

Henrik Benjaminsen
Gina Matilde Holseth

En empirisk undersøkelse av effekten av en IKEA-etablering på varehandelssektoren

Masteroppgave i Samfunnsøkonomi

Veileder: Bjarne Strøm

Trondheim, juni 2019

Norges Teknisk-Naturvitenskapelige Universitet

Fakultet for økonomi

Institutt for samfunnsøkonomi

Forord

Denne masteroppgaven markerer slutten på vår fem-årige utdanning ved institutt for samfunnsøkonomi ved NTNU. Vi ønsker å takke vår veileder Bjarne Strøm for gode tilbakemeldinger, oppfølging og tilgjengelighet under skrivingen.

Vi vil også takke våre medstudenter på lesesal 6B for gode samtaler, mye latter og motiverende klapp på skulderen i tunge stunder.

(En del av) de data som er benyttet i denne publikasjonen er hentet fra kommunedatabasen til NSD — Norsk senter for forskningsdata AS. NSD er ikke ansvarlig for analyse av dataene eller for de tolkninger som er gjort her.

Alle synspunkter og eventuelle feil i oppgaven er våre egne.

Norges Teknisk-Naturvitenskapelige Universitet

Trondheim, juni 2019

Henrik Benjaminsen

Gina Matilde Holseth

Sammendrag

Formålet med denne oppgaven er å undersøke hvordan en IKEA-etablering påvirker omsetningen og sysselsettingen i varehandelssektoren i kommunen der IKEA etableres. Oppgaven tar utgangspunkt i et paneldatasett for et utvalg av norske kommuner i tidsperioden 1977 til 1992. Utfallsvariablene er omsetning i varehandelssektoren per innbygger og sysselsetting i varehandelssektoren som andel av befolkning.

Vi benytter to ulike estimeringsmetoder for å finne etableringseffekten, differanse i differanse-metoden (DID) og den syntetiske kontrollmetoden (SCM). Vi får tvetydige resultater for etableringseffektene. Resultatene fra DID-metoden indikerer en negativ effekt på omsetningen, mens SCM finner en positiv effekt assosiert med etableringen. Begge metodene finner en positiv effekt av IKEA-etableringen på sysselsettingen.

Eksterne hendelser i form av et oljeprissjokk og en ny åpningstidslov gjør det utfordrende å skille ut en behandlingseffekt som ikke er påvirket av disse hendelsene.

Utfordringer med statistisk inferens, i begge metodene vi benytter, gjør det vanskelig å konkludere med hvorvidt effektene vi observerer er troverdige estimater for de sanne etableringseffektene.

Nøkkelord – Masteroppgave, IKEA-etablering, Differanse i differanse-metoden, Den syntetiske kontrollmetoden, behandlingseffekt.

Abstract

The purpose of this thesis is to investigate the impact of a new IKEA establishment on revenue and employment in the durable goods retail sector in the entry municipality. The thesis is based on a panel dataset for a selection of Norwegian municipalities in the timeperiod 1977 to 1992. The variables of interest are revenue in the durable goods retail sector per capita, and employment in the durable goods retail sector as a share of the population.

We use two different estimation methods to find the effect of the IKEA establishment, the difference in difference method (DID) and the synthetic control method (SCM). We get ambiguous results from the two methods. The results from the DID method indicate a negative effect on revenue, while SCM finds a positive effect. Both methods find a positive effect on employment.

External events in the form of an oil price shock and a new opening hours law make it difficult to distinguish a treatment effect that is not affected by these events.

The challenges with statistical inference, in both of the methods that we use, make it difficult to conclude whether the effects we observe are credible estimates for the true effects of a new IKEA establishment.

Keywords – Master thesis, IKEA-entry, The difference-in-difference method, The Synthetic Control method, Entry effect.

Innhold

1	Innledning	1
1.1	Problemstilling og hypotese	1
2	Utgangspunkt for analyse	3
2.1	Tidligere forskning	3
2.2	Teoretisk rammeverk	5
2.2.1	Direkte effekter	5
2.2.2	Indirekte effekter	6
3	Økonometrisk rammeverk	9
3.1	Innledning	9
3.2	Differanse i differanse-metode (DID)	9
3.2.1	Modell med koblet data for enhet og år	10
3.2.2	Modell med paneldata	12
3.3	Den syntetiske kontrollmetoden (SCM)	13
3.3.1	Modellen	14
3.3.2	Implementering av SCM	16
4	Datamaterialet	19
4.1	Behandlingsenheter	19
4.2	Definering av utvalget	22
4.2.1	Alternativ kontrollgruppe	22
4.3	Variabler	23
4.3.1	Utfallsvariabler	23
4.3.2	Forklaringsvariabel	25
4.3.3	Kontrollvariabler	26
4.4	Deskriptiv statistikk	29
4.5	Svakheter ved datamaterialet	30
4.6	Oppsummering	31
5	Metodiske utfordringer	33

5.1	Utfordringer ved DID-metoden	33
5.1.1	Forutsetninger	33
5.1.2	Økonometriske utfordringer	34
5.2	Utfordringer og begrensninger ved SCM	35
6	Resultater	37
6.1	DID	37
6.1.1	DID: kontrollgruppe 1.	38
6.1.2	DID: kontrollgruppe 2.	40
6.1.3	DID: Placeboanalyse.	42
6.1.4	Oppsummering	45
6.2	SCM	45
6.2.1	Effekt på omsetning	45
6.2.2	Effekt på sysselsetting	47
6.2.3	SCM: Placebo og statistisk inferens.	49
6.2.4	Placebotesting	50
6.2.5	RMSPE-rangering	53
6.2.6	Oppsummering	55
7	Oppsummering og diskusjon	57
7.1	Mulige utvidelser	60
8	Konklusjon	63
	References	65
	Appendiks	69
A	Restleddsegenskaper ved bruk av paneldata	69
B	Utvidelse av det økonometriske rammeverket til SCM	70
C	Utelatte kommuner fra datamaterialet grunnet manglende observasjoner.	72
D	Kommunebestemte åpningstidsreguleringer i 1982	73
E	Kommuner i kontrollgruppe 1	75
F	Differanse i differanse-metode.	76
G	Den syntetiske kontrollmetoden	82

Figurer

1	Grafisk fremstilling av differanse i differanse-estimatoren fra tabell 1. . . .	11
2	Utvikling over tid i omsetningen og sysselsettingen i varehandelssektoren i Sandnes/Stavanger.	21
3	Nasjonal utvikling over tid, med trendlinje, i omsetning i varehandelssektoren per innbygger fra 1977 til 1992.	24
4	Nasjonal utvikling over tid, med trendlinje, i sysselsetting i varehandelssektoren som andel av befolkning fra 1977 til 1992.	25
5	Grafisk fremstilling av utviklingen og etableringseffekten på omsetning i varehandelssektoren per innbygger.	46
6	Grafisk fremstilling av utviklingen og etableringseffekten på sysselsetting i varehandelssektoren som andel av befolkning.	48
7	In place-placebotest for effekten på omsetning.	50
8	To-sidige p-verdier for effekten på omsetning.	52
9	To-sidige p-verdier for effekten på sysselsetting.	52
10	RMSPE-raten for Bergen og kontrollgruppe 1.	54
11	RMSPE-raten for Bergen og kontrollgruppe 1.	54
12	In place-placebotest for effekten på sysselsetting.	83

Tabeller

1	Illustrasjon av DID-estimatoren (Wooldridge, 2015)	11
2	Deskriptiv statistikk for utfallsvariablene.	29
3	Gjennomsnittsverdi for utfallsvariablene i perioden før etableringen, 1977-1983.	38
4	Estimerte effekter av IKEA-etableringen på omsetning i varehandelssektoren per innbygger. Kontrollgruppe 1.	38
5	Estimerte effekter av IKEA-etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 1.	39
6	Estimerte effekter av IKEA-etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 2.	40
7	Estimerte effekter av IKEA-etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 2.	41
8	Estimerte effekter av den falske etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 1 og 2.	43
9	Estimerte effekter av den falske etableringen på sysselsettingen som andel av befolkning. Kontrollgruppe 1 og 2.	44
10	Utviklingen i omsetning i varehandelssektoren per innbygger for Bergen og den syntetiske kontrollenheten. Differansen i utviklingen angir effekten av IKEA-etableringen.	47
11	Utviklingen i sysselsettingen i varehandelssektoren som andel av befolkning for Bergen og den syntetiske kontrollenheten. Differansen i utviklingen angir effekten av IKEA-etableringen.	49
12	Kommuner med manglende observasjoner som dermed er utelatt fra datamaterialet.	72
13	Kommunebestemte stengetider på hverdager for kommunene i kontrollgruppe 1, i tillegg til Bergen, Sandnes og Stavanger, før Åpningstidsloven trådte i kraft i 1985.	73
14	Kommunebestemte stengetider på lørdager for kommunene i kontrollgruppe 1, Bergen, Sandnes og Stavanger, før åpningstidsreguleringen trådte i kraft i 1985.	74

15	Kommunene som inngår i kontrollgruppe 1	75
16	Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 1. .	76
17	Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 1.	77
18	Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 2. .	78
19	Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på sysselsetting i varehandelssektoren som andel av befolkning. Kontrollgruppe 2.	79
20	Fullstendige resultater fra estimeringen av effekten av den falske etableringen på omsetning i varehandelssektoren per innbygger. Kontrollgruppe 1 og 2.	80
21	Fullstendige resultater fra estimeringen av effekten av den falske etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 1 og 2.	81
22	Vekting av kommuner og modellspesifikasjon i den syntetiske kontrollgruppen til Bergen. Omsetning per innbygger.	82
23	Vekting av kommuner og modellspesifikasjon i den syntetiske kontrollgruppen til Bergen. Sysselsetting som andel av befolkning.	82

1 Innledning

I de siste tiårene har vi sett en gradvis utvikling i varehandelssektoren. Fra hovedsaklig å bestå av små kjeder og familiebedrifter har sektoren opplevd en vridning mot større aktører og såkalte big box-enheter.¹ Den svenske møbelkjeden IKEA er et kjent eksempel på en slik big box-enhet. IKEA har blitt én av verdens største hjemmeinnredningskjeder med over 200 000 ansatte, 422 varehus og nesten en milliard besøkende på verdensbasis.² Med en gjennomsnittlig omsetning på 900 millioner i året per varehus, er det rimelig å anta at en ny IKEA-etablering kan føre til stor regional vekst i etableringsområdet (IKEA, 2017b). Den faktiske effekten av en slik etablering er imidlertid usikker, ettersom det er gjennomført få undersøkelser på effektene av en IKEA-etablering.

Det finnes et stort antall studier som har studert effekten av big box-etableringer, med fokus på hvordan denne påvirker omsetningen og sysselsettingen i varehandelssektoren. Hoveddelen av disse studiene undersøker etableringseffekter av Walmart i USA. Resultatene de finner er tvetydige, og gir ingen klar indikasjon om effekten har en positiv eller negativ innvirkning på omsetning og sysselsetting. På grunn av grunnleggende forskjeller mellom IKEA og Walmart er det problematisk å generalisere effektene som er funnet i studier på Walmart til å gjelde for IKEA-etableringer.

Det finnes et lite antall studier som ser på effekten av IKEA-etableringer i Sverige, hvor samtlige benytter data fra 2000-tallet. Så vidt vi kjenner til er det ingen empiriske studier som undersøker effekten av IKEA-etableringer i Norge. Med denne oppgaven ønsker vi derfor å bidra til litteraturen ved å undersøke effekten av en IKEA-etablering i Norge på 1980-tallet.

1.1 Problemstilling og hypotese

Målet med oppgaven er å undersøke effekten av en IKEA-etablering på omsetningen og sysselsettingen i varehandelssektoren i kommunen der IKEA etableres. Undersøkelsen baserer seg på paneldata for årene 1977 til 1992. I denne tidsperioden ble det etablert to IKEA-varehus i Norge. Ett i Bergen i 1984 og ett i Sandnes i 1988. En mulig effekt er

¹Big box-enheter refererer til kjedebutikker med stort utvalg, ofte med et utseende som "en stor boks".

²Tall hentet fra IKEA, "IKEA by numbers, 2018".

at en IKEA-etablering, med behov for mange ansatte og potensiale for stor omsetning, vil øke omsetningen og sysselsettingen i varehandelssektoren direkte. Det kan også gi indirekte effekter på omsetningen og sysselsetting via effekten etableringen har på allerede etablerte bedrifter i varehandelssektoren. Nettoeffekten på omsetning og sysselsetting i det regionale området er derfor et empirisk spørsmål.

For å undersøke og estimere effektene av etableringen tar vi utgangspunkt i to ulike estimeringsmetoder. Vi bruker omsetningen og sysselsettingen i varehandelssektoren som et mål på hvordan varehandelssektoren påvirkes av IKEA-etableringen.

Oppgaven består av syv kapitler og er strukturert på følgende måte. Kapittel 2 legger utgangspunktet for analysen og presenterer både resultater fra tidligere forskning og det teoretiske rammeverket i oppgaven. Videre blir metodene som benyttes i de empiriske analysene redegjort for i kapittel 3. I kapittel 4 presenteres og defineres datamaterialet som brukes i oppgaven. Kapittel 5 belyser ulike metodiske utfordringer ved analysene vi skal gjennomføre, mens kapittel 6 presenterer resultatene fra analysene og undersøker robustheten ved ulike metoder for statistisk inferens. Kapittel 7 oppsummerer resultatene fra analysen, sammenligner disse på tvers av metoder og kobler det opp mot tidligere forskning. Helt til slutt presenteres konklusjonen i kapittel 8.

2 Utgangspunkt for analyse

I dette kapittelet presenteres tidligere forskning på effekter av etablering av IKEA og andre big box-enheter. Deretter presenterer vi et teoretisk rammeverk som gjennomgår mulige direkte og indirekte effekter på varehandelssektoren som følge av en IKEA-etablering.

2.1 Tidligere forskning

Det er gjennomført et stort antall studier som undersøker effektene av big box-etableringer, hvorav de fleste har sett på detaljistkjeden Walmart i USA. Store deler av den tidligere forskningen som kan være relevant med tanke på en IKEA-etablering er gjennomgått i artikkelen til Pedersen (2014), og er studier som ser på effekter av Walmart-etablering.³ De finner tvetydige resultater for effekten på omsetning, men forskningen indikerer at det er en positiv totaleffekt i etableringsområdet (Pedersen, 2014). For sysselsettingen er det ingen klar effekt.

Det er rimelig å anta at etableringseffektene som følge av en IKEA-etablering skiller seg fra effektene av en Walmart-etablering. En grunn til dette kan være at IKEA konsentrerer seg mer om salg av varige forbruksvarer som møbler, interiør og lignende. Konsumenter har en tendens til å reise lengre for å kjøpe varige fremfor ikke-varige forbrugsgoder⁴ (Brown, 1993). Walmart kan dermed tilfredsstille nærliggende etterspørsel i større grad enn hva IKEA kan. Dette gjør at vi fokuserer på tidligere forskning som undersøker effekter av IKEA-etableringer, siden dette er mer relevant i forbindelse med vår problemstilling.

De siste 10 årene er det publisert flere studier som ser på effektene av IKEA-etableringer i Sverige. Daunfeldt, Mihaescu, Nilsson & Rudholm (2014) studerer effektene av IKEA-etablering på omsetningen og sysselsettingen i varehandelssektoren. De undersøker også tilførselen av kjøpekraft i etableringskommunen for å kontrollere for tilreisende fra andre kommuner som handler på IKEA. Daunfeldt, Mihaescu, Nilsson & Rudholm (2015) ser på hvordan en IKEA-etablering påvirker bedrifter i

³Pedersen (2014) er en rapport som ble laget som del av et forprosjekt i forkant av etableringen av IKEA på Danebo i Vestfold. Målet var å utvikle kunnskap om hvordan positive effekter kan oppnås og negative unngås i varehandelssektoren som følge av en big box-etablering. Etablering på IKEA på Danebo ble aldri realisert.

⁴Varige forbrugsgoder er møbler, kjøkkenmaskiner, og lignende, og er varer med lengre holdbarhet enn mat, sminke, med flere (Ekberg, 2018).

varehandelssektoren i kommuner der det etableres en IKEA. Rudholm, Li & Carling (2018) undersøker effekten på produktiviteten innenfor varehandelssektoren i kommuner som får en IKEA-etablering. Vi tar nå for oss de tre studiene i rekkefølgen de ble presentert og gjennomgår resultatene.

Daunfeldt *et al.* (2014) undersøker effekten på omsetning, sysselsetting og endringen i kjøpekraft i etterkant av en etablering. De finner en økning i omsetning i varehandelssektoren på omtrent 20% i etableringskommunen. Videre finner de at etableringen assosieres med en økning i sysselsettingen i varehandelssektoren på mellom 10% og 17%. For nærliggende kommuner finner de en negativ effekt på omsetning på -1.88%, men ingen effekt på sysselsettingen.

Daunfeldt *et al.* (2015) undersøker effektene på omsetning og sysselsetting hos nærliggende bedrifter i varehandelssektoren som følge av en IKEA-etablering. De finner at bedrifter i varehandelssektoren i kommunen der IKEA etableres har en gjennomsnittlig økning i omsetningen på 11%. For bedrifter som befinner seg mindre enn 1 km unna der IKEA etableres var effekten over dobbelt så stor, og førte til en økning i omsetningen på 26%. Videre indikerte resultatene at den positive effekten av IKEA-etableringen på omsetningen ble mindre jo lengre unna bedriftene befant seg fra IKEA. For bedrifter som lå 5 til 10 km unna IKEA, og for bedrifter som befinner seg i bysentrum, fant de ikke statistisk signifikante effekter på omsetning. De undersøkte også effekten på sysselsetting for nærliggende bedrifter, men fant ingen statistisk signifikante resultater.

Rudholm *et al.* (2018) undersøker effekten på produktivitet i fire ulike kommuner som fikk etablert en IKEA. De finner at produktiviteten i Haparanda øker med 80%. Dette skyldes trolig at Haparanda er en kommune med et lite antall innbyggere og har en lav omsetning i varehandelssektoren før etableringen (Rudholm, 2018). Den kraftige økningen i produktiviteten kan knyttes til en 722% økning i omsetningen 5 år etter IKEA-etableringen. I Kalmar finner de en økning i produktiviteten på 8%. For Göteborg kommune finner de ingen statistisk signifikant effekt som følge av etableringen.

Resultatene fra studiene som har undersøkt effekten av en IKEA-etablering på varehandelssektoren indikerer at etableringen er assosiert med en økning i omsetningen. For effektene på sysselsettingen i varehandelssektoren var resultatene mer tvetydige. De ga likevel en indikasjon på at etableringen kunne assosieres med en positiv effekt.

2.2 Teoretisk rammeverk

I dette delkapittelet presenteres et teoretisk rammeverk som beskriver mulige effekter på varehandelssektoren som følge av en IKEA-etablering. Det presenteres ikke en formell teoretisk modell, men vi diskuterer mer generelt mulige direkte og indirekte effekter. Vi tar utgangspunkt i det teoretiske rammeverket presentert i Rudholm *et al.* (2018). De benytter et tankeeksperiment hvor de ser for seg to identiske kommuner. Det etableres en IKEA i den ene kommunen, mens den andre kommunen forblir uendret. Hvilke effekter på varehandelssektoren kan vi forvente å observere i kommunen der IKEA etableres, men ikke i kommunen som forblir uendret?

2.2.1 Direkte effekter

Direkte effekt på omsetning

Ett enkelt IKEA-varehus hadde i 2017 en årlig gjennomsnittlig omsetning på over 900 millioner kroner (IKEA, 2017b). Justert for inflasjon tilsvarer dette over 350 millioner kroner i 1984.⁵ Vi har ikke funnet informasjon om omsetningen til IKEA-Norge på 1980-tallet, men det er ikke urimelig å anta at omsetningen til et IKEA-varehus i Norge kan ha vært i denne størrelsesordenen. Kommunen som får et IKEA-varehus kan dermed forvente en positiv direkte effekt på omsetningen som følge av etableringen. Dette støttes av resultatene fra tidligere forskning på effekten av IKEA-etablering.

Direkte effekt på sysselsetting

IKEA defineres som en big box-enhet, og det er derfor rimelig å anta at det er nødvendig med et betydelig antall ansatte for å betjene alle avdelingene i varehuset. I 1985 hadde ett IKEA-varehus i gjennomsnitt 166 ansatte (IKEA, 2019). Hvis vi antar at antall sysselsatte ved de to IKEA-varehusene i Norge også befant seg i dette sjiktet, eventuelt noe lavere, kan vi forvente en positiv direkte effekt på sysselsettingen i varehandelssektoren i kommunen som får en IKEA-etablering.

⁵Regnet om til 1984 kroner med priskalkulatoren til SSB.

2.2.2 Indirekte effekter

Det er flere indirekte effekter som kan påvirke omsetningen og sysselsettingen i varehandelssektoren ved en IKEA-etablering. Hvorvidt de indirekte effektene har en positiv eller negativ totaleffekt på omsetningen og sysselsettingen avhenger av flere faktorer, og det kan være vanskelig å konkludere med retningen på effektene. Under presenteres ulike indirekte effekter.

Økt produktivitet per sysselsatte

Rudholm *et al.* (2018) finner en økning i produktiviteten per sysselsatte i varehandelssektoren i etableringskommunen. Funnene deres indikerer at IKEA-varehus har en høyere produktivitet enn allerede etablerte aktører i varehandelssektoren. Den indirekte effekten av etableringen vil dermed avhenge av tilpasningsevnen til de andre aktørene i sektoren.

Bedrifter i direkte konkurranse med IKEA, og som har lavere produktivitet, vil kunne få svekket konkurranseevne dersom de ikke klarer å utjevne differansen i produktivitet per ansatt. Det er rimelig å anta at ikke alle bedrifter vil klare denne omstillingen, og som følge av dette kan bli nødt til å legge ned. Den indirekte effekten avhenger dermed av hvor stor del av omsetningen, og de sysselsatte fra bedriftene som legges ned, som absorberes av andre aktører i varehandelssektoren. Dersom alt absorberes kan vi forvente å ikke observere en effekt, da dette vil være en ren omfordeling innad i sektoren. Ved delvis eller ingen absorpsjon vil den indirekte effekten på omsetning og sysselsetting forventes å være negativ.

