han kalder ”politiske eliter”. I en Zallersk forstand er Trump således også en politisk elite uafhængigt af hans anti-establishment tilgang til politik (Zaller 1992: 13). Ifølge Zallers teori influerer Trumps holdninger dermed befolkningens holdninger igennem den information, han tilvejebringer

Hvordan befolkningen i demokratiet reagerer på informationsstrømmen fra politiske eliter er et omdiskuteret emne i litteraturen. Den økonomiske tradition fastslår, at den rationelle borger vil vurdere, om politikernes arbejde øger hendes nytte og former sin holdning til politikerne herefter (Carmines og Huckfeldt 1998); den sociologiske tradition argumenterer for, at borgernes holdninger er kraftigt præget af opvækst og miljø (Zuckerman 2015; Jennings 2009) og politisk psykologi lægger blandt andet vægt på, hvordan mental bias og prædispositioner påvirker holdninger (Taber og Lodge 2006).

Charles S. Taber og Milton Lodge opstiller en teori om en sådan mental bias kaldet ”motivated reasoning”. De argumenterer for, at folk blandt andet er præget af ”confirmation bias”, hvilket betyder, at de vil søge afsendere, som tilbyder information, der er i overensstemmelse med deres eksisterende holdninger (Taber og Lodge 2006: 757). Samtidig er det et element i motivated reasoning-teorien, at folk er ukritiske over for argumenter, de er enige i, samtidig med at de er

overkritiske over for argumenter, der ikke passer ind i deres verdenssyn (Taber og Lodge 2006: 755-757).

På baggrund af teorien om motivated reasoning og Zallers teori om politiske elitors indflydelse på holdningsdannelse, kan man således stille spørgsmålstejn ved, om Trumps aversion mod medierne medfører, at de, der støtter ham politisk, får samme holdning og fravælger de ”falske” medier.

Denne opgaves problemstilling lyder derfor:

*I hvilket omfang medfører politisk støtte til Trump forud for præsidentvalget i 2016, at vælgerne fulgte mindre med i nyheder om national politik?*

Og opgavens hypotese er: Personer, der havde intentioner om at stemme på Trump, ville følge mindre med i politiske nyheder end personer, der ikke havde intention om at stemme på Trump.

Den underliggende mekanisme er her, at da Trump mener, at medierne som helhed behandler ham uretfærdigt, vil hans støtter – påvirket af den information han giver dem i kraft af sin position som politisk elite – mene det samme. De vil derfor på grund af confirmation bias undgå selvsamme medier.

Dette puzzle er relevant, fordi medierne er en afgørende aktør i det moderne samfund. Potentiel underminering af mediernes troværdighed i den brede befolkning kan derfor have vidtrækkende konsekvenser for, hvordan den politiske dagsorden sættes, samt for hvilke informationer der videregives og hvordan.

Dette spørgsmål vil blive undersøgt via to supplerende metoder med forskelligt sigte: Coarsened Exact Matching (CEM) og instrumentvariabelanalyse (IV). I det følgende afsnit gennemgår jeg de to metoders funktioner og antagelser, samt hvordan de hver kan bidrage til at undersøge problemstillingen sammenlignet med ren multivariat OLS-regression. I afsnittet gennemgår jeg også det anvendte data, samt hvordan problemstillingen operationaliseres i relevante variable. Dernæst analyserer jeg problemstillingen med de to metoder, hvorefter jeg diskuterer resultaterne og designet samt hvordan andre metoder kan supplere analysen.

## **2. Metoder og data**

I det følgende afsnit gennemgår jeg først det anvendte datasæt og opgavens centrale variable, hvorefter jeg illustrerer udfordringen med kausal inferens i et potential outcomes framework.

Dernæst følger en gennemgang af først Coarsened Exact Matching og så instrumentvariable, samt hvordan metoderne anvendes i undersøgelsen.

## 2.1. Data og centrale variable

Til at undersøge om og i hvilket omfang intention om at stemme på Trump forud for det amerikanske præsidentvalg i 2016 påvirker opmærksomhed på politiske nyheder (fremover: politisk opmærksomhed), anvender jeg data fra American National Election Studies (ANES) fra 2016. ANES-datasættet indeholder data fra en række interviews og computerbesvarelser med et n på 4.271 respondenter gennemført i perioden 7. september til 7. november 2016. Der er således tale om observationsdata snarere end eksperimentdata. Spørgsmålene i surveyet omhandler en række temaer, heriblandt politisk observans, holdninger forud for valget, medieforbrug m.fl. (ANES 2016).

Den afhængige variabel – politisk opmærksomhed – er undersøgt ved at spørge respondenterne, hvor meget de følger med i nyheder om national politik i tv, radio, avis eller på internettet.

Respondenterne kunne svare på en 5-trins skala hvor 1 er ”A great deal” og 5 er ”None at all”.

Spørgsmålet omkodes, sådan at 5 på skalaen indikerer, at man følger meget med og 1 indikerer, at man slet ikke følger med, da dette er mere intuitivt.

Spørgsmålet blev stillet til de respondenter, der i et tidligere spørgsmål havde tilkendegivet, at de på en typisk uge fulgte med i nyhederne (dog ikke sport) flere end nul gange i enten tv, radio, avis eller på internettet. En enkelt respondent havde undladt at besvare spørgsmålet og 77 respondenter fulgte ikke med i nyhederne mere end nul gange om ugen. Disse 78 respondenter kodes som ”missing”.

Den uafhængige, treatment-variabel udtrykker, hvorvidt respondenterne forud for valget havde intention om at stemme på Trump. Spørgsmålet vedrører, hvem man vil stemme på og blev stillet til respondenter, der enten var eller havde intention om at registrere sig som vælgere, og som ikke brevstemte eller på anden vis stemte før valgdagen d. 8. november 2016. På den baggrund er 903 besvarelser kodet som missing. Variablen er derefter omkodet til en dikotom variabel, hvori alle besvarelser, der peger på noget andet end Trump, er slået sammen. Dette skyldes at det interessante forhold i denne undersøgelse ikke er, hvad man havde intention om at stemme, men blot om man ville stemme på Trump. Således indikerer ’1’, at man havde intention om at stemme på Trump (1.357 respondenter), og ’0’ indikerer, at man ikke havde intention om at stemme på Trump (2.011 respondenter).

Derudover kontrolleres for baggrundsvariablene køn, alder, uddannelsesniveau, husstandens samlede årsindkomst før skat, samt en dikotom variabel for om respondenter ved præsidentvalget i 2012 angiver at have stemt på Barack Obama (0) eller på Mitt Romney (1). Som vil blive uddybet

på s. 11, anvender jeg også en variabel for respondenternes ideologiske selvplacering på en skala fra 1 (ekstremt liberal) til 7 (ekstremt konservativ) som instrumentvariabel. En oversigt over denne opgaves centrale variable fremgår af bilag 1.

## 2.2. Det fundamentale problem med kausal inferens

Guldstandarden for estimation af kausal effekt er et eksperimentelt design, da eksperimentel data pr. definition indebærer tilfældig inddeling af observationer i henholdsvis en eksperimentgruppe, der får treatment, og i en kontrolgruppe, der ikke får treatment. Eventuelle outcome-effekter må således som udgangspunkt skyldes treatment, da de to grupper i gennemsnit er ens (Justesen og Klemmensen 2014: 61).

Når man arbejder med observationsdata, er det imidlertid ikke tilfældigt, om observationerne befinder sig i treatmentgruppen eller i kontrolgruppen, da mennesker har en tendens til systematisk at selvselektere sig ind i visse sociale konstellationer frem for andre (Justesen og Klemmensen 2014: 62). Når vi sammenligner empiriske og observerbare forhold hos én gruppe personer med empiriske og observerbare forhold hos en anden i en analyse med observationsdata, kan det derfor være vanskeligt at afdække kausale effekter af treatment  $X$  på outcome  $Y$  (Justesen og Klemmensen 2014: 60). Da denne undersøgelse anvender observationsdata, er det derfor relevant at overveje, hvordan man alligevel kan tilnærme sig kausalitet. For at være konkret vil den specifikke treatment-variabel i de følgende afsnit blive betegnet  $D_i$  i stedet for  $X$  (som kan betegne uafhængige variable generelt).

I et hypotetisk samfundsvidenskabeligt studie kan observationerne være individer<sup>1</sup>,  $i$ . Hvis treatment for individet  $i$  i dette studie er udtrykt ved  $D_i$  og antager formen 1 (treatment) eller 0 (fravær af treatment), kan man inddele observationerne i en treated gruppe, når  $D_i = 1$  og i en utreated gruppe når  $D_i = 0$ . Den variabel, vi ønsker at undersøge, er udtrykt ved  $Y_i$ , som kan have to potentielle outcomes afhængigt af, om et individ er treated eller ej, hvilket sammenfattes i nedenstående udtryk:

$$Y_i = \begin{cases} Y_{1i} \text{ når } D_i = 1 \\ Y_{0i} \text{ når } D_i = 0 \end{cases} \Leftrightarrow$$
$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})D_i$$

---

<sup>1</sup> Observationer behøver ikke være på individniveau.

