

Emil Øverby

Sammenligning av agglomerasjonselastisiteten i privat og offentlig sektor- en studie med bakgrunn i norsk regionsdata mellom 1993 og 2010

Masteroppgave i Master i samfunnsøkonomi

Veileder: Hildegunn Ekroll Stokke

September 2019

Emil Øverby

Sammenligning av agglomerasjonselastisiteten i privat og offentlig sektor- en studie med bakgrunn i norsk regionsdata mellom 1993 og 2010

Masteroppgave i Master i samfunnsøkonomi
Veileder: Hildegunn Ekroll Stokke
September 2019

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi

Forord

Jeg vil først og fremst takke Hildegunn Ekroll Stokke for verdifull veiledning gjennom hele prosessen. Ønsker også å takke andre personer som føler de har bidratt. Eventuelle håndtrykk til disse kan gis ut på forespørsel.

«Se mamma, jeg er ferdig!» - Emil

Emil Øverby

Drammen, 2. september 2019, kl. 13.32

Sammendrag

Denne oppgaven har studert sammenhengen mellom lønn og befolkningsstørrelse for privat og offentlig sektor i Norge. Dette gjøres ved å sammenligne agglomerasjonselastisiteten. Tidligere forskning gir klare indikasjoner på at agglomerasjonselastisiteten er positiv for privat sektor, mens bidragene for den på offentlig har vært få. Hensikten med denne oppgaven er dermed å sammenligne den estimerte elastisiteten for de to sektorene. Dette har blitt gjennomført med norsk data på regionsnivå fra tidsperioden 1993-2010.

Agglomerasjonselastisiteten virker tilsynelatende å være høyere i privat enn i offentlig sektor over hele tidsperioden. Dette blir også konklusjonen etter å ha undersøkt utviklingen dens over hvert enkelt år. På bakgrunn av at denne sammenhengen ikke har vært tilstand for så mange undersøkelser, er dette relevante opplysninger for senere forskning. Videre ble disse resultatene robusthetstestet ved å sammenligne lønnsnivået i syv storregioner med de resterende regionene. Resultatet av dette viser en signifikant lønnsforskjell for privat sektor, mens det ikke finnes noen for offentlig sektor. Avslutningsvis ble historisk befolkningsdata benyttet for å instrumentere for eventuelle simultanitetsproblemer mellom lønn og befolkningsstørrelse, hvor resultatene gir ulike indikasjoner for de to sektorene.

Abstract

This study has investigated the correlation between income and population size for the private and public sectors in Norway. This is done by comparing the agglomeration elasticity. Previous research in this field gives strong indications that the agglomeration elasticity has a positive value for the private sector, while there has not been that much contribution on the research of the public sector. The purpose of this study is therefore to compare the estimated elasticity for both sectors. This has been executed with Norwegian data on region level from the time period 1993-2010. The agglomeration elasticity seems to be higher in the private than the public sector during this time period. This is also the conclusion after investigating its development over each year. Because this correlation has not been the subject for that many investigations, this is relevant information for further research. The results were then tested for robustness by comparing the level of income in seven big regions with the rest of the regions. The results show a significant income difference in the private sector, while there is none for the public sector. Finally, historical population data was used to instrument for eventual simultaneity problems between income and population size, where the results give different indications for the two sectors.

Innholdsfortegnelse

1. Innledning.....	1
1.1 Problemstilling.....	1
1.2 Disposisjon.....	1
2. Teori.....	2
2.1 Deling.....	2
2.1.1 Deling av udelelige goder og fasiliteter.....	2
2.1.2 Deling av gevinster fra variasjon.....	2
2.1.3 Gevinster fra individuell spesialisering.....	6
2.1.4 Deling av risiko.....	6
2.2 Matching.....	7
2.2.1 Bedre kvalitet på matchen.....	7
2.2.2 Økt sannsynlighet for matching.....	8
2.2.3 Minsker kostnaden knyttet til hold-up problemer.....	8
2.3 Læring.....	8
2.3.1 Kunnskapsgenerering.....	9
2.3.2 Kunnskapsdiffusjon.....	9
2.3.3 Akkumulasjon av kunnskap.....	9
2.4 Potensielle svakheter.....	10
3. Tidligere forskning.....	11
3.1 Statiske agglomerasjonseffekter.....	11
3.1.1 Identifiseringsproblemet.....	12

3.1.2 Sorteringsproblemet.....	13
3.2 Dynamiske agglomerasjonseffekter.....	14
4. Empirisk strategi.....	16
4.1 Metode.....	16
4.2 Økonometriske utfordringer.....	17
4.2.1 Brudd på antakelsen knyttet til eksogen variabel.....	18
4.2.1.1 Utelatt variabelproblem.....	18
4.2.1.2 Målefeil.....	20
4.2.1.3 Simultanitet.....	20
4.2.2 Håndtering av brudd på antakelsene.....	22
4.2.2.1 Instrumentvariabel.....	22
5. Data.....	24
5.1 Avhengig variabel.....	24
5.2 Forklaringsvariabler.....	25
5.2.1 Befolkning.....	25
5.2.2 Høyere utdanning.....	25
5.2.3 Kvinner.....	26
5.2.4 Innvandrere.....	26
5.2.5 Offentlig administrasjon.....	27
5.2.6 Finansiering og forsikring.....	27
5.3 Instrumentvariabel.....	28
5.4 Regioner.....	28

5.5 Deskriptiv statistikk.....	30
6. Analyse.....	32
6.1 Paneldata.....	32
6.1.1 Privat sektor.....	33
6.1.2 Offentlig sektor.....	34
6.1.3 Storregion-dummy.....	36
6.1.4 Konklusjon paneldatamodell.....	37
6.2 Tverrsnittsdata.....	38
6.2.1 IV-estimering.....	39
6.2.2 Konklusjon tverrsnittsdatamodell.....	43
7. Konklusjon.....	44
7.1 Videre forskning.....	44
8. Referanser.....	45
Appendiks.....	I
Appendiks A – Utledninger teoridel.....	I
Appendiks B - Tabeller.....	VI
Appendiks C – Liste over økonomiske regioner SSB 2000.....	VIII

1 Innledning

Teoretisk kan sammenhengen mellom produktivitet og befolkningsstørrelse trekkes helt tilbake til Smith (1776). At denne sammenhengen er positiv, slik at produktivitetsnivået er større i byer har senere fått bred aksept av fagmiljøet. Carlsen et al. (2016) sine resultater viste at lønningene i Oslo lå 17% over gjennomsnittsnivået ellers i Norge. Ved dagens kommunesammenslåinger er dette et interessant fenomen, fordi arbeidsmarkedsregionene i landet antas å minke i størrelse.

Felles for den tidligere forskningen innenfor dette feltet er at den utelukkende har benyttet seg av data for privat sektor. I denne sammenhengen er denne oppgaven et unikt bidrag til å belyse om effektene i privat sektor også kan gjenfinnes i offentlig sektor. Dette vil kunne gi en bedre forståelse av mekanismene bak agglomerasjon, ved at den sammenligner resultater fra både privat og offentlig sektor.

1.1 Problemstilling

Oppgavens mål er å sammenligne agglomerasjonselastisiteten for offentlig og privat sektor. Dette gjøres ved å benytte seg av 89 økonomiske regioner basert på SSB sin inndeling fra 2000 (SSB 2000). Resonnementet er at på bakgrunn av lavere lønn i offentlig sektor enn i privat sektor vil produktivitetsfordelene knyttet til agglomerasjon være lavere her.

1.2 Disposisjon

Oppgaven er delt inn slik at kapittel to vil være en teoridel som legger grunnlaget for valget av problemstillingen. Kapittel tre vil se på tidligere forskning på området som blant annet innebefatter forskjellen mellom en statisk og en dynamisk agglomerasjonseffekt. Kapittel fire ser på den empiriske strategien som vil bli brukt, samt de økonometriske utfordringene som er knyttet til den. Dataen vil bli beskrevet og framlagt i kapittel fem for å gi et innblikk i hva analysen i kapittel seks vil inneholde. Analysedelen vil inneholde resultatene fra de gjennomførte regresjonene med en liten tolkning av implikasjonen disse gir. Avslutningsvis vil kapittel syv komme med konklusjoner på funnene som har blitt gjort i oppgaven.

2 Teori

Denne delen vil basere seg på Duranton og Puga (2004) sin artikkel som ser på teoretiske mikrogrunnlag for urbant agglomererte økonomier. Effektene som forekommer gjennom agglomerasjonen deles i deres artikkel inn i tre kanaler; deling, matching og læring. Alle tre forklarer produktiviteter knyttet til en økning av befolkningsstørrelsen innenfor et geografisk område.

Artikkelforfatterne formaliserer flere av de teoretiske effektene. På grunn av omfanget vil kun en av disse effektene formaliseres i denne oppgaven; gevinsten oppnådd ved en større variasjon av mellomliggende goder. Forklaringen bak de resterende effektene vil være en oppsummering basert på det som ansees som mest sentralt.