Stordriftsfordeler

IKEA drar nytte av stordriftsfordeler i produksjon og innkjøp.⁶ Dette fører til lavere marginalkostnader per enhet, og IKEA har dermed mulighet til å senke prisene sine til et lavere nivå enn det konkurrerende bedrifter potensielt kan. Bedrifter som ikke klarer å konkurrere på pris, vil kunne bli nødt til å legge ned. Retningen på den indirekte effekten vil avhenge av hvorvidt hele eller deler av omsetningen og sysselsettingen blir absorbert av andre aktører i sektoren.

Substitutter og komplementære goder

⁶Stordriftsfordeler omtaler fordelene ved å produsere store kvantum ved at kostnadene per produserte enhet er lavere enn ved små kvantum. Dette skjer siden gjennomsnittskostnadene på lang sikt reduseres for hver ekstra produserte enhet (Idsø, 2014).

Økonomisk teori deler goder og varer inn i flere ulike kategorier. Med tanke på indirekte effekter på omsetning og sysselsetting er det relevant å skille mellom substitutter og komplementære varer. To varer er substitutter for hverandre dersom varene dekker det samme behovet hos konsumenten. To varer er komplementære dersom de utfyller hverandre, eller hvis den ene varen øker nytten av den andre.

Retningen på de indirekte effektene på omsetningen og sysselsettingen avhenger av om varene IKEA selger er substitutter eller komplementære til varene som selges av andre aktører i varehandelssektoren.

IKEA vil kunne tilby varer til relativt lavere priser enn andre mindre aktører på grunn av stordriftsfordeler. Bedrifter som selger substitutter vil dermed kunne oppleve nedgang i omsetningen dersom de ikke klarer å tilpasse seg til det nye prisnivået og vil potensielt bli nødt til å legge ned. Dette kommer av at konsumentene trolig vil vri etterspørselen sin mot tilbyderer med den laveste prisen på en vare. Igjen vil absorpsjongraden av omsetning og sysselsetting av andre aktører i varehandelssektoren avgjøre retningen på den indirekte effekten.

Funnene fra gjennomgangen til Pedersen (2014) indikerer at big box-etableringer reduserer prisnivået på varekategoriene de tilbyr. Hvordan denne prisreduksjonen påvirker omsetningen avhenger av om prisfølsomheten til konsumentene er høy eller lav.⁷ Ved høy prisfølsomhet vil lavere priser hos big box-enheten kunne føre til økt salg. Dette kan også føre til vekst i omsetningen i komplementære varekategorier (Pedersen, 2014). Ved lav prisfølsomhet vil kvantum som selges kunne forbli uendret, men reduksjonen i prisene kan føre til lavere total omsetning i sektoren.

Teorien om substitutter og komplementære varer gir ingen klar indikasjon på retningen til effektene på omsetning og sysselsetting. Vi antar likevel at den positive effekten fra salg av komplementære varer trolig vil dominere en eventuell negativ effekt fra salg av substitutter. Funnene i artikkelen til Han (2018) støtter opp denne antagelsen. De finner at det i hovedsak er bedrifter som selger komplementære varer til IKEA som påvirkes av en etablering.

⁷Prisfølsomhet er et mål på hvor følsom etterspørselen (tilbudet) er overfor prisendringer. Dersom følsomheten er høy, vil etterspørselen (tilbudet) endres mer ved en prisendring enn dersom den er lav.

Bedriftsklynger

Litteraturen som omtaler ringvirkningene fra bedriftsklynger finner effekter som indikerer lavere kostnader, bedre tilgang på kvalifisert arbeidskraft og spillover i form av kunnskapsoverføring mellom aktører (Rudholm, 2018). I Sverige etablerer IKEA seg typisk i områder med 10 til 20 andre aktører innenfor varehandelssektoren (Rudholm, 2018). Vi mener det er rimelig å anta at etableringer i Norge er av samme art. Dette tyder på at etablering av en IKEA i en allerede eksisterende bedriftsklynge vil kunne føre til flere positive indirekte effekter på omsetningen og sysselsettingen i varehandelssektoren.

3 Økonometrisk rammeverk

I dette kapitlet presenteres det økonometriske rammeverket som ligger til grunn for de empiriske analysene som utføres i denne oppgaven. Delkapittel 3.2 presenterer differanse i differanse-metoden. Den syntetiske kontrollmetoden gjøres rede for i delkapittel 3.3.

3.1 Innledning

I denne oppgaven ønsker vi å undersøke om det er en kausal sammenheng mellom en IKEA-etablering og omsetningen og sysselsettingen i varehandelssektoren. Vi betrakter en etablering som en behandling og ønsker å undersøke effekten av denne, gitt at det er en effekt. Den ideelle forskningsmetoden for å finne denne effekten er å benytte et randomisert kontrollert studie (RCT). Denne metoden går ut på at man tilfeldig (randomisert) fordeler deltakere til enten en behandlingsgruppe eller en kontrollgruppe. Behandlingsgruppen blir eksponert for behandlingen, mens kontrollgruppen ikke blir det. Ved tilfeldig fordeling vil de ulike gruppene være like *før* behandlingen og eventuelle forskjeller *etter* behandlingen antas å kunne si noe om effekten av denne (Svartdal, 2018). Slike kontrollerte studier er vanskelig å gjennomføre grunnet praktiske utfordringer, etiske problemstillinger og høye kostnader.

I vårt tilfelle er det rimelig å anta at 'tildelingen' av en IKEA ikke er tilfeldig, men at det er basert på visse karakteristika ved behandlingsenheten. Dette kan gjøre det vanskelig å finne en tilfredsstillende kontrollgruppe for den behandlede enheten. Vi skal nå presentere de to metodene vi benytter for å estimere effekten av en IKEA-etablering.

3.2 Differanse i differanse-metode (DID)

Differanse i differanse-metoden, heretter kalt DID, er én av de mest brukte metodene for å undersøke effekter av big box-etableringer. Denne metoden kan, dersom alle forutsetningene holder, være en tilfredsstillende strategi tilsvarende *så godt som* randomisert fordelt.

Utgangspunktet for metoden er at datamaterialet består av to ulike grupper med observasjoner, en behandlingsgruppe og en kontrollgruppe, og at disse gruppene observeres på ulike tidspunkter. Estimering ved DID kontrollerer for systematiske forskjeller mellom

behandlings- og kontrollgruppen ved å bruke data fra to ulike tidsperioder. Data fra perioden før behandling og data fra perioden etter behandling. Navnet forklarer i korte trekk metoden ved at den tar for seg endringen i to differanser: differansen mellom behandlings- og kontrollgruppen, og differansen i deres respektive utvikling over tid. En egenskap ved metoden er at den første differansen eliminerer systematiske ulikheter mellom behandlings- og kontrollgruppen som antas å være konstante over tid. Den andre differansen eliminerer forskjellen i utviklingen over tid ved å anta at den er identisk for de to gruppene i fravær av behandling. En kausal effekt kan dermed identifiseres under hypotesen om at de to gruppene ville utviklet seg identisk dersom behandlingen ikke hadde funnet sted.

Vi presenterer først en klassisk to-periode differanse i differanse-modell med koblet data for enhet og år. Denne viser mekanismen bak metoden på en enkel måte. Deretter utvider vi modellen til å gjelde for paneldata.

3.2.1 Modell med koblet data for enhet og år

La oss starte med en klassisk modell med koblede data for enhet og år, som beskrevet i kapittel 13.2 i Wooldridge (2015). Her antar vi at vi har data for 2 perioder, der periode 1 er før behandlingen og periode 2 er etter behandlingen. Modellen som benyttes er en regresjonsmodell der vi introduserer dummy-variabler for å fange opp effekten av behandlingen. Lar vi y definere utfallsvariabelen, kan modellen formuleres som:

$$y = \beta_0 + \delta_0 D2 + \beta_1 Treat + \delta_1 Treat \times D2 + restledd \quad (3.1)$$

Her er $D2=1$ for observasjoner etter behandlingen og null ellers. $Treat=1$ for observasjoner som tilhører enhetene i behandlingsgruppen og null ellers. Her vil $\hat{\delta}_1$ være differanse i differanse-estimatoren som angir behandlingseffekten. Dersom det ikke inkluderes forklaringsvariabler vil $\hat{\delta}_1$ være gitt ved:

$$\hat{\delta}_1 = (\bar{y}_{behandling,2} - \bar{y}_{kontroll,2}) - (\bar{y}_{behandling,1} - \bar{y}_{kontroll,1})$$

Den første fotskriften angir gruppen, og den andre fotskriften angir perioden. Streken over y -ene indikerer at det er den gjennomsnittlige verdien til utfallsvariabelen. Ved bruk av koblede data for enhet og år omtales δ_1 ofte som *the average treatment effect* siden den måler effekten av behandlingen på gjennomsnittsutfallet til y (Wooldridge, 2015).

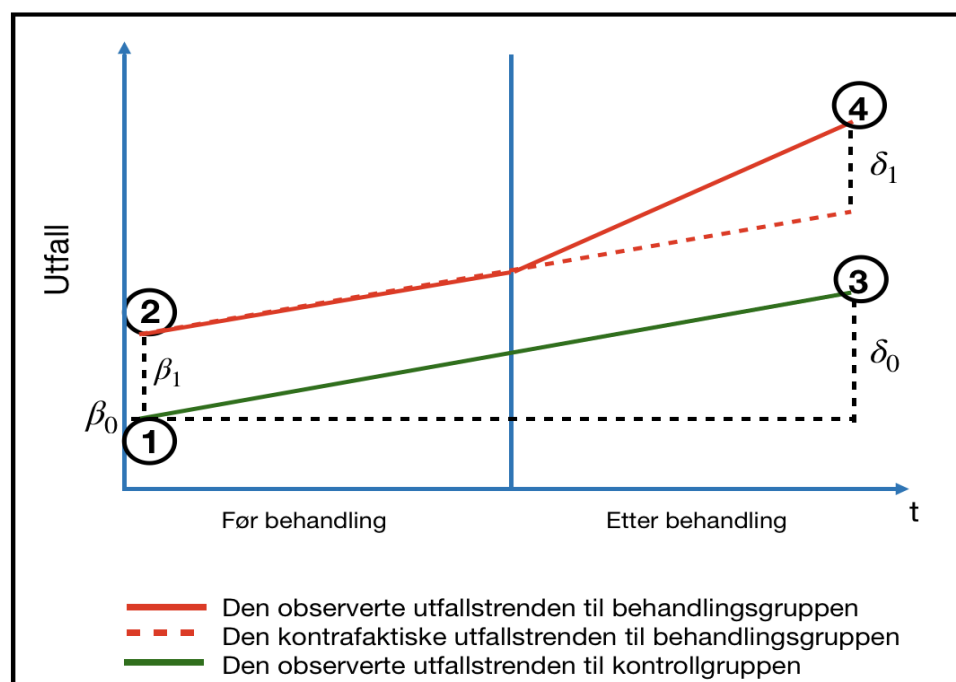
Tabell 1: Illustrasjon av DID-estimatoren (Wooldridge, 2015)

	Før	Etter	Etter - Før
Kontroll	①: β_0	③: $\beta_0 + \delta_0$	δ_0
Behandling	②: $\beta_0 + \beta_1$	④: $\beta_0 + \delta_0 + \beta_1 + \delta_1$	$\delta_0 + \delta_1$
Behandling - Kontroll	β_1	$\beta_1 + \delta_1$	δ_1

Merk: Se figur 1 for grafisk fremstilling av punkt ①-④.

Tabell 1 viser hvordan DID-estimatoren utledes. Vi ser at δ_1 tilsvarer endringen i differansen mellom behandlings- og kontrollgruppen før og etter behandling.

DID-metoden baserer seg på en forutsetning om parallelle trender. Dette innebærer at trenden i utfallsvariabelen i perioden før behandling må være lik for behandlings- og kontrollgruppen. Nivået på utfallsvariabelen må ikke nødvendigvis være likt, så lenge utviklingen over tid er den samme.⁸ Dette illustreres i figur 1.

Figur 1: Grafisk fremstilling av differanse i differanse-estimatoren fra tabell 1.

Vi ser at den observerte utfallstrenden til behandlings- og kontrollgruppen er parallelle i perioden før behandling. Ved fravær av behandling antar vi at utfallstrenden til behandlingsgruppen ville fulgt den samme utvikling over tid som kontrollgruppen. Dette omtales som det kontrafaktiske utfallet, og er representert ved den stiplede linjen. Etter behandling ser vi at den observerte utfallstrenden til behandlingsgruppen avviker fra den kontrafaktiske utfallstrenden. Det er differansen mellom den stiplede linjen og den

⁸Se kapittel 5.1.1 for mer detaljert om forutsetninger ved DID-metode.

observerte utfallstrenden til behandlingsgruppen som er DID-estimatoren, og som tolkes som effekten av behandling. Dersom Gauss-Markov-forutsetningene holder vil estimering av modellen ved minste kvadraters metode (OLS) gi den beste lineære forventningsrette estimatoren (Wooldridge, 2015).

3.2.2 Modell med paneldata

Dersom et datasett inneholder informasjon om flere enheter over en tidsperiode kalles dette et paneldatasett. Dette innebærer at vi har variasjon over to dimensjoner, enhet og tid, med henholdsvis fotskrift i og t . Ved estimering av DID-estimatoren med et slikt datasett vil modellen formuleres som:

$$\begin{aligned} y_{it} &= \beta_0 + \delta_0 D_t + \beta_1 Treat_i + \delta_1 Treat_i \times DTreat_t + u_{it} \\ u_{it} &= \eta_i + \epsilon_{it} \end{aligned} \tag{3.2}$$

Der $DTreat_i = 1$ for alle år etter behandlingen og null for alle år før behandlingen. D_t er årsdummier. Årsdummier blir inkludert i modellen for å kontrollere for makroøkonomiske fenomener og sjokk som varierer over tid, men som er like for alle kommuner.

Vi ser av ligning 3.2 at restleddet er dekomponert i to deler, en enhetsspesifikk del, η_i , og en idiosynkratisk del, ϵ_{it} . Den enhetsspesifikke restleddskomponenten inneholder uobserverte enhetsfaste effekter som ikke varierer over tid. Den idiosynkratiske delen av restleddet representerer uobserverte faktorer som kan variere over tid⁹(Wooldridge, 2015).

Ved å anvende OLS direkte på modellen, ofte omtalt som *pooled OLS*, utnytter vi all variasjon i datasettet, både within og between.¹⁰ Dersom forutsetningene for OLS er oppfylt, og alle restleddsegenskapene gjelder, vil estimering ved pooled OLS gi forventningsrette og konsistente estimatorene. Pooled OLS vil dermed foretrekkes fremfor andre metoder estimeringsmetoder som kun utnytter deler av variasjonen i datamaterialet.¹¹

Det er rimelig å anta at det kan være flere enhetspesifikke faktorer som er uobserverbare, og dermed plukkes opp i restleddet. Hvis én eller flere av disse er korrelert med forklaringsvariablene, vil eksogenitetsforutsetningen knyttet til den

⁹Se avsnitt A i appendiks for liste over restleddsegenskaper som gjelder ved OLS-estimering med paneldata.

¹⁰Whitin-variasjon beskriver variasjon innenfor en enhet over tid, mens between-variasjon beskriver variasjonen mellom enheter (Wooldridge, 2015).

¹¹Andre metoder som kan benyttes er *Fixed Effects* og *Random Effects*.

enhetsspesifikke restleddskomponenten være brutt. Dette medfører en utelatt variabelskjevhet.¹² Pooled OLS anvendt direkte på modellen vil dermed gi estimatorer som hverken er forventningsrette eller konsistente.

En løsning på dette problemet er å utvide modellen med faste enhetseffekter. Dette gjør at vi får individspesifikke konstantledd som fanger opp uobserverbar enhetsspesifikk heterogenitet. De faste enhetseffektene vil fange opp effekten av $Treat_i$ i ligning (3.2) og vi får dermed ikke identifisert β_1 . Ved å inkludere faste enhetseffekter vil en variant av ligning (3.2) se slik ut:

$$\begin{aligned} y_{it} &= \alpha_i + \delta_0 D_t + \delta_1 Treat_i \times DTreat_t + \epsilon_{it} \\ \alpha_i &= \beta_0 + \beta_1 Treat_i + \eta_i \end{aligned} \tag{3.3}$$

Ligning (3.3) viser at den enhetsspesifikke restleddskomponenten inngår i konstantleddet. Vi har valgt å benytte oss av en *fixed effects*-transformasjon, også omtalt som *within*-transformasjon, for å fjerne de faste enhetseffektene når vi estimerer ved DID.¹³ Transformasjonen gjør at variasjonen i datamaterialet mellom, *between*, ulike enheter fjernes. Vi står dermed igjen med informasjon om endring innad i de ulike enhetene over tid, *within*. En naturlig utvidelse av ligning 3.3 er å inkludere kontrollvariabler, x_{it} . Ved bruk av kontrollvariabler kan vi kontrollere for at det kan være systematiske forskjeller mellom behandlings- og kontrollgruppen over tid. Modellen vil da se slik ut:

$$y_{it} = \alpha_i + \delta_0 D_t + \delta_1 Treat_i \times DTreat_t + x_{it} \times \beta + \epsilon_{it} \tag{3.4}$$

I kapittel 6.1 presenterer vi resultatene fra DID-metoden med modellspesifikasjonene gitt ved ligning (3.3) og (3.4).

3.3 Den syntetiske kontrollmetoden (SCM)

En annen metode som kan benyttes for å undersøke den kausale effekten av en IKEA-etablering er den syntetiske kontrollmetoden. Metoden ble først tatt i bruk for å undersøke den økonomiske effekten av terrorkonflikten i Baskerland på utviklingen i

¹²Se kapittel 5.1 for mer om utfordringer knyttet til utelatte variabler.

¹³Andre metoder som kunne vært benyttet er inkludering av (n-1) enhetsspesifikke dummyvariabler eller en *first difference*-transformasjon. Metodene implementeres på ulike måter, men gir identiske verdier på koeffisientene ved estimering.

BNP fra 1970 til 1990 (Abadie & Gardeazabal, 2003). Siden Baskerland var det eneste området med terrorkonflikt, var det vanskelig å finne enheter som kunne benyttes i en kontrollgruppe. Det var også store økonomiske forskjeller mellom regionene i Spania i denne tidsperioden, noe som problematiserte valget av kontrollenheter ytterligere. Metoden ble deretter videreutviklet gjennom en studie på effekten av tobakksregulering i California (Abadie, Diamond & Hainmueller, 2010). Den siste artikkelen i serien om den syntetiske kontrollmetoden er fra Abadie, Diamond & Hainmueller (2015) som undersøker hvordan sammenslåingen av Øst- og Vest-Tyskland påvirket utviklingen i BNP i Vest-Tyskland etter sammenslåingen.

Prinsippet med metoden er å lage en syntetisk versjon av den behandlede enheten som skal benyttes som kontrollenhet. Den syntetiske kontrollenheten skal i teorien imitere den kontrafaktiske utviklingen i utfallsvariabelen. I likhet med DID-metoden benytter vi oss av en kontrollgruppe bestående av enheter som ikke har mottatt behandling. SCM skiller seg derimot fra DID-metoden ved at den ikke vekter de ulike kontrollenhetene likt, men tildeler en vekt basert på likheter mellom kontrollenhetene og den behandlede enheten i perioden før behandling.

3.3.1 Modellen

Vi benytter modellen formulert i Carling & Li (2017) for å presentere det økonometriske rammeverket for metoden.

Vi lar den observerte utfallsvariabelen av interesse betegnes som y_{jt} for enhet j på tidspunkt t . 0 betegner den behandlede enheten og $j = 1, \dots, J$ refererer til de J tilgjengelige kontrollenhetene i utvalget. Videre antar vi at alle enheter i utvalget observeres over hele tidsperioden $t = 1, \dots, T_0, \dots, T$, der T_0 er tidspunktet for behandlingen. η_{0t} betegner det uobserverbare kontrafaktiske utfallet til utfallsvariabelen av interesse for den behandlede enheten. $\hat{\eta}_{0t}$ er estimatoren for det kontrafaktiske utfallet og er utfallsvariabelen til den syntetiske kontrollenheten. Den er konstruert av et vektet snitt av kontrollenheter som dermed simulerer utviklingen til den behandlede enheten i fravær av behandling:

$$(1) \quad \hat{\eta}_{0t} = \sum_{j=1}^J w_j y_{jt}$$

Ligning (1) er pålagt restriksjonene $\sum_{j=1}^J w_j = 1$ og $0 \leq w_j \leq 1$. Dette betyr at vekten som tildeles de ulike kontrollenhetene skal være positiv, mindre enn 1, og de skal totalt summere

seg til 1. $w = (w_1, \dots, w_J)'$ er parametervektoren som indikerer de J kontrollenhetenes vekter.

Vi antar at behandlingen ikke påvirker utfallsvariabelen i tiden før den inntreffer, slik at $y_{0t} - \eta_{jt} = 0$ i tidsperioden $t < T_0$. I perioden etter behandlingen $t \geq T_0$ vil parameteret til behandlingseffekten være gitt ved:

$$(2) \quad \alpha_t = y_{0t} - \eta_{0t}.$$

Siden η_{0t} ikke kan observeres kan vi heller ikke estimere α_t . Ved å benytte den observerbare estimatoren $\hat{\eta}_{0t}$ istedenfor, kan vi derimot estimere behandlingseffekten, $\hat{\alpha}$, på tidspunkt t :

$$(3) \quad \hat{\alpha}_t = y_{0t} - \hat{\eta}_{0t} = y_{0t} - \sum_{j=1}^J w_j y_{jt}$$

Vekten som tildeles de ulike kontrollkommunene, w_j , bestemmes ved å optimere en objektfunksjon som minimerer avviket mellom den observerte behandlede enheten og enhetene i kontrollgruppen, i tiden før behandling. Dette gjøres med hensyn på lag i utfallsvariabelen og kontrollvariabler. Med lag i utfallsvariabelen menes det at verdien på utfallsvariabelen i de ulike årene før behandling kan benyttes som en kontrollvariabel i modellspesifikasjonen. Videre i oppgaven vil vi omtale lag i utfallsvariabelen kun ved bruk av ordet lag.