De to potentielle outcomes for  $Y_i$  er således  $Y_{1i}$  hvis individet er treated, og  $Y_{0i}$  hvis individet ikke er treated (Angrist og Pischke 2009: 13-14). I dette setup defineres den kausale effekt af treatment udtrykt ved  $\delta_i$  som forskellen mellem  $Y_i$  når  $D_i = 1$  og  $Y_i$  når  $D_i = 0$ . Det vil sige, at følgende udtryk gælder:  $\delta_i = Y_{1i} - Y_{0i}$  (Gerber og Green 2012: 23).

Udfordringen ved denne indsigt er, at man aldrig kan observere både  $Y_{1i}$  og  $Y_{0i}$  for den samme person, da en observation logisk nok ikke både kan befinde sig i den treatede og i den utreatede gruppe, og derfor kan vi ikke direkte observere  $\delta_i$  (Gerber og Green 2012: 23-24). Hvordan kan vi så vide, hvordan en person, der ikke har fået treatment, ville reagere på treatment – og omvendt? Dette betegnes som det fundamentale problem med kausal inferens, hvilket man kan argumentere for, at matching som metode kan være behjælpelig med. Derfor anvendes matching i denne opgave.

## 2.3. Coarsened Exact Matching

### 2.3.1. Baggrund for matching

Matching er ikke i sig selv en metode til estimation. Det er i stedet en måde at behandle data på, så estimatorne bliver mindre model-afhængige (Iacus et al. 2016: 2). I en analyse med matching finder man par af observationer, der er så ens som muligt på baggrundsvARIABLE med den undtagelse, at den ene observation er treated ( $D_i = 1$ ), og den anden ikke er ( $D_i = 0$ ) (Justesen og Klemmensen 2014: 62). Vi finder altså observationer, der minder om treatede observationer, men som ikke er treatede, sådan at man tilnærmelsesvis kan observere forskellen på treatment og fravær af treatment ( $\delta_i$ ) på ”den samme” observation.

På den måde kan matching simulere det eksperimentelle design ved at observationer, der har fået treatment, så vidt muligt er magen til observationer, der ikke har fået treatment, på relevante observerbare kontrolvariable (Justesen og Klemmensen 2014: 61). Forskelle i outcomes må derfor skyldes treatment, da treatment i princippet er det, der varierer mellem de to grupper.

Kausal inferens via matching forudsætter imidlertid ligesom OLS-regression antagelse om fravær af uobserveret heterogenitet og dermed fravær af omitted variable bias. Det betyder, at de variable der påvirker, om man er treated eller ej skal være kendte og observerbare (Justesen og Klemmensen 2014: 68; Iacus et al. 2016: 2-3).

Matching og OLS er på den måde begge kontrolstrategier med ”selection on observables”, og deres estimators kausale fortolkning er afhængig af kontrol for confoundere. Matching er således ligesom OLS følsom over for systematisk påvirkning af uobserverede eller uobserverbare faktorer på  $X$  eller  $Y$  (Justesen og Klemmensen 2014: 69). Dermed kan man også argumentere for, at den empiriske

forskel mellem estimater med ren OLS-regression kontra matching er lille (Angrist og Pischke 2009: 69-70).

En styrke ved matching er imidlertid, at man eksplicit undersøger graden af såkaldt common support. Man konfronteres dermed med spørgsmålet om den overordnede balance mellem treatment- og kontrolgruppen udtrykt ved  $\mathcal{L}_1$  (Iacus et al. 2016: 4). Bevidsthed om  $\mathcal{L}_1$ -målet er dermed med til at sikre sammenlignelighed mellem observationerne (Justesen og Klemmensen 2014: 63-64). Endvidere medfører matching, at man dropper observationer, der ikke kan sammenlignes, hvilket gør treatment- og kontrolgrupperne mere heterogene, hvormed bias i estimaterne reduceres (Justesen og Klemmensen 2014: 65; Nall 2015: 396). På den måde kan matching af datasættes bruges til at reducere bias i OLS-regressioner.

Matching kan gennemføres på flere måder, hvoraf eksakt matching i udgangspunktet er den mest simple og intuitive. Her matches observationer i treatmentgruppen med observationer i kontrolgruppen, der har den eksakt samme værdi på alle kontrolvariable.

Hvis man har mange kontrolvariable med mange forskellige værdier, støder man imidlertid på dimensionalitetsproblemet, hvor det er vanskeligt at finde observationer i de to grupper, der kan matches eksakt på alle forhold (Justesen og Klemmensen 2014: 63). En løsning på dette problem er at benytte Coarsened Exact Matching (CEM).

Essensen i CEM er, at man dimensionalitetsreducerer ved at opdele baggrundsvariablenes værdier i færre, grovere (coarsened) men meningsfulde kategorier. Når der er færre dimensioner i de baggrundsvariable, man matcher på, er der færre kombinationer af værdier, der skal stemme overens, før man har et eksakt match mellem en observation i treatmentgruppen og en observation i kontrolgruppen (Nall 2015: 397-398).

Dimensionalitetsreduktion fungerer under antagelse af, at der tilnærmelsesvis ingen substantiel forskel er på hvilken af to ensartede værdier, en observation har på baggrundsvariable. Derfor kan en treated 32-årig mand eksempelvis matches med en untreated 33-årig mand – når aldersvariablen gøres grovere således at begge personer falder i den mere grovkornede alderskategori ”30-40 år” – fordi der i gennemsnit ikke burde være nogen substantiel forskel på en 32-årigs og 33-årigs livssituation, alt andet lige.

I nærværende opgave er treatment-variablen,  $D_i$ , hvorvidt respondenterne havde intention om at stemme på Trump eller ej. Kandidatvalg er ikke noget, som man kan tildele individet, da det er en



beslutning, som man (i udgangspunktet) træffer selv. Dette gør betegnelsen "treatment" abstrakt. Variablen fungerer ikke desto mindre som treatment, fordi den netop inddeler respondenterne i to grupper – dem der vil stemme på Trump (treatmentgruppe,  $D_i = 1$ ) kontra dem, der ikke vil (kontrolgruppe,  $D_i = 0$ ).

### 2.3.2. Coarsening af baggrundsvARIABLE

BaggrundsvARIABLENE for køn og kandidatvalg i 2012 kan ikke *coarses*, da de er dikotome. Alder inddeles derimod i følgende syv kategorier: under 30, 30-39 år, 40-49 år, 50-59 år, 60-69 år, 70-79 år, samt 80 år og opefter. Variablen er kategorisk. Fordelingen på kategorierne fremgår af nedenstående tabel 1:

<b>Tabel 1: Fordeling på aldersvariablen</b>		
<b>Alder</b>	<b>Antal</b>	<b>Andel af n, %</b>
Under 30	227	9,4
30-39	416	17,3
40-49	366	15,2
50-59	497	20,6
60-69	508	21,1
70-79	274	11,4
80 og derover	119	4,9
<b>SUM</b>	<b>2.407</b>	<b>99,9</b>

Da ANES-datasættet er indsamlet i en amerikansk kontekst, indeholder uddannelses-variablen 15 kategorier helt fra respondenter, der kun har gennemført første til fjerde klasse til og med ph.d.-niveau. I CEM-undersøgelsen kodes variablen om, så den indeholder følgende tre kategorier: Folkeskole og gymnasium, undergraduate, samt master og derover. Det er almindeligt i USA at færdiggøre sin uddannelse med den firårige undergraduate, hvorfor denne kategori er meningsfuld uden kandidat/master i dette tilfælde i modsætning til f.eks. dansk kontekst. Variablen er intervalskaleret, og dens fordeling fremgår af tabel 2.

<b>Tabel 2: Fordeling på uddannelsesvariablen</b>		
<b>Uddannelsesniveau</b>	<b>Antal</b>	<b>Andel af n %</b>
1: Folkeskole + gymnasium	973	40,4
2: Undergraduate	1.008	41,9
3: Master og derover	426	17,7
<b>SUM</b>	<b>2.407</b>	<b>100,0</b>

Indkomst-variablen indeholder 28 forskellige niveauer, som på samme måde som de to foregående kodes om til syv kategorier: Under 14.999 USD, 15.000-24.999, 25.000-39.999, 40.000-59.999, 60.000-79.999, 80.000-124.999 samt 125.000 USD og opefter i hustanden om året før skat.