2.1 Deling

2.1.1 Deling av udelelige goder og fasiliteter

Udelelige goder eller fasiliteter er noe som er praktisk vanskelig, eller umulig, å dele ut til enkeltindivider. Eksempler på dette kan være en flyplass, et rettsvesen eller et sykehus. I disse tilfellene vil det være fordelaktig å la flere benytte seg av det samme godet på samme tid. Slike goder vil ofte kreve en stor investering i startfasen, før framtidige tjenester vil kunne tilbys til en lav marginalkostnad. Det at flere benytter seg av godet fører til at investeringskostnadene blir lavere per person.

Når mange personer benytter seg av et gode vil det bli en trade-off ved at flere deler på kostnadene knyttet til oppstartsinvesteringen, og at flere benytter seg av godet samtidig som det gir kostnader for de som benytter seg av det.

2.1.2 Deling av gevinster fra variasjon

For å vise produktivitetsfordelene som oppstår ved en variert produksjon av mellomliggende goder vil det i denne oppgaven benyttes en modell formulert i artikkelen til Duranton og Puga (2004).

Vi har m sektorer med indeks $j = 1, \dots, m$. Alle sektorer inneholder konkurrerende bedrifter som produserer sluttgoder som senere vil bli konsumert. Produksjonen av disse har konstant skalaavkastning, med konstant substitusjonselastisitet lik $\frac{(1+\epsilon^j)}{\epsilon^j}$ hvor $\epsilon^j > 0$. Her vil en høyere ϵ^j gi en lavere elastisitet. Produsentene av sluttgodene vil benytte seg av mellomliggende goder, som er spesifikke for hver sin sektor. Videre har vi at den aggregerte totale produksjonen i sektor j er:

$$Y^j = \left\{ \int_0^{n^j} [x^j(h)]^{\frac{1}{1+\epsilon^j}} dh \right\}^{1+\epsilon^j} \quad (1)$$

hvor $x^j(h)$ er aggregert produksjon av mellomliggende gode h i sektor j og n er antall variasjoner av disse. Produksjonen av de mellomliggende godene gjøres av monopolistiske konkurransedrevne bedrifter som har følgende produktfunksjon:

$$x^j(h) = \beta^j l^j(h) - \alpha^j \quad (2)$$

med $l^j(h)$ som antall arbeidere i sektor j og β som en koeffisient på marginalproduktiviteten av arbeid. α er de faste kostnadene i den samme sektoren. Det er her tiltakende avkastning på produksjonen av de mellomliggende godene samt uendelig antall variasjoner som kan bli produsert. Disse to faktorene fører til at hver bedrift kun produserer en variant av godet som ingen andre bedrifter vil kopiere. Kostnadene knyttet til produksjonen av sluttgodene er:

$$\int_0^{n^j} q^j(h) x^j(h) dh \quad (3)$$

hvor q^j er prisen på det mellomliggende godet. Ved å minimere denne gitt produksjonsfunksjonen (1) får man:¹

$$x^j(h) = \frac{\left([q^j(h)]^{-\frac{1+\epsilon^j}{\epsilon^j}} Y^j \right)}{\left\{ \int_0^{n^j} [q^j(h')]^{-\frac{1}{\epsilon^j}} dh' \right\}^{1+\epsilon^j}} \quad (4)$$

som viser etterspørselen etter det mellomliggende godet. Her kan det observeres at en økning i prisen på det mellomliggende godet vil senke produsert mengde. Dette vil avhenge av substitusjonselastisiteten hvor en lavere elastisitet vil føre til at mengden reduseres mer. Tanken bak dette er at en lavere elastisitet gjør det lettere å flytte over til andre produkter. Profitten til produsentene av de mellomliggende godene er:

$$\pi^j = q^j x^j - w^j l^j \quad (5)$$

hvor q^j er prisen på vare x^j for sektor j og w^j er lønnskostnaden til en arbeider og l er antall arbeidere i samme sektor, alt for produsentene av de hjemmehørende godene. For å finne optimal pris for produsentene maksimeres denne med hensyn på q^j :²

$$q^j = \frac{1 + \epsilon^j}{\beta^j} w^j \quad (6)$$

Prisen på innsatsvarer er en mark-up over de marginale kostnadene $\frac{w^j}{\beta^j}$. Videre kan det observeres at høyere ϵ^j gir høyere pris. Gitt at det er full tilgang i markedet blir profitten for

¹ Fullstendig utledning i appendiks A.1.

² Fullstendig utledning i appendiks A.2.

produsentene dratt ned til 0. Ved å sette $\pi^j = q^j x^j - w^j l^j = 0$ samt bruke (2) og (6) kommer det fram at produksjonsnivået for hver sektor blir:³

$$x^j = \frac{\alpha^j}{\epsilon^j} \quad (7)$$

Ved å ta bruk i (2) og (7) blir antall ansatte i hver sektor:⁴

$$l^j = \frac{\alpha^j(1 + \epsilon^j)}{\beta^j \epsilon^j} \quad (8)$$

som igjen gir totalt antall produsenter av mellomliggende goder lik:

$$n^j = \frac{L^j}{l^j} = \frac{\beta^j \epsilon^j}{\alpha^j(1 + \epsilon^j)} L^j \quad (9)$$

hvor L^j er det totale tilbudet av arbeidere i sektor j . Ved å sette denne og (7) inn i (1) får man:⁵

$$Y^j = \left[n^j (x^j)^{\frac{1}{1 + \epsilon^j}} \right]^{1 + \epsilon^j} = (L^j)^{1 + \epsilon^j} \quad (10)$$

som gir tiltakende skalautbytte på sektornivå. Tanken bak dette er at en økning i arbeidskraften vil føre til flere produsenter av de mellomliggende godene. Dette gjør at sluttprodusentene får en større variasjon av goder å velge mellom som vil gjøre dem mer produktive. En økning i

³ Fullstendig utledning i appendiks A.3.

⁴ Fullstendig utledning i appendiks A.4.

⁵ Fullstendig utledning i appendiks A.5.

produksjonen av sluttgoder vil da med andre ord trenge en mindre proporsjonal økning av innsatsfaktorene.

Denne modellen har fått kritikk på bakgrunn av at den virker litt mekanisk; det at en større arbeidsstyrke øker produksjonen av mellomliggende goder, som gjør at sluttproduksjonen øker mer enn proporsjonalt, kommer av den konstante substitusjonselastisiteten. I tillegg vil økningen i produksjonen av mellomliggende goder komme av at det kommer flere tilbydere, og ikke av operasjonsskalaen til hver tilbyder. Dette gjør at en økning av arbeidsstyrken kun skifter den ekstensive marginen til produksjonen.

2.1.3 Gevinster fra individuell spesialisering

Adam Smith (1776) mente at en økning av antall arbeidere førte til at produksjonen økte mer enn proporsjonalt ettersom arbeiderne da kunne få mer spesialiserte oppgaver. Årsaken bak dette delte han inn i tre; ved å gjøre samme oppgave flere ganger blir du bedre på å gjøre den, kjent som «learning by doing», det vil være kostnadsbesparende med grunnlag i at arbeidere som har faste arbeidsoppgaver ikke trenger nye maskiner eller nye lokaler, og en deling av arbeiderne vil øke antall arbeidsbesparende innovasjoner som gjør at de enklere oppgavene kan mekaniseres.

Spesialisering av arbeidsoppgaver er ikke kun verdiskaping for bedriften på grunn av kostnadsbesparelser. Bedriften vil også kunne gi arbeidstakerne insentiv gjennom økt lønn for den spesialiserte arbeidsoppgaven. Størrelsen på arbeidsstyrken vil også påvirke hvor spesialisert den enkelte arbeider er.

2.1.4 Deling av risiko

I et større arbeidsmarked vil det være slik at bedrifter har bedre tilgang på nye arbeidstakere, en effekt som blir omtalt som pooling. Dette baserer seg på Marshall (1890) hvor ideen er at det innenfor en lokalisert industri vil være en konstant tilgang på kunnskap. Bidraget fra dette er at det vil bli enklere for bedrifter å få inn nye arbeidere hvis det skal oppstå et idiosynkratisk sjokk. Krugman (1991) har formalisert dette ved en modell hvor arbeidere mottar en lønn lik deres marginale produkt. Dette fører til at en økning av antall bedrifter vil øke lønnen, fordi dette gir en nedgang i antall arbeidere per bedrift, som igjen øker den marginale produktiviteten til de resterende. Videre vil det være slik at forventet lønn synker med en økning i den lokale

arbeidsstyrken og med intensiteten til avtakende avkastning. Profittfunksjonen til bedriftene vil være konvekse for idiosynkratiske sjokk ettersom bedriftene vil endre produksjonen gitt utfallet av sjokket.

Ved en større usikkerhet knyttet til sjokkene blir effekten av pooling sterkere ettersom bedriftene opplever mindre risiko. Effekten blir også sterkere for en økning av antall bedrifter siden det er større valgmuligheter hvis det trengs ekstra arbeidskraft. Det er også slik at hvis to lokale arbeidsmarkeder kunne valgt lokalisering før et eventuelt sjokk ville de ha agglomerert sammen selv om forventet avkastning hadde vært lik begge steder.