Utfallsvariabelen bestemmes av kontrollvariabler i en lineær faktormodell. Den lineære faktormodellen for kontrollenheter og den kontrafaktiske til den behandlede enheten er gitt ved henholdsvis ligning (4) og (5):

$$(4) \quad y_{jt} = \delta_t + \theta_t z_j + \lambda_t \mu_j + \epsilon_{jt}$$

$$(5) \quad \eta_{0t} = \delta_t + \theta_t z_0 + \lambda_t \mu_0 + \epsilon_{0t}$$

der δ_t er en ukjent fellesfaktor med konstant faktorbelastning på tvers av enheter. z_j er en $(r \times 1)$ vektor med r observerte kontrollvariabler som ikke påvirkes av behandlingen, og tilhørende parametervektor θ_t . μ_j er en $(F \times 1)$ vektor av uobserverte kontrollvariabler med tilhørende parametervektor λ_t . ϵ_{jt} og ϵ_{0t} er restledd med fellesvarians σ^2 . Se appendiks avsnitt B for en detaljert gjennomgang av hvordan objektfunksjonene minimeres.

Når objektfunksjonene er minimert, og vektene er tildelt de ulike kontrollenhetene, er det kontrafaktiske utfallet estimert. Vi kan dermed undersøke effekten av behandling ved å se på differansen mellom den behandlede enheten og den syntetiske kontrollenheten i perioden etter behandling.

3.3.2 Implementering av SCM

Før man utfører den syntetiske kontrollmetoden må man velge en kombinasjon av kontrollvariabler og lag som utgjør modellspesifikasjonen. Målet er å finne en syntetisk kontrollenhet som følger utviklingen i den behandlede enheten i perioden *før* etableringen så godt som mulig. Det finnes få retningslinjer for hvilke kontrollvariabler som bør benyttes og hvor mange lag som bør inkluderes i modellspesifikasjon.

Det er imidlertid viktig med balanse i vektingen mellom kontrollvariabler og antall lag ved konstruksjonen av den syntetiske kontrollenheten (Botosaru & Ferman, 2019). Med et begrenset antall år før behandling er det spesielt viktig å være oppmerksom på vektingen som blir gitt. Ved bruk av mange lag, og med få år tilgjengelig før behandlingen, vil vektingen av disse kunne bli dominerende. Vi vil dermed miste mye av innflytelsen fra kontrollvariablene og oppnår ikke en god balanse i vektingen.

Artikkelen til Ferman, Pinto & Possebom (2016) tar opp problemet rundt valg av modellspesifikasjon og hvordan mangelen på retningslinjer kan føre til spesifikasjonsleting.¹⁴ Metoden er konstruert på en slik måte at det er mulig å lete seg frem til den kombinasjonen som gir ønsket resultat. Ferman *et al.* (2016) presenterer flere mulige løsninger på dette problemet. Én av disse er å benytte en metode som baserer valget av modellspesifikasjon på et objektivt kriterium. Dette gjør at forskeren ikke kan ta subjektive valg i like stor grad, og på den måten forhindres spesifikasjonsleting. Vi benytter oss av et slikt kriterium når vi skal bestemme modellspesifikasjonen vi bruker i denne oppgaven.

Proseduren vi benytter oss av baserer i stor grad valget av modellspesifikasjon på pre-Root Mean Square Prediction Error (pre-RMSPE). Pre-RMSPE er et positivt tall som beskriver størrelsen på avviket i tiden før behandlingen mellom den syntetiske kontrollenheten og den behandlede enheten. En lav verdi på pre-RMSPE betyr at avviket er lite, mens en høy verdi på pre-RMSPE betyr at avviket er stort. Verdien på pre-RMSPE til en modellspesifikasjon må sammenlignes med verdien til andre modellspesifikasjoner for å kunne benyttes som et mål på 'styrken' til den syntetiske kontrollenheten. Ved å teste ulike kombinasjoner av kontrollvariabler og lag finner man den modellspesifikasjonen

¹⁴Med spesifikasjonsleting menes det at man aktivt leter etter, og velger, kombinasjonen av kontrollvariabler og lag som gir det ønskede resultatet etter behandling.

som har den laveste verdien på pre-RMSPE. Formelen for pre-RMSPE er gitt ved:

$$pre - RMSPE = \sqrt{\frac{1}{T_0} \sum_{t=1}^{T_0} \left(y_{0t} - \sum_{j=1}^J w_j^* y_{jt} \right)^2}$$

der w_j^* er vekten til kontrollenhetene som minimerer avviket mellom den behandlede enheten og kontrollenhetene i tiden før behandling.

Prosessen vi følger for å velge modellspesifikasjon består av to steg. Det første steget er å finne de kombinasjonene av kontrollvariabler og lag som gir de laveste verdiene på pre-RMSPE. Vi starter med en så enkel modell som mulig, der kun utfallsvariabelen og to kontrollvariabler benyttes.¹⁵ Deretter legger vi til kontrollvariabler, én etter én, og ser på hvordan dette endrer pre-RMSPE. Ved høyere verdi på pre-RMSPE etter at en variabel er lagt til fjernes variabelen fra modellspesifikasjonen. Ved lavere verdi beholder vi variabelen. Vi gjennomfører denne prosessen helt til vi har testet alle kontrollvariablene. Modellspesifikasjon består nå av X antall kontrollvariabler. Deretter legger vi til lag i den valgte modellspesifikasjonen. Dette gjør vi ved å først legge til lag enkeltvis for hvert år i tiden før behandling. Videre tester vi ut ulike kombinasjoner av lag, og velger de modellspesifikasjonene som gir den laveste verdien på pre-RMSPE.

Når vi har kommet frem til et utvalg modellspesifikasjoner med lav pre-RMSPE, går vi videre til neste steg. Det andre steget går ut på å finne den modellspesifikasjonen som har den beste balansen i vektingen mellom kontrollvariabler og lag.¹⁶ Vi ønsker ikke å benytte oss av en modellspesifikasjon som har svært lav pre-RMSPE-verdi hvis det viser seg at vektingen mellom kontrollvariabler og lag er dårlig balansert. Til slutt undersøker vi derfor vektingen til kontrollvariablene og lag i de ulike modellspesifikasjonen, og velger den med best balanse i vektingen blant de med lavest pre-RMSPE-verdi når vi utfører analysen.

¹⁵Metoden krever minst to kontrollvariabler, inkludert egne lag i forklaringsvariabelen.

¹⁶Med 'den beste balansen', i vektingen mellom kontrollvariabler og lag, menes det at ingen av de to får tildelt vektorer som gjør at de er dominerende i konstruksjonen av den syntetiske kontrollenheten

4 Datamaterialet

I dette kapittelet presenteres datamaterialet vi benytter i de empiriske analysene. Data er hentet fra kommunedatabasen til Norsk Senter for Forskningsdata (NSD) som tverrsnittsdata, og er deretter omgjort til et sterkt balansert paneldatasett for tidsperioden 1977 til 1992. Vi har benyttet kommunestrukturen fra 1992 som bestod av 439 kommuner.¹⁷ Grunnet manglende data for noen kommuner står vi igjen med 395 kommuner.¹⁸ I løpet av tidsperioden 1977 til 1992 ble det etablert to IKEA-varehus i Norge, ett i Bergen i 1984 og ett i Sandnes i 1988. På grunn av den geografiske nærheten mellom Sandnes og Stavanger har vi besluttet å slå sammen de to kommunene.¹⁹

4.1 Behandlingsenheter

Vi har i utgangspunktet to potensielle behandlingsenheter som kan benyttes for å undersøke den kausale effekten av en IKEA-etablering, Bergen og Sandnes/Stavanger. I forkant av en analyse av etableringseffekten i de to kommunene, er det naturlig å undersøke om det er potensielle eksterne sjokk eller politiske avgjørelser som kan påvirke resultatene. Det er spesielt to hendelser som kan ha påvirket omsetningen og sysselsettingen i varehandelssektoren i tidsperioden for vårt datamateriale.

Åpningstidsloven, 1985

I april 1985 ble en ny åpningstidslov godkjent i Stortinget, videre omtalt som Åpningstidsloven. Loven innskrenket mulighetene kommunene hadde til å regulere åpningstider i varehandelssektoren selv. Kommunene kunne fortsatt begrense åpningstidene, men de kunne ikke pålegge butikker å stenge tidligere enn kl. 20.00 på hverdager, og kl. 18.00 på lørdager (Bensnes & Strøm, 2018).

Vi har hentet ut data fra NSD om åpningstidsbestemmelser for Bergen og Sandnes/Stavanger, i tillegg til kommunene som inngår i kontrollgruppen som defineres i delkapittel 4.2. Datamaterialet er fra 1982 og inneholder informasjon om kommunenes

¹⁷For kommuner som ble sammenslått i løpet av tidsperioden 1977 til 1992, vil den *nye* kommunen gjelde for hele tidsperioden, uavhengig av når sammenslåingen fant sted.

¹⁸Se liste over kommuner med manglende data i appendiks avsnitt C.

¹⁹På grunn av tett geografisk plassering vil befolkningen kunne utnytte aktører i begge kommuner når de skal handle varige forbruksgoder. Det er derfor trolig at begge kommunene vil bli påvirket av en IKEA-etablering.

egne åpningstidsreguleringer i varehandelssektoren.

De kommunebestemte åpningstidsreguleringene på hverdager og lørdager i hver kommune i 1982, er presentert i henholdsvis tabell 13 og 14 i avsnitt D i appendiks.

Tabell 13 og 14 viser at bedriftene i varehandelssektoren i Bergen måtte stenge senest kl. 17.00 på hverdager og kl. 15.00 på lørdager. Den nye loven førte dermed til en potensiell 18 timers forlengelse av åpningstidene hver uke for den enkelte bedrift i varehandelssektoren i Bergen.

Videre ser vi av tabellene at bedrifter i Sandnes kunne holde åpent til senest kl. 23.00 på hverdager, og at vi mangler informasjon om lørdager. For Stavanger ser vi av tabell 13 at bedriftene kunne holde åpent til kl. 22.00 på hverdager, mens tabell 14 viser at de kunne holde åpent til kl. 20.00 på lørdager.

Dette kan indikere at implementeringen av Åpningstidsloven ville påvirket varehandelssektoren i Bergen i større grad enn i Sandnes/Stavanger. IKEA-etableringen i Bergen i 1984 sammenfaller nærmest med at Åpningstidsloven trådte i kraft i 1985. Dette kan gjøre det utfordrende å skille ut effekten av IKEA-etableringen. Vi ser av tabell 13 og 14 at majoriteten av de øvrige kommunene har lignende reguleringer for stengetider som Bergen i 1982. Ved å inkludere årsummier ved bruk av DID-metoden vil dette kunne kontrollere for effekten av Åpningstidsloven.

Oljekrisen fra 1986 til 1990

I 1986 kom det et negativt oljeprissjokk som førte til en halvering av normprisen på olje. Denne halveringen varte frem til 1990 og førte til at inntektene fra oljesektoren ble kraftig redusert. Oljeprisen sank fra over 27 dollar fatet til under 15 dollar per fat i 1986.²⁰ Rapporten til Andersen (2000) gir et innblikk i effekten av tilbudssjokket i oljesektoren for denne perioden:

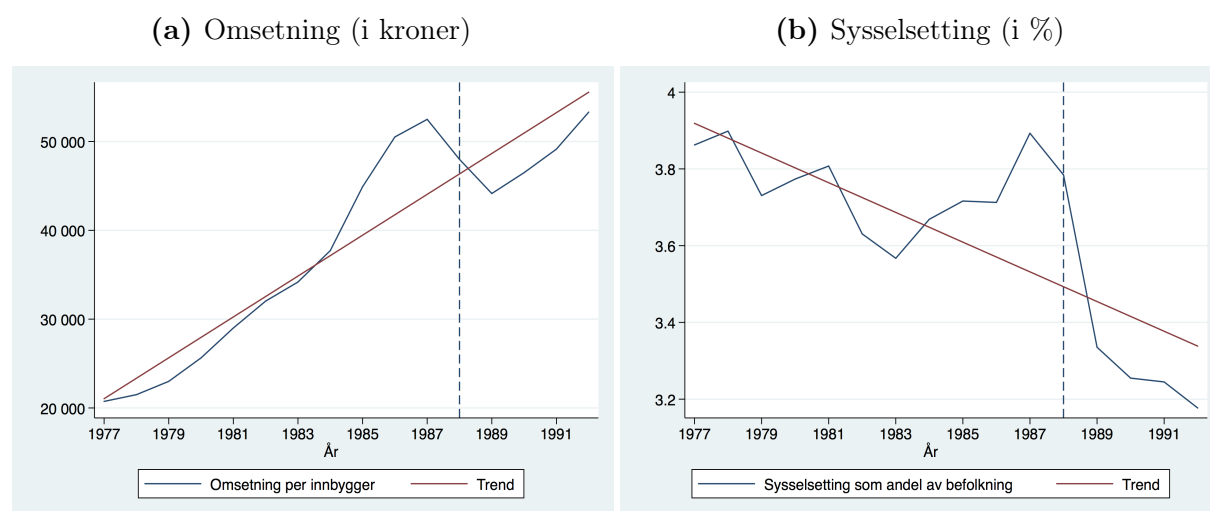
Oljeprisfallet i 1985/86 senket lønnsomheten særlig for høykostprodusenter av petroleum, som Norge, Storbritannia og USA. For Norge førte prisfallet til redusert årlig produksjonsverdi med ca. 40 milliarder kroner (korrigert for etterslepet i prisene i gasskontraktene). Petroleumssektorens betydning i økonomien falt dramatisk. Et overskudd i handelsbalansen på 41 milliarder kroner i 1985 ble på ett år snudd til et underskudd på 18 milliarder i 1986. Sektorens andel av BNP sank til rundt 10 prosent. [...] Dette gjorde

²⁰Tall hentet fra Statistisk sentralbyrå. <https://www.ssb.no/a/histstat/aarbok/ht-0801-324.html>

at Brundtland-regjeringen, som overtok etter Willoch i 1986, måtte foreta store innstramminger i offentlige budsjetter. De offentlige budsjetter ble holdt lave og lønnsøkningen var moderat. Innstrammingerne bidro til at arbeidsledighetstallene etter hvert ble høye i norsk målestokk (Andersen, 2000).

Prissjokket i oljesektoren fikk dermed store innvirkninger på handelsbalansen og sysselsettingen. Figur 2 viser hvordan utviklingen i varehandelssektoren til Sandnes/Stavanger ble påvirket av sjokket.

Figur 2: Utvikling over tid i omsetningen og sysselsettingen i varehandelssektoren i Sandnes/Stavanger.



Figur 2a viser utviklingen i omsetningen. Den stiplede linjen viser tidspunktet for etableringen av IKEA i Sandnes. På grunn av det negative oljeprissjokket vil det være vanskelig å skille ut effekten på omsetning av IKEA-etableringen i Sandnes/Stavanger.

Figur 2b viser utviklingen i sysselsettingen. Vi ser at det er en positiv utvikling fra 1983 som varer frem til 1987. Da snur utviklingen brått, og den blir kraftig negativ. Dette sammenfaller med tidspunktet for oljeprissjokket. Utviklingen i sysselsettingen er dermed kraftig negativ når IKEA etableres i 1988, og i årene etter etableringen. Dette indikerer at den negative effekten av oljeprissjokket på varehandelssektoren i Sandnes/Stavanger dominerer eventuelle andre effekter, herunder effekten av IKEA.

Inkludering av årsummier ved bruk av DID-metoden vil kunne kontrollere for effekten av oljeprissjokket, som antas å være lik for hele landet. Den kraftige negative effekten på Sandnes/Stavanger, som vi derimot antar er ulik fra resten av landet, vil

dermed trolig ikke, eller bare delvis, kontrolleres for med årsummier.

På grunn av vanskeligheten med å skille ut effekten av IKEA-etableringen i Sandnes/Stavanger i 1988, velger vi å ikke benytte denne som behandlingsenhet. I tillegg velger vi å fjerne Sandnes/Stavanger fra datamaterialet siden den heller ikke vil være en egnet kontrollkommune for Bergen. Vi utfører dermed alle analyser med Bergen som eneste behandlingsenhet.

4.2 Definerings av utvalget

Vi bruker to ulike empiriske metoder i denne oppgaven som begge benytter seg av en kontrollgruppe. Ved bruk av kontrollgrupper i komparative analyser er det viktig at kontrollgruppen som benyttes kan legge grunnlaget for et troverdig kontrafaktisk utfall for den behandlede enheten. Det er dermed trolig ikke ideelt å benytte alle de 392 tilgjengelige kommunene i datamaterialet som kontrollgruppe.

Det kan argumenteres for at kommuner av veldig liten størrelse, i forhold til behandlingskommunen, ikke vil være gode kontrollenheter siden det er lite trolig at de vil ha en sammenlignbar varehandelssektor. Vi velger derfor å benytte kommuner som er så lik den behandlede enheten som mulig i kontrollgruppen. Man kan forvente at like kommuner vil ha relativt lik økonomisk utvikling over tid ved fravær av behandling (Rudholm, 2018). Carling (2018) oppgir 30 kontrollenheter som et veiledende minimumsantall når man skal definere kontrollgruppen ved bruk av den syntetiske kontrollmetoden. Vi valgte derfor de 30 kommunene med høyest gjennomsnittlig befolkning, omsetning og sysselsetting i varehandelssektoren, i perioden før etableringen, til å utgjøre kontrollgruppen for Bergen.²¹ Liste over kommunene som inngår i kontrollgruppen, heretter omtalt som kontrollgruppe 1, finnes i avsnitt E i appendiks. Vi benytter også kontrollgruppe 1 ved bruk av DID-metoden.

4.2.1 Alternativ kontrollgruppe

Etter 1992, som er det siste året vi har informasjon om, har det blitt etablert IKEA i fire andre kommuner i Norge. Av de fire kommunene det gjelder har vi kun informasjon

²¹Det ble etablert en IKEA i Asker i 1975, og Asker kan dermed ikke benyttes som kontrollkommune for Bergen.

om to, Kristiansand og Trondheim.²² Vi lager en alternativ kontrollgruppe bestående av Kristiansand og Trondheim. Motivasjonen bak dette er at de 30 kommunene i kontrollgruppe 1 ikke nødvendigvis representerer en tilfredsstillende kontrollgruppe for Bergen. Det er rimelig å anta at IKEA ser etter spesifikke karakteristika ved en kommune når de bestemmer hvor de skal etablere seg. Dette gjør at det trolig finnes likheter mellom kommunene der IKEA etableres og slike *etableringskommuner* vil dermed egne seg godt i en kontrollgruppe. Videre i oppgaven vil den alternative kontrollgruppen omtales som kontrollgruppe 2.

4.3 Variabler

4.3.1 Utfallsvariabler

Vi benytter oss av to utfallsvariabler når vi skal undersøke effekten av IKEA-etableringen på varehandelssektoren i Bergen.²³

Omsetning i varehandelssektoren per innbygger

Den ene utfallsvariabelen vi benytter er omsetning i varehandelssektoren per innbygger. Omsetning, når den hentes ut fra NSD, er oppgitt i 100 000 kroner og omfatter kontantsalg og salg på kreditt. Formelt er omsetning i varehandelssektoren per innbygger formulert som:

$$oms_and_be_{it} = \frac{omsetning_{it}}{befolkning_{it}} \times 100000 \quad (4.1)$$

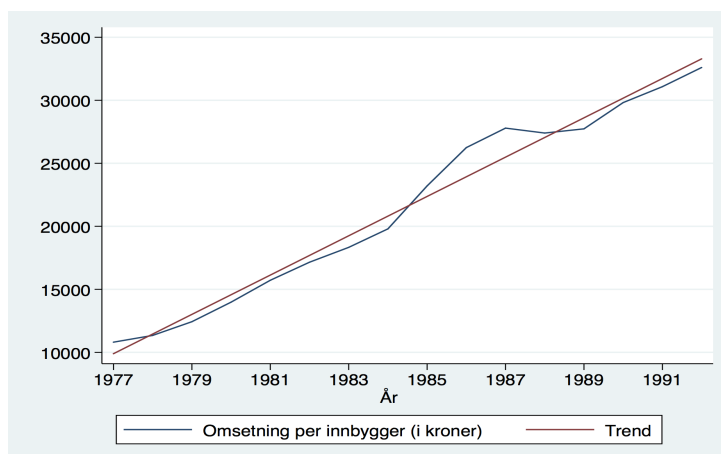
der $omsetning_{it}$ er omsetningen i varehandelssektoren, gitt i 100 000 kroner, for kommune i på tidspunkt t . $befolkning_{it}$ er antall innbyggere i kommune i på tidspunkt t . Vi multipliserer med 100 000 for å få omsetning i varehandelssektoren per innbygger oppgitt i hele kroner og ikke i desimaler.

I figur 3 presenteres den nasjonale utviklingen i omsetning per innbygger i varehandelssektoren for hele analyseperioden.

²²Det har blitt etablert en IKEA i Ringsaker og Oslo også, men disse er blant kommunene med mangelfull data og inngår derfor ikke i datasettet.

²³Varehandel, *detaljhandel*, omfatter virksomheter som driver salg av nye og brukte varer i eget navn og for egen regning, vesentlig til personlig bruk eller til private husholdninger, fra fast utsalgssted eller fra torgplass, ved omførsel eller ved postordre. Her grupperes også virksomheter som driver auksjonshandel og kommisjonssalg ellers til privatpersoner og husholdninger, dvs. salg for annens regning, men i eget navn. Definisjonen er hentet fra variabelforklaring til NSD.

Figur 3: Nasjonal utvikling over tid, med trendlinje, i omsetning i varehandelssektoren per innbygger fra 1977 til 1992.



Merk: Figuren er basert på datamaterialet vårt med 395 kommuner.

Vi ser av figuren at omsetningen har en jevn vekst som nærmest sammenfaller med trendlinjen hele tidsperioden. Unntaket er fra og med midten av 1980-tallet. Da ser vi en økning i omsetning per innbygger på landsbasis. Denne økningen sammenfaller med tidspunktet for IKEA-etableringen i Bergen. Vi kontrollerer for denne økningen på landsbasis ved å inkludere årsummier i modellene vi estimerer ved bruk av DID-metoden.