Kategorierne er kodet med hensyn til, at de skal være meningsfulde, snarere end at der for enhver pris skal være lige langt imellem hver kategori. Variablen er således ordinal, og dens fordeling fremgår af tabel 3.

<b>Tabel 3: Fordeling på indkomstvariablen</b>		
<b>Husstandsindkomst pr. år før skat, USD</b>	<b>Antal</b>	<b>Andel af n %</b>
1: Under 14.999	253	10,5
2: 15.000-24.999	176	7,3
3: 25.000-39.999	303	12,6
4: 40.000-59.999	386	16,0
5: 60.000-79.999k	375	15,6
6: 80.000-124.999	512	21,3
7: 125.000 og derover	402	16,7
<b>SUM</b>	<b>2,407</b>	<b>100,0</b>

CEM anvendt på denne opgaves problemstilling muliggør således sammenligning af politisk opmærksomhed hos respondenter i den ene gruppe i forhold til tilnærmelsesvis ens respondenter i den anden. Resultatet af matching-processen bliver dermed et matched datasæt, som undersøgelsen anvender multivariat OLS-regression på, hvilket som nævnt på s. 8 reducerer bias i forhold til OLS-regression alene. Resultaterne herfra bliver gennemgået i afsnit 3.1, hvor  $\mathcal{L}_1$ -målene for det originale henholdsvis matchede datasæt også behandles.

Fravær af uobserveret heterogenitet er imidlertid en utestbar antagelse. For at identificere kausale effekter fungerer matching derfor godt sammen med eksempelvis et kvasiekperimentelt design, som kan være med til at eliminere den forstyrrende effekt af potentielle uobserverede eller udeladte variable (Justesen og Klemmensen 2014: 69; 74). Derfor vil nærværende analyse også blive suppleret af en instrumentvariabelanalyse.

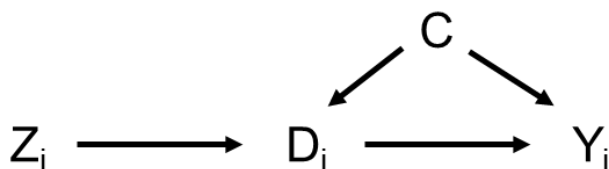
## 2.4. Instrumentvariable

Som nævnt kan OLS-regression med eller uden matching medføre uoverbevisende forsøg på kausalestimation. Med en instrument-variabel ( $Z$ ), som også blot kaldes et instrument, kan treatment-variablen imidlertid manipuleres på en sådan måde, at kontrol for baggrundsvARIABLE, som er afgørende for OLS-estimation, får mindre betydning. Instrumentvariabelanalyse (IV) fungerer

som et kvasiekperimentelt design, hvor tildeling af treatment er betinget af noget andet end tilfældighed (Angrist og Pischke 2009: 122). På den måde kan man estimere den rene effekt af  $D_i$ , der i IV-sammenhæng betegnes den endogene regressor, på outcome,  $Y$ .

Denne rene effekt af  $D_i$  på outcome kan være sløret af confoundere,  $C$ , der kan påvirke både treatment og outcome. Såfremt instrumentet,  $Z_i$ , er uafhængigt af disse confoundere, kan  $Z_i$  derfor bruges til at "instrumentere" variation i outcome  $Y_i$  igennem  $D_i$ . IV fungerer derfor som en kausal kædereaktion, hvorigennem  $Z_i$  påvirker  $D_i$ , hvilket giver outcome  $Y_i$  uafhængigt af eventuelle udeladte variable,  $C$  (Angrist og Pischke 2015: 113). På den måde løser IV problemet med omitted variable bias, som der kan være risiko for, hvis undersøgelsen udelukkende anvendte CEM (Angrist og Pischke 2009: 115). Intuitionen bag IV fremgår af nedenstående figur 1.

**Figur 1**



Til estimation af kausale effekter fungerer IV under forudsætning af, at instrumentet opfylder tre kriterier. For det første skal instrumentet opfylde relevanskriteriet, hvilket betyder, at instrumentet skal have en klar og tydelig effekt på  $D_i$  (Angrist og Pischke 2015: 106). Denne effekt kaldes "first stage", fordi den betegner effekten i det først led fra  $Z_i$  til  $D_i$  i figur 1 (Angrist og Pischke 2015: 107). Hvis relevanskriteriet ikke er opfyldt, er kæden brudt, og instrumentet er derfor svagt.

For det andet skal instrumentet opfylde uafhængighedsantagelsen; at instrumentet er urelateret til eventuelle udeladte variable. For det tredje skal instrumentet alene påvirke outcome igennem den endogene regressor. Dette forhold kaldes eksklusionsrestriktionen (Angrist og Pischke 2015: 106).

Som nævnt benyttes i denne opgave en variabel for respondenternes ideologiske selvplacering på en skala fra 1 (ekstremt liberal) til 7 (ekstremt konservativ) som instrument. Relevanskriteriet testes med en simpel bivariat logistisk regression i tabel 4, som viser en tydelig, signifikant positiv sammenhæng mellem ideologi ( $Z_i$ ) og kandidatvalg ( $D_i$ ). Jo mere konservativ, jo mere er man altså tilbøjelig til at ville stemme på Trump. Instrumentet er således ikke svagt.

**Tabel 4: Test af relevanskriteriet, first stage**

<i>Dependent variable:</i>	
Intention om at stemme Trump, $D_i$	
Ideologi, $Z_i$	1.234*** (0.046)
Constant	-5.759*** (0.215)
Observations	2,735
Log Likelihood	-1,177.072
Akaike Inf. Crit.	2,358.145
Note:	* $p < 0,1$ ** $p < 0,05$ *** $p < 0,01$

Analysen med instrumentet vedholder kontrol for de førnævnte baggrundsvariable, som ifølge Michigan-skolens kausaltrakt påvirker ideologi, og dermed forventes det, at uafhængighedsantagelsen også er opfyldt (Hansen og Stubager 2017: 386-387).

Eksklusionsrestriktionen forudsætter, at ideologi kun påvirker politisk opmærksomhed, såfremt det påvirker kandidatvalg. Denne antagelse synes opfyldt, idet der ikke umiddelbart er noget, der taler for, at konservative personer systematisk pga. ideologi skulle have større eller mindre opmærksomhed på politiske nyheder end liberale personer, eller omvendt (om de så også benytter de samme mediekilder er en anden diskussion, som ligger uden for perspektivet i denne undersøgelse).

Selvom instrumentet ikke er dikotomt, kan de teoretiske forventninger til effekten af et dikotomt instrument bruges til at uddybe IV-kæden. Forholdet mellem leddene i IV-kæden er nemlig afgørende parametre, når man skal estimere den rene effekt af  $D_i$  på  $Y_i$ . Den førnævnte first stage-effekt kan betegnes  $\Phi$  (Phi), og er i denne sammenhæng forskellen i kandidatvalg mellem liberale og konservative (Angrist og Pischke 2015: 109-110).

Et andet vigtigt parameter er den såkaldte "reduced form" med betegnelsen  $\rho$  (Rho), der dækker den direkte effekt af  $Z_i$  på  $Y_i$  igennem  $D_i$ , altså effekten ideologi på politisk opmærksomhed igennem kandidatvalg. I denne undersøgelse viser  $\rho$  dermed forskellen i den gennemsnitlige politiske opmærksomhed mellem liberale og konservative (Angrist og Pischke 2015: 109-110).

Den kausale effekt, som det er målet at afdække med IV, er som nævnt den rene effekt af  $D_i$  på  $Y_i$ , og denne betegnes "LATE" (Local Average Treatment Effect),  $\lambda$  (Lambda). LATE defineres som

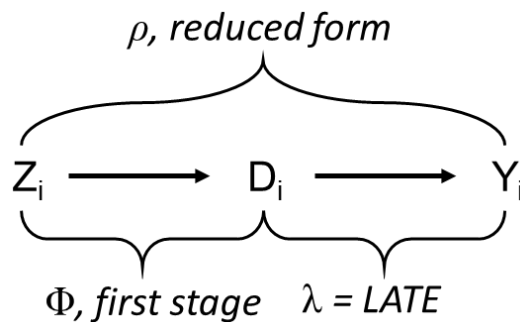
forholdet mellem  $\rho$  og  $\Phi$ . I denne opgave vil det sige, at LATE er forskellen i politisk opmærksomhed mellem liberale og konservative divideret med forskellen i kandidatvalg mellem liberale og konservative (Angrist og Pischke 2015: 107-110). Altså er effekten, opgaven undersøger, LATE, givet ved

$$LATE = \lambda = \frac{\rho}{\phi}, \text{ dvs.}$$

$$\text{Effekt af kandidatvalg på politisk opmærksomhed} = \frac{\text{Effekt af ideologi på politisk opmærksomhed}}{\text{Effekt af ideologi på kandidatvalg}}$$

Figur 2 illustrerer, hvor i IV-kæden first stage, reduced form og LATE findes.