2.2 Matching

2.2.1 Bedre kvalitet på matchen

Denne effekten går hovedsakelig gjennom to kanaler; en økning i antall agenter øker kvaliteten på matchingen (Helsley og Strange 1990), og det fører til en sterkere konkurranse mellom bedriftene, noe som gir lavere faste kostnader på bakgrunn av at bedriftene øker mindre enn proporsjonelt med arbeidsstyrken.

Hvis en bedrift ansetter en person som ikke er en perfekt match for jobben vil bedriften påføre en kostnad. Denne kostnaden, som kan sees på som en kostnad knyttet til opplæring, er bedriften klar over og de vil følgelig prøve å minimere den. Det er her en større total arbeidskraft gir en agglomerasjonseffekt ved at antall bedrifter øker, noe som fører til at det vil være lettere for en gjennomsnittlig arbeider å finne en jobb som passer. Grunnen til at de faste kostnadene går ned med en økt arbeidsstyrke er at antall bedrifter øker mindre proporsjonal enn arbeidsstyrken, på bakgrunn av større konkurranse i arbeidsmarkedet. Dette fører til at utbyttet per arbeider øker med tanke på de faste kostnadene bedriftene står ovenfor.

Kim (1989) viser at økt arbeidskraft fører til mer investering i spesifikk human kapital i forhold til generell humankapital. Dette bidrar til økt produktivitet hos den enkelte arbeider. Tolkningen av dette er at økt befolkning fører til økt spesialisering. Venables (2002) argumenter for at det i byer er høyere produktivitet ettersom arbeidsstyrken kan antas å inneholde en større andel av høyt produktive personer. Argumentet ligger i at alle arbeidere får en økt produktivitet av å samarbeide med en høyt produktiv person, hvor effekten vil være størst for de som allerede er høyt produktive selv. Det strides videre hvorvidt høyere levekostnader i byene fører til at det

ikke er lønnsomt for mindre produktive arbeidere å bosette seg der, i motsetning til de allerede høyt produktive som får en høyere effekt av å jobbe med andre høyt produktive arbeidere.

2.2.2 Økt sannsynlighet for matching

Denne effekten går hovedsakelig ut på at jo flere arbeidssøkere som prøver å finne en passende jobb, jo større er sannsynligheten for at noen finner en match. Hvis det finnes flere valgmuligheter for en arbeidssøker vil det være mindre sannsynlig at ingen av disse vil passe. En proporsjonal økning av antall jobbsøkere og ledige stillinger vil føre til at antall matcher øker mer enn proporsjonalt. Det vil også være større sannsynlighet for match hvis det finnes flere jobbmuligheter i et marked enn det er mulig å få oversikt over simultant. Det at det finnes flere valgmuligheter for arbeidstakeren vil kunne føre til at den blir mer kresen i valg av jobb. Videre er det også ofte slik at to personer som verken er helt like eller helt forskjellige jobber best sammen. To helt forskjellige personer vil ofte ikke komme best overens, mens to helt like personer vil ha lite å lære av hverandre.

2.2.3 Minsker kostnaden knyttet til hold-up problemer

Hvis det er mistillit mellom to parter, for eksempel mellom en arbeidstaker og en arbeidsgiver eller mellom en tilbyder og en etterspørter, kan det føre til at kontrakten mellom disse ikke fullføres. Dette kan skje når kontrakten skal reforhandles etter at investeringen er gjennomført. Siden tilbyderen vet at kontrakten skal reforhandles etter at investeringen har funnet sted, vil det kunne oppstå tvil knyttet til om investeringen skal gjennomføres eller ikke. I en større by vil dette være et mindre problem ettersom det vil være enklere for investoren å finne nye samarbeidspartnere. Hvis investoren vet at det er mer sannsynlig at kontrakten opprettholdes vil det kunne investeres mer i første steg. Slik vil også profitten til bedriften bli høyere, noe som fører til en bedre situasjon for begge aktører.

2.3 Læring

Læring må sees på som en viktig aktivitet på bakgrunn av at det benyttes mye ressurser på det. Læring skjer først og fremst gjennom interaksjoner mellom personer. Mesteparten av disse interaksjonene vil foregå ansikt til ansikt, noe det er større sannsynlighet for å oppnå i byer hvor

flere personer er samlet. Det snakkes om ulike effekter av disse interaksjonene. Jovanovic (1997) mener at over 20% av en bedrift sine ressurser brukes på læring for de ansatte.

2.3.1 Kunnskapsgenerering

En nyetablert bedrift vil ofte trenge tid før den finner ut hvilken retning den skal ta i forhold til fremtiden og bedriftens utvikling. Selv om ledelsen har et mål for prosjektet, vil mangelen på kunnskap og erfaring føre til at det vil ta tid før den beste strategien utvikles. Denne søken vil være en læringsprosess som vil virke som en kostnad. Det vil allikevel virke som en investering for fremtiden i og med at bedriften tilslutt vil finne hvilken strategi som er best for den. Fram til bedriften finner ut hvordan strategien er mest lønnsom vil det være fordelaktig å være inne i et diversifisert miljø hvor det vil finnes flere valgmuligheter. Når bedriften finner sin plass vil dette være innenfor et spesialisert miljø, hvor det er lettere for den å tilegne seg ny kunnskap. På denne måten vil det diversifiserte miljøet virke som en positiv medvirkende faktor for bedriften. Et slikt miljø vil lettere kunne oppstå i større byer.

2.3.2 Kunnskapsdiffusjon

Levekostnadene er høyere i byer enn i utkantsområder. Men det er også i byer det finnes størst mulighet for læring. Dette gjør at unge personer som bosetter seg i byer gjør en investering for fremtiden ved å flytte dit.

Etterspørselen er mer stabil i store byer, noe som gjør det lettere for bedrifter å vurdere om de skal foreta en investering eller ikke. I tillegg vil det være slik at mulighet for mangel på oversikt over fremtiden i en stor by vil påvirke en bedrifts beslutninger. Dette vil igjen påvirke konkurrerende bedrifters avgjørelser. Det vil også ofte være slik at ingen bedrifter ønsker ta den første avgjørelsen angående investeringer, da dette kan være en fordel for konkurrenten.

2.3.3 Akkumulasjon av kunnskap

Denne delen bygger på relasjoner som beskriver produksjonen av ulike goder og akkumuleringen av innsatsfaktorer. Akkumuleringen av humankapital vil for byen som helhet ikke få avtakende utbytte, selv om dette er tilfellet for den enkelte person. Årsaken bak dette er at i byer vil flere ha den samme kunnskapen, noe som gjør den til avtakende for det enkelte

individet. Byen som helhet vil derimot ha et bredere spekter av kunnskap som gir en økonomisk effekt.

2.4 Potensielle svakheter

Det finnes ulike mikroøkonomiske mekanismer som i seg selv kan forklare eksistensen til byer. Selv om disse mekanismene er robuste mot ulike spesifikasjoner er det også slik at det er vanskelig å skille mekanismene fra hverandre empirisk. Dette er viktig å få på plass fra et politisk perspektiv, da slike grunnlag kan gi viktige nye perspektiver.

Siden aktører i praksis ikke har fullstendig informasjon kan det være at fordelene fra byer er større enn beskrevet her. Dette forekommer fordi bedrifter som oppholder seg i byer er i større grad i kontakt med kunder, noe som gjør det lettere å finne nye løsninger gjennom ulike eksperimenter.

Det bør også legges merke til at mekanismene bak læringseffektene ikke er like godt utviklet som for deling og matching. Duranton og Puga (2004) forklarer at dette gjelder spesielt effektene av spill-overs fra kunnskap og at på grunn av viktigheten dette har for vekst og innovasjon bør dette legges mer vekt på i framtidig arbeid.

3 Tidligere forskning

Denne delen vil presentere tidligere forskning på sammenhengen mellom produktivitet og geografi. Den første delen vil gå i dybden på statiske agglomerasjonseffekter, mens den andre vil beskrive dynamiske agglomerasjonseffekter. Grunnlaget for forskjellig dybdenivå er at analysen utført i denne oppgaven har undersøkt de statiske effektene. På de statiske effektene vil det også bli vist til to historiske utfordringer man har stått ovenfor når disse skal bevises empirisk. Dette gjelder utfordringene knyttet til identifisering og sortering av arbeidere.

Til felles for tidligere forskning er at de i hovedsak har benyttet data for privat sektor. Her skiller denne oppgaven seg ved at den også analyserer data fra offentlig sektor.

3.1 Statiske agglomerasjonseffekter

Sammenhengen mellom geografi og produktivitet ble allerede undersøkt av Smith (1776). Marshall (1920) forklarte hvordan agglomerasjonsøkonomier får fordeler gjennom tre effekter han kalte for input sharing, labor market pooling og knowledge spillovers. Eksempler fra de tre effektene er muligheten til å skaffe billige innsatsvarer fra nærstående bedrifter, raskt ansette personer fra nærområdet i ledige stillinger og hvordan kunnskap i en bedrift lett kan overføres til en annen nærliggende bedrift. Disse tre teoretiske effektene ses fortsatt på som gjeldende i dag.