Sysselsetting i varehandelssektoren som andel av befolkning

Den andre utfallsvariabelen er sysselsetting i varehandelssektoren som andel av befolkning. Sysselsetting er målt i antall årsverk og rundet opp til nærmeste heltall.²⁴ Formelt er sysselsetting i varehandelssektoren som andel av befolkning formulert som:

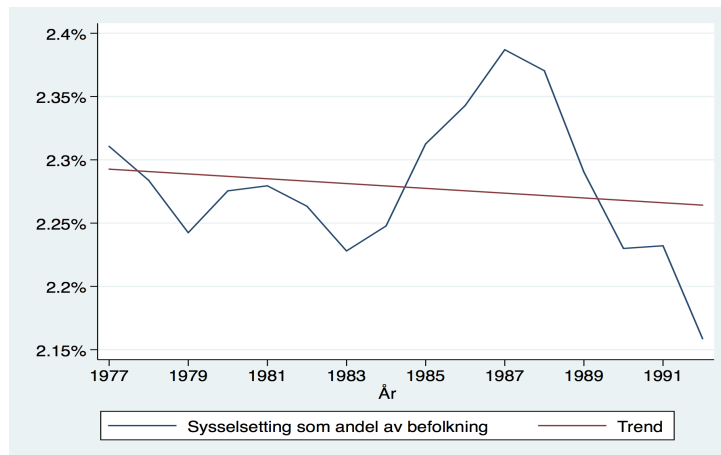
$$sys_and_be_{it} = \frac{sysselsetting_{it}}{befolkning_{it}} \times 100 \quad (4.2)$$

der $sysselsetting_{it}$ er antall hele årsverk i varehandelssektoren i kommune i på tidspunkt t . $befolkning_{it}$ er antall innbyggere i kommune i på tidspunkt t . Vi multipliserer med 100 for å få sysselsetting i varehandelssektoren som andel av befolkning oppgitt i prosent.

Vi ser av figur 4 at sysselsettingen som andel av befolkning på nasjonalt nivå er mer volatil enn utviklingen i omsetning. Den gjennomsnittlige utviklingen er svakt negativ og reduseres med omlag 0.04 prosentpoeng totalt. Vi observerer at det er en stigning i

²⁴Sysselsetting blir målt i antall årsverk. Tallene omfatter antall årsverk utført av funksjonærer, arbeidere, eiere og familiemedlemmer som arbeider i bedriftene, og tallene blir avrundet eller forhøyet til nærmeste hele tall. Varehandelen sysselsetter mange deltidsansatte. Tallet på utførte årsverk vil derfor være lavere enn tallet på sysselsatte. Definisjonen er hentet fra NSD sin kommunedatabase.

Figur 4: Nasjonal utvikling over tid, med trendlinje, i sysselsetting i varehandelssektoren som andel av befolkning fra 1977 til 1992.



Merk: Figuren er laget basert på datamaterialet med 395 kommuner.

utviklingen fra 1983 som fortsetter frem til 1986, hvor den når sitt toppunkt. Deretter synker det. Vi kontrollerer for nasjonale endringer i sysselsettingen som andel av befolkning ved å inkludere årsummier ved bruk av DID-metoden.

4.3.2 Forklaringsvariabel

Vi har ingen spesifikk informasjon om IKEA i datamaterialet vårt, og vi må dermed lage en variabel som antas å kunne plukke opp eventuelle effekter av IKEA-etableringen i Bergen.

Det er rimelig å anta at en eventuell effekt av IKEA kun vil påvirke varehandelssektoren i perioden etter etableringen. Vi definerer dermed en årsummyvariabel som er lik 1 for alle år etter behandlingen, inkludert behandlingsåret. Videre definerer vi en kommunespesifikk variabel som er lik 1 for kommunen som får behandling, i dette tilfellet Bergen. Ved å lage en interaksjonsdummyvariabel er tanken at vi skal kunne skille ut en eventuell effekt av å være Bergen i perioden etter etableringen, og denne vil dermed tolkes som etableringseffekten.

Formelt vil forklaringsvariabelen formuleres som:

$$IKEA_i84_t = IKEA_i \times etter84_t \quad (4.3)$$

der $IKEA_i$ er lik 1 hvis kommune i mottar behandling på et vilkårlig tidspunkt, null ellers.

$etter84_t$ er lik 1 for alle tidspunkt, t , etter behandlingen i 1984, null ellers. Det er den estimerte koeffisienten til forklaringsvariabelen som er av interesse når vi skal undersøke etableringseffekten av IKEA på varehandelssektoren. Denne benyttes kun ved estimering med DID-metoden.

4.3.3 Kontrollvariabler

I dette delkapittelet skal vi presentere de ulike kontrollvariablene vi benytter oss av i analysen. Vi har valgt å benytte oss av kontrollvariabler som har blitt ansett som relevant for lignende problemsstillinger i tidligere forskning. Rudholm *et al.* (2018) benytter blant annet befolkningsantall og andelen av befolkningen med universitetsutdanning som kontrollvariabler i deres artikkel om effekten av IKEA-etableringer i Sverige. Vi har i tillegg hentet ut data om alderssammensetningen i de ulike kommunene.

Alderssammensetning

All data om alderssammensetning er hentet fra NSD sin kommunedatabase for menn og kvinner i separate aldersgrupper på fem år. I vår analyse er det aldersgrupper med større aldersintervaller som er av interesse. Vi la derfor sammen informasjonen om menn og kvinner til én gruppe. Deretter la vi sammen tre og tre av de femårige aldersgruppene, som NSD hadde definert, slik at vi endte opp med syv aldersgrupper med årsspenn på 15 år hver. De syv aldersgruppene er definert som følger:

Aldersgruppe 1 består av personer som er mellom 0 og 14 år gamle.

Aldersgruppe 2 består av personer som er mellom 15 og 29 år gamle.

Aldersgruppe 3 består av personer som er mellom 30 og 44 år gamle.

Aldersgruppe 4 består av personer som er mellom 45 og 59 år gamle.

Aldersgruppe 5 består av personer som er mellom 60 og 74 år gamle.

Aldersgruppe 6 består av personer som er mellom 75 og 89 år gamle.

Aldersgruppe 7 består av personer som er over 90 år gamle.

Omsetningen og sysselsettingen i en kommune kan være påvirket av alderssammensetningen i kommunen. Det er rimelig å anta at personer, avhengig av alder, har forskjellige handlemønstre. Omsetningen i varehandelssektoren kan dermed

påvirkes ulikt avhengig av størrelsen på de ulike aldersgruppene i kommunen. Det kan også være en sammenheng mellom sysselsettingen som andel av befolkning i en kommune og andelen personer i arbeidsfør alder. Dermed kan sysselsettingen i varehandelssektoren påvirkes avhengig av størrelsen på aldersgruppene i de ulike kommunene.

Vi vekter de ulike aldersgruppene med befolkningen i kommunen. Dette gjør at vi kan se hvor stor andel en aldersgruppe utgjør av den totale befolkningen. Formelt er aldersgruppe 1 som andel av befolkning formulert som:

$$ald_en_be_{it} = \frac{aldersgruppe\ en_{it}}{befolkning_{it}} \times 100 \quad (4.4)$$

der *aldersgruppe en_{it}* er antall personer som inngår i aldersgruppe 1 i kommune *i* på tidspunkt *t*. *befolkning_{it}* er antall innbyggere i kommune *i* på tidspunkt *t*. Vi multipliserer med 100 for å få andelen av befolkningen som inngår i aldersgruppe 1 oppgitt i prosent.

Variablene for resten av aldersgruppene beregnes på samme måte som aldersgruppe 1.

Ved estimering med DID-metoden benytter vi variablene for aldersgruppe 2-5, som andel av befolkning, som kontrollvariabler. Aldersgruppe 1, 6 og 7 utgjør dermed referansekategoriene. I analysen med den syntetiske kontrollmetoden vil prosessen forklart i delkapittel 3.3.2 avgjøre hvilke kontrollvariabler som inngår i modellspesifikasjonen.

Utdanningsnivå

Data for utdanningsnivå er hentet fra kommunedatabasen til NSD og er basert på antall personer over 16 år. Det mangler data for årene 1977 til 1979 for alle kommuner. Dette medfører at det er begrensninger knyttet til bruken av utdanningsnivå som kontrollvariabel, noe vi tar hensyn til ved gjennomføring av analysene våre.

NSD definerer de ulike utdanningsnivåene som følger:

Ingen utdanning inkluderer alle personer over 16 år med uoppgitt eller ingen fullført utdanning.

Grunnskole inkluderer alle personer over 16 år med grunnskoleutdanning som høyeste fullførte utdanning.

Videregående skole inkluderer alle personer over 16 år som har fullført videregående utdanning på to år eller mer. Personer som har fullført færre enn 120 studiepoeng ved et

universitet eller høyskole inngår også.

Kort universitetsutdanning inkluderer personer over 16 år som har fullført mer enn 120 studiepoeng i universitets- og høyskolesystemet, men som ikke har oppnådd en høyere grad fra høyere utdanning.

Lang universitetsutdanning inkluderer personer som har fullført en universitetsutdanning på 4 år eller mer ved et universitet eller høyskole. Dette inkluderer også de som har fullført forskerutdanning.

Formelt er personer med grunnskolenivå som andel av befolkning formulert som:

$$grunn_and_be_{it} = \frac{grunnskole_{it}}{befolkning_{it}} \times 100 \quad (4.5)$$

der $grunnskole_{it}$ er antall personer over 16 år med grunnskole som høyeste fullførte utdanning i kommune i på tidspunkt t . $befolkning_{it}$ er antall innbyggere i kommune i på tidspunkt t . Vi multipliserer med 100 for å få andelen av befolkningen over 16 år med fullført grunnskole oppgitt i prosent.

Variablene for de øvrige utdanningsnivåene beregnes på samme måte som vist for kontrollvariabelen $grunn_and_be_{it}$.

Vi benytter variablene for andelen av befolkningen med fullført videregående skole, kort- og lang universitetsutdanning som kontrollvariabler ved bruk av DID-metoden. Ingen utdanning og grunnskole utgjør dermed referansekategoriene. Ved bruk av SCM er alle variablene for utdanningsnivå tilgjengelige som potensielle kontrollvariabler, men alle inkluderes ikke nødvendigvis i modellspesifikasjonen.

Befolkning

Vi benytter ikke befolkning som kontrollvariabel ved estimering med DID-metoden. Bakgrunnen for dette er at siden alle variablene er vektet med befolkning, vil inkludering av befolkning som kontrollvariabel kunne føre til et endogenitetsproblem. Befolkning benyttes derimot som kontrollvariabel ved bruk av SCM, siden estimeringen der er av en annen art.

4.4 Deskriptiv statistikk

For å få et inntrykk av variasjonen i utfallsvariablene i tiden før og etter IKEA-etableringen har vi satt opp en tabell med deskriptiv statistikk som viser gjennomsnitt, minimums- og maksimumsverdier nedenfor. Vi skiller mellom Bergen, kontrollgruppe 1 og kontrollgruppe 2. I tillegg til utfallsvariablene, inkluderer vi også deskriptiv statistikk for befolkning, sysselsetting og omsetning for å se på eventuelle forskjeller mellom Bergen og kontrollgruppene. Deskriptiv statistikk er presentert i tabell 2 og tallene er rundet opp til nærmeste heltall.

Tabell 2: Deskriptiv statistikk for utfallsvariablene.

	Bergen		Kontrollgr. 1		Kontrollgr. 2	
	Før	Etter	Før	Etter	Før	Etter
Sysselsetting/befolkning						
Gj.snitt	3.68	3.57	4.00	3.83	3.87	3.68
Min	3.57	3.25	2.03	1.86	3.66	3.31
Max	3.78	3.85	6.39	6.15	4.17	3.98
Omsetning/befolkning						
Gj.snitt	23596	46516	26673	49740	26519	49325
Min	17878	34359	890	17897	21032	37362
Max	31016	54589	50607	80307	34169	57315

Merk: Sysselsetting som andel av befolkning er oppgitt i prosent.

Omsetning per innbygger er oppgitt i kroner.

	Bergen		Kontrollgr. 1		Kontrollgr. 2	
	Før	Etter	Før	Etter	Før	Etter
Befolkning						
Gj.snitt	209492	210405	34328	35231	97912	99938
Min	207292	207332	15631	16903	60037	61704
Max	212755	216066	135558	139630	135558	139630
Sysselsetting						
Gj.snitt	7705	7508	1342	1313	3746	3640
Min	7405	6939	468	528	2324	2333
Max	8033	8005	5281	5122	5281	5122
Omsetning						
Gj.snitt	49342	97993	9031	17370	25925	49610
Min	38036	71238	2711	5284	12798	23054
Max	64294	117948	46014	80029	46014	80029

Merk: Omsetning er oppgitt i 100 000 kroner.

Vi ser at Bergen har den laveste andelen sysselsatte i varehandelssektoren sammenlignet med kontrollgruppe 1 og 2, både i perioden før og etter etableringen. Videre

ser vi at kontrollgruppe 1 har mest variasjon mellom kommunene med 4.36 prosentpoeng differanse. Bergen har også det laveste gjennomsnittet i begge periodene når det kommer til omsetning per innbygger, med 23596 og 46516 kroner i gjennomsnitt i henholdsvis perioden før og etter behandling. For de to kontrollgruppene er gjennomsnittlig omsetning per innbygger omlag 3000 kroner høyere enn i Bergen i begge perioder. Dette indikerer at det har vært en lignende utvikling i Bergen som i kontrollgruppe 1 og 2.

I den nedre delen av tabellen ser vi gjennomsnittet for befolkning, omsetning og sysselsetting i perioden før og etter etableringen. Dette gjør at vi kan se de faktiske forskjellene mellom kontrollgruppene og Bergen. Vi ser at gjennomsnittlig befolkning i Bergen er omtrent seks ganger så stor som i kontrollgruppe 1, og dobbelt så stor som kontrollgruppe 2. Differansene er av tilsvarende størrelse når det kommer til sysselsetting og omsetning. Dette kan forklare hvorfor det ikke er betydelige forskjeller i omsetning per innbygger og sysselsetting som andel av befolkning mellom Bergen og de ulike kontrollgruppene.

4.5 Svakheter ved datamaterialet

Utelatte kommuner

Grunnet manglende data har vi fjernet 44 kommuner fra det opprinnelige datamaterialet, bestående av 439 kommuner. Dette kan føre til en svakhet ved datasettet vårt dersom kommuner som kunne vært relevant for analysen er blant de kommunene som ble fjernet.

Av de 44 kommunene som ble fjernet var det 8 kommuner som hadde et innbyggertall på over 15 000 i 1977. Dette indikerer at de 8 kommunene kunne blitt inkludert i utvalget som utgjør kontrollgruppe 1.

Hovedsvakheten knyttet til utelatte kommuner er at Oslo og Ringsaker, som har fått IKEA i senere tid, er blant de som ble fjernet. Dette fører til at kontrollgruppe 2 mister mye variasjon og representerer ikke en like tilfredsstillende kontrollgruppe for Bergen. Oslo er dessuten den eneste kommunen som er større enn Bergen i folketall, noe som kunne vært nyttig å ha i en kontrollgruppe.

Within-variasjon

En svakhet ved datamaterialet vårt er at alle tilgjengelige variabler er demografiske. Det kan argumenteres for at demografiske faktorer, ved fravær av store sjokk, ikke varierer mye over

tid innad i en kommune. Dette gjør at det er en risiko for at det er liten within-variasjon i variablene våre. I delkapittel 3.2.2 forklarte vi at vi benytter en modell med kommunefaste effekter når vi estimerer DID-estimatoren, og utfører en within-transformasjon. En slik transformasjon innebærer at vi kun utnytter variasjonen innad i en kommune over tid når vi estimerer modellen. Liten within-variasjon i kontrollvariablene vil dermed kunne føre til at den estimerte modellspesifikasjonen kun forklarer en liten del av den totale uforklarte variasjonen i utfallsvariabelen.

4.6 Oppsummering

I dette kapitlet har vi presentert de to tilgjengelige behandlingsenhetene, og forklart hvorfor vi kun benytter oss av Bergen i de videre analysene. Vi har definert to kontrollgrupper, den ene består av 30 utvalgte kommuner og den andre består av to kommuner som i senere tid har fått etablert IKEA.

Vi benytter oss av omsetning i varehandelssektoren per innbygger og sysselsetting i varehandelssektoren som andel av befolkning som utfallsvariabler. Videre i oppgaven vil de omtales som henholdsvis omsetning og sysselsetting.

5 Metodiske utfordringer

I dette kapitlet vil vi belyse mulige utfordringer ved gjennomføringen av analysene i denne oppgaven. Vi gjør først rede for generelle forutsetninger for estimering ved bruk av DID-metoden. Deretter presenterer vi mulige økonomiske utfordringer vi står overfor ved estimering av DID-estimatoren og hvordan disse kan tas hensyn til. Til slutt gjør vi rede for utfordringer knyttet til bruk av den syntetiske kontrollmetoden.

5.1 Utfordringer ved DID-metoden

5.1.1 Forutsetninger

Parallell trend i utfallsvariabelen

Ved bruk av DID-metoden er parallelle trender hovedforutsetningen som må holde for å kunne finne den kausale effekten av behandlingen. Dette innebærer at trenden i utfallsvariabelen i perioden før behandling må være lik for behandlings- og kontrollgruppen. Nivået kan være ulikt, så lenge utviklingen over tid er den samme.

Prinsippet med forutsetningen er at DID-estimatoren tolker enhver endring i utviklingen over tid, mellom behandlings- og kontrollgruppen, i perioden etter behandlingen som behandlingseffekten. Dersom trendene ikke er parallelle i perioden før, vil DID-estimatoren kunne under- eller overestimere behandlingseffekten. I hvilken retning, og i hvor stor grad, effekten vil kunne feilestimeres avhenger av om utfallstrendene til behandlings- og kontrollgruppen konvergerer eller divergerer.

Vi gjennomfører placeboanalyser i kapittel 6 for å teste om forutsetningen om parallelle trender mellom Bergen og kontrollgruppe 1 og 2 er oppfylt.

Tildeling uavhengig av utfallsvariabelen

En annen forutsetning for at DID-estimatoren skal være forventningsrett og kunne angi den kausale effekten av en IKEA-etablering, er at tildelingen må være gjort uavhengig av verdien på utfallsvariablene i perioden før etableringen finner sted Columbia University.

Vi har tidligere antatt at det ikke er tilfeldig hvor det velges å etablere en IKEA. Det er dermed rimelig å anta at omsetningen, eller omsetningsutviklingen, innen varehandelssektoren i en potensiell etableringskommune tas med i betraktningen når

plasseringen av et IKEA-varehus skal bestemmes. Dersom det er tilfellet vil dette indikere brudd på forutsetningen om tildeling uavhengig av utfallsvariabelen. Vi har ingen informasjon om faktorene ved en kommune som IKEA vektlegger i valget av plasseringen av et nytt varehus. Det er dermed uvisst om forutsetningen er brutt.

5.1.2 Økonometriske utfordringer

Utelatte variabler

Utelatte variabler omtaler de variablene som påvirker utfallsvariabelen, men som ikke er inkludert i modellen og dermed inngår i restleddet. Dette kan skyldes at man glemmer å inkludere observerbare data, det er mangel på data, eller det er uobserverbare variabler som det kan være vanskelig å kvantifisere.

Utelatte variabler skaper problemer dersom én eller flere av disse, som da inngår i restleddet, er korrelert med en forklaringsvariabel. Dette vil medføre utelatt variabelskjevhet. Det er rimelig å anta at restleddet inkluderer flere utelatte variabler som er konstante over tid innad i en kommune og som er korrelert med forklaringsvariabelen vår. Dette bryter med forutsetningen om eksogene forklaringsvariabler og estimering ved pooled OLS vil gi forventningsskjevne estimatorer. For å løse dette problemet benytter vi oss av en modell med kommunefaste effekter. Dette er ekvivalent med å utføre en within-transformasjon som måler variablene som avvik fra kommunespesifikke gjennomsnitt. En slik transformasjon kontrollerer for utelatte variabler som er faste over tid innad i en kommune.

Seriekorrelasjon og heteroskedastisitet

Ved estimering med paneldata kan restleddene være seriekorrelerte innenfor en kommune. Selv om vi kontrollerer for dette ved estimering med faste kommuneeffekter, kan vi ikke utelukke at restleddet fremdeles er seriekorrelert. Seriekorrelasjon vil ikke påvirke estimatene våre, som fremdeles vil være forventningsrette, men formlene for utregning av standardavvik og varians vil ikke lenger være gyldige. Statistisk inferens basert på ukorrigerede standardavvik vil dermed være ugyldige og kan gi villedende resultater. Vår løsning på dette problemet er å benytte cluster-robuste standardavvik som korrigerer standardavvikene for vilkårlig seriekorrelasjon over tid innad i en kommune.

Liten within-variasjon

I delkapittel 4.5 presenterte vi en mulig svakhet ved datamaterialet vårt dersom det er liten within-variasjon i kontrollvariablene våre. Ved liten within-variasjon vil estimering av en modell med kommunefaste effekter, *fixed effects*-modell, fremdeles gi forventningsrette estimatorer, men de vil ikke være effisiente. Dette skyldes at når within-transformasjonen utføres på en variabel med liten within-variasjon vil variansen til estimatene kunne gå mot uendelig. Den estimerte within-estimatoren vil dermed bli upålitelig (Plümper, 2007).

Med bakgrunn i det begrensede antallet kontrollvariabler vi har tilgjengelig, har vi tidligere antatt at det er meget sannsynlig at det finnes kommunespesifikke variabler som er korrelert med forklaringsvariablen, men som er blitt utelatt fra modellen. Vi har derfor valgt å benytte kommunefaste effekter i modellen som estimeres for å oppnå en forventningsrett estimator, selv om dette potensielt kan gå på bekostning av effisiensen til estimatoren.