**Figur 2**



Det er i den forbindelse nødvendigt kort at forholde sig til hvem LATE angiver effekter for. Dette spørgsmål knytter sig til problemer med "noncompliance" med treatment hos respondenterne. Respondenter kan nemlig inddeles i de fire grupper: compliers (tager treatment hvis de treats, ellers ikke), never-takers (afviser treatment hvis de treats), defiers (afviser treatment hvis de treats, tager treatment hvis de ikke treats), samt always-takers (tager treatment uanset om de treats eller ej) (Angrist og Pischke 2015: 111).

Som nævnt er treatment i denne undersøgelse lidt abstrakt om man har intention om at stemme på Trump eller ej. Hvis vi for argumentets skyld antager, at instrumentet, ideologi, også er dikotomt, kan noncompliance-situationen sammenfattes i nedenstående tabel 5<sup>2</sup>.

<sup>2</sup> Tabel 5 er baseret på Angrist og Pischke 2015: 112.

**Tabel 5: Compliance og noncompliance**

		Konservativ, $Z_i = 1$	
		Stemmer ikke Trump, $D_i = 0$	Stemmer Trump, $D_i = 1$
Liberal, $Z_i = 0$	Stemmer ikke Trump, $D_i = 0$	Never-taker	Complier
	Stemmer Trump, $D_i = 1$	Defier	Always-taker

I denne undersøgelse er eksempler på noncompliance konservative amerikanere, der ikke ville stemme på Trump, eller liberale amerikanere, der foretrak Trump frem for Clinton eller andre kandidater. Disse mennesker tager enten ikke treatment, selvom instrumentet burde få dem til det, eller også tager de treatment uanset, hvad instrumentet angiver. For disse mennesker kan vi ikke estimere LATE, og noncompliance gør det i stedet mere realistisk at estimere "CACE" (Complier Average Causal Effect), dvs. den gennemsnitlige effekt af treatment for respondenter, der er compliers (Gerber og Green 2012: 142).

Under antagelse af monotonicitet er der fravær af defiers, hvilket vil sige, at selvom instrumentet ikke påvirker alle, vil de, der påvirkes, blive påvirket på samme måde – instrumentet virker kun én vej (Angrist og Pischke 2009: 154). I så fald gælder, at  $CACE = LATE$ . LATE udtrykker med andre ord at instrumentet – hvis identificeret korrekt – identificerer den gennemsnitlige kausale effekt for den subpopulation af observationer, hvis treatment-status ændres af instrumentet. Effekten er med andre ord lokal for denne subpopulation<sup>3</sup> (Samii 2016: 950).

Med ovenstående i baghovedet bygger de substantielle resultater i den følgende IV-analyse på 2SLS-regression (Two-Stage Least Squares), som angiver statistisk signifikant korrelation mellem variable, og som muliggør kontrol for kovariater, hvormed risikoen for omitted variable bias mindskes, hvilket som nævnt indledningsvis er en af styrkerne ved instrumentvariable (Angrist og Pischke 2015: 132; 135).

### 3. Analyse

I dette afsnit analyseres effekten af intention om at stemme Trump,  $D_i$ , på politisk opmærksomhed,  $Y_i$ , først med CEM og dernæst med IV. Analysen tager udgangspunkt i en simpel bivariat sammenhæng, og en multivariat OLS-regression bliver også præsenteret.

<sup>3</sup> På samme måde estimerer matching effekter, der er lokale for respondenter med common support.

### 3.1. Analyse med CEM

Som motivation for analysen indleder jeg med en bivariat OLS-regression i tabel 6, der viser en sammenhæng mellem intention om at stemme på Trump og politisk opmærksomhed, som formentlig er biased af confoundere.

Tabellen viser, i modstrid med hypotesen, at vælgere, der havde intention om at stemme på Trump med et 95 % konfidensinterval har signifikant større politisk opmærksomhed end vælgere, der ikke vil stemme på Trump. Dette tyder på, at Trumps elitestatus og vælgernes egen confirmation bias ikke får dem til at vende medierne ryggen.

**Tabel 6: Bivariat sammenhæng mellem kandidatvalg og politisk opmærksomhed**

	<i>Dependent variable:</i>
	Politisk opmærksomhed
Intention om at stemme Trump	0,086** (0,036)
Constant	3,570*** (0,023)
Observations	3.330
R <sup>2</sup>	0,002
Adjusted R <sup>2</sup>	0,001
Residual Std. Error	1,031 (df = 3328)
F Statistic	5,587** (df = 1; 3328)
Note:	* p<0,1 ** p<0,05 *** p<0,01

Matching-analysen mindsker den potentielle bias. Forud for matching fjernes missing observationer, og jeg foretager en imbalance-test, der måler den globale balance i forhold til fordelingen af alle interaktionerne mellem kovariater i treatmentgruppen og kontrolgruppen angivet ved  $\mathcal{L}_1$ . Perfekt balance medfører, at  $\mathcal{L}_1 = 0$ , hvorimod  $\mathcal{L}_1 = 1$  angiver, at der ingen overensstemmelse er mellem de to grupper (Iacus et al. 2016: 4-5). Den indledende imbalance-test viser, at  $\mathcal{L}_1 = 0,890$ . De univariate balance-mål før matching fremgår af bilag 7.2.

Baggrundsvariablene alder, uddannelse og indkomst opdeles i grovere kategorier jævnfør afsnit 2.3.2, og dernæst genereres et ”pruned” datasæt, hvor respondenter i de nye, grovere kategorier i treatmentgruppen matches eksakt med respondenter, der har samme værdier på baggrundsvariable i kontrolgruppen. Dette medfører, at  $\mathcal{L}_1 = 0,573$ , hvilket viser, at der er bedre balance end før, hvilket

skulle give mindre modelafhængighed og mindre biased estimater, da der er mindre heterogenitet i datasættet (Iacus et al. 2016: 2). De univariate balance-mål efter matching fremgår af bilag 7.3. Matching-processen medfører imidlertid også, at datasættet trimmes ned til 1.318 sammenlignelige respondenter. De færre respondenter giver potentielt større usikkerhed, men da datasættet fortsat indeholder et højt antal observationer, vurderes dette ikke at være et problem.

Tabel 7 på næste side viser en multivariat OLS-regression, der tester effekten af intention om at stemme på Trump på politisk adfærd i det prunede datasæt inklusiv kontrol for de tidligere nævnte kontrolvariable (model 2). Tabellen viser også den tilsvarende multivariate OLS-regression på det originale, umatchedede datasæt (model 1).

Model 2 viser, at sammenhængen mellem at ville stemme på Trump og politisk opmærksom både er blevet højsignifikant og stærkere efter matching end i den bivariate sammenhæng. Model 1 tegner et tilsvarende billede, om end koefficienten ikke er lige så stor som matching-analysens koefficient. Analysen med CEM viser dermed, at vælgere, der havde intention om at stemme på Trump, havde højere politisk opmærksomhed end vælgere, der var tilnærmelsesvis ens ift. køn, alder, uddannelse, indkomst og kandidatvalg i 2012, men som ikke havde intention om at stemme på Trump.

Både model 1 og 2 viser også af signifikante sammenhænge, at mænd er mere politisk opmærksomme end kvinder, at politisk opmærksomhed stiger med alderen (med undtagelse af aldersgruppen 30-39, som er insignifikant) og med øget uddannelsesniveau, og at den politiske opmærksomhed er mindre for respondenter, der stemte på Mitt Romney snarere end Barack Obama i 2012.



**Tabel 7: Unmatched kontra matched sammenhæng**

	<i>Dependent variable:</i>	
	Politisk opmærksomhed	
	(1)	(2)
Intention om at stemme Trump	0.159*** (0.059)	0.171*** (0.065)
Køn: 1. kvinde	-0.222*** (0.040)	-0.205*** (0.056)
Alder	0.017*** (0.001)	
Alder 30-39		-0.021 (0.124)
Alder 40-49		0.273** (0.131)
Alder 50-59		0.358*** (0.122)
Alder 60-69		0.576*** (0.121)
Alder 70-79		0.741*** (0.144)
Alder 80 og derover		0.493** (0.211)
Uddannelse	0.069*** (0.010)	0.078*** (0.027)
Indkomst	0.003 (0.003)	0.021 (0.018)
Valg 2012: 1. Mitt Romney	-0.185*** (0.058)	-0.185*** (0.066)
Constant	2.081*** (0.132)	3.143*** (0.132)
Observations	2,407	1,318
R <sup>2</sup>	0.107	0.086
Adjusted R <sup>2</sup>	0.104	0.078
Residual Std. Error	0.957 (df = 2400)	0.969 (df = 1306)
F Statistic	47.730*** (df = 6; 2400)	11.133*** (df = 11; 1306)
Note:	* p<0,1 ** p<0,05 *** p<0,01	

### 3.2. Analyse med IV

Med ideologi som instrument er det målet, at instrumentvariabelanalysen skal supplere CEM med estimer, der er mindre afhængige af fravær af omitted variable bias. Fordi det er højst sandsynligt, at ideologi påvirker, hvad respondenterne stemte i 2012, indeholder 2SLS-regressionen kontrol for denne variabel for at undgå brud med eksklusionsrestriktionen. Endvidere indeholder analysen kontrol for opgavens øvrige baggrundsvariable<sup>4</sup> for at undgå brud med uafhængighedsantagelsen, da det er en klassisk indsigt, at disse sociodemografiske faktorer kan påvirke ideologi (Hansen og Stubager 2017: 386-387).