Nyere forskning har hatt en fordel knyttet til tilgang på mer raffinert data. Dette har åpnet for muligheter til å besvare flere spørsmål, samt besvare tidligere spørsmål med bedre presisjon. Blant disse har Sveikauskas (1975) sin studie hvor det ble brukt output per arbeider som målenhet på produktivitet vært fremgangsgivende. Nakamura (1985) fant både en lokalisierende og en urbaniserende effekt fra agglomerasjon i japanske storbyer. Henderson (1986) fant at produktiviteten til bedrifter økte sammen med størrelsen på industrien. Krugman (1991) viste at agglomerasjon fungerer selv med lave transportkostnader, gitt at de fleste arbeiderne er mobile. Dekle og Eaton (1999) så på effekten ved å benytte seg av leiekostnadene bedriftene sto ovenfor. Tanken her var at høyere leiepriser i byene ville medføre at bedriftene også var nødt til å være mer produktive. Glaser og Maré (2001) fant et lønnspremium i storbyene, og at dette var høyere for arbeidere som hadde jobbet over en lengre periode. Ciccone og Hall (1996) fant en positiv sammenheng mellom befolkningstetthet og arbeidsproduktiviteten til arbeiderne i amerikanske county. Combes et al. (2010) fant ved fransk data at sysselsettingstettheten har

en positiv innvirkning på den lokale produktiviteten. Carlsen et al. (2016) fant ved sin norske data at det urbane lønnspremiumet er høyere for personer med høyt utdanningsnivå.

Melo et al. (2009) lagde en metaanalyse hvor de sammenligner 729 estimater for agglomerasjonselastisiteten fra 34 ulike studier. Tidsspranget for disse studiene er fra 1965 til 2002. Det totale gjennomsnittet for estimatene gir en lønnsøkning på 5,8% når befolkningsstørrelsen har doblet seg. Estimater fra studier som har brukt befolkningsstørrelse som forklaringsvariabel har et gjennomsnittlig resultat på 3,2%. Forklaringen på dette avviket kan ha mange årsaker, men historisk forklarer Moomaw (1981) at studier som har benyttet seg av en produktfunksjon gav høyere estimater.

Combes et al. (2010) påpekte at elastisiteten til den lokale produktiviteten med tanke på sysselsettingstettheten typisk ligger mellom 4 og 10%. Nakamura (1985) fant en agglomerasjonselastisitet i Japan på 3,4%. Sveikaukas (1975) sine estimater viste at en dobling av befolkningsstørrelsen økte output per arbeider med cirka 6%. Ciccone og Hall (1996) fant at en fordobling av befolkningstettheten i amerikanske county øker arbeidsproduktiviteten med 6%. Combes et al. (2010) fant med sin franske data at lønnen økte med på 5% når de doblet sysselsettingstettheten. Carlsen et al. (2016) kom fram til at arbeidere i Oslo i gjennomsnitt har 17% høyere lønn enn arbeidere andre steder i landet. De la Roca og Puga (2017) forklarte hvordan arbeidere i Madrid tjener 21% mer enn arbeidere i den tredje største byen i Spania, Valencia, 46% mer enn arbeidere i en by på medianstørrelsen og 55% mer enn arbeidere på landsbygda. På samme måte fant Glaeser (2011) at amerikanske arbeidere i området på mer enn 1 million innbyggere tjener 30% mer enn arbeidere på landsbygda.

I 3.1.1 og 3.1.2 vil historiske utfordringer knyttet til estimeringen av agglomerasjonselastisiteten, hvordan disse utfordringene har blitt løst samt hvorvidt det påvirker resultatene bli beskrevet.

3.1.1 Identifiseringsproblemet

Identifiseringsproblemet handler om at lønn og befolkningsstørrelse kan være simultant bestemt. Med dette menes at det at en økt befolkningsstørrelse ikke nødvendigvis gir høyere lønn gjennom agglomerasjonsmekanismene, men at høyere lønn kan tiltrekke seg flere

arbeidere. Resultatet av dette er at den kausale sammenhengen mellom lønn og befolkning som det blir prøvd estimert ikke vil være korrekt.

Ciccone og Hall (1996) løste dette problemet ved å bruke historiske befolkningsdata som en instrumentvariabel. Tanken her er at det er en sammenheng mellom historisk og nyere befolkningsstørrelse, samtidig som den ikke korrelerer med nyere lønnsnivå. Gitt dette vil historisk befolkningsdata fungere som en instrumentvariabel. Denne metoden har senere blitt benyttet av blant annet Carlsen et al. (2016), Combes et al. (2010) og Glaser og Maré (2001). Combes et al. (2010) har også brukt ulike geologiske faktorer som kvalitet på jorden og meter over havet som instrument for nyere befolkningsnivå. Argumentasjonen her er at dette kan være faktorer som har historisk påvirket hvor folk har bosatt seg, samtidig som det antas at disse ikke påvirker det nyere lønnsnivået.

Estimater fra tidligere analyser gir klare indikasjoner på at identifiseringsproblemet er reelt. Ved å inkludere historisk befolkningsdata i analysen fant Combes et al. (2010) at agglomerasjonselastisiteten minket fra cirka 5% til cirka 4%. Dette tyder på at OLS overestimerer effekten befolkningsstørrelsen har på lønnen. Konklusjonen av dette blir derfor at effekten går begge veier, økt lønn fører også til økt befolkning.

3.1.2 Sorteringsproblemet

Sorteringsproblemet eller det kvalitative endogenitetsproblemet går ut på at arbeidere i ulike regioner kan ha ulike egenskaper. Dette kan for eksempel forekomme hvis noen regioner tiltrekker seg mer produktive arbeiderne enn andre. Det finnes ulike karakteristikkene som kan være med på å forklare dette. Noen av disse karakteristikkene har gjennom tidligere forskning blitt empirisk bevist, slik at det i dag kan kontrolleres for dem. Eksempler på dette er at personer med høyere utdanning blir ansett som mer produktive enn personer uten høyere utdanning. Etersom dette er en karakteristikk som er målbar kan vi kontrollere for den ved å inkludere den som en forklaringsvariabel. Problemet oppstår i det disse karakteristikkene ikke kan måles, når de er uobserverbare. Når disse korrelerer med noen variabler i analysen vil det oppstå skjevheter i estimatene. Eksempler på dette kan være arbeiderens ambisjonsnivå eller arbeidsinnsats.

En løsning på sorteringsproblemet som blant annet har blitt benyttet av Combes et al. (2010) og Carlsen et al. (2016) har vært å bruke individfaste effekter. Hensikten med dette er å kun se på endringene i lønnen til personer som migrerer fra et sted til et annet. Gjennom dette vil man

kunne utnytte seg av tidsdimensjonen til dataen, slik at det vil være fordelaktig å ha en lang tidsperiode på datasettet. Combes et al. (2010) bruker også lokale faste effekter for å skille de lokale fra de individeffektene.

Relevansen til sortering av arbeidere kan det fås et innblikk i ved å se på endringen i estimatene ved bruk av løsningene beskrevet ovenfor. Combes et al. (2010) fant først en agglomerasjonselastisitet på 4,8% ved å kun benytte seg av modell uten kontrollvariabler. Når de inkluderte andre observerbare karaktrestikker økte estimatet marginalt til 5,1%. Når individfaste effekter inkluderes faller estimatet til 3,3%. På samme måte faller den estimerte elastisiteten til De la Roca og Puga (2017) fra 4,6 til 2,4% ved inkludering av individfaste effekter. Disse resultatene tilser at arbeidergruppene i de ulike geografiske områdene ikke er homogene, noe som har ført til en skjevhet på estimatene ved å ikke inkludere de individfaste effektene. Combes et al. (2010) menes dette skyldes at de mer produktive arbeiderne sorteres inn i de store byene. Carlsen et al. (2016) fant at sortering i de store byene er reelt med tanke på de med høy utdanning. De fant også at denne effekten er sterkere hos de yngre arbeidstakerne.

3.2 Dynamiske agglomerasjonseffekter

Økningen av produktivitet hos arbeiderne kan komme på ulike tidspunkt, hvor agglomerasjonseffekten enten kan være statisk eller dynamisk eller en kombinasjon av begge. Ved en statisk effekt vil produktivitetsøkningen skje med en gang individet starter arbeidet i den større byen. På samme måte vil produktiviteten raskt avta ved utflytting fra den større byen. Den dynamiske effekten endrer produktiviteten over en lengre tidsperiode, ved at arbeideren blir mer produktiv gjennom læringsmekanismer. Dette medfører at den økte produktiviteten vil vedvare for arbeideren selv om den flytter ut fra byen. Basert på de ulike effektene har tidligere analyser på agglomerasjonseffektene gitt en todelt inndeling.

Glaeser og Maré (2001) benytter seg av tre ulike amerikanske datasett for å se om lønnspremiumet som oppnås i byer kommer av en «wage level» eller en «wage growth»-effekt. Wage level-effekten gir en rask lønnsøkning ved innflytting til byen, samtidig som den synker ved utflytting. Wage growth-effekten øker derimot produktiviteten over lengre tid, men lønnen vil fortsatt være høy hvis arbeideren flytter ut igjen av byen. For å undersøke om hvilke effekter det er som spiller inn har Glaeser og Maré samlet data for enkeltindivider over en lengre

tidsperiode, hvor fokusområdet er hvordan lønnen endrer seg både før og etter innflytningen til en region med en større by⁶. Resultatene deres viser at arbeidere som har bodd i en region med en større by også opplever høyere lønn hvis de senere bosetter seg utenfor regionen. De finner også at lønnen øker ved innflytting til en region med en større by. De konkluderer med det at lønnspremiumet i byene påvirkes både av wage growth og wage level-effekten.