5.2 Utfordringer og begrensninger ved SCM

I kapittel 3.3 beskriver vi den syntetiske kontrollmetoden. Etter introduksjonen av SCM av Abadie *et al.* (2003, 2010, 2015) har metoden blitt et nytt og populært verktøy for økonometrikere til bruk i komparative analyser (Athey, 2017). Selv om metoden har blitt benyttet en rekke ganger har den også fått kritikk. Ferman (2016) viser hvordan mangelen på retningslinjer ved valg av modellspesifikasjon fører til muligheter for spesifikasjonsleting. Mangelen på retningslinjer har ført til at flere har brukt alle lag i modellspesifikasjonen de har benyttet.²⁵ Det viser seg at når alle lag er inkludert vil SCM kun ta hensyn til disse når kontrollenheten konstrueres (Kaul, 2018). Kaul *et al.* (2015) viser at kontrollvariablene som utelates vil kunne føre til en skjevhet i estimatene, men at dette avhenger av lengden på tidsperioden før behandling og forklaringskraften til de utelatte kontrollvariablene.

Botosaru og Ferman (2019) ser nærmere på hvordan brudd i antagelsen om perfekt balanse mellom observerte kontrollvariabler og lag påvirker grensene i intervallet til

²⁵En ikke utfyllende liste over publikasjoner som benytter alle lag ved bruk av den syntetiske kontrollmetoden: (Cavallo, 2012), (Bilgel, 2015), (Billmeier, 2013), (Bohn, 2014), (Hinrichs, 2012), (Kreif, 2016), (Liu, 2015), (O'Neill, 2016), (Nannicini, 2011) og (Stearns, 2015).

forventningskjevheten i estimatoren.²⁶ De viser, med data fra Abadie *et al.* (2015), at kun bruk av alle lag gir nesten identiske resultater som ved bruk av kontrollvariabler. De argumenterer likevel for at en kombinasjon av kontrollvariabler, som gir god balanse før behandling, sammen med lag vil gi den laveste forventningskjevheten i resultatene. Spesielt i tilfeller hvor det kun er data for et lite antall år før behandlingen vil denne metoden gi mer robuste resultater (Botosaru, 2019). Dette er tilfellet for vårt datasett og en av grunnene til at vår prosess for valg av modellspesifikasjon tar hensyn til balanse i vektingen mellom kontrollvariabler og lag.

²⁶Med perfekt balanse menes det at det finnes et sett med vektorer, for hver av variablene, som gjør at verdiene til den behandlede gruppen er lik det vektete snittet til kontrollgruppen når vektene brukes (Botosaru, 2019).

6 Resultater

I dette kapittelet presenteres resultatene av analysene vi har utført ved bruk av DID-metoden og SCM. Delkapittel 6.1 tar for seg resultatene fra den empiriske analysen med differanse i differanse-metoden. Alle regresjonene er gjennomført i programmet *STATA 15* og inkluderer kommunefaste effekter, i tillegg til årsummier. Vi anvender cluster-robuste standardavvik i alle estimeringene. Delkapittel 6.2 presenterer resultatene fra analysen med den syntetiske kontrollmetoden. Alle analysene er gjennomført i programvaren *STATA 15* med tilleggspakkene *synth* og *synth_runner*.²⁷

6.1 DID

Modellspesifikasjonene vi benytter for å estimere DID-estimatoren er presentert ved ligning (3.3) og (3.4) i delkapittel 3.2.2. Vi estimerer DID-estimatoren med tre ulike modellspesifikasjoner. Resultatene fra estimeringene blir presentert i samme tabell. Ligning (3.3) er modellspesifikasjonen uten kontrollvariabler og resultatene fra estimeringen blir rapportert i kolonne (1). Vi estimerer to varianter av modellspesifikasjonen med kontrollvariabler. Den ene, som rapporteres i kolonne (2), inkluderer kun variablene for alderssammensetning. Siden vi mangler informasjon om variablene for utdanningsnivå i årene før 1980, begrenser vi tidsperioden for den siste modellspesifikasjonen. Vi kan dermed inkludere både utdanningsnivå og alderssammensetning som kontrollvariabler når vi ser på perioden 1980 til 1992. Resultatet fra denne modellspesifikasjonen presenteres i kolonne (3).

Ved estimering med faste individeffekter vil ikke konstantleddene kunne tolkes på vanlig måte som gjennomsnittet for referansekategoriene.²⁸ Dette gjør at vi ikke kan tolke effektene opp mot en gjennomsnittsverdi for å finne den relative størrelsen på effekten. Vi har derfor valgt å benytte oss av gjennomsnittsverdien for utfallsvariablene våre i perioden før etableringen som referanseverdier. Dette vil ikke vise den sanne effekten, men vil kunne gi oss en numerisk verdi på som vi kan sammenligne på tvers av kontrollgruppene. Vi

²⁷Programtillegget *synth* er utviklet av (Abadie, 2011) og *synth_runner* er en utvidelse av tillegget *synth* laget av (Galiani, 2016)

²⁸Ved bruk av within-transformerte data vil konstantleddet som oppgis av *STATA* vise gjennomsnittlig verdi på de faste enhetseffektene.

oppgir den relative størrelsen på effekten i parentes.

Gjennomsnittene er tatt fra tabell 4 om deskriptiv statistikk og repeteres her:

Tabell 3: Gjennomsnittsverdi for utfallsvariablene i perioden før etableringen, 1977-1983.

	Omsetning	Sysselsetting
Kontrollgruppe 1	26 673 kroner	4%
Kontrollgruppe 2	26 519 kroner	3.9%

Vi ser av tabell 3 at omsetning er oppgitt i kroner og at sysselsetting er oppgitt i prosent. Dette gjør at de estimerte koeffisientene tolkes forskjellig avhengig av om det er omsetning eller sysselsetting som er utfallsvariabel. En endring i omsetning per innbygger er oppgitt i kroner, mens en endring i sysselsetting som andel av befolkning er oppgitt i prosentpoeng.

6.1.1 DID: kontrollgruppe 1.

Resultatene fra DID-analysen med kontrollgruppe 1 presenteres i tabell 4 og 5. Fullstendig tabell for estimeringen finnes i avsnitt F i appendiks.

Tabell 4: Estimerte effekter av IKEA-etableringen på omsetning i varehandelssektoren per innbygger. Kontrollgruppe 1.

Utfallsvariabel: Omsetning per innbygger, 1977 til 1992.			
	(1)	(2)	(3)
IKEA84	-147 (950)	-108 (1060)	-1290 (1290)
_cons	21300*** (757)	124800** (54400)	86000* (49900)
Årsdummier	JA	JA	JA
Kommunefaste effekter	JA	JA	JA
Kontrollvariabler	NEI	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
Observasjoner	496	496	403
Antall kommuner	31	31	31
Within R^2	0.942	0.946	0.939

Merk: Cluster-robuste standardavvik i parentes.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Omsetning

Tabell 4 rapporterer en gjennomgående negativ effekt av IKEA-etableringen på omsetningen. Den minste effekten, i absoluttverdi, gir en gjennomsnittlig reduksjon på 108 kroner (-0.4%). Den største gir en reduksjon på 1290 kroner (-4.8%). Ingen av effektene er imidlertid statistisk signifikante. Dette er et noe overraskende resultat da vi forventet å se en positiv effekt på omsetning som følge av etableringen.

Tabell 5: Estimerte effekter av IKEA-etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 1.

Utfallsvariabel: Sysselsetting som andel av befolkning, 1977 til 1992.

	(1)	(2)	(3)
IKEA84	0.07 (0.0433)	0.14*** (0.0328)	0.13*** (0.0376)
_cons	4.12*** (0.0347)	2.49 (2.10)	5.61** (2.23)
Årsdummier	JA	JA	JA
Kommunefaste effekter	JA	JA	JA
Kontrollvariabler	NEI	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
Observasjoner	496	496	403
Antall kommuner	31	31	31
Within R^2	0.528	0.606	0.616

Merk: Cluster-robuste standardavvik i parentes.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Sysselsetting

Resultatet fra grunnmodellen viser at sysselsettingen økte med gjennomsnittlig 0.07 prosentpoeng (+1.8%) etter etableringen av IKEA. Når vi inkluderer kontrollvariabler ser vi at etableringseffekten indikerer en gjennomsnittlig økning i sysselsettingen med 0.14 prosentpoeng (+3.5%) i modellen med kontrollvariabler for alderssammensetning, og 0.13 prosentpoeng (+3.3%) i modellen med alle kontrollvariablene. Begge resultatene er statistisk signifikante til 1% signifikansnivå.

En positiv etableringseffekt på sysselsettingen er i tråd med det vi forventet å finne, og fremstår som et troverdig resultat. Det står i kontrast til den estimerte effekten på omsetningen som var negativ, og dermed motsatt av det vi forventet.

6.1.2 DID: kontrollgruppe 2.

Vi gjennomfører de samme analysene med kontrollgruppe 2. Fullstendig tabell for estimeringen finnes i avsnitt F i appendiks. Kontrollgruppe 2 består av kommuner som fikk etablert IKEA på et senere tidspunkt enn Bergen. Motivasjonen bak å bruke denne kontrollgruppen er forklart mer utdypende i kapittel 4. En begrensning ved denne kontrollgruppen er at den kun består av to kommuner. Dette kan gjøre statistisk inferens vanskelig, ved at de estimerte standardavvikene blir mindre pålitelige ved få tilgjengelige observasjoner.

Omsetning

Tabell 6: Estimerte effekter av IKEA-etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 2.

Utfallsvariabel: Omsetning per innbygger, 1977 til 1992.			
	(1)	(2)	(3)
IKEA84	115 (857)	-938 (851)	-3000** (321)
_cons	20100*** (324)	-15200 (14300)	593900* (145600)
Årsdummier	JA	JA	JA
Kommunefaste effekter	JA	JA	JA
Kontrollvariabler	NEI	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
Observasjoner	48	48	39
Antall kommuner	3	3	3
Within R^2	0.996	0.996	0.996

Merk: Cluster-robuste standardavvik i parentes.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Vi ser i kolonne (1), i tabell 6, at IKEA-etableringen assosieres med en økning i omsetningen med 115 kroner (+0.4%) i gjennomsnitt. Effekten er ikke statistisk signifikant. Ved inkludering av kontrollvariabler endrer effekten av IKEA på omsetningen fortegn, og indikerer en negativ effekt. Omsetningen reduseres med 938 kroner (-3.5%) og 3000 kroner (-11.3%) i gjennomsnitt for modellspekifikasjonene som rapporteres i henholdsvis kolonne (2) og (3). Sistnevnte effekt er signifikant til 5% signifikansnivå.

Resultatene fra estimering med kontrollgruppe 2 viser, i likhet med resultatene for kontrollgruppe 1, en negativ effekt på omsetningen etter etableringen av IKEA. Effekten har dermed motsatt fortegn av det vi forventet.

Vi ser at den eneste statistisk signifikante effekten indikerer en gjennomsnittlig reduksjon i omsetningen med 11.3%. Dette er en kraftig reduksjon, både isolert sett og sammenliknet med de andre estimerte effektene.

Sysselsetting

Tabell 7: Estimerte effekter av IKEA-etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 2.

Utfallsvariabel: Sysselsetting som andel av befolkning, 1977 til 1992. Kontrollgruppe 2.

	(1)	(2)	(3)
IKEA84	0.08** (0.017)	0.18* (0.055)	0.14** (0.017)
_cons	3.92*** (0.012)	3.21 (2.55)	35.5 (15)
Årsdummier	JA	JA	JA
Faste kommuneeffekter	JA	JA	JA
Kontrollvariabler	NEI	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
Observasjoner	48	48	39
Antall kommuner	3	3	3
Within R^2	0.928	0.933	0.945

Merk: Cluster-robuste standardavvik i parentes.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Resultatene i tabell 7 indikerer at sysselsettingen øker etter etableringen av IKEA. Estimeringen av grunnmodellen gir en gjennomsnittlig økning i sysselsettingen med

0.08 prosentpoeng (+2.1%), signifikant til 5% signifikansnivå. Ved inkludering av kontrollvariabler impliserer resultatene at etableringen bidrar til, i gjennomsnitt, en 0.18 (+4.7%) og 0.14 (+3.7%) prosentpoengs økning i sysselsettingen. Effektene er statistisk signifikante til henholdsvis 10% og 5% signifikansnivå.

De estimerte effektene vi finner ved bruk av kontrollgruppe 2 er forholdsvis like de vi finner med kontrollgruppe 1. Negativ effekt av etableringen på omsetning og positiv effekt på sysselsetting. Igjen får vi et uventet resultat, noe som gjør at vi stiller spørsmål ved egnetheten til kontrollgruppene våre ved bruk av DID-metoden.

Vi ønsker å teste om kontrollgruppene vi har benyttet for å estimere DID-estimatoren kan sies å representere tilfredsstillende kontrollgrupper for Bergen. Dette kan testes ved å utføre en placeboanalyse, noe vi presenterer i neste delkapittel.

6.1.3 DID: Placeboanalyse.

Ved bruk av DID-metoden kan man gjennomføre en placeboanalyse for å undersøke om forutsetningen om parallelle trender holder. Placeboanalysen baseres på et tankeeksperiment. Vi antar at behandlingen inntreffer på et tidligere tidspunkt enn det som er tilfellet, fra nå av omtalt som falsk behandling. Ved parallelle trender mellom behandlingsenheten og kontrollgruppen i tiden før behandling, skal ikke en estimering av den falske behandlingen gi effekter som er statistisk signifikante. Fullstendig tabell for estimeringen finnes i avsnitt F i appendiks.

Vi utfører placeboanalysen ved å begrense tidsperioden vi ser på til kun å gjelde for årene før IKEA ble etablert i Bergen, altså 1977 til 1983. Den falske behandlingen inntreffer i 1980. Perioden 1977 til 1979 vil da være tiden før den falske behandlingen, og perioden 1980 til 1983 vil være den *falske* behandlingsperioden. Siden vi ikke har informasjon om utdanningsnivå for årene 1977 til 1979 kan ikke de benyttes som kontrollvariabler. Vi inkluderer derfor kun alderssammensetning som kontrollvariabler i placeboanalysen.

Placeboanalysen utføres med omsetning og sysselsetting som utfallsvariabler, og vi estimerer modellene uten og med kontrollvariabler. Vi gjør dette for både kontrollgruppe 1 og kontrollgruppe 2.

Placeboanalyse: omsetning.**Tabell 8:** Estimerte effekter av den falske etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 1 og 2.

Utfallsvariabel: Omsetning per innbygger, 1977 til 1983.

	Kontrollgruppe 1		Kontrollgruppe 2	
	(1)	(2)	(3)	(4)
IKEA80	-4.41 (0.00435)	-549 (0.00524)	-136** (0.000227)	-639 (0.00586)
_cons	21300*** (0.00323)	100800* (0.514)	20100*** (0.00240)	-193000* (0.524)
Årsdummier	JA	JA	JA	JA
Faste kommuneeffekter	JA	JA	JA	JA
Kontrollvariabler	NEI	JA	NEI	JA
Estimeringsperiode	1977-1983	1977-1983	1977-1983	1977-1983
Observasjoner	217	217	21	21
Antall kommuner	31	31	3	3
Within R^2	0.898	0.912	0.995	0.998

Merk: Cluster-robuste standardavvik i parentes.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Kolonne (1) og (2) viser resultatene fra placeboanalysen med kontrollgruppe 1. Resultatet fra kolonne (1) indikerer en svak negativ effekt av den falske behandlingen på omsetningen, men den er ikke statistisk signifikant. Kolonne (2) viser resultatene for modellspesifikasjonen med kontrollvariabler, og vi ser at effekten er negativ. Den er ikke statistisk signifikant.

For kontrollgruppe 2 ser vi av kolonne (3) at den falske behandlingen er assosiert med en negativ effekt på omsetningen. Effekten er statistisk signifikant til 5% signifikansnivå. Kolonne (4) viser også en negativ effekt, men den er ikke statistisk signifikant.

Et statistisk signifikant resultat i kolonne (3) indikerer et brudd på forutsetningen om parallelle trender ved bruk av kontrollgruppe 2 med omsetning som utfallsvariabel.

Placeboanalyse: sysselsetting.

Tabell 9: Estimerte effekter av den falske etableringen på sysselsettingen som andel av befolkning. Kontrollgruppe 1 og 2.

Utfallsvariabel: Sysselsetting som andel av befolkning, 1977 til 1983.				
	Kontrollgruppe 1		Kontrollgruppe 2	
	(1)	(2)	(3)	(4)
IKEA80	0.06**	0.04	0.05	-0.03
	(0.0277)	(0.0361)	(0.0277)	(0.0544)
_cons	4.12***	-3.38	3.92***	-13.4
	(0.0216)	(2.56)	(0.009)	(10.7)
Årsdummier	JA	JA	JA	JA
Kommunefaste effekter	JA	JA	JA	JA
Kontrollvariabler	NEI	JA	NEI	JA
Estimeringsperiode	1977-1983	1977-1983	1977-1983	1977-1983
Observasjoner	217	217	21	21
Antall kommuner	31	31	3	3
Within R^2	0.434	0.547	0.826	0.925

Merk: Cluster-robuste standardavvik i parentes.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tabell 9 viser placeboanalysen for sysselsetting. Kolonne (1) og (2) viser resultatene fra placeboanalysen med kontrollgruppe 1. Modellen i kolonne (1) er estimert uten kontrollvariabler og indikerer en positiv statistisk signifikant effekt til 3.6% signifikansnivå. Modellen i kolonne (2) inneholder kontrollvariabler og indikerer også en positiv effekt, men resultatet er ikke statistisk signifikant.

Kolonne (3) og (4) viser resultatene ved bruk av kontrollgruppe 2. Kolonne (3) indikerer en positiv effekt, som kolonne (1) og (2), men den er ikke statistisk signifikant. I motsetning til resultatene fra de andre placeboanalysene indikerer kolonne (4) en negativ effekt på sysselsettingen av den falske behandlingen. Den er ikke statistisk signifikant.

En statistisk signifikant effekt på sysselsettingen ved bruk av kontrollgruppe 1 kan tyde på brudd på forutsetningen om parallelle trender.

6.1.4 Oppsummering

Resultatene fra estimeringen med DID-metoden gir forholdsvis like, og noe overraskende, resultater ved bruk av begge kontrollgruppene. Vi finner en negativ effekt på omsetningen og en positiv effekt på sysselsettingen. Dette er et noe motstridende resultat da vi kan forvente at omsetningen og sysselsettingen vil følge hverandres utvikling til en viss grad.

Resultatene fra placeboanalysen indikerer at det er et mulig brudd på forutsetningen om parallelle trender mellom Bergen og kontrollgruppe 1 og 2, basert på analysen med henholdsvis sysselsetting og omsetning som utfallsvariabel.

Den motsatte effekten av IKEA-etableringen på omsetning sammenlignet med sysselsetting, i tillegg til indikasjonen om mulig brudd på forutsetningen om parallelle trender, antyder at de estimerte effektene ikke angir den sanne effekten av IKEA-etableringen.

6.2 SCM

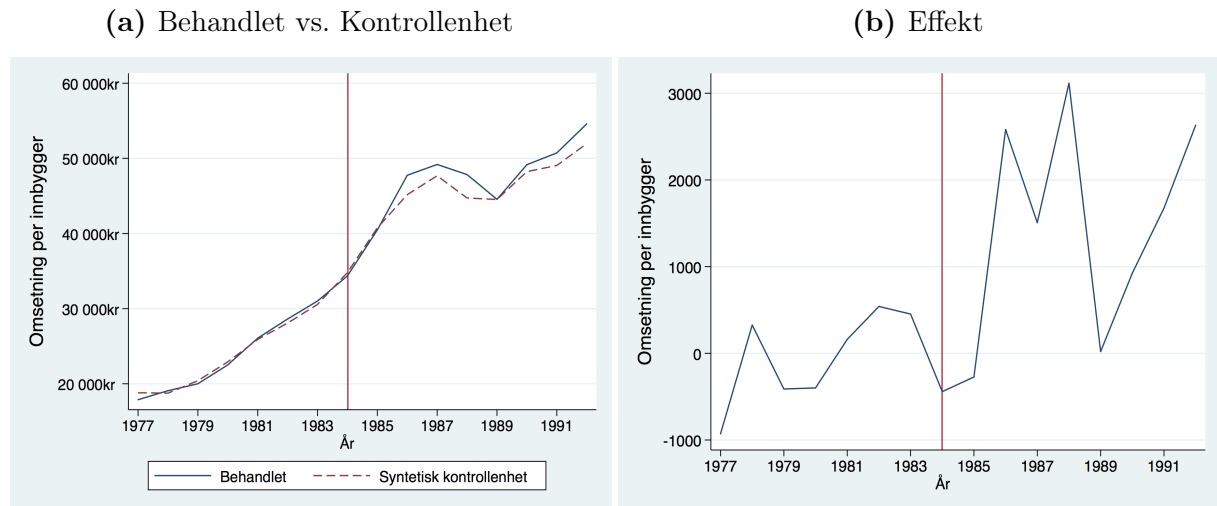
I dette delkapittelet presenteres resultatene fra den syntetiske kontrollmetoden. Valg av modellspesifikasjon er gjennomført med metoden forklart i avsnitt 3.3.2 om implementering.

6.2.1 Effekt på omsetning

Vektingen av kontrollkommuner og kontrollvariabler som ligger til grunn for estimeringen av den syntetiske kontrollenheten er rapportert i avsnitt G i appendiks. Tabell 22a i appendiks viser kun kommunene som har blitt tildelt vekt i den syntetiske kontrollenheten. Tabell 22b i appendiks viser hvilke kontrollvariablene og lag som er inkludert i modellspesifikasjonen. Den viser også vektingen til de ulike variablene. Vi observerer at Trondheim har blitt tildelt størst vekt av kommunene i kontrollgruppen, med en vekting på 78.5%. Dette er ikke overraskende med tanke på at Trondheim er en av de største byene i kontrollgruppen. I tabell 11b i appendiks ser vi at lag for omsetningen i 1977 og 1980 har fått en total vekting på 64.8%, mens kontrollvariablene har en totalvekt på 35.2%. Dette er ikke en perfekt balanse, men i avveiningen mellom lav pre-RMSPE og balansert vekting ble denne modellspesifikasjonen ansett som den beste.

Figur 5 viser resultatene fra den syntetiske kontrollmetoden med omsetning som utfallsvariabel.

Figur 5: Grafisk fremstilling av utviklingen og etableringseffekten på omsetning i varehandelssektoren per innbygger.