Tabel 6 viser ved 2SLS-regression effekten af intention om at ville stemme Trump på politisk opmærksomhed instrumenteret af ideologi. Estimerne er udregnet med robuste standardfejl, da 2SLS ikke tager højde for usikkerheden i first stage.

I modsætning til koefficienten fra matching-analysen viser IV-analysen i tabel 8 med de rette kontroller den signifikante sammenhæng, at respondenter, der havde intention om at ville stemme på Trump, havde betydeligt mindre politisk opmærksomhed end respondenter, der havde intention om at ville stemme på en anden kandidat (95 % konfidensinterval). Dette indikerer, at antagelsen om fravær af uobserveret heterogenitet i matching-analysen ikke er overholdt, hvorfor matching-koefficienten har været biased og misvisende.

Resultatet af IV-analysen understøtter dermed opgavens hypotese og giver en indikation af, at Trump i kraft af sin position som en politisk elite i Zallers framework kan videregive sit syn på medierne til vælgerne, der så pga. confirmation bias bliver mindre opmærksomme på politiske nyheder. Videre undersøgelser kan undersøge hvordan og i hvilket omfang denne mekanisme mere specifikt gør sig gældende.

---

<sup>4</sup> Køn, alder, uddannelsesniveau og indkomstniveau.

**Tabel 8: Instrumenteret effekt af kandidatvalg på politisk opmærksomhed**

	<i>Dependent variable:</i>
	Politisk opmærksomhed
Intention om at stemme Trump	-0.519** (0.245)
Køn: 1. kvinde	-0.238*** (0.045)
Alder	0.019*** (0.001)
Uddannelse	0.043*** (0.012)
Indkomst	0.004 (0.003)
Valg 2012: 1. Mitt Romney	0.302* (0.183)
Constant	2.427*** (0.163)
Observations	2,002
R <sup>2</sup>	0.058
Adjusted R <sup>2</sup>	0.055
Residual Std. Error	0.964 (df = 1995)
<i>Note:</i>	*p<0,1 **p<0,05 ***p<0,01

Det fremgår imidlertid af tabel 9, at hvis 2SLS-regressionen køres uden de nævnte kontroller, har koefficienten samme fortegn, men sammenhængen er betydeligt svagere og insignifikant, jf. tabel 9. Dette indikerer, at estimaterne er inkonsistente, og man skal derfor være påpasselig med at slutte en kausal sammenhæng, hvilket jeg vil diskutere nærmere i afsnit 4.

**Tabel 9: Instrumenteret bivariat effekt**

	<i>Dependent variable:</i>
	Politisk opmærksomhed
Intention om at stemme Trump	-0.043 (0.062)
Constant	3.715*** (0.032)
Observations	2,707
R <sup>2</sup>	-0.002
Adjusted R <sup>2</sup>	-0.002
Residual Std. Error	1.013 (df = 2705)
<i>Note:</i>	*p<0,1 **p<0,05 ***p<0,01

#### 4. Diskussion af resultater, kausalitet og andre designs

I dette afsnit diskuterer jeg først analyseresultaterne i forhold til mulighederne for at drage kausale slutninger samt udfordringer med det anvendte design. Dernæst præsenterer jeg alternative designs hvormed borgernes opmærksomhed på politiske nyheder kan undersøges.

##### 4.1. Analysens resultater og kausalitet

Inkonsistens i analyseresultaterne leder naturligt til en diskussion af kausalitet, og hvordan klarere resultater kan opnås. Udover de fire krav til kausalitet, som ofte anvendes, kan man argumentere for, at kausale slutninger må leve op til strengere krav: For det første må der ikke være omvendt kausalitet, sådan at der er en effekt af Y på X. For det andet må der ikke være korrelation mellem confoundere, som er korreleret med X og confoundere, som er korreleret med Y (Hariri 2012: 7).

En logistisk regression med politisk opmærksomhed som uafhængig variabel og intention om at stemme Trump som afhængig variabel viser en svag om end signifikant sammenhæng, hvilket indikerer, at det første krav ikke er overholdt ( $\beta = 0,081$ ;  $P = 0,182$ ). Det andet krav er sværere at teste men det synes ikke utænkeligt, at det heller ikke er overholdt problemerne med omitted variable bias i matching analysen og analyseresultaternes inkonsistens taget i betragtning. Det er derfor muligt, at resultaterne ikke har en kausal fortolkning på trods af tydelig korrelation (Hariri 2012: 1). Dette betyder, at man ikke nødvendigvis kan påvise en kausal sammenhæng mellem intention om at stemme på Trump og politisk opmærksomhed.

Hvis man mener, at det er mediernes opgave at fungere som en demokratiets vagt, kan man argumentere for, at dette i en demokratisk sammenhæng er positivt, fordi det betyder, at politikere,

der drager tvivl om mediedækningens relevans, ikke nødvendigvis får folk til at være mindre opmærksomme på politiske nyheder. Stadig er forholdet mellem politiske eliter, medierne, vælgerne og deres confirmation bias et emne, der må undersøges nærmere, så klarere konklusioner kan drages end dem, der følger af denne analyse.

Foruden et eksperimentelt design, der kan teste, hvordan politiske kandidater påvirker vælgernes politiske opmærksomhed eller medieforbrug generelt, kunne nærværende undersøgelse gentages med et andet og potentielt stærkere instrument. Hvor ideologi åbenlyst påvirker kandidatvalg (Zi skaber selektion ind i  $D_i$ ), hvormed relevanskriteriet er opfyldt, er antagelserne om overholdelse af uafhængighedsantagelsen og eksklusionsrestriktionen mere tvivlsomme, selvom der er blevet forsøgt kontrolleret for variable, der kan påvirke instrumentet.

Det kvasieksperimentelle potentiale i IV er større, når instrumentet mere åbenlyst er tilfældigt tildelt, f.eks. i form af et eksogent chok (Lundborg et al. 2017: 1616; 1620). I et emne så salient og præget af high politics som et amerikansk præsidentvalg har det imidlertid vist sig svært at identificere variable, der med konceptuel og teoretisk klarhed kan forventes at påvirke hinanden uden at være gensidigt påvirket af andre faktorer, som er en del af den massive maskine, et valg udgør.

Uanset hvilke empiriske kontroller og teoretiske forventninger, man tager forbehold for, er den kausale effekt af IV et spørgsmål om sammenligninger af reduced form-effekter mellem grupper defineret af instrumentet og justeret af first stage (Angrist og Pischke 2015: 131). Hvis ikke instrumentet påvirker disse grupper tilstrækkeligt tydeligt og uforstyrret, udfordres instrumentets kausale effekt simpelthen.

#### **4.2. Alternative designs i undersøgelser af politisk opmærksomhed**

Ovennævnte udfordringer leder til spørgsmålet om, hvordan IV kan perspektiveres til andre metoder. IV-analysen har som beskrevet på s. 14 en lokal karakter, fordi LATE er effekten for subpopulationen af compliers – effekten er tilstedeværende for disse mennesker alene, da det er dem, hvis status ændres af instrumentet. Denne lokale effekt på en subpopulation er ligeledes gældende i regressionsdiskontinuitets-designs (RDD), hvor effekten er lokal for respondenter omkring en cut-off tærskelværdi, der udløser et fænomen og dermed definerer en treatmentgruppe og en kontrolgruppe (Samii 2016: 950).

Med RDD måler man således effekt hos personer, med en værdi på en relevant variabel, der ligger lige over cut-off-værdien ift. personer, hvis værdi ligger lige under. F.eks. kan personer, der er 18 år

og 1 dag på valgdagen stemme hvorimod personer, der er 17 år og 364 dage ikke kan, selvom personerne praktisk talt er lige gamle. De to grupper er dermed tilnærmelsesvis ens og er derfor i høj grad sammenlignelige, men kun den gruppe, hvis alder er akkurat højere end cut-off (18 år) oplever fænomenet at kunne stemme for første gang (de "treates" tilfældigt med stemmeret) (Angrist og Pischke 2009: 251-253).