De la Roca og Puga (2017) brukte spansk registerdata for å undersøke både den statiske og den dynamiske agglomerasjonseffekten. For å se på forskjellen mellom de to effektene så de først på en statisk modell hvor de kun inkluderte by-spesifikke effekter og ulike individ- og jobbrelevante karaktrestikker. Ved den dynamiske modellen inkluderte de individfaste effekter samt erfaringen til arbeiderne. Ved å se på erfaringen til de ulike arbeiderne kunne de undersøke om lønnen deres øker over tid gitt hvor de har arbeidet. Avkastningen av erfaringen tillates også til å variere geografisk. For å estimere de ulike effektene så de på elastisiteten på kort og mellomlang sikt. Med en elastisitet på mellomlang sikt på 0,051 og en på kort sikt på 0,0223 trakk de konklusjonen at den statiske og den dynamiske effekten bidrar like mye. Videre fant de at lønnen til arbeidere fortsatt var høye etter utflytting fra de store byene, som de mener tyder på at det finnes viktige læringsmekanismer i de store byene.

Carlsen et. al (2016) undersøkte både den statiske og den dynamiske effekten ved å dele deres norske individdata inn i tre; Oslo, seks andre storbyer og resten av landet. Dette ble gjort for å sammenligne agglomerasjonsmekanismen i ulike storbyer. Årsaken bak dette er at disse områdene er unike da de har stor pendlerandel. Denne metoden vil bli benyttet ved en robusthetssjekk i avsnitt 6.1.3 i denne oppgaven. Videre estimerte de agglomerasjonseffekten for ulike utdanningsgrupper. Resultatene deres tilsier at både statiske og dynamiske agglomerasjonseffekter inntreffer. Den totale agglomerasjonseffekten for arbeiderne i Oslo er på 17%, hvor den statiske står for 7% og den dynamiske for 10%.

For å kunne undersøke om det forekommer en dynamisk agglomerasjonseffekt er det essensielt å benytte individdata. Årsaken til dette er at man må undersøke lønnsutviklingen til et bestemt individ over tid. Dette medfører at resultatene kan bli mer detaljerte når man ser på de dynamiske effektene. Basert på at tilgjengelig data for denne oppgaven er på et regionsnivå vil dette derimot ikke være gjennomførbart.

⁶ County som har minst en by med mer enn 500.000 innbyggere

4 Empirisk strategi

I denne delen av oppgaven vil valg av metode for best mulig å kunne besvare problemstillingen bli framlagt. Videre vil eventuelle utfordringer knyttet til metoden bli diskutert. Diskusjonen vil hovedsakelig gå ut på å vise eventuelle problemer knyttet til resultatene. Både metoden og de eventuelle økonometriske utfordringene dette kan medføre baserer seg på teori fra Wooldridge (2016).

4.1 Metode

For å best kunne forklare produktivitetseffekten befolkningsstørrelsen innehar har det blitt valgt to ulike modeller for de to sektorene:

$$\lnwagepub_{it} = \alpha_0 + \alpha_1 \lnbepolk_{it} + \alpha_2 hoyereutd_{it} + \gamma x_{it} + \eta_t + \varepsilon_{it} \quad (4.1)$$

$$\lnwagepriv_{it} = \beta_0 + \beta_1 \lnbepolk_{it} + \beta_2 hoyereutd_{it} + \mu x_{it} + \eta_t + \varepsilon_{it} \quad (4.2)$$

hvor \lnwagepub og \lnwagepriv er den naturlige logaritmen til gjennomsnittlig dagslønn i offentlig og privat sektor i region i i år t , \lnbepolk er den naturlige logaritmen til befolkningsstørrelsen i region i i år t og $hoyereutd$ er andelen av befolkningen i region i i år t som har fullført minst fireårig høyere utdanning. I tillegg til dette inneholder modellene en vektor med kontrollvariabler, her nevnt som x_{it} , med påfølgende koeffisienter γ og μ . To av kontrollvariablene; andelen kvinner og andel personer med innvandrerbakgrunn, vil være med i begge modellene. Andelen av heltidsansatte i offentlig sektor innenfor offentlig administrasjon vil være med i modell (4.1), mens andelen av heltidsansatte i privat sektor innenfor finans og forsikring vil være med i modell (4.2). En nærmere beskrivelse av alle variablene vil bli gjort i kapittel 5. η_t indikerer en årsummy som skal kontrollere for faste effekter for det bestemte året. ε_{it} er et stokastisk feilledd som fanger opp all variasjon modellen ikke klarer å beskrive. Metoden som vil bli brukt for å estimere koeffisientene vil være Ordinary Least Square (OLS). Denne metoden minimerer summen av de kvadrerte avvikene. For offentlig sektor gjøres dette ved å minimere følgende uttrykk:

$$\sum_{i=1}^{89} (\lnwagepub_i - \hat{\alpha}_0 - \hat{\alpha}_1 \lnbefolk_i - \hat{\alpha}_2 hoyereutd_i)^2 \quad (4.3)$$

hvor \lnwagepub er den observerte verdien for gjennomsnittlig logaritmisk lønn i offentlig sektor i region i . $\hat{\alpha}_0$ er den estimerte koeffisienten for konstantleddet og $\hat{\alpha}_2$ er det samme for andelen med høyere utdanning i regionen. $\hat{\alpha}_1$ er estimatet for agglomerasjonselastisiteten som er interesseområdet knyttet til denne oppgaven.

For den private sektoren får vi følgende uttrykk:

$$\sum_{i=1}^{89} (\lnwagepriv_i - \hat{\beta}_0 - \hat{\beta}_1 \lnbefolk_i - \hat{\beta}_2 hoyereutd_i)^2 \quad (4.4)$$

hvor eneste forskjell er koeffisientene foran variablene. Oppgaven vil ha til hensikt å sjekke om $\hat{\alpha}_1$ og $\hat{\beta}_1$ er like, som vil si at agglomerasjonseffekten er lik i de to sektorene.

4.2 Økonometriske utfordringer

For at OLS skal gi forventningsrette estimatorer må følgende fire antakelser holde: Modellen må ha linearitet i parameterne, observasjonene må være tilfeldig trukket ut fra populasjonen, ingen variabler kan ha perfekt multikollinearitet og forventet verdi på feilleddet må være null gitt verdiene på forklaringsvariablene. I tillegg til dette sikrer andre antakelser at inferensen er til å stole på. Gitt at alle antakelsene ovenfor holder vil estimatene fra OLS være BLUE (Best Linear Unbiased Estimator).

Gitt modellene (4.1) og (4.2) vil det være linearitet i parameterne. Observasjonene har blitt innhentet over en lengre tidsperiode og er tilfeldig trukket ut. Antakelsen om ingen perfekt multikollinearitet handler om at ingen variabler verken er konstante eller har en eksakt lineær sammenheng med en av de andre. For variablene offentlig administrasjon og finans og

forsikring har det ikke vært mulig å oppdrive tall for tidsperioden 1993 til 1999. Løsningen på dette ble å erstatte disse verdiene med verdiene fra år 2000. Dette medfører at det vil være knyttet en høyere multikollinearitet til disse variablene enn hvis det hadde fantes data for alle årene. Resultatet av dette er en økt varians på estimatene (Wooldridge 2016). Dette kommer av at ved en høy korrelasjon blir det vanskelig å skille den kausale effekten til de ulike variablene. Den siste antakelsen om at forventningsverdien til feilledet skal være null er nødt til å holde for at forklaringsvariablene ikke skal være endogene. Det finnes flere feilkilder som kan gjøre at den ikke holder. Dette vil nå bli utdypet i 4.2.1.

4.2.1 Brudd på antakelsen knyttet til eksogen variabel

Hvis antakelsen om at forventningsverdien til feilledet ikke er lik null gitt verdiene på forklaringsvariablene ikke holder vil forklaringsvariablene være endogent bestemt. Dette gjør at estimatorene ikke er forventningsrette. Wooldridge (2016) beskriver tre tilfeller hvor dette kan skje. For enkelthetsskyld legges det til grunn at alle andre antakelser holder til enhver tid.