Figur 5a viser den faktiske utviklingen i omsetningen i Bergen og utviklingen i den syntetiske kontrollenheten. Den røde linjen er satt for å markere tidspunktet for etableringen. Tilpasningen i perioden før etableringen gir en indikasjon på troverdigheten til den syntetiske kontrollenheten som et kontrafaktisk utfall til Bergen i perioden etter etableringen. Vi ser i figur 5a at tilpasningen virker å være meget god i perioden før etableringen. For å tydeliggjøre effekten på omsetningen viser figur 5b differansen i utfallsvariabelen mellom Bergen og den syntetiske kontrollenheten. Tabell 10 viser tallverdiene som korresponderer med effekten vi observerer i figur 5b.

Den gjennomsnittlige effekten fra 1984 til 1992 på omsetningen er på 1304 kroner. Dette tilsvarer en 2.89% økning i omsetningen i Bergen sammenlignet med gjennomsnittlig verdi på omsetningen til den syntetiske kontrollenheten for samme tidsperiode.

Vi observerer at det er en negativ effekt på omsetningen i 1984. Dette tilsvarer en reduksjon i omsetningen i Bergen på 1.26% sammenlignet med den syntetiske kontrollenheten. I 1985 ser vi at effekten fremdeles er svakt negativ, men fra og med 1986 er den strengt positiv. Vi forventet en positiv direkte effekt av etableringen og dette er dermed et noe overraskende resultat.

Tabell 10: Utviklingen i omsetning i varehandelssektoren per innbygger for Bergen og den syntetiske kontrollenheten. Differansen i utviklingen angir effekten av IKEA-etableringen.

Utfallsvariabel: Omsetning per innbygger, 1977 til 1992.

År	Bergen (1)	Syntetiske Bergen (2)	Effekten (3)
1984	34359	34801	-441
1985	40477	40749	-273
1986	47752	45171	2581
1987	49191	47684	1507
1988	47842	44727	3115
1989	44550	44529	20
1990	49161	48236	924
1991	50727	49052	1675
1992	54589	51958	2631

I 1988 er effekten sterkt positiv og indikerer en økning i omsetningen i Bergen på 6.96% sammenlignet med den syntetiske kontrollenheten dette året. Året etter er effekten derimot tilnærmet lik null. Dette kan ha en sammenheng med oljeprissjokket som også finner sted i denne tidsperioden.

Siden oljeprissjokket kan tenkes å ha påvirket omsetningen i Bergen etter 1987, undersøker vi den gjennomsnittlige effekten for årene 1984 til 1987. For denne tidsperioden er den gjennomsnittlige effekten på omsetningen 843 kroner. Dette tilsvarer en 2% økning i omsetningen i Bergen sammenlignet med gjennomsnittet til den syntetiske kontrollenheten.

Vi observerer dermed at effekten på omsetningen for hele perioden er større enn for perioden i forkant av oljeprissjokket. Dette er motsatt av hva vi forventet å observere.

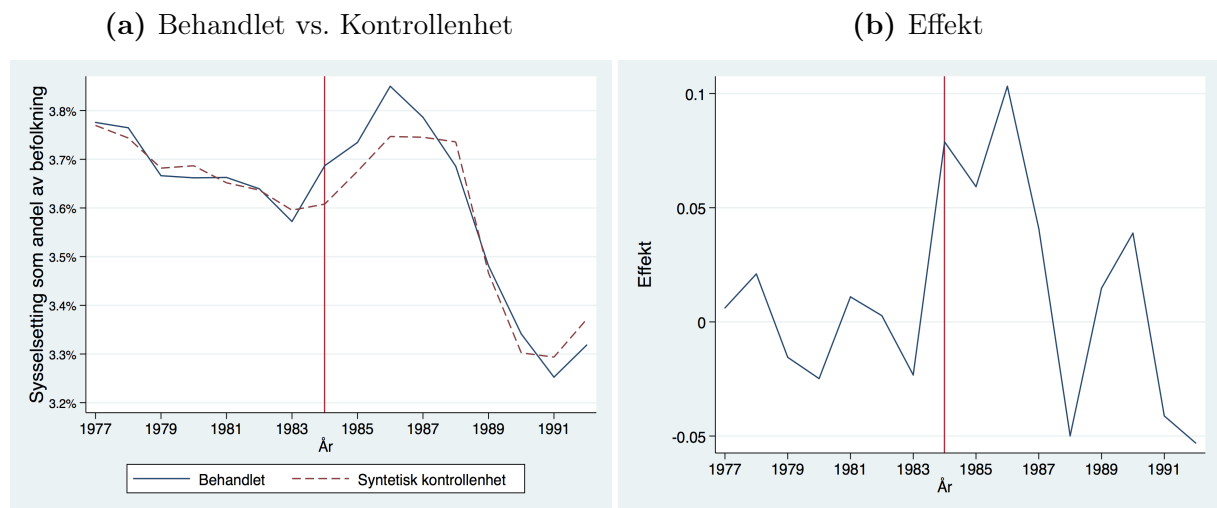
6.2.2 Effekt på sysselsetting

Vektingen av kontrollkommuner og kontrollvariabler som ligger til grunn for estimeringen av den syntetiske kontrollenheten er rapportert i tabell 23 i appendiks i avsnitt G. Tabell 23a viser at Trondheim og Moss har fått størst vekt på henholdsvis 31% og 29.6%. Igjen observerer vi at Trondheim har fått størst vekt, men i dette tilfellet er den ikke like dominerende som i den syntetiske kontrollenheten for omsetning. Tabell 23b viser at den totale vektingen til lag er på 37.8%, mens den totale vektingen av kontrollvariablene er på

62.2%. Variablene vgs og aldersgruppe 1 har fått størst vekt på henholdsvis 19% og 38%. Det er ikke en optimal balanse i vektingen, men den beste blant alternativene.

Figur 6 viser resultatene fra den syntetiske kontrollmetoden med sysselsetting som utfallsvariabel.

Figur 6: Grafisk fremstilling av utviklingen og etableringseffekten på sysselsetting i varehandelssektoren som andel av befolkning.



Vi observerer at tilpasningen i perioden før etableringen ikke er like god som den vi observerer i analysen med omsetning. Videre ser vi i figur 6b at effekten i etterkant av etableringen er volatil. Dette undersøker vi nærmere i tabell 11.

Den gjennomsnittlige effekten på sysselsetting for hele tidsperioden etter etableringen er på 0.02 prosentpoeng. Sammenlignet med gjennomsnittet for den syntetiske kontrollenheten tilsvarer dette en økning på 0.6%.

Vi observerer en positiv direkte effekt på sysselsettingen i 1984 på 2.19% sammenlignet med den syntetiske kontrollenheten. Dette er i tråd med forventningen om en positiv effekt av etableringen. Videre ser vi at den største effekten er i 1986 og indikerer en økning i sysselsettingen på 2.75%.

De første årene etter etableringen er effekten positiv. I 1988 observerer vi imidlertid at effekten på sysselsettingen er blitt negativ. Denne negative effekten sammenfaller med tidsperioden for oljeprissjokket og det kan være en mulig sammenheng her.

Tabell 11: Utviklingen i sysselsettingen i varehandelssektoren som andel av befolkning for Bergen og den syntetiske kontrollenheten. Differansen i utviklingen angir effekten av IKEA-etableringen.

Utfallsvariabel: Sysselsetting som andel av befolkning, 1977 til 1992.

År	Bergen (1)	Syntetiske Bergen (2)	Effekten (3)
1984	3.687	3.608	0.079
1985	3.735	3.675	0.059
1986	3.850	3.747	0.103
1987	3.786	3.745	0.041
1988	3.686	3.736	-0.050
1989	3.481	3.467	0.015
1990	3.341	3.303	0.039
1991	3.252	3.294	-0.041
1992	3.318	3.371	-0.053

På grunn av oljeprissjokket ønsker vi å undersøke effekten for årene 1984 til 1987 isolert. For denne tidsperioden er den gjennomsnittlige effekten på sysselsettingen på 0.07 prosentpoeng. Dette tilsvarer en økning mot gjennomsnittet for den syntetiske kontrollenheten på 1.91%.

Vi finner at gjennomsnittseffekten for 1984 til 1987 er 1.3 prosentpoeng høyere enn for hele tidsperioden. Dette stemmer overens med forventningen om en positiv direkte effekt på sysselsettingen.

6.2.3 SCM: Placebo og statistisk inferens.

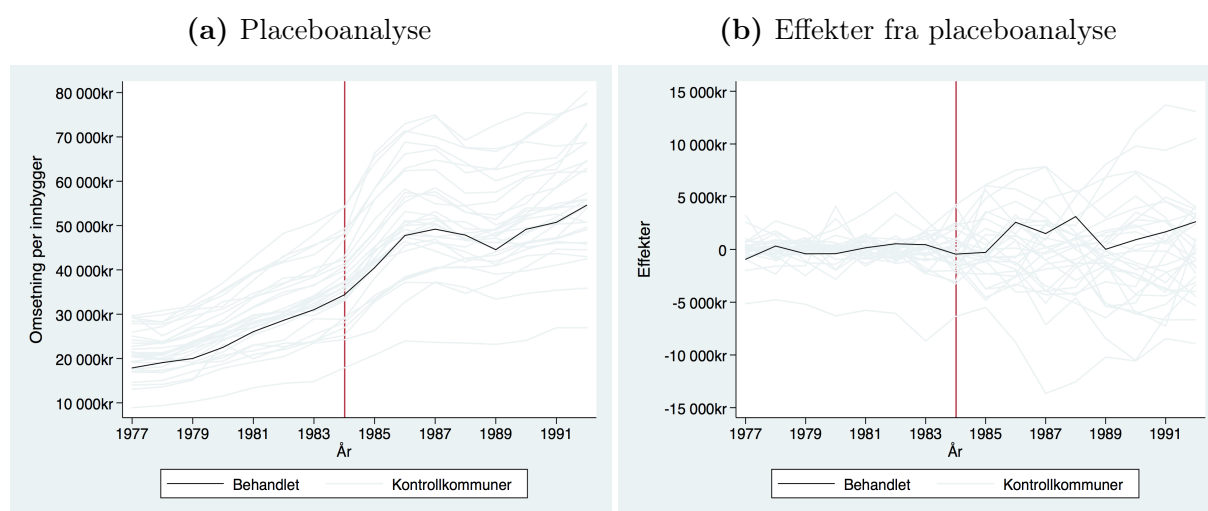
Hovedkritikken rettet mot den syntetiske kontrollmetoden er knyttet til mangelen på gode metoder for å undersøke om resultatene er statistisk signifikante. Dette kommer av at det kun er én eller få behandlede enheter, noe som fører til at de klassiske metodene for statistisk inferens utelukkes. Abadie *et al.* (2010, 2015) presenterer flere mulige metoder for statistisk inferens ved bruk av den syntetiske kontrollmetoden. Vi benytter oss av to av metodene de presenterer for å undersøke om resultatene våre er statistisk signifikante, placebo testing og RMSPE-rangering.

6.2.4 Placebotesting

Placebotesting gjennomføres ved å utføre SCM på alle kommunene i kontrollgruppe 1, som om de hadde fått en IKEA-etablering i 1984. Modellspesifikasjonen som benyttes er den samme som ble brukt for å undersøke effekten på omsetning og sysselsetting i Bergen. Ved å gjennomføre analysen for alle de ubehandlede kommunene får vi en *in place*-placebotest (Galiani, 2016). I fravær av eksterne sjokk eller behandling skal kommunene i kontrollgruppe 1 i teorien ikke oppleve noen store endringer i utfallsvariabelen etter 'behandlingstidspunktet'. Vi kan sammenligne effektene fra placebotesten med effekten fra den behandlede enheten for å undersøke om resultatet er signifikant eller ikke. Hvis effekten fra den behandlede enheten skiller seg ut fra placeboestene, er det stor sannsynlighet for at effekten vi observerer ikke er tilfeldig.

Metoden

Figur 7: In place-placebotest for effekten på omsetning.



Figur 7 viser resultatene fra en *in place*-placebotest med omsetning som utfallsvariabel. Figur 7a viser utviklingen i utfallsvariabelen for alle kommunene i kontrollgruppe 1. Den uthevede linjen tilhørende Bergen. Det er figur 7b som er av interesse. Den viser de estimerte effektene fra *in place*-placebotestene. Ved å undersøke om effekten vi finner for Bergen skiller seg ut fra placeboeffektene vi finner for de andre kommunene, kan vi få en indikasjon på om resultatet vi har fått er statistisk signifikant.

Galiani *et al.* (2016) presenterer en metode hvor det lages p-verdier for effekten i hver enkelt år etter behandling, basert på en placeboanalyse. Tolkningen av p-verdiene blir

andelen av kommuner i kontrollgruppen som har en minst like stor effekt som Bergen i den tidsperioden. Den estimerte effekten på utfallsvariabelen for den behandlede kommunen i periode t er $\hat{\alpha}_{1t}$, mens den estimerte effekten på utfallsvariabelen for kommunene i kontrollgruppe 1 i periode t er $\hat{\alpha}_{1t}^{PL}$. Den to-sidige p-verdien kan formuleres som:

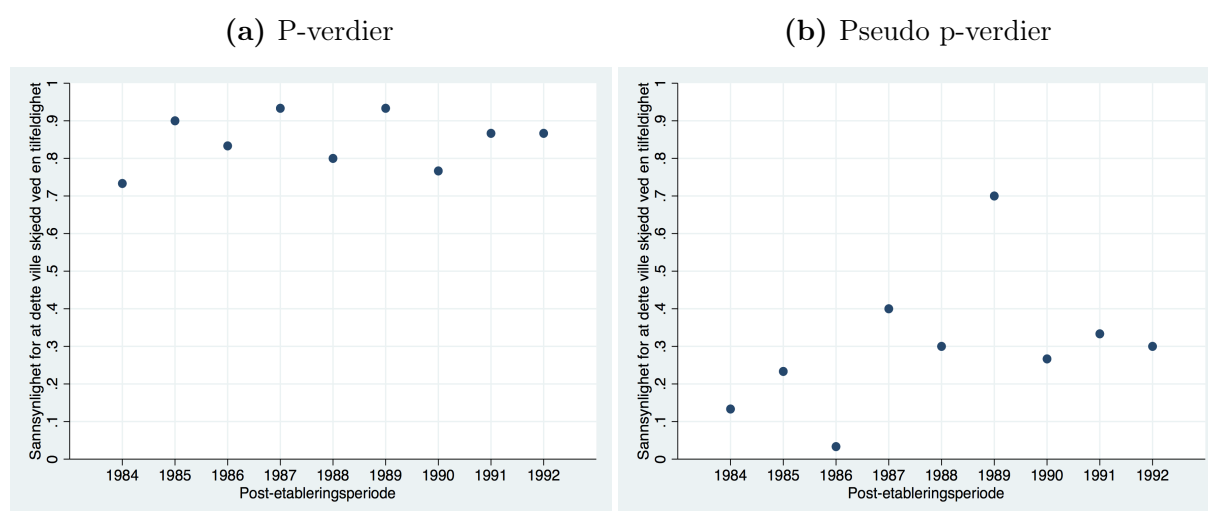
$$p - verdi = Pr(|\hat{\alpha}_{1t}^{PL}| \geq |\hat{\alpha}_{1t}|) \\ = \frac{\sum_{j \neq 1} 1(|\hat{\alpha}_{jt}| \geq |\hat{\alpha}_{1t}|)}{J}$$

Figur 8 viser de to-sidige p-verdiene for effekten på omsetning i Bergen. Figur 8a viser to-sidig p-verdi. Den kalkuleres ved å finne antallet kommuner fra kontrollgruppe 1 som har en effekt med absoluttverdi som er høyere enn effekten for Bergen. Antallet kommuner med høyere absoluttverdi deles deretter på totalt antall kommuner i kontrollgruppen. Dette gjøres for hvert år, t , etter behandlingen og gir sannsynligheten for at effekten vi observerer er tilfeldig. Et potensielt problem med p-verdiene i figur 8a er at de ikke tar hensyn til pre-RMSPE verdien. Det kan dermed kan det eksistere kommuner med en dårlig tilpasning som også opplever en stor effekt i perioden etter. En dårlig tilpasning vil kunne gi villedende resultater ved bruk av SCM. Kommuner med en slik tilpasning burde ikke tillegges vekt når p-verdiene skal kalkuleres. Ved å kalkulere pseudo p-verdier tar vi hensyn til pre-RMSPE. Disse er presentert i figur 8b. To-sidige pseudo p-verdier kalkuleres ved å først dividere absoluttverdien til effektene, $\hat{\alpha}_{1t}$ og $\hat{\alpha}_{1t}^{PL}$, med pre-RMSPE. Deretter rangeres de, og antallet kommuner med høyere verdi på pseudo-effekten divideres på det totale antallet kommuner i kontrollgruppe 1.²⁹

Placebotest: omsetning

Ingen av de to-sidige p-verdiene rapportert i figur 8a er statistisk signifikante til 5% signifikansnivå. Siden vi ønsker å ta hensyn til pre-RMSPE fokuserer vi på figur 8b. Her er effekten i 1988 den med lavest to-sidig pseudo p-verdi, men denne er kun statistisk signifikant til 10% signifikansnivå. Dette indikerer at effekten på omsetning ikke er statistisk signifikant og kan ikke utelukkes å være tilfeldig.

²⁹Pseudo-effekten er absoluttverdien til effekten dividert på pre-RMSPE.

Figur 8: To-sidige p-verdier for effekten på omsetning.**Placebotest: sysselsetting****Figur 9:** To-sidige p-verdier for effekten på sysselsetting.

Figur 9a og 9b viser henholdsvis de to-sidige p-verdiene og de to-sidige pseudo p-verdiene for effekten på sysselsettingen. Vi observerer at de to-sidige p-verdiene ikke er statistisk signifikante. De to-sidige pseudo p-verdiene er derimot vesentlig lavere. Dette indikerer at pre-RMSPE for Bergen er lav relativt til kommunene i kontrollgruppe 1 ved estimering av *in place*-placebotesten. Pseudo p-verdien for 1986 er på 0.333 og indikerer at effekten på sysselsettingen i Bergen for dette året er statistisk signifikant til 3.33% signifikansnivå. Ingen effekter for noen andre år er statistisk signifikante til 10% signifikansnivå for de to-sidige pseudo p-verdiene.

6.2.5 RMSPE-rangering

Abadie *et al.* (2010) foreslår å benytte forholdet mellom post-RMSPE og pre-RMSPE som en test for å finne sannsynligheten for at effekten, som observeres i utfallsvariablene til den behandlede enheten, er tilfeldig. Pre-RMSPE er et mål på avviket mellom den behandlede og den syntetiske enheten i perioden før behandling, og er forklart i delkapittel 3.3.2 om implementering. Post-RMSPE er på samme måte et mål på avviket mellom den syntetiske og den behandlede enheten, men for tiden etter behandlingen. En høy verdi på post-RMSPE vil i tilfellet for den behandlede enheten tyde på at det er en effekt av behandlingen.

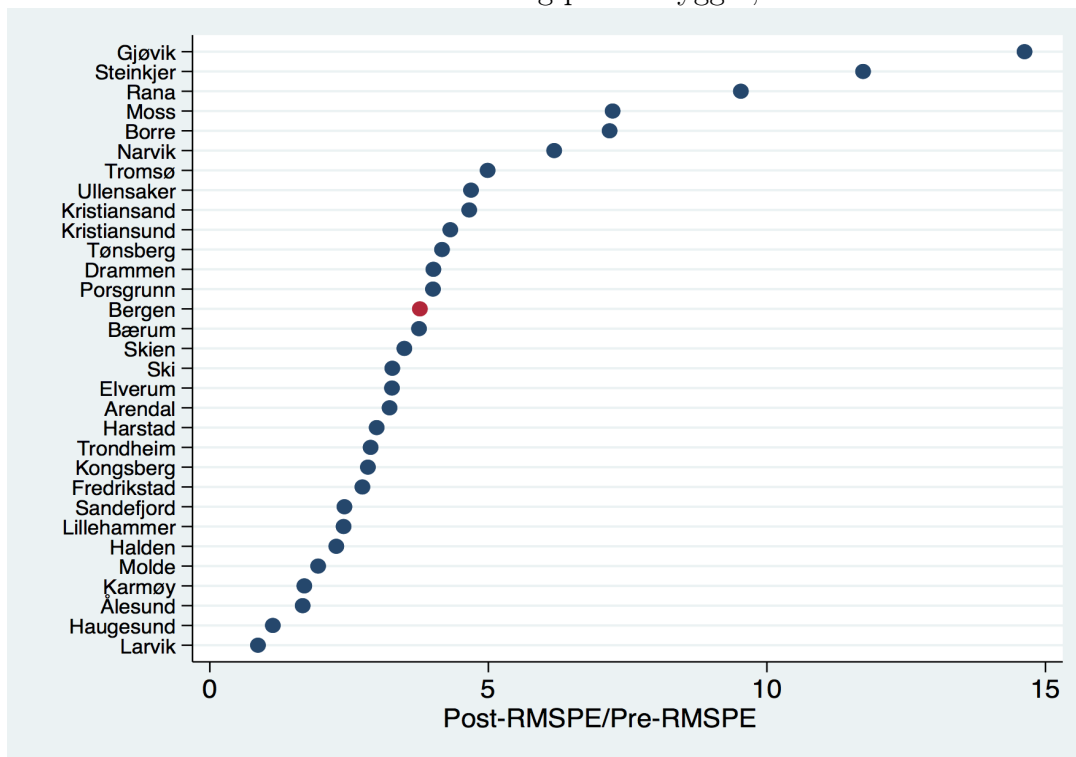
Ved å benytte resultatene fra *in place*-placebotesten kan vi lage en RMSPE-rangering med raten mellom post- og pre-RMSPE for hver kommune. Ved å dividere post-RMSPE på pre-RMSPE vil kommuner med høy pre-RMSPE få et lavere rate, og dermed en lavere rangering. RMSPE-rangeringen gir sannsynligheten for å finne en tilsvarende effekt som den vi observerer i utfallsvariabelen til den behandlede enheten, hvis behandlingen skulle vært gitt til en kommune i kontrollgruppe 1 tilfeldig.

RMSPE-rangering. Resultater.