I studiet af demokratiske borgeres politiske opmærksomhed kunne RDD anvendes til at undersøge den langsigtede effekt på politisk opmærksomhed af lige akkurat at få stemmeret op til et valg, sammenlignet med personer, der lige akkurat ikke fik stemmeret, og derfor ikke blev eksponeret for den demokratiske proces på samme måde. Et eksempel på et sådant studie er Alexander Coppock og Donald P. Greens undersøgelse af effekten af lige at få stemmeret op til et valg på den langsigtede valgdeltagelse (Coppock og Green 2016). Selvom alle personer får stemmeret før eller senere, er der en særlig begejstring forbundet med lige at nå det, som kan påvirke, hvordan man forholder sig til politik, og hvor meget man følger med senere i livet. Running-variablen i en sådan undersøgelse er dermed alder, og tærskelværdien til cut-off er 18 år.

Differences-in-Differences design (DiD) udnytter ligeledes et cut-off til at undersøge effekter. I modsætning til RDD skal cut-off imidlertid være tidsligt, og effekter undersøges på aggregeret niveau, f.eks. grupper eller stater mv. Intuitionen bag DiD er, at man har to tilnærmelsesvis ens enheder, f.eks. landene  $s_1$  og  $s_0$ , hvoraf kun  $s_1$  treates med (oplever) et eksogent fænomen, f.eks. voldelige optøjer (Enos et al. 2017). Man vil undersøge, hvordan treatment påvirker  $s_1$ , så man observerer outcome  $Y_i$  for  $s_1$ , før tidspunktet for treatment ( $t_0$ ) og efter tidspunktet for treatment ( $t_1$ ), og trækker de to outcomes fra hinanden.

Men da det jævnfør det fundamentale problem med kausal inferens (afsnit 2.2) ikke kan observeres, hvordan  $s_1$  ville se ud ved  $t_1$  uden treatment, kan det ikke udelukkes, at resultatet er biased af andre faktorer, der også ændrer sig over tid. Man observerer derfor ligeledes outcome for den tilnærmelsesvis ens men utreatede enhed  $s_0$  ved både  $t_0$  og  $t_1$  og trækker disse to outcomes fra hinanden. Forskellen mellem de to enheder før og efter treatment udtrykker således den kausale effekt af treatment  $\delta_i$ , som vi søger (Angrist og Pischke 2009: 228-229). Intuitionen i DiD er således udtrykt ved ligningen:

$$\delta_i = (Y_i|s_1, t_1 - Y_i|s_1, t_0) - (Y_i|s_0, t_1 - Y_i|s_0, t_0)$$

Forudsætningen for at estimere kausale effekter med DiD er antagelsen om parallelle trends, som siger, at udviklingen over tid – trenden – i de to enheder ville være den samme med fravær af treatment (Angrist og Pischke 2009: 230). Så meget desto vigtigere er det, at de to enheder er tilnærmelsesvis ens.

DiD kunne bruges til at undersøge, hvordan fremkomsten af højrenationale politiske bevægelser påvirker borgernes politiske opmærksomhed på aggregeret niveau. I Frankrig tabte Marine Le Pen det seneste præsidentvalg, hvorimod Alternative für Deutschland (AfD) oplevede stor fremgang og potentiale for øget indflydelse i Tyskland (Christensen og Kristensen 2017). Tyskland blev på den måde treated med højrenationalisme. Landene er nabolande, og begge er store, europæiske økonomier, og det er derfor muligt at antagelsen om parallelle trends holder. På den måde kan man undersøge, hvilken effekt det højrenationale AfD's gode valgresultat har på tyskernes politiske opmærksomhed i forhold til den politiske opmærksomhed i Frankrig.

Flytter man fokus fra kausale sammenhænge til prædiktion af outcome-værdier,  $\hat{Y}$ , kan superviseret maskinlæring med udgangspunkt i data hentet via bl.a. Twitters API eksempelvis anvendes til at undersøge, hvordan og i hvilket omfang folk interagerer med nyhedssider og politikere på sociale medier (Theocharis et al. 2016). En sådan måling af folks politiske vaner kunne endvidere sammenlignes med politiske tilbøjeligheder.

Disse potentielle designs kræver naturligvis yderligere uddybning af antagelser og designs samt teoretisk underbygninger for at fungere i praksis, men eksemplerne tjener til at illustrere, hvordan forskellige designs potentielt kan anvendes til at undersøge forskellige vinkler på det samme fænomen: politisk opmærksomhed.

## 5. Konklusion

Denne opgave har søgt at undersøge, hvordan intention om at ville stemme på Trump ved det amerikanske præsidentvalg i 2016 påvirkede vælgerens opmærksomhed på nyheder om national politik. Med udgangspunkt i Zallers teori om politisk holdningsdannelse og Tabers og Lodges teori om motivated reasoning og confirmation bias, var det forventningen, at Trumps holdning til medierne ville få vælgerne til også at vende politiske nyheder ryggen. Forud for analysen var hypotesen derfor, at respondenter, der havde intention om at ville stemme på Trump ville blive mindre opmærksomme på politiske nyheder end respondenter, der ikke havde intention om at stemme på Trump.

Med data fra American National Election Survey (ANES) fra 2016 blev sammenhængen først undersøgt med Coarsened Exact Matching og OLS-regression, og dernæst med en instrumentvariabel og 2SLS-regression. Matching-analysen viste, i modstrid med hypotesen, at respondenter, der ville stemme på Trump, blev signifikant mere opmærksomme på politiske nyheder. Fravær af uobserveret heterogenitet er imidlertid en forudsætning for, at denne korrelation har en kausal fortolkning. Derfor blev denne analyse suppleret med en IV-analyse.

IV-estimerne med kontrol for kovariater, der kunne påvirke instrumentet, viste det modsatte resultat: Respondenter, der havde intention om at stemme på Trump blev signifikant mindre opmærksomme på politiske nyheder, hvilket stemmer overens med hypotesen og de teoretiske forventninger. Dette indikerer, at antagelsen om fravær af uobserveret heterogenitet ikke er opfyldt i matching-analysen. Uden kontroller blev IV-sammenhængen imidlertid svag og insignifikant.

Dette betyder, at det ikke på baggrund af denne undersøgelse er givet, at intention om at stemme på Trump havde en kausal negativ effekt på vælgernes opmærksomhed på politiske nyheder, selvom der er tegn på korrelation. De inkonsistente resultater kan foruden de førnævnte problemer med uobserveret heterogenitet også opstå, hvis instrumentet ikke opfylder uafhængighedsantagelsen og eksklusionsrestriktionen tilstrækkeligt tydeligt. I så fald kan instrumentet ikke effektivt estimere den rene effekt af X på Y.

Opmærksomhed på politiske nyheder kan derfor yderligere undersøges med et instrument, der tydeligere overholder disse kriterier, eller man kan undersøge, hvordan andre forhold påvirker opmærksomhed på politiske nyheder med andre designs. Man kan anvende regressionsdiskontinuitet til f.eks. at se på, hvordan det lige at nå at få valget op til et valg påvirker opmærksomhed på politiske nyheder downstream.

Alternativt kan et differences-in-differences design undersøge, hvordan opmærksomhed på politiske nyheder påvirkes på aggregeret niveau af højrenationale strømninger i et land, eller maskinlæring og data fra sociale medier kan anvendes til at prædiktere folks forhold til politiske nyheder, politikere og andre aktører.



## 6. Kilder

### 6.1. Data

ANES (2016). "ANES 2016 Time Series Study." *American National Election Studies*. Lokaliseret d. 1. december 2017 på:

[http://www.electionstudies.org/studypages/download/datacenter\\_all\\_NoData.php](http://www.electionstudies.org/studypages/download/datacenter_all_NoData.php)

### 6.2. Litteraturliste

Angrist, Joshua D. og Jörn-Steffen Pischke (2009). *Mostly Harmless Econometrics – An Empiricist's Companion*. Princeton University Press. Princeton og Oxford.

Angrist, Joshua D. og Jörn-Steffen Pischke (2015). *Mastering 'Metrics – The Path from Cause to Effect*. Princeton University Press. Princeton og Oxford.

Christensen, Lisa Maria Frendved og Lars Hjorth Kristensen (24. september 2017). "Succes til højernationalt parti forandrer tysk politik - valg kaldes et chok for Merkel." *TV2*. Lokaliseret d. 21. december 2017 på: <http://nyheder.tv2.dk/udland/2017-09-24-succes-til-hoejernationalt-parti-forandrer-tysk-politik-valg-kaldes-et-chok-for>.

Gerber, Alan S. og Donald P. Green (2012). *Field Experiments – Design, Analysis, and Interpretation*. W. W. Norton & Company. New York og London.