4.2.1.1 Utelatt variabelproblem

La oss si at vi undersøker effekten befolkningsstørrelsen har på lønningene i offentlig sektor. Modellen blir:

$$\ln wage_{pub_i} = \tilde{\beta}_0 + \tilde{\beta}_1 \ln befolk_i + v_i \quad (4.4)$$

$$\text{hvor } v_i = u_i + \beta_2 \text{hoyereut}_i$$

Det antas her at den sanne modellen er:

$$\ln wage_{pub_i} = \hat{\beta}_0 + \hat{\beta}_1 \ln befolk_i + \hat{\beta}_2 \text{hoyereut}_i + u_i \quad (4.5)$$

Modell (4.4) vil her være en underspesifisert modell ettersom en relevant variabel, *hoyereutd*, er utelatt. Dette vil føre til at estimatet på agglomerasjonselastisiteten, $\tilde{\beta}_1$, vil avvike fra dens faktiske verdi $\hat{\beta}_1$. Hvor stort dette avviket er avhenger av to faktorer; relevansen til den utelatte variabelen og korrelasjonen mellom estimert og utelatt variabel. Dette kan bli vist matematisk ved å først se på forventningsverdien til den estimerte koeffisienten $\tilde{\beta}_1$ i en multippel regresjon:

$$E(\tilde{\beta}_1) = E(\hat{\beta}_1 + \hat{\beta}_2 \tilde{\delta}_1) = E(\hat{\beta}_1) + E(\hat{\beta}_2) \tilde{\delta}_1 = \beta_1 + \beta_2 \tilde{\delta}_1 \quad (4.6)$$

hvor $\tilde{\delta}_1$ er korrelasjonen mellom den utelatte variabelen og interessevariablene. Videre gir dette en feilestimering av $\tilde{\beta}_1$ lik:

$$\text{Feilestimering}(\tilde{\beta}_1) = E(\tilde{\beta}_1) - \beta_1 = \beta_2 \tilde{\delta}_1 \quad (4.7)$$

Graden av feilestimering vil her avhenge av verdien på β_2 og $\tilde{\delta}_1$. Det vil forekomme en overestimering av $\tilde{\beta}_1$ hvis verdiene på begge variablene enten er sammenfallende positive eller negative. Ved ulike verdier vil det oppstå en underestimering. Det kan også legges merke til at dersom den utelatte variabelen ikke er signifikant i forhold til den avhengige variabelen, $\beta_2 = 0$, vil utelatelsen ikke ha innvirkning på estimatet. Det samme vil skje hvis det ikke er korrelasjon mellom utelatt variabel og interessevariabel, $\tilde{\delta}_1 = 0$. Antakelsen om at restleddet har null i forventningsverdi brytes kun hvis den/de utelatte variablene er relevante og har korrelasjon med interessevariabelen. Hvis ikke vil det ikke oppstå feilestimering, men det vil forekomme en høyere varians for estimatet.

Når lønn skal undersøkes vil det alltid være mange faktorer som spiller inn, samtidig som det vil være vanskelig å kontrollere dem. Dette, sammen med mangel på datagrunnlag, kan ha ført til at det har blitt utelatt variabler fra regresjonen som både kan være relevante samt korrelerer med interessevariabelen.

4.2.1.2 Målefeil

Med målefeil menes at variabelens observerte verdi ikke sammenfaller med dens faktiske verdi. En upresis verdi på variabelen vil ha ulik innvirkning på resultatet avhengig av hvor den har funnet sted.

Wooldridge (2016) viser hvordan en målefeil i avhengig variabel fører til økt varians for OLS-estimatorene kun hvis den ikke er korrelert med noen av forklaringsvariablene. Dette er erfaringsmessig en rimelig antakelse å ta. Gitt at denne antakelsen holder vil eneste implikasjon være at konstantleddet får en feilaktig koeffisient hvis målefeilen er systematisk.⁷

Erfaringsmessig er målefeil i forklaringsvariablene mer problematisk. Med en antakelse på at gjennomsnittsverdien på målefeilen er null, vil variansen til estimatoren bli påvirket dersom målefeilen skjer i forklaringsvariablene samtidig som den er ukorrelert med observerbar verdi. Estimatorene vil derimot ikke være forventingsrette hvis målefeilen er ukorrelert med den uobserverbare verdien på forklaringsvariablene. Da vil man systematisk underestimere effekten hvis den uobserverbare variabelen har en positiv koeffisient og overestimere effekten dersom koeffisienten er negativ.

Basert på at det meste av dataen er innhentet vil det være mulighet for at verdiene for noen av variablene har blitt feil. Det antas likevel at hvis dette har hendt har det ikke blitt gjort systematisk slik at eneste resultat blir høyere varians i variablene.

4.2.1.3 Simultanitet

Dette bygger på at den avhengige variabelen kan være simultant bestemt med forklaringsvariabelen. Implikasjonen av dette er at den korrelerer med feilleddet som vil gi inkonsistente estimatorene. Dette kan vises matematisk ved å først se på en modell for lønn:

$$\lnwagepub = \alpha_1 \lnbefolk + \beta_1 hoyereutd + u_1 \quad (4.8)$$

⁷ Gjennomsnittsverdien til målefeilen ikke er lik null.

hvor lønn også påvirker befolkningsstørrelsen i regionen:

$$\lnbepolk = \alpha_2 \lnwagepub + \beta_2 temperatur + u_2 \quad (4.9)$$

Det er også tatt med en temperaturvariabel som antas å være ukorrelert med u_2 , på samme måte som *hoyereutd* antas å være ukorrelert med u_1 . Begge antas videre å være relevante for hver sin avhengige variabel. Tilslutt antas det at u_1 og u_2 også er ukorrelerte.

Ved å sette *lnwagepub* fra (4.8) inn for *lnwagepub* i (4.9) får vi:

$$\lnbepolk = \pi_{21} hoyereutd + \pi_{22} temperatur + v_2 \quad (4.10)$$

hvor

$$\pi_{21} = \frac{\alpha_2 \beta_1}{(1 - \alpha_2 \alpha_1)}, \pi_{22} = \frac{\beta_2}{(1 - \alpha_2 \alpha_1)} \text{ og } v_2 = \frac{\alpha_2 u_1 + u_2}{1 - \alpha_2 \alpha_1}$$

som er den reduserte ligningen for *lnbepolk*. Selv om denne vil kunne gi konsise estimater for π_{21} og π_{22} , er det slik at *lnbepolk* må være ukorrelert med u_1 for at eksogenitetsantakelsen ikke skal bli brutt. Siden det her er antatt at u_1 og u_2 er ukorrelerte, samt at v_2 er en lineær funksjon av både u_1 og u_2 , vil *lnbepolk* og u_2 være korrelerte med mindre $\alpha_2 = 0$. Hvis den hadde vært null ville ikke lønn ha hatt noen påvirkning på befolkningsstørrelsen og det hadde ikke vært et simultanitetsproblem knyttet til variablene.

Som nevnt i 3.1.1 er det fare for at det er simultanitet mellom lønn og befolkningsstørrelse. I denne analysen vil dette kunne oppstå hvis det er slik at befolkningsnivået i regionen har blitt påvirket av lønnsnivået i en av sektorene. Dette vil føre til at variabelen er endogen bestemt som er et brudd på antakelsen. I tillegg er det ikke tatt høyde for faktorer som påvirker valget om å ta høyere utdanning slik at denne også kan være endogen.

4.2.2 Håndtering av brudd på antakelsene

På grunn av et mulig simultanitetsproblem mellom lønn og befolkningsstørrelse i regionene vil det bli benyttet instrumentvariabler. Hvorfor og hvordan dette kan løse problemet vil bli forklart i 4.2.2.1.

4.2.2.1 Instrumentvariabel

Vi ser på den samme modellen som tidligere:

$$\lnwagepub = \beta_0 + \beta_1 \lnbefolk + \beta_2 hoyereutd + u_1 \quad (4.12)$$

Basert på at det er et mulig simultanitetsproblem knyttet til lønn og befolkningsstørrelse vil det være en fare for at \lnbefolk og u_1 korrelerer som vil gi skjeve estimatorer. Antakelsen om at $hoyereutd$ er en eksogen forklaringsvariabel gjelder fortsatt, slik at det kun er \lnbefolk det er nødt til å korrigeres for. Dette gjøres ved å instrumentere variabelen på følgende måte:

$$\lnbefolk = \pi_0 + \pi_1 z_1 + v_2 \quad (4.13)$$

hvor z_1 er instrumentvariabelen og v_2 er feilleddet. For at instrumentvariabelen skal benyttes må den inngå i ligning (4.13) men ikke i ligning (4.12). I tillegg til dette må to andre kriterier holde:

$$\text{cov}(z_1, u_1) = 0 \quad (4.14)$$

$$\pi_1 \neq 0 \quad (4.15)$$

Det første kriteriet kalles ekskluderingskriteriet, som går ut på at instrumentvariabelen ikke korrelerer med feilledet i den opprinnelige modellen. Hvis den hadde gjort det hadde den ikke vært en eksogen variabel og hadde ikke løst problemet vi står ovenfor.

Det andre kriteriet er at instrumentvariabelen er relevant for den endogene variabelen den skal instrumentere for. Årsaken til dette er at ved høyere korrelasjon mellom dem vil man kunne komme nærmere den endogene variabelen sin sanne verdi. Siden det her benyttes like mange instrumentvariabler som endogene forklaringsvariabler vil den endogene variabelen være eksakt identifisert. Stock og Yogo (2005) argumenterer for at det ikke kun holder for instrumentvariabelen å ha statistisk inferens på et 5%-signifikansnivå, de anbefaler å en T-statistikk på minst $\sqrt{10} \approx 3,2$ eller en F-statistikk på minst 10.