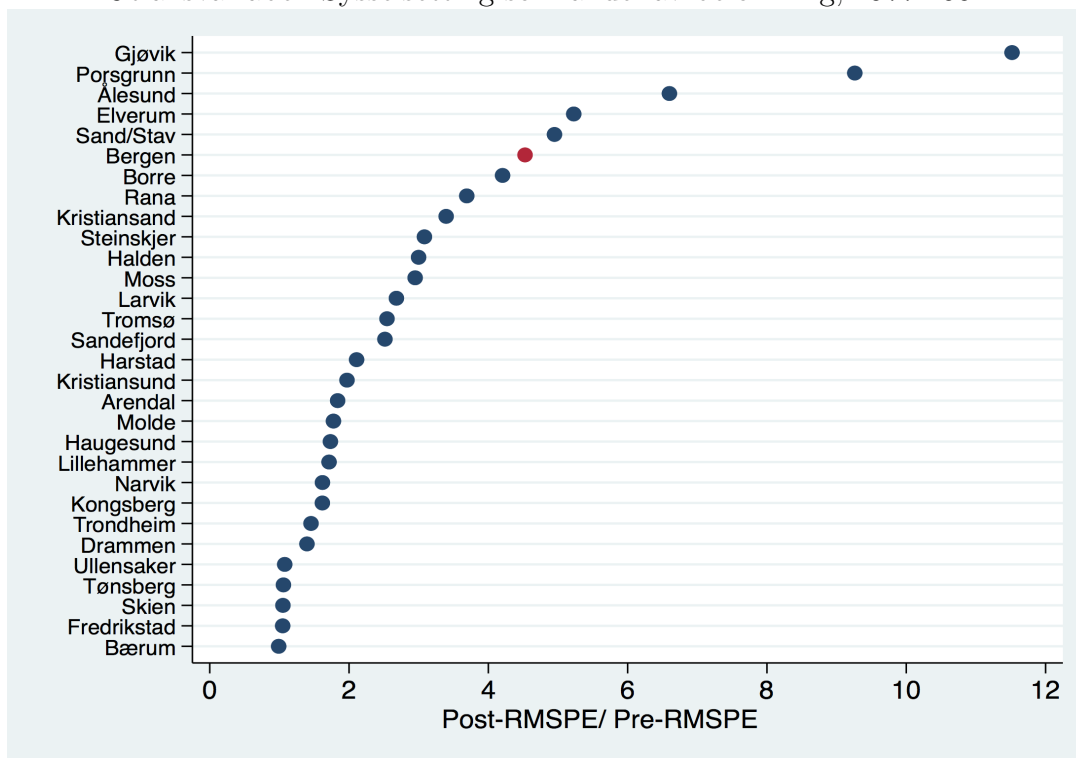
Figur 10 og 11 viser RMSPE-rangeringene med henholdsvis omsetning og sysselsetting som utfallsvariabel. Alle 30 kommuner, i tillegg til Bergen, er presentert med kulepunkter. Bergen er markert i rødt. For omsetningen har Bergen en RMSPE-rate på under 5 og er rangert som nummer 14 fra toppen. Det er dermed 13 kommuner i kontrollgruppen som har en høyere RMSPE-rate enn Bergen. Sannsynligheten for at effekten vi observerer er tilfeldig er dermed $13/31 = 0.419$. Som tilsvarer en sannsynlighet på 41.9%, og er ikke statistisk signifikant. For sysselsettingen er sannsynligheten for at effekten vi observerer er tilfeldig, 16.1%. Den dermed ikke statistisk signifikant. Resultatene fra RMSPE-rangeringen indikerer at effektene vi observerer for Bergen over hele tidsperioden etter etableringen, for begge utfallsvariabler, ikke er statistisk signifikante.

Figur 10: RMSPE-raten for Bergen og kontrollgruppe 1.

Utfallsvariabel: Omsetning per innbygger, 1977-1992.

**Figur 11:** RMSPE-raten for Bergen og kontrollgruppe 1.

Utfallsvariabel: Sysselsetting som andel av befolkning, 1977-1992.



6.2.6 Oppsummering

Resultatene fra den syntetiske kontrollmetoden indikerer en gjennomsnittlig positiv effekt på både omsetningen og sysselsettingen av IKEA-etableringen. Vi finner en økning i omsetningen på 1304kr. Dette tilsvarer en relativ økning på 2.89%. For sysselsettingen finner vi en gjennomsnittlig effekt på 0.02 prosentpoeng. Sammenlignet med den syntetiske kontrollenheten tilsvarer dette en økning på 0.6%.

På grunn av oljeprissjokket i årene 1986 til 1990 isolerte vi gjennomsnittlig effekt på både omsetning og sysselsetting for årene 1984 til 1987. For omsetningen finner vi en gjennomsnittlig effekt på 843 kroner. Dette tilsvarer en 2% økning. Vi finner en gjennomsnittlig effekt på 0.07 prosentpoeng, tilsvarende en 1.91% økning i sysselsettingen i Bergen sammenlignet med den syntetiske kontrollenheten. For 1988 indikerer den to-sidige pseudo p-verdien en statistisk signifikant effekt på omsetningen til 10% signifikansnivå. De to-sidige pseudo p-verdiene indikerer en statistisk signifikant effekt på sysselsettingen til 3.33% signifikansnivå i 1986.

RMSPE-rangeringen for omsetningen finner at det er 41.9% sannsynlighet for at effekten vi observerer i Bergen er tilfeldig, mens for sysselsettingen er sannsynligheten på 16.1%. Dette indikerer ingen statistisk signifikant effekt på omsetning eller sysselsettingen av IKEA-etableringen i Bergen.

7 Oppsummering og diskusjon

Målet med oppgaven var å undersøke effekten en IKEA-etablering har på omsetningen og sysselsettingen i varehandelssektoren. Tidligere studier av effektene, som assosieres med en etablering av IKEA i Sverige, ga en indikasjon på at vi kunne forvente å finne en økning i omsetningen i varehandelssektoren. For effekten på sysselsetting er resultatene fra den tidligere forskningen mer tvetydige, der Daunfeldt *et al.* (2014) finner en økning i sysselsettingen, mens Daunfeldt *et al.* (2015) ikke finner en statistisk signifikant effekt. Ettersom oppgaven vår bygger på et unikt datasett og vi har gjennomført flere ulike empiriske undersøkelser, vil ikke resultatene våre nødvendigvis være de samme som resultatene fra tidligere forskning.

Vi finner, i motsetning til Daunfeldt *et al.* (2014) og Daunfeldt *et al.* (2015), en negativ effekt av IKEA-etableringen på omsetningen ved bruk av differanse i differanse-metode. Dette er et noe overraskende resultat, siden det er det motsatte av hva vi forventet å finne som en effekt av etableringen.

Figur 2 fra kapittel 4 viser at det er en jevn vekst i den nasjonale utviklingen i omsetning, med en periode med kraftigere vekst fra 1984. Den tiltakende veksten i omsetningen sammenfaller med tidspunktet for etableringen av IKEA i Bergen. En nasjonal positiv utvikling i omsetning, i tillegg til resultatene fra tidligere forskning og forventningene om en positiv direkte effekt, gjør at vi stiller spørsmål ved troverdigheten til de negative estimatene fra DID-estimatoren. Vi fant at én av de estimerte effektene ved bruk av kontrollgruppe 2, var statistisk signifikant og assosiert med en relativ reduksjon i omsetningen på 11.3%. Det er lite troverdig at en så stor negativ, og samtidig statistisk signifikant, effekt på omsetningen er et riktig mål på den sanne effekten.

Resultatene fra estimering av DID-estimatoren for effekten på omsetning, peker i retning av at DID-metoden ikke nødvendigvis er den best egnede estimeringsmetoden med vårt datamateriale. Det er derimot for tidlig å trekke en slutning uten å ha vurdert troverdigheten til estimatene av effekten på sysselsetting.

Den estimerte etableringseffekten på sysselsetting er positiv og i tråd med det vi forventet å finne, både med tanke på retningen til effekten og basert på resultatene til Daunfeldt *et al.* (2014). Med unntak av én estimator, er alle de estimerte effektene av

IKEA-etableringen på sysselsetting statistisk signifikante. Dette står i kontrast til det Daunfeldt *et al.* (2015) finner, som er at det ikke er en statistisk signifikant effekt.

Resultatene fra placebotesten indikerer at det kan være brudd på forutsetningen om parallelle trender med begge kontrollgruppene, noe som kan tyde på at effektene feilestimeres. Det er derimot kun ett av resultatene for hver kontrollgruppe som er statistisk signifikant, noe som gjør det vanskelig å kunne konkludere med at det faktisk er et brudd eller ikke. Dersom det er tilfellet at det er et brudd på denne forutsetningen, kan dette være med på forklare den negative effekten på omsetning som vi fant ved estimering med DID-metoden.

Ved å inkludere demografiske kontrollvariabler, er det en risiko for at det er liten within-variasjon. Estimering med kommunefaste effekter vil derfor kunne gi upålitelige standardavvik og DID-estimatoren blir ineffisient. Dette kan være én av grunnene til at vi observerer en sterkt negativ og statistisk signifikant effekt på omsetningen ved estimering med kontrollgruppe 2. Statistisk inferens vil være vanskelig ved liten within-variasjon i kontrollvariablene og vi kan ikke trekke noen slutninger basert på resultatene.

Med bakgrunn i de estimerte effektene og de ulike potensielle problemene som er blitt drøftet, er det grunn til å tro at DID-metoden ikke er den best egnede estimeringsmetoden for å predikere den sanne effekten av en IKEA-etablering. Hovedargumentet bak denne antagelsen er mangelen på en tilfredsstillende kontrollgruppe for Bergen. Selv om placeboanalysene ikke indikerte et tydelig brudd på forutsetningen om parallelle trender, var det nok til å så tvil om hvorvidt DID-estimatoren klarer å si noe om den kausale effekten av IKEA-etableringen på varehandelssektoren i Bergen i 1984.

Den syntetiske kontrollmetoden løser eventuelle problemer knyttet til manglende tilfredsstillende kontrollgruppe, ved at den konstruerer en syntetisk versjon av behandlingsenheten. Dette gir oss muligheten til å kunne sammenligne Bergen med det vi antar ville vært Bergen ved fravær av behandling, representert ved den syntetiske kontrollenheten. Det er derfor rimelig å anta at resultatene fra SCM, i større grad enn DID-metoden, vil kunne si noe om den sanne effekten av etableringen.

Den positive gjennomsnittseffekten på omsetning over hele tidsperioden vi finner ved bruk av SCM, er i tråd med resultater fra tidligere forskning og med hva vi forventet skulle være retningen på effekten. Den gjennomsnittlige effekten gir en økning i omsetningen på 2.89% i forhold til den syntetiske kontrollenheten.

Vi forventet å se en positiv direkte effekt på omsetningen ved etablering av IKEA, men dette var ikke tilfellet. Resultatene indikerer derimot at den direkte effekten er negativ i de første årene etter etableringen. Dette er overraskende, men effektene er imidlertid ikke statistisk signifikante.

Den syntetiske kontrollmetoden finner en direkte økning i sysselsettingen i de første årene etter etableringen. Den positive direkte effekten er i henhold til det vi forventet med bakgrunn i det teoretiske rammeverket. Det samsvarer også med funnene gjort i Daunfeldt (2014). For hele tidsperioden etter etableringen finner vi en svak gjennomsnittlig økning i sysselsettingen på 0.6%. Dette resultatet er mer i tråd med artikkelen til Daunfeldt *et al.* (2015) som ikke finner en statistisk signifikant effekt på sysselsetting i kommuner som får en IKEA-etablering.

Placeboanalysen indikerer en statistisk signifikant effekt på sysselsettingen i 1986. Åpningstidsloven som trådte i kraft i 1985 kan være årsaken til at vi observerer en så kraftig effekt dette året. Bergen kommune hadde en relativt streng åpningstidsbestemmelse før i 1982. Hvis kommunen ikke endret åpningstidene i forkant av 1985, kan effekten vi observerer i 1986 skyldes de utvidede åpningstidene og ikke IKEA-etableringen. Dette underbygges av artikkelen til Bensnes (2018) som finner at det er en økning i antall sysselsatte i varehandelssektoren mellom 16 til 19 år, for kommuner som blir påvirket av Åpningstidsloven. Dette gjør at vi stiller spørsmål ved hvorvidt effekten på sysselsettingen som vi observerer i 1986, kan knyttes til etableringen av IKEA.

Omsetningen kan også ha blitt påvirket av Åpningstidsloven. De utvidede åpningstidene til bedriftene i varehandelssektoren kan ha gjort det mulig for å flere å handle utenfor normal arbeidstid på hverdager og lørdager. Dette kan ha ført til en økning i omsetningen i sektoren.

Placeboanalysen og RMSPE-rangeringen indikerer at resultatet for omsetningen og sysselsettingen over hele tidsperioden etter etableringen ikke er statistisk signifikant. Rudholm *et al.* (2018) finner at relativ størrelse på kommunen, med hensyn på antall innbyggere og størrelsen på varehandelssektor før etablering, henger sammen med en eventuell økning i produktivitet. Små kommuner som får et IKEA-varehus opplever en kraftigere effekt av etableringen enn større kommuner. Selv om Rudholm *et al.* (2018) undersøker effekten på produktivitet og dermed ikke er direkte sammenlignbar med vår problemstilling, kan vi benytte denne informasjonen. Bergen er den største kommunen i

vårt datasett med tanke på antall innbyggere, omsetning og sysselsetting. Det er dermed ikke urimelig at effektene av en IKEA-etablering ikke vil være like kraftige i Bergen som i en mindre kommune.

Fra 1987 til 1990 observerer vi en nasjonal nedgangsperiode i sysselsettingen og omsetningen i varehandelssektoren. Dette er forventet med tanke på oljeprissjokket i samme tidsperiode. Det er mulig at indirekte effekter, knyttet til stordriftfordeler og høyere produktivitet hos IKEA, forsterker en allerede vanskelig periode for mange aktører i varehandelssektoren. Bedrifter som opplever et fall i omsetningen som følge av oljeprissjokket, vil kunne bli nødt til å legge ned på grunn av den mulige kombinerte negative effekten. En eventuell motsatt effekt er også mulig, der positive indirekte effekter fra bedriftsklynger og økt produktivitet hos andre aktører, som følge av etableringen, motvirker noe av nedgangen vi kan forvente å observere som følge av oljeprissjokket.

For å oppsummere resultatene fra analysene utført med de to ulike estimeringsmetodene, finner vi en motstridende effekt på omsetning. Der DID-metoden indikerer en negativ effekt på omsetning som følge av etableringen, mens SCM finner en positiv effekt. Siden metodene estimerer effekter med ulike fortegn, vil én av de nødvendigvis angi feil retning på effekten. Problemer med statistisk inferens i begge metodene gjør at det er vanskelig å konkludere med hvilken metode som har feilestimert retningen på effekten.

Begge metodene finner en positiv effekt på sysselsettingen som følge av IKEA-etableringen. Dette gjør at vi anser det som troverdig at den sanne effekten av IKEA-etableringen på sysselsettingen er positiv.

7.1 Mulige utvidelser

Gjennom arbeidet med denne oppgaven har vi kommet over flere interessante mulige utvidelser. Daunfeldt *et al.* (2014) undersøker hvordan omsetningen og sysselsettingen i nærliggende kommuner påvirkes av en IKEA-etablering. De finner en svak negativ effekt i nærliggende kommuner og ingen statistisk signifikant effekt på sysselsettingen. Det er ikke sikkert at en studie som undersøker effekten av en IKEA-etablering i Norge vil få de samme resultatene. Det kan derfor være interessant å undersøke dette nærmere for IKEA-etableringer i Norge.

Det er også mulig at en IKEA-etablering kan påvirke omsetningen og sysselsettingen i større økonomiske regioner enn bare i kommuner som får en IKEA-etablering. Ved å slå sammen kommuner til økonomiske regioner er det mulig å undersøke mulige effekter av etableringen utenfor spesifikke kommuner.

Bensnes (2018) benytter endringen i Åpningstidsloven for å undersøke effekten av økt etterspørsel i varehandelssektoren på andelen som fullfører videregående utdanning. Dette er også en mulig utvidelse hvor det kan undersøkes om en IKEA-etablering påvirker andelen som fullfører videregående utdanning.

8 Konklusjon

Formålet med denne oppgaven var å undersøke effekten av en IKEA-etablering på varehandelssektoren. Nærmere bestemt så vi på hvordan effekten av IKEA Åsane, som ble etablert i Bergen i 1984, påvirket omsetning i varehandelssektoren per innbygger og sysselsetting i varehandelssektoren som andel av befolkning. For å undersøke disse effektene benyttet vi to ulike empiriske estimeringsmetoder på et paneldatasett for tidsperioden 1977 til 1992.

Metodene vi har benyttet i denne oppgaven er differanse i differanse-metoden og den syntetiske kontrollmetoden. Vi får tvetydige resultater for effektene ved bruk av de to metodene. Resultatene fra DID-metoden indikerer en negativ effekt på omsetningen, mens SCM finner en positiv effekt assosiert med etableringen. Begge metodene finner en positiv effekt av IKEA-etableringen på sysselsettingen i Bergen.

De motstridende effektene på omsetningen ved bruk av de ulike estimeringsmetodene, indikerer at én av de har feilestimert retningen på den sanne effekten av IKEA-etableringen. Vi har ikke det empiriske grunnlaget som er nødvendig for å kunne konkludere med hvilken av de to metodene som feilestimerer retningen på effekten.

Siden begge metodene finner en positiv effekt på sysselsettingen assosiert med IKEA-etableringen, gir dette en god indikasjon på at dette kan være den sanne retningen på effekten.

Eksterne hendelser i form av oljeprissjokket og den nye åpningstidsloven gjør det utfordrende å skille ut en troverdig effekt, siden disse sammenfaller med tidsperioden etter IKEA-etableringen i Bergen.

Utfordringer med statistisk inferens i begge metodene gjør det vanskelig å konkludere med hvorvidt effektene vi observerer er troverdige estimater for de sanne effektene som følge av IKEA-etableringen i Bergen.

Referanser

- Abadie, Alberto. Diamond, A. H. J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.
- Abadie, Alberto. Diamond, A. H. J. (2011). Synth: An r package for synthetic control methods in comparative case studies. *Journal of Statistical Software*, 42(13).
- Abadie, Alberto. Diamond, A. H. J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.
- Abadie, Alberto. Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *The American Economic Review*, 93(1):113–132.
- Andersen, Svein S. Austvik, O. G. (2000). Nasjonal handlefrihet - nye internasjonale rammebetingelser. petroleum, makt og demokrati. *Makt- og demokratiutredningens rapportserie*, Rapport 21.
- Athey, Susan. Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32.
- Basker, E. (2005b). Selling a cheaper mousetrap: Wal-mart's effect on retail prices. *Journal of Urban Economics*, 58(2):203–229.
- Beck, Nathaniel % Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review*, 89(3):634–647.
- Beck, N. og Katz, J. N. (2001). Throwing out the baby with the bath water: A comment on green, kim, and yoon. *International Organization*, 55(2):487–495.
- Bensnes, Simon Søbstad. Strøm, B. (2018). Earning or learning? how extending closing time in the retail sector affects youth employment and education. *Oxford Bulletin of Economics and Statistics*, 81(2):299–327.
- Bilgel, Firat. Galle, B. (2015). Financial incentives for kidney donation: A comparative case study using synthetic controls. *Journal of Health Economics*, 43:103–117.
- Billmeier, Andreas. Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *The Review of Economics and Statistics*, 95(3):983–1001.
- Bohn, Sarah. Lofstrom, M. R. S. (2014). Did the 2007 legal arizona workers act reduce the state's unauthorized immigrant population? *The Review of Economics and Statistics*, 96(2):258–296.
- Botosaru, Irene. Ferman, B. (2019). On the role of covariates in the synthetic control method. *The Econometrics Journal*, <https://doi.org/10.1093/ectj/utz001>.
- Brown, S. (1993). Retail location theory: evolution and evaluation. *The International Review of Retail, Distribution and Consumer Research*, 3(2):185–229.
- Carling, Kenneth. li, Y. (2018). The power of the synthetic control method. *Working papers in transport, tourism, information technology and microdata analysis*.

- Cavallo, Eduardo. Galiani, S. N. I. P. J. (2012). Catastrophic natural disasters and economic growth. *The Review of Economics and Statistics*, 95(5):1549–1561.
- Columbia University, M. S. o. P. H. *Difference-in-Difference Estimation*. Hentet 11.05.2019 fra, <https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation>.
- Daunfeldt, Sven-Olov; Mihaescu, O. N. H. R. N. (2014). What happens when ikea comes to town? *Regional Studies*, 51(2):313–323.
- Daunfeldt, Sven-Olov. Mihaescu, O. N. H. R. N. (2015). When ikea enters: Do local retailers win or lose? *HUI Research Working Paper Nr. 109*.
- Ekberg, E. (5. april 2018). *IKEA*. I Store norske leksikon. Hentet 07.05.2019 fra, <https://snl.no/IKEA>.
- Ferman, Bruno. Pinto, C. (2018). Inference in differences-in-differences with few treated groups and heteroskedasticity. *Review of Economics and Statistics*, (2):1–16.
- Ferman, Bruno. Pinto, C. P. V. (2016). Cherry picking with synthetic controls. *Working paper*, <https://ideas.repec.org/p/fgv/eesptd/420.html>.
- Frank, R. H. (2003). *Microeconomics and behavior (5. utg.)*. Boston: McGraw-Hill.
- Galiani, Sebastian. Quistorff, B. (2016). The synth_runner package: Utilities to automate synthetic control estimation using synth. *The Stata Journal*, 17(4):834–849.
- Han, Mengjie. Mihaescu, O. L. Y. R. N. (2018). Comparison and one-stop shopping after big-box retail entry: A spatial difference-in-difference analysis. *Journal of Retailing and Consumer Services*, 40:175–187.
- Hinrichs, P. (2012). The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. *The Review of Economics and Statistics*, 94(3):712–722.
- Idsø, J. (9.juni 2014). *stordriftsfordeler*. I Store norske leksikon. Hentet 30.04.2019 fra, <https://snl.no/stordriftsfordeler>.
- IKEA (2017a). *IKEA® 2017 by numbers*. 2019-04-22, <https://highlights.ikea.com/2017/facts-and-figures/>.
- IKEA (2017b). *Yearly Summary FY17*. 2019-04-23, https://www.ikea.com/ms/en_AU/media/pdf/global/FY17_Yearly_Summary.pdf.
- IKEA (2019). *IKEA history - how it all began*. 2019-04-23, https://www.ikea.com/ms/en_AU/about_ikea/the_ikea_way/history/.
- Kaul, Ashok. Klobner, S. P. G. S. M. (2018). Synthetic control methods: Never use all pre-intervention outcomes together with covariates. *MPRA Paper No. 83790, posted 12 January 2018*, (<https://mpra.ub.uni-muenchen.de/83790/>).
- Kreif, Noémi. Grieve, R. H. D. T. A. J. N. S. S. M. (2016). Examination of the synthetic control method for evaluating health policies with multiple treated units. *Health Economics*, 25:1514–1528.