Enos, Ryan D., Aaron Russell Kaufman og Melissa L. Sands (2017). "Can Violent Protest Change Local Policy Support? Evidence from the Aftermath of the 1992 Los Angeles Riot." Udkast under udarbejdelse fra 8. januar 2017.

Carmines, Edward G. og Robert Huckfeldt (1998). "Political Behavior: An Overview." I: Robert E. Goodin og Hans-Dieter Klingemann (red.). *New Handbook of Political Science*. Oxford University Press. Oxford. S. 223-305.

Coppock, Alexander og Donald P. Green (2016). "Is Voting Habit Forming? New Evidence from Experiments and Regression Discontinuities." *American Journal of Political Science*. Vol. 60, No. 4. S. 1044-1062.

Hansen, Kasper Møller og Rune Stubager (2017). "Konklusion – en samlet vælgeradfærdsmodel" I: Kasper Møller Hansen og Rune Stubager (red.). *Oprør fra udkanten – Folketingsvalget 2015*. (1. udgave, 2. oplag). Jurist- og Økonomforbundets Forlag. S. 385-413.

Iacus, Stefano M., Gary King og Giuseppe Porro (2016). "CEM: Software for Coarsened Exact Matching." *Journal of Statistical Software*.

Jennings, M. Kent, Laura Stoker og Jake Bowers (2009). "Politics across Generations: Family Transmission Reexamined." *Journal of Politics*. Vol. 71, No. 3. S. 782-799.

Justesen, Mogens K. og Robert Klemmensen (2014). "Sammenligninger af observationer: kausalitet, matching og observationsdata." *Politica*. Vol. 46, No. 1. S. 60-78.

Lundborg, Peter, Erik Plug og Astrid Würtz Rasmussen (2017). "Can Woman Have Children and a Career? IV Evidence from IVF Treatments." *American Economic Review*. Vol. 107, No. 6. S. 1611-1637.

Nall, Clayton (2015). "The Political Consequences of Spatial Policies: How Interstate Highways Facilitated Geographic Polarization." *The Journal of Politics*. Vol. 77, No. 2. S. 394-406.

Samii, Cyrus (2016). "Causal Empiricism in Quantitative Research." *The Journal of Politics*. Vol. 76, No. 3. S. 941-955.

Taber, Charles S. og Milton Lodge (2006). "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science*. Vol. 50, No. 3. S. 755-769.

Zaller, John R. (1992). *The nature and origins of mass opinion*. Cambridge University. Cambridge. S. 1-39.

Zuckerman, Alan S. (2015). "Personal networks as contexts for political behavior" I: Zuckerman Allan S. (red.). *The social logic of politics*. Temple University Press. Philadelphia. S. 3-20.

## 7. Bilag

### 7.1. Centrale variable

Variabel	Type	Beskrivelse ( <b>kilde: ANES 2016</b> )
<b>polop</b> <b>V161009</b>	Outcome, Y <sub>i</sub>	How much attention do you pay to news about national politics on TV, radio, printed newspapers, or the Internet? [Skala fra 1: "None at all" til 5: "A great deal"]
<b>trump</b> <b>V161031</b>	Treatment, D <sub>i</sub> / X	Who do you think you will vote for? [1. Trump / 0. Anden kandidat]
<b>ideologi</b> <b>V161126</b>	Instrument, Z <sub>i</sub>	We hear a lot of talk these days about liberals and conservatives. Here is a seven-point scale on which the political views that people might hold are arranged from extremely liberal (1) to extremely conservative (7). Where would you place yourself on this scale, or haven't you thought much about this? [Skala fra 1 til 7]
<b>køn</b> <b>V161342</b>	Kontrol	Respondent's self-identified gender. [1. Kvinde / 0. Mand]
<b>alder</b> <b>V161267</b>	Kontrol	Respondent's age.
<b>uddannelse</b> <b>V161270</b>	Kontrol	What is the highest level of school you have completed or the highest degree you have received?
<b>indkomst</b> <b>V161361x</b>	Kontrol	The next question is about the total income of all the members of your family living here / your total income in 2015, before taxes. This figure should include income from all sources, including salaries, wages, pensions, Social Security, dividends, interest, and all other income.
<b>valg2012</b> <b>V161006</b>	Kontrol	In 2012 Barack Obama ran on the Democratic ticket against Mitt Romney for the Republicans. Do you remember for sure whether or not you voted in that election? (HVIS JA): Which one did you vote for? [1. Mitt Romney / 0. Barack Obama]

### 7.2. Imbalance output for matching

Multivariate Imbalance Measure: L1=0.890  
Percentage of local common support: LCS=8.4%

Univariate Imbalance Measures:

	statistic	type	L1	min	25%	50%	75%	max
køn	20.0502108	(Chi2)	0.093656494	NA	NA	NA	NA	NA
alder	4.5716916	(diff)	0.003626112	3	6	5	3	0
uddannelse	-0.3830247	(diff)	0.012263883	2	0	-1	0	0
indkomst	0.4743523	(diff)	0.055381664	0	1	1	0	0
valg2012	1283.0445952	(Chi2)	0.735468990	NA	NA	NA	NA	NA

### 7.3. Imbalance output efter matching

Multivariate Imbalance Measure: L1=0.573  
Percentage of local common support: LCS=100.0%

Univariate Imbalance Measures:

	statistic	type	L1	min	25%	50%	75%	max
køn	9.23044552	(Chi2)	0.08570190	NA	NA	NA	NA	NA
alder	29.68121967	(Chi2)	0.11135978	NA	NA	NA	NA	NA
uddannelse	0.05819299	(diff)	0.04303512	0	0	0	0	0
indkomst	0.42288648	(diff)	0.09792238	0	1	0	0	0
valg2012	395.19656379	(Chi2)	0.55104534	NA	NA	NA	NA	NA