Estimeringen av variabelen når man benytter seg av en instrumentvariabel gjøres gjennom en to-steps metode kalt two stage least square (2SLS). Ved 2SLS kjøres først en OLS-regresjon hvor den endogene forklaringsvariabelen er avhengig variabel. Forklaringsvariablene vil her være instrumentvariabler som best mulig skal kunne forklare den endogene forklaringsvariabelen. Videre benyttes verdiene fra den endogene forklaringsvariabelen i steg en når OLS-regresjonen kjøres på den opprinnelige modellen vår igjen. Dette kan løse det potensielle simultanitetsproblemet.

5 Data

Dataen brukt i denne oppgaven vil først bli beskrevet her, før den senere vil bli framlagt deskriptivt. Innhenting av dataen har blitt gjort fra Statistisk Sentralbyrå (SSB) sin statistikkbank og fra Norsk senter for forskningsdata (NSD) sin kommunedatabase. Tidsperioden for dataen er fra 1993 til 2010, med unntak av dataen for forklaringsvariablene hentet fra NSD hvor det kun fantes tall fra år 2000 og utover. De økonomiske regionene som ligger til grunnlag for analysen er basert på SSB sin inndeling fra 2000 (SSB 2000). Lønnstallene brukt for å se på relevansen til variablene er hentet fra SSB sin lønnsstatistikk (SSB 2018a).

5.1 Avhengig variabel

Den beste måten for å kunne forklare endringen i produktiviteten i privat- og offentlig sektor vil være å se på deres respektive lønnsnivå. Lønnen er hentet ut ifra SSB sin individdata, som senere har blitt aggregert opp på regionsnivå. Metoden for aggregeringen er den samme som ble benyttet i Carlsen et al. (2016). Den er gitt som gjennomsnittlig dagslønn, gitt i 2010-priser⁸. Individene inkludert i denne oppgaven har alle hatt fulltidsstillinger hvor de har jobbet for kun en arbeidsgiver i minst tre måneder det samme året. Personer hvor to arbeidskontrakter har maks tre måneders overlapping har også blitt inkludert. Dataen består av både menn og kvinner, uansett etnisk opprinnelse, med alder mellom 20 og 66 år. For å minimere påvirkning fra ekstremobservasjoner har personer med lønn innenfor 0,5 prosentgrensen i topp og bunn blitt utelatt. Arbeidere innenfor primærnæringene, jordbruk, skog og fiske, har ikke blitt tatt med ettersom dette er yrker som ikke er optimale for å måle produktivitet. For å best kunne måle produktiviteten til arbeiderne er arbeidsgiveravgiften⁹ inkludert i lønnen. Grunnet dette er at det gir et bedre bilde for kostnadene bedriften har rundt arbeidstakeren.

Dataen fra SSB definerer ikke bedrifter som offentlige eller private. Dette medfører at det ikke er mulig ut ifra individdataen å vite om personen jobbet i offentlig eller privat sektor. Skillet mellom offentlig og privat sektor har dermed blitt gjort ved å dele de ulike næringene inn i de

⁸ Justert med SSB sin konsumprisindeks

⁹ Denne er ulik for regionene men siden den er lik for ansatte i offentlig og privat sektor blir det ikke noe problem.

to sektorene. Offentlig sektor har da blitt fulltidsansatte i følgende tre næringer: Offentlig administrasjon og forsvar, undervisning og helse- og sosialtjenester. Resten av de fulltidsansatte har blitt oppført som arbeidere i privat sektor. Resultatet av dette er at for eksempel ansatte på private skoler, barnehager og sykehjem vil regnes som ansatte i offentlig sektor.

5.2 Forklaringsvariabler

Her vil det komme en beskrivelse av alle forklaringsvariablene som vil bli benyttet i analysen. Relevansen og grunnlaget for at de har blitt tatt med vil også bli beskrevet.

5.2.1 Befolkning

Befolkningsstørrelsen i de ulike regionene er basert på tall fra SSB. Det har her vært mulig å finne tall på regionsnivå, slik at det ikke har blitt gjennomført noe aggregering. Befolkningen innebefatter personer registrert som boende i regionen, og det har ikke blitt gjort noen restriksjon på verken alder eller kjønn. Ettersom lønnsdataen er basert på fulltidsansatte i regionen i aldersgruppen mellom 20 og 66 år gir ikke dette en nøyaktig beskrivelse av tilstanden. Men siden dette vil være likt for alle regioner antas det at det ikke har noe innvirkning på fremtidige resultater.

Under robusthetstesting for paneldatamodellen vil denne variabelen bli erstattet av en storregion-dummy med verdi «1» for regioner som i år 2010 hadde en befolkning på over 150.000. Alle regioner som hadde et befolkningstall under 150.000 samme år vil få verdi «0». Totalt var det sju regioner som hadde over 150.000 innbyggere i 2010.¹⁰

5.2.2 Høyere utdanning

Denne variabelen viser til andelen av personer over 16 år i regionen med høyere utdanning, hvor høyere utdanning er oppnådd hvis personen har fullført minst 4 år på universitets- og høyskolenivå, samt forskerutdanning. Gitt skolesystemet i Norge vil dette ikke kunne oppnås

¹⁰ Dette gjaldt regionene Oslo, Bergen, Trondheim, Stavanger/Sandnes, Lillestrøm, Drammen og Bærum/Asker.

før vedkommende er 22 år ¹¹, noe som fører til en underestimering av andelen med høy utdanning i regionen for de fulltidsansatte. I og med at denne underestimeringen kan antas å være lik for alle regioner vil dette allikevel ikke ha noen innvirkning i den fremtidige analysen. Videre mangler SSB fullstendige opplysninger om utdanningsnivået til innvandrere før 2014. Dette kan medføre et mer upresist estimat.

5.2.3 Kvinner

Andelen kvinner i regionen kan være av interesse for analysen i og med at menn statistisk sett har høyere lønn enn kvinner¹². Utregningen av denne variabelen har blitt gjort ved å dele antall kvinner over 18 år på antall personer over 18 år i den respektive regionen. Det kan her påpekes at de med heltidsstilling mest sannsynlig er eldre enn 18 år, men på samme måte som ved utregningen av variabelen med høyere utdanning vil dette være likt for alle regioner og derfor ikke bli tatt hensyn til.

5.2.4 Innvandrere

Definisjonen på en innvandrer i denne sammenhengen er en person født utenfor Norge hvor både foreldre og besteforeldre også er utenlandsfødte. Årsaken til at dette kan være relevant for analysen begrunnes i at gjennomsnittlig månedslønn mottatt av en innvandrer er lavere enn for en etnisk norsk person¹³. Det kunne vært hensiktsmessig å skille personene med innvandrerbakgrunn mellom personer med vestlig og ikke-vestlig bakgrunn, siden personer med ikke-vestlig bakgrunn har et enda lavere gjennomsnittlig lønnsnivå. Det har derimot ikke vært mulig å oppdrive tall for dette for hele den aktuelle tidsperioden. Gitt lønnsforskjellen mellom etnisk norske og personer med innvandrerbakgrunn vil variabelen likevel være av interesse. Variabelen viser til andelen av regionens totale befolkning med innvandrerbakgrunn.

¹¹ Gitt at personen ikke har hoppet over noen år og dermed ligger foran sitt normerte skoleløp.

¹² Gjennomsnittlig månedslønn for menn og kvinner i 2018 lå på henholdsvis 48.420 kroner og 42.170 kroner.

¹³ Mens gjennomsnittlig månedslønn for alle personer i Norge i 2018 var på 45.610 kroner var den på 40.180 kroner for personer med innvandrerbakgrunn.

5.2.5 Offentlig administrasjon

Av de tre næringene som utgjør offentlig ansatte i denne analysen er gjennomsnittlig månedslønn høyest i næringen for offentlig administrasjon og forsvar, og trygdeordninger underlagt offentlig forvaltning¹⁴. På bakgrunn av dette er det blitt forsøkt å kontrollere for andelen heltidsansatte i regionen som jobber innenfor denne næringsgruppen, som har næringskode 84. I NSD sin kommunedatabase kalles næringen for offentlig administrasjon, forsvar, sosialforsikring, med næringskode 75 for årene 2000 til 2008 og næringskode 84 for 2009 og 2010. Tallene gjelder for sysselsatte mellom 15 og 74 år med minst 30 timer avtalt arbeidstid i uka. Kommunedataen har senere blitt aggregert opp på regionsnivå før den har blitt delt på totalt antall personer i regionen med minst 30 timer avtalt arbeidstid i uka innenfor de tre næringene som innebefatter offentlig sektor. Denne variabelen blir dermed andelen av personer med fulltidsstillinger i offentlig sektor i regionen som jobber innenfor næringen offentlig administrasjon, forsvar, sosialforsikring. SSB har endret betegnelsen for næringen i løpet av perioden 2010 til 2018, men i denne oppgaven er det lagt til grunn at de er like på bakgrunn av at de har like næringskoder. Endringen av næringskode fra 2008 til 2009 er heller ikke lagt til grunn som påvirkning til oppgaven, da antall ansatte i de to ulike årene er relativt lik¹⁵.