- Liu, S. (2015). Spillovers from universities: Evidence from the land-grant program. *Journal of Urban Economics*, 87(0):25–41.
- Meyer, B. D. (April 1995). Natural and quasi-experiments in economics. *Journal of Business Economic Statistics*, 13(2):151–161.
- Nannicini, Tommaso. Billmeier, A. (2011). Economies in transition: How important is trade openness for growth? *Oxford Bulletin of Economics and Statistics*, 73(3):287–314.
- O'Neill, Stephen. Kreif, N. G. R. S. M. S. J. S. (2016). Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Services and Outcomes Research Methodology*, 16(0):1–21.
- Pedersen, P. E. (2014). Når ikea kommer til by'n- en gjennomgang av forskningslitteraturen om big boxetableringer. *Skriftserien fra Høgskolen i Buskerud og Vestfold*, (1).
- Plümper, Thomas og Troeger, V. E. (2007). Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects. *Political Analysis*, 15(2):124–139.
- Rudholm, Niklas. Li, Y. C. K. (2018). How does big-box entry affect labor productivity in durable goods retailing? a synthetic control approach. *HUI Working Paper Series 130*.
- Stearns, J. (2015). The effects of paid maternity leave: Evidence from temporary disability insurance. *Journal of Health Economics*, 43(0):85–102.
- Svartdal, F. (7. juni 2018). *randomisert kontrollstudie*. I Store norske leksikon. Hentet 20.05.2019 fra, https://snl.no/randomisert_kontrollstudie.
- Thorsnæs, G. (20. mars 2019). *Sandnes*. I Store norske leksikon. Hentet 05.05.2019 fra, <https://snl.no/Sandnes>.
- Verbeek, M. (2017). *A guide to modern econometrics (5. utg.)*. John Wiley & Sons, Inc..
- Wooldridge, J. (2015). *Introductory econometrics: A modern approach (6. utg.)*. Boston: South-Western Cengage Learning.

Appendiks

A Restleddsegenskaper ved bruk av paneldata

Restleddene i en modell med paneldata må ha følgende egenskaper for at estimering ved OLS skal kunne gi forventningsrette og konsistente estimasjoner:

$$\text{i) } E(\epsilon_{it}|x_{it}, z_{i1}) = 0$$

$$\text{ii) } E(\epsilon_{it}\epsilon_{js}) = \begin{cases} \sigma_\epsilon^2 & \text{i=j og t=s} \\ 0 & \text{ellers} \end{cases}$$

$$\text{iii) } E(\eta_i\eta_j) = \begin{cases} \sigma_\eta^2 & \text{i=j} \\ 0 & \text{ellers} \end{cases}$$

$$\text{iv) } E(\epsilon_{it}\eta_j) = 0 \text{ for alle } i, j \text{ og } t$$

$$\text{v) } E(\eta_i|x_{it}, z_{i1}) = 0$$

Restleddsegenskapene forklart med ord:

- i) Forklaringsvariablene er ukorrelert med med den idiosynkratiske restleddskomponenten.
- ii) Det idiosynkratiske restleddet er ukorrelert innenfor, *within*, og på tvers, *between*, av enheter, samt at det har konstant varians.
- iii) Det enhetsspesifikke restleddet er ukorrelert på tvers av enheter og har konstant varians.
- iv) Det idiosynkratiske og det enhetsspesifikke restleddet er ukorrelerte.
- v) Forklaringsvariablene er ukorrelert med den enhetsspesifikke restleddskomponenten.

B Utvidelse av det økonometriske rammeverket til SCM

Vi ser i det følgende bort i fra kontrollvariablene.

Den estimerte behandlingseffekten, $\hat{\alpha}_t = y_{0t} - \sum_{j=1}^J w_j y_{jt}$, vil nå være en stokastisk variabel siden den er en funksjon av de ukjente variablene, y_{jt} , som er betinget ved faktormodellen gitt ved ligning (4). De estimerte vektene bestemmes av data i pre-intervensjonsperioden, tiden før behandlingen inntreffer, og er dermed stokastisk urelatert til $\hat{\alpha}_t$. Vektoren for vektene til de ulike kontrollenhetene, W , oppnås ved å minimere avviket i utfallsvariabelen mellom den behandlede enheten og den kontrollenhetene i pre-intervensjonsperioden. Imidlertid blir et slikt avvik, i tilstedeværelsen av kontrollvariabler, målt i form av utfallsvariabelen og kontrollvariablene simultant. For å kunne skille mellom dette formuleres to separate objektfunksjoner, f_V og f_W for henholdsvis utfallsvariabelen og kontrollvariablene. Vi tar for oss objektfunksjonen til utfallsvariabelen først:

$$f_V = \|y_0 - Y_1 w\|^2 = (y_0 - Y_1 w)'(y_0 - Y_1 w)$$

der $y_0((T_0 - 1) \times 1)$ er utfallsvariabelen til den behandlede enheten i pre-intervensjonsperioden, $Y_1((T_0 - 1) \times J)$ er utfallsvariabelen til de J kontrollenhetene i pre-intervensjonsperioden. Mer detaljert har vi:

$$y_0 = \begin{pmatrix} y_{0,1} \\ y_{0,2} \\ \vdots \\ y_{0,T_0-1} \end{pmatrix} \quad \text{og} \quad Y_1 = \begin{pmatrix} y_{11} & y_{21} & \dots & y_{J1} \\ y_{12} & y_{22} & \dots & y_{J2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1,T_0-1} & y_{2,T_0-1} & \dots & y_{J,T_0-1} \end{pmatrix}$$

Objektfunksjonen til kontrollvariablene er gitt ved:

$$f_W = \|z_0 - V Z_1 w\|^2 = (z_0 - Z_1 w)'V(z_0 - Z_1 w)$$

der $z_0(r \times 1)$ er de r kontrollvariablene til den behandlede enheten i pre-intervensjonsperioden antatt tidskonstant, $Z_1(r \times J)$ er r kontrollvariablene til kontrollenhetene i pre-intervensjonsperioden.³⁰ Mer detaljert:

³⁰Ved bruk av metoden i STATA blir alle kontrollvariabler omgjort til gjennomsnitt i pre-intervensjonsperioden

$$z_0 = \begin{pmatrix} z_{01} \\ z_{02} \\ \vdots \\ z_{0r} \end{pmatrix} \quad \text{og} \quad Z_1 = \begin{pmatrix} z_{11} & z_{21} & \dots & z_{J1} \\ z_{12} & z_{22} & \dots & z_{J2} \\ \vdots & \vdots & \ddots & \vdots \\ z_{1r} & z_{2r} & \dots & z_{Jr} \end{pmatrix}$$

Elementet z_{jk} er målt ved en lineær funksjon av kontrollvariabelen $x_{jk,t}$ gjennom pre-intervensjonsperioden: $z_{jk} = g(x_{jk,t})$ der $x_{jk,t}$ er den j-ende enhetens k-ende kontrollvariabel på tidspunkt t og $g(x_{jk,t}) = \sum_{t=1}^{T_0-1} \iota_t x_{jk,t}$ med restriksjoner $\iota_t \in [0, 1]$ og $\sum_{t=1}^{T_0-1} \iota_t = 1$.

Vi har V som er en $(r \times r)$ symmetrisk og positiv semidefinit matrise som tilegner vekt til alle kontrollvariablene.

$$V = \begin{pmatrix} v_1 & & & \\ & v_2 & & \\ & & \ddots & \\ & & & v_r \end{pmatrix}$$

Vi skal nå skissere optimeringsprosedyren som er en nøstet loop av de to objektfunksjonene implementert i følgende 3 steg :

- [1] En initialiserende matrise for V settes inn i f_W for å optimeres, for deretter å oppnå en lokal optimumsverdi for f_W og w .
- [2] Setter deretter den oppdaterte w inn i f_V for å oppnå en verdi for f_V .
- [3] Itererer deretter steg [1] og [2] til f_V oppnår sitt minimum.

Matrisen V initialiseres som $(Z_1' Z_1)^{-1} Z_1' z_0$ hvis matrisen $Z_1' Z_1$ er reversibel, ellers bruker man $(r \times r)$ identitetsmatrisen skalert ved $\frac{1}{r}$.

C Utelatte kommuner fra datamaterialet grunnet manglende observasjoner.

Tabell 12: Kommuner med manglende observasjoner som dermed er utelatt fra datamaterialet.

Kommunenavn	
Askvoll	Modalen
Birkenes	Mosvik
Bodø	Måsøy
Eid	Nordkapp
Eidskog	Nordreisa
Fedje	Oslo
Fjaler	Rakkestad
Flekkefjord	Ringerike
Fosnes	Ringsaker
Froland	Rømskog
Gaular	Røyrvik
Gloppen	Sarpsborg
Granvin	Sirdal
Hamar	Skedsmo
Hole	Skjervøy
Kongsvinger	Songdalen
Kvitsøy	Suldal
Kåfjord	Sørfold
Leikanger	Utsira
Lyngen	Vennesla
Lærdal	Vik
Lørenskog	Vindafjord

D Kommunebestemte åpningstidsreguleringer i 1982

Tabell 13: Kommunebestemte stengetider på hverdager for kommunene i kontrollgruppe 1, i tillegg til Bergen, Sandnes og Stavanger, før Åpningstidsloven trådte i kraft i 1985.

Kode	Kommuner	Antall
Kode 1 (Kl. 17.00)	Bergen, Drammen, Elverum, Harstad, Kongsberg, Kristiansand, Skien.	7
Kode 2 (Kl. 18.00)	Fredrikstad, Gjøvik, Halden, Moss, Narvik, Porsgrunn, Rana, Trondheim, Ullensaker, Ålesund.	10
Kode 3 (Kl. 19.00)	Haugesund, Karmøy, Tromsø.	3
Kode 4 (Kl. 20.00)	Bærum, Molde, Sandefjord, Ski.	4
Kode 5 (Kl. 21.00)		0
Kode 6 (Kl. 22.00)	Steinkjer, Stavanger.	2
Kode 7 (Kl. 23.00)	Kristiansand, Lillehammer, Sandnes.	3
Mangler info	Arendal, Borre, Larvik, Tønsberg.	4

Tabell 14: Kommunebestemte stengetider på lørdager for kommunene i kontrollgruppe 1, Bergen, Sandnes og Stavanger, før åpningstidsreguleringen trådte i kraft i 1985.

Kode	Kommuner	Antall
Kode 1 (Ikke senere enn kl. 13.00)	Drammen, Elverum, Harstad, Kongsberg, Skien, Tromsø.	6
Kode 2 (Ikke senere enn kl. 14.00)		0
Kode 3 (Ikke senere enn kl. 15.00)	Bergen, Fredrikstad, Gjøvik, Halden, Kristiansand, Moss, Narvik, Trondheim.	8
Kode 4 (Ikke senere enn kl. 16.00)	Haugesund, Karmøy, Porsgrunn, Ullensaker, Ålesund.	5
Kode 5 (Ikke senere enn kl. 17.00)	Bærum, Rana.	2
Kode 6 (Ikke senere enn kl. 18.00)	Ski	1
Kode 7 (Ikke senere enn kl. 19.00)		0
Kode 8 (Ikke senere enn kl. 20.00)	Kristiansund, Lillehammer, Molde, Sandefjord, Stavanger, Steinkjer.	6
Kode 9 (Manglende datakode)	Arendal, Borre, Larvik, Tønsberg, Sandnes.	5

E Kommuner i kontrollgruppe 1

Tabell 15: Kommunene som inngår i kontrollgruppe 1

Kommunenavn	
Arendal	Lillehammer
Borre	Molde
Bærum	Moss
Drammen	Narvik
Elverum	Porsgrunn
Fredrikstad	Rana
Gjøvik	Sandefjord
Halden	Ski
Harstad	Skien
Haugesund	Steinskjer
Karmøy	Tromsø
Kongsberg	Trondheim
Kristiansand	Tønsberg
Kristiansund	Ullensaker
Larvik	Ålesund

F Differanse i differanse-metode.

Tabell 16: Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 1.

Utfallsvariabel: Omsetning per innbygger, 1977 til 1992.			
	(1)	(2)	(3)
IKEA84	-146.6 (949.9)	-108.1 (1056.0)	-1289.1 (1287.4)
aldersgruppe 2		-993.0 (944.0)	-189.3 (985.0)
aldersgruppe 3		-2159.3* (1094.1)	-1654.3 (1241.2)
aldersgruppe 4		-1834.0* (947.5)	-1966.3* (1076.1)
aldersgruppe 5		-795.7 (917.3)	-195.7 (963.2)
vgs			-108.4 (546.6)
kort universitetsutdanning			2694.6 (1692.7)
lang universitetsutdanning			-2096.2 (3110.7)
_cons	21320.2*** (757.0)	124839.4** (54428.6)	86004.0* (49884.0)
Årsdummier	JA	JA	JA
Kommunefaste effekter	JA	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
N	496	496	403
R^2	0.942	0.946	0.939

Merk: Robuste standardavvik i parentes. Kontrollvariablene er vektet med befolkning.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tabell 17: Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 1.

Utfallsvariabel: Sysselsetting som andel av befolkning, 1977 til 1992.			
	(1)	(2)	(3)
IKEA84	0.0701 (0.0433)	0.135*** (0.0328)	0.127*** (0.0376)
aldersgruppe 2		0.0345 (0.0400)	0.0273 (0.0471)
aldersgruppe 3		-0.0599 (0.0505)	-0.101* (0.0571)
aldersgruppe 4		0.000862 (0.0414)	-0.0582 (0.0473)
aldersgruppe 5		0.141*** (0.0393)	0.137** (0.0524)
vgs			-0.0261 (0.0339)
kort universitetsutdanning			-0.0626 (0.0766)
lang universitetsutdanning			0.0120 (0.141)
_cons	4.122*** (0.0347)	2.488 (2.099)	5.608** (2.225)
Årsdummier	JA	JA	JA
Kommunefaste effekter	JA	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
N	496	496	403
R^2	0.528	0.606	0.616

Merk: Robuste standardavvik i parentes. Kontrollvariablene er vektet med befolkning.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tabell 18: Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på omsetningen i varehandelssektoren per innbygger. Kontrollgruppe 2.

Utfallsvariabel: Omsetning per innbygger, 1977 til 1992.			
	(1)	(2)	(3)
IKEA84	114.5 (856.6)	-937.6 (850.7)	-3004.2** (320.5)
aldersgruppe 2		1225.6* (405.8)	-6460.1 (2736.6)
aldersgruppe 3		958.8* (260.2)	-8085.0* (2302.7)
aldersgruppe 4		285.4 (588.6)	-8363.8*** (693.9)
aldersgruppe 5		-1292.3 (840.8)	-5113.5 (2730.2)
vgs			-2840.5 (1057.1)
kort universitetsutdanning			(3896.9)
lang universitetsutdanning			13696.9** (3081.8)
_cons	20146.7*** (324.1)	-15214.5 (14291.0)	593899.0* (145620.5)
Årsdummier	JA	JA	JA
Kommunefaste effekter	JA	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
N	48	48	39
R^2	0.996	0.996	0.996

Merk: Robuste standardavvik i parentes. Kontrollvariablene er vektet med befolkning.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tabell 19: Fullstendige resultater fra estimeringen av effekten av IKEA-etableringen på sysselsetting i varehandelssektoren som andel av befolkning. Kontrollgruppe 2.

Utfallsvariabel: Sysselsetting som andel av befolkning, 1977 til 1992.			
	(1)	(2)	(3)
IKEA84	0.0819** (0.0166)	0.181* (0.0546)	0.138** (0.0167)
aldersgruppe 2		-0.130 (0.0767)	-0.531 (0.288)
aldersgruppe 3		0.0388* (0.0114)	-0.386 (0.164)
aldersgruppe 4		0.0859 (0.0853)	-0.257 (0.250)
aldersgruppe 5		0.140 (0.109)	-0.193 (0.0785)
vgs			-0.124 (0.0910)
kort universitetsutdanning			-0.315* (0.0794)
lang universitetsutdanning			0.701 (0.302)
_cons	3.921*** (0.0116)	3.210 (2.551)	35.50 (14.59)
Årsdummier	JA	JA	JA
Faste kommuneeffekter	JA	JA	JA
Estimeringsperiode	1977-1992	1977-1992	1980-1992
N	48	48	39
R^2	0.928	0.933	0.945

Merk: Robuste standardavvik i parentes. Kontrollvariablene er vektet med befolkning.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tabell 20: Fullstendige resultater fra estimeringen av effekten av den falske etableringen på omsetning i varehandelssektoren per innbygger. Kontrollgruppe 1 og 2.

Utfallsvariabel: Omsetning per innbygger, 1977 til 1983.

	Kontrollgruppe 1		Kontrollgruppe 2	
	(1)	(2)	(3)	(4)
	sys	oms	sys	oms
IKEA80	-4.405 (434.8)	-549.1 (523.9)	-136.2** (22.74)	-638.9 (586.4)
aldersgruppe 2		-1470.9 (888.8)		337.9 (1950.8)
aldersgruppe 3		-2419.9** (994.0)		3140.6** (611.3)
aldersgruppe 4		-653.5 (926.6)		6665.1*** (504.7)
aldersgruppe 5		649.4 (1200.7)		2904.3* (887.4)
_cons	21320.2*** (323.0)	100816.7* (51374.3)	20146.7*** (239.7)	-193036.8* (52433.3)
Årsdummier	JA	JA	JA	JA
Faste kommuneeffekter	JA	JA	JA	JA
Estimeringsperiode	1977-1983	1977-1983	1977-1983	1977-1983
N	217	217	21	21
R^2	0.898	0.912	0.995	0.998

Merk: Robuste standardavvik i parentes. Kontrollvariablene er vektet med befolkning.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Tabell 21: Fullstendige resultater fra estimeringen av effekten av den falske etableringen på sysselsettingen i varehandelssektoren som andel av befolkning. Kontrollgruppe 1 og 2.

Utfallsvariabel: sysselsetting som andel av befolkning, 1977 til 1983.

	Kontrollgruppe 1		Kontrollgruppe 2	
	(1)	(2)	(3)	(4)
IKEA80	0.0608** (0.0277)	0.0370 (0.0361)	0.0480 (0.0277)	-0.0281 (0.0544)
aldersgruppe 2		-0.00634 (0.0375)		0.0555 (0.213)
aldersgruppe 3		0.121* (0.0664)		0.369 (0.128)
aldersgruppe 4		0.200*** (0.0398)		0.620* (0.191)
aldersgruppe 5		0.154*** (0.0542)		-0.0807 (0.0723)
_cons	4.122*** (0.0216)	-3.385 (2.559)	3.921*** (0.00853)	-13.44 (10.73)
Årsdummier	JA	JA	JA	JA
Kommunefaste effekter	JA	JA	JA	JA
Estimeringsperiode	1977-1983	1977-1983	1977-1983	1977-1983
N	217	217	21	21
R^2	0.434	0.547	0.826	0.925

Merk: Robuste standardavvik i parentes. Kontrollvariablene er vektet med befolkning.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

G Den syntetiske kontrollmetoden

Vekting av kontrollenheter, kontrollvariabler og lag ved bruk av SCM.

Tabell 22: Vekting av kommuner og modellspesifikasjon i den syntetiske kontrollgruppen til Bergen. Omsetning per innbygger.

(a) Kommuner		(b) Variabler*	
Kommune	Vekt	Variabelnavn	Vekt
Bærum	0.054	omsetning(1977)	0.542
Karmøy	0.16	omsetning(1980)	0.106
Trondheim	0.785	befolkning	0.046
		grunnskole	0.023
		vgs	0.011
		kort universitetsutdanning	0.132
		lang universitetsutdanning	0.127
		aldersgruppe 1	0.004
		aldersgruppe 3	0.01

*Alle variabler er som andel av befolkning.

Tabell 23: Vekting av kommuner og modellspesifikasjon i den syntetiske kontrollgruppen til Bergen. Sysselsetting som andel av befolkning.

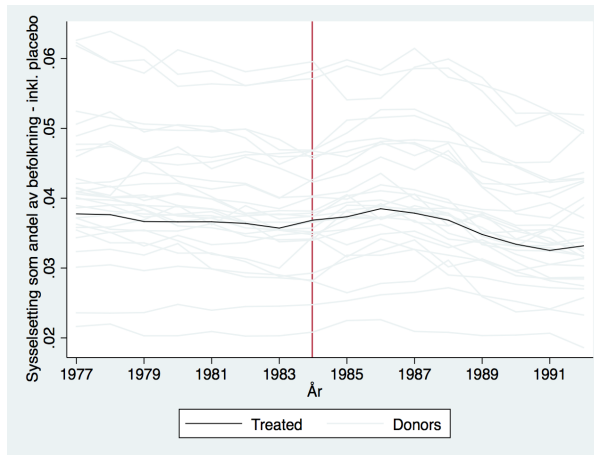
(a) Kommuner		(b) Variabler*	
Kommune	Vekt	Variabelnavn	Vekt
Bærum	0.191	sysselsetting(1980)	0.135
Halden	0.113	sysselsetting(1982)	0.243
Karmøy	0.011	befolkning	0.0001
Kongsberg	0.074	vgs	0.19
Moss	0.296	kort universitetsutdanning	0.008
Tromsø	0.006	lang universitetsutdanning	0.001
Trondheim	0.31	ingen utdanning	0.001
		aldersgruppe 1	0.381
		aldersgruppe 4	0.002
		aldersgruppe 5	0.038

*Alle variabler er som andel av befolkning

Placebotest: sysselsetting.

Figur 12: In place-placebotest for effekten på sysselsetting.

(a) Placebotest



(b) Effekter, placebo test