## 8. R-script

```
library("Hmisc")
library("dplyr")
library("car")
library("stargazer")
library("foreign")
library("cem")
library("AER")
library("sandwich")

getwd()
list.files()
anesdatadta <- read.dta("anes_timeseries_2016_Stata12.dta")
saveRDS(anesdatadta, "ANES 2016.rds")

anesdalle <- readRDS("ANES 2016.rds")

anesd <- select(anesdalle,
  V161003,V161006,V161008,V161009,V161031,V161126,V161139,V161153,
  V161154,V161171,V161205,V161219,V161267,V161270,V161297,V161298,
  V161342,V161361x)

rm(anesdatadta)

##INSPEKTION AF DATA
glimpse(anesd)

str(anesd$V161031) #Numerisk
describe(anesd$V161031) #Intention om at stemme Trump eller andre, X

str(anesd$V161009) #Faktor
describe(anesd$V161009) #Politisk opmærksomhed, Y

##BEARBEJDNING AF DATA
#Omkodning af stemmeintention-variablen
anesd <- rename(anesd, "trump" = "V161031")
anesd$trump <- Recode(anesd$trump,
  "-9=NA;-8=NA;-1=NA;1=0;2=1;3=0;4=0;5=0;6=0;7=0;8=0")
anesd$trump <- as.factor(anesd$trump)
anesd$trump <- Recode(anesd$trump,"0='0. Ikke Trump';1='1. Trump'")

str(anesd$trump)
describe(anesd$trump) #Variablen er omkodet korrekt

#Omkodningen af politisk opmærksomhed-variablen Y
```

```
anesd <- rename(anesd, "polop" = "V161009")
```

```
levels(anesd$polop) #1. Finder kategoriernes factor levels
```

```
anesd$polop <- as.numeric(anesd$polop) #2. Omdan til numerisk (angiver faktor levels, ikke  
værdier)
```

```
#3. Omkod factor levels til værdier vha. informationen fra 1.
```

```
anesd$polop <- Recode(anesd$polop,  
  "1=NA;2=NA;3=NA;4=5;5=4;6=3;7=2;8=1")
```

```
str(anesd$polop)
```

```
describe(anesd$polop) #Omkodet korrekt
```

```
##INDLEDENDE SAMMENHÆNGE
```

```
#Bivariat
```

```
olsbi <- lm(polop~trump, data=anesd)
```

```
summary(olsbi)
```

```
stargazer(olsbi, type="text",
```

```
  title="Tabel 6: Bivariat sammenhæng mellem kandidatvalg og politisk opmærksomhed",
```

```
  dep.var.labels="Politisk opmærksomhed",
```

```
  covariate.labels="Intention om at stemme Trump",
```

```
  out="1. Bivariat X på Y.htm")
```

```
#Multivariat
```

```
##Navngiver og omkoder kontrolvariable
```

```
anesd <- anesd %>%
```

```
  rename("køn"="V161342",
```

```
    "alder"="V161267",
```

```
    "uddannelse"="V161270",
```

```
    "indkomst"="V161361x",
```

```
    "valg2012"="V161006" #Hvad stemte du ved præsidentvalget 2012?
```

```
  )
```

```
describe(anesd$køn)
```

```
describe(anesd$alder)
```

```
describe(anesd$uddannelse)
```

```
describe(anesd$indkomst)
```

```
describe(anesd$valg2012)
```

```
str(anesd$køn) #Factor
```

```
levels(anesd$køn)
```

```
anesd$køn <- Recode(anesd$køn,"
```

```
  '-9. Refused'=NA;'-8. Don\'t know (FTF only)'=NA;'3. Other'=NA;
```

```
  '1. Male'=0. Mand;'2. Female'=1. Kvinde''')
```

```
str(anesd$alder) #Numerisk
anesd$alder <- Recode(anesd$alder, "-9=NA; -8=NA")

str(anesd$uddannelse) #Faktor
levels(anesd$uddannelse)
anesd$uddannelse <- as.numeric(anesd$uddannelse)
anesd$uddannelse <- Recode(anesd$uddannelse, "
    1=NA;2=NA;3=1;4=2;5=3;6=4;7=5;8=6;9=7;10=8;11=9;12=10;
    13=11;14=12;15=13;16=14;17=NA;18=15;19=9;20=NA")

head(anesd$indkomst)
str(anesd$indkomst) #Factor=!numerisk som ved Y
levels(anesd$indkomst)
anesd$indkomst <- as.numeric(anesd$indkomst)
anesd$indkomst <- Recode(anesd$indkomst, "
    1=NA;2=NA;3=1;4=2;5=3;6=4;7=5;8=6;9=7;10=8;11=9;12=10;
    13=11;14=12;15=13;16=14;17=15;18=16;19=17;20=18;21=19;
    22=20;23=21;24=22;25=23;26=24;27=25;28=26;29=27;30=28")

str(anesd$valg2012) #Factor
levels(anesd$valg2012)
anesd$valg2012 <- as.numeric(anesd$valg2012)
anesd$valg2012 <- Recode(anesd$valg2012, "
    1=NA;2=NA;3=NA;6=NA;7=NA;8=NA")
anesd$valg2012 <- as.factor(anesd$valg2012)
anesd$valg2012 <- Recode(anesd$valg2012, "
    4='0. Barack Obama';5='1. Mitt Romney'")

###Multivariat
olsmu <- lm(polop~trump+køn+alder+uddannelse+indkomst+valg2012,data=anesd)
summary(olsmu)

##ANALYSE MED COARSE EXACT MATCHING, CEM
#Reducerer datasættet til de relevante variable og fjerner missing
anesd_red <- anesd %>%
  select(polop,trump,køn,alder,uddannelse,indkomst,valg2012) %>%
  na.omit(anesd_red)

#1. Evaluerer balance
imbalance(anesd_red$trump,data=anesd_red,drop=c("trump","polop")) #L1=0,890

#2. Coarsening variable
anesd_red$alder <- cut(anesd_red$alder,
  breaks = c(18,30,40,50,60,70,80,Inf),
```

```
labels = c("Under 30","30-39","40-49","50-59","60-69","70-79","80+"),
right = FALSE)
describe(anesd_red$alder)

anesd_red$uddannelse <- Recode(anesd_red$uddannelse,"
1=1;2=1;3=1;4=1;5=1;6=1;7=1;8=1;9=1;10=1;
11=3;12=3;13=3;
14=4;15=4;")
describe(anesd_red$uddannelse) #Folkeskole+gymnasium, undergraduate, master og over

anesd_red$indkomst <- Recode(anesd_red$indkomst,"
1=1;2=1;3=1;4=1;
5=2;6=2;7=2;8=2;
9=3;10=3;11=3;12=3;
13=4;14=4;15=4;16=4;
17=5;18=5;19=5;20=5;
21=6;22=6;23=6;24=6;
25=7;26=7;27=7;28=7")
describe(anesd_red$indkomst) #Under 14.999 USD, 15-25, 25-40, 40-60, 60-80, 80-125, 125+

#3. Kører matching
cemmatch <- cem(treatment="trump",data=anesd_red,drop="polop")

#4. Laver pruned datasæt
anesmatch <- anesd_red[cemmatch$matched,]

#5. Evaluerer balance igen
imbalance(anesmatch$trump,data=anesmatch,drop=c("trump","polop")) #L1=0,573

#6. Estimerer effekter med pruned data
olsprunbi <- lm(polop~trump,data=anesmatch)
summary(olsprunbi)
stargazer(olsprunbi,type="text")

olsprunmu <-
lm(polop~trump+køn+alder+uddannelse+indkomst+valg2012,data=anesmatch)
summary(olsprunmu)

stargazer(olsmu,olsprunmu,type="text",
title="Tabel 7: Matched og unmatched sammenhæng",
dep.var.labels="Politisk opmærksomhed",
covariate.labels=c("Intention om at stemme Trump","Køn: 1. kvinde",
"Alder","Alder 30-39","Alder 40-49", "Alder 50-59",
"Alder 60-69", "Alder 70-79", "Alder 80 og +",
"Uddannelse","Indkomst","Valg 2012: 1. Mitt Romney"),
```



```
out="2. Multivariat og matching.htm")
```

```
#INSTRUMENT VARIABLE ANALYSE, IV
```

```
#Omkoder instrumentet: ideologisk selvplacering (H-V), Z
```

```
anesd <- rename(anesd,"ideologi"="V161126")
```

```
str(anesd$ideologi) #Faktor
```

```
describe(anesd$ideologi)
```

```
levels(anesd$ideologi)
```

```
anesd$ideologi <- as.numeric(anesd$ideologi)
```

```
anesd$ideologi <- Recode(anesd$ideologi,"  
1=NA;2=NA;3=1;4=2;5=3;6=4;7=5;8=6;9=7;10=NA")
```

```
#Tester relevanskriteriet, at Z påvirker X
```

```
firststage <- glm(trump~ideologi,data=anesd,family="binomial")
```

```
summary(firststage) #Højsignifikant positiv effekt af ideologi Z på Trump-vote X. Jo mere  
konservativ, jo mere Trump.
```

```
stargazer(firststage,type="text",  
  title="Tabel 4: Test af relevanskriteriet, first stage",  
  dep.var.labels="Intention om at stemme Trump, Di",  
  covariate.labels="Ideologi, Zi",  
  out="3. Relevanskriteriet.htm")
```

```
#IVreg
```

```
ivbi <- ivreg(polop~trump,~ideologi,data=anesd)
```

```
ivbir <- coeftest(ivbi,vcovHC(ivbi))
```

```
summary(ivbi)
```

```
stargazer(ivbi,type="text",  
  title="Tabel 7: Instrumenteret bivariat sammenhæng",  
  dep.var.labels="Politisk opmærksomhed",  
  covariate.labels="Intention om at stemme Trump",  
  out="4. Instrument bivariat standard.htm")
```

```
stargazer(ivbir,type="text",  
  title="Tabel 7: Instrumenteret bivariat sammenhæng",  
  dep.var.labels="Politisk opmærksomhed",  
  covariate.labels="Intention om at stemme Trump",  
  out="5. Instrument bivariat robust.htm")
```

```
ivmu <- ivreg(polop~trump+køn+alder+uddannelse+indkomst+valg2012,
```

```
~ideologi+køn+alder+uddannelse+indkomst+valg2012,data=anesd)
ivmur <- coeftest(ivmu,vcovHC(ivmu))

summary(ivmu)

stargazer(ivmu,type="text",
  title="Tabel 6: Instrumenteret effekt af kandidatvalg på politisk opmærksomhed",
  dep.var.labels="Politisk opmærksomhed",
  covariate.labels=c("Intention om at stemme Trump",
    "Køn: 1. kvinde", "Alder", "Uddannelse",
    "Indkomst","Valg 2012: 1. Mitt Romney"),
  out="6. Instrument multivariat standard.htm")

stargazer(ivmur,type="text",
  title="Tabel 6: Instrumenteret effekt af kandidatvalg på politisk opmærksomhed",
  dep.var.labels="Politisk opmærksomhed",
  covariate.labels=c("Intention om at stemme Trump",
    "Køn: 1. kvinde", "Alder", "Uddannelse",
    "Indkomst","Valg 2012: 1. Mitt Romney"),
  out="7. Instrument multivariat robust.htm")

#HARIRI-TEST
#A) Påvirker Y på X
summary(glm(trump~polop,data=anesd,family="binomial"),type="text")

#Køn alder uddannelse indkomst valg2012
summary(lm(ideologi~køn,data=anesd))
summary(lm(ideologi~alder,data=anesd))
summary(lm(ideologi~uddannelse,data=anesd))
summary(lm(ideologi~indkomst,data=anesd))
summary(lm(ideologi~valg2012,data=anesd))
```