5.2.6 Finansiering og forsikring

Innenfor det som blir ansett som privat sektor er lønnsnivået nokså variert. Tall fra 2018 tilsier at næringen med høyest lønnsnivå var bergverksdrift og utvinning. Endringen av næringer fra 2008 til 2009 har derimot gjort identifiseringen av denne næringen vanskelig da grupperingen av næringene har endret seg betraktelig. Det blir derfor i denne oppgaven brukt tall fra næringen med nest høyest gjennomsnittlig månedslønn, finansierings- og forsikringsvirksomhet, med næringskodene 64-66. Det har blitt hentet inn tall fra NSD for næringen finansiell tjenesteyting

¹⁴ Gjennomsnittlig månedslønn for 2018 for henholdsvis offentlig administrasjon og forsvar, og trygdeordninger underlagt offentlig forvaltning, undervisning og helse- og sosialtjenester lå på 48.600 kroner, 44.730 kroner og 42.230 kroner.

¹⁵ Antall sysselsatte mellom 15 og 74 år med minst 30 timer avtalt arbeidstid i uka for de to årene var på henholdsvis 131.890 og 130.757.

fra år 2000 til 2008 og for næringen finansiering og forsikring for år 2009 og 2010. Tallene fra 2000 til 2008 er fra næringer med næringskode 65-67, mens tallene for næringene for 2009 og 2010 har kode 64-66. Variabelen vil vise til andelen i den private sektoren som jobber innenfor finansiering og forsikringsnæringen. Næringskodene for 2018 er like som fra 2009 og 2010, noe som kan tyde på at dette omhandler de samme personene. Selv med ulike næringskoder er antall fulltidssatte i næringen relativt lik i 2008 og 2009, noe som kan tyde på at de samme personene ble tatt med i beregningene.¹⁶

5.3 Instrumentvariabel

På bakgrunn av mulige simultanitetsproblemer, beskrevet i avsnitt 3.1.1, er det tatt med en variabel som har som hensikt å instrumentere for dagens befolkningsnivå. Gitt spesifikasjonene en instrumentvariabel skal ha, relevant for variabelen den skal instrumentere, samt ingen relasjon til avhengig variabel, har valget falt på å bruke befolkningsnivået for regionene i 1875. Disse befolkningstallene kommer fra en folketelling 31. desember 1875. Metoden som ble benyttet var at å se på antall «hjemmehørende personer», det vil si personer som på det tidspunktet hadde fast bosted innenfor tellingsenheten. Personer som var midlertidig fraværende ble også medregnet.

5.4 Regioner

For å best kunne analysere befolkningens effekt på produktiviteten i den offentlige og private sektoren har kommunedataen blitt aggregert opp på et regionsnivå med totalt 89 enheter¹⁷. Inndelingen av regionene følger fra EUs NUTS 4-nivå, et nivå mellom kommune og fylke, med krav om at regionene består av hele sammenhengende kommuner som ikke krysser fylkesgrenser. Disse regionene er knyttet til områdenes økonomiske forhold (SSB 2000) ved blant annet å studere handels- og arbeidsmarkedet. Dette har blitt gjort ved å se på pendlingstabeller for å se sammenhengen mellom bosteds- og arbeidskommune, hvor

¹⁶ Antall sysselsatte mellom 15 og 74 år med minst 30 timer avtalt arbeidstid i uka for de to årene var på henholdsvis 44.859 og 44.335.

¹⁷ I SSB (2000) er det blitt benyttet 90, men Søndre og Nordre Sunnhordland har senere blitt slått sammen.

omsetningstall har blitt brukt som et supplement. I tillegg til dette har det blitt sett på innbyggertall i kommunens største tettsteder, noe som har blitt ansett som et svakere kriterium. Ved tilfeller der kommuner fortsatt ikke har blitt tilegnet en region, har lokalkunnskap blitt benyttet for å avgjøre en kommunes tilhørighet.

Siden SSB sin inndeling ikke godtar at regioner innehar kommuner fra mer enn et fylke, har Bhuller (2009) kommet opp med en ny inndeling av regioner. De samme senterkommunene blir brukt, bare med en liten justering av kriteriene. Hovedkriteriet er at kommuner med mer enn 10 prosent utpendling til en naboregion slås sammen med regionen. Det kan også bli en sammenslåing hvis kommunen har tilstrekkelig utpendling til en kommune som allerede har slått seg sammen med naboregionen. Det blir benyttet et mer utvidet regionsbegrep ved at kommuner tilknyttet en småregion kan slås sammen med en større region hvis et av sentrene i småregionen har betydelig utpendling til et av sentrene i storregionen. I tillegg må hver arbeidsmarkedsregion ha minst 17.500 bosatte i år 2000. Disse kriteriene fører til en ny inndeling på 46 regioner.

Basert på kriteriene blir noen av regionene til Bhuller (2009) ganske store, noe som kan svekke funksjonaliteten. I tillegg var det i 2000 30 regioner som hadde færre enn 17.500 innbyggere, noe som også svekker funksjonaliteten. Med bakgrunn i dette ble SSB sin inndeling fra 2010 benyttet i denne oppgaven.

5.5 Deskriptiv statistikk

Tabellen under viser deskriptiv statistikk for variablene som vil bli benyttet i den fremtidige analysen. Alle lønnsdata er i norske kroner.

Tabell 5.1. Deskriptiv statistikk. Tabellen viser en oversikt over gjennomsnittlig verdi, standardavvik og høyeste og laveste verdi på variablene som ble inkludert i analysen.

	Gj. snitt	St. avvik	Min	Maks
Dagslønn off. sektor (1993-2010)	1099,50	162,12	764,84	1588,88
Dagslønn priv. sektor (1993-2010)	1175,77	209,32	748,45	2038,92
Befolkning (1993-2010)	50905,8	74420,49	5219	586860
Andel høyere utdanning (1993-2010)	0,181	0,053	0,08	0,447
Andel offentlig adm. (2000-2010)	0,255	0,069	0,146	0,634
Andel finans og forsikring (2000-2010)	0,022	0,015	0,004	0,134
Andel kvinner (1993-2010)	0,505	0,008	0,478	0,534
Andel innvandrere (1993-2010)	0,039	0,023	0,006	0,209
Befolkning 1875	20095,51	18162,23	1628	107833

Ut ifra tabell 5.1 kan det observeres at gjennomsnittlig dagslønn er høyere i privat sektor enn i offentlig, 1175,77 mot 1099,50. Videre kan det også observeres at lønnsnivået varierer mer i privat sektor som følge av et høyere standardavvik. I 2010 var det høyeste lønnsnivået for begge sektorene i regionen Bærum/Asker, hvor offentlige dagslønninger var i snitt 1588,88 mens de var i snitt 2038,92 i den private sektoren. Regionen med den laveste gjennomsnittlige dagslønnen i offentlig sektor i 2010 var i Nord-Troms, med 1159,10. Lavest for privat sektor var i Nord-Gudbrandsdalen med 1175,87. Videre har gjennomsnittlig lønn i privat sektor økt fra 939,19 i 1993 til 1442,33 i 2010. For offentlig sektor har det gått fra 888,03 til 1358,10.

Totalt gjennom hele tidsperioden lå gjennomsnittlig befolkningsnivå i regionene på ca. 50.906. Oslo hadde størst befolkningsnivå i 2010 med 586.860, mens lavest hadde Grong på 5219. I 2010 var det sju regioner; Bergen, Bærum/Asker, Drammen, Lillestrøm, Oslo, Stavanger/Sandnes og Trondheim, som hadde over 150.000 innbyggere. Vi kan her se at fire av regionene, Oslo, Bergen, Stavanger/Sandnes og Trondheim, er de fire største tettstedene i Norge (SSB 2018b), mens resten av regionene er i pendlingsavstand til Oslo. Gjennomsnittlig befolkningsstørrelse i alle regioner har økt fra 48305 til 54531 fra 1993 til 2010.

For andelen med høyere utdanning lå gjennomsnittsnivået på 18,1% i alle regionene over hele tidsperioden. I 2010 var det høyest i Bærum/Asker med 44,7%. Frøya/Hitra hadde lavest det samme året med 14,5%. I 1993 var andelen personer som tok høyere utdanning på gjennomsnittlig 13,8%, mens tallet for 2010 var på 22,2%.

I snitt jobbet 25,5% av offentlig ansatte innenfor næringen offentlig administrasjon, forsvar og sosialforsikring. Den høyeste andelen som ble målt i 2010 var i Brekstad på 43%. Laveste andel samme år var i 14,7% i Ørsta/Volda.

For den private sektoren sto næringen finansering og forsikring for gjennomsnittlig 2,2% av de heltidsansatte over hele tidsperioden. Snittet har endret seg fra 1,9% i 1993 til 2,3% i 2010. Den høyeste andelen som ble målt i 2010 var i Bærum/Asker på 7,1%. Den laveste andelen var i Frøya/Hitra på 0,5%.

Andelen kvinner i norske regioner lå gjennomsnittlig over hele tidsperioden på 50,5%. Andelen var størst i 2010 i Rjukan på 51,8%, mens den var lavest i Frøya/Hitra 48,5% samme år.

Gjennomsnittlig hadde 3,9% av befolkningen over hele tidsperioden innvandrerbakgrunn. Dette gjennomsnittet har endret seg fra 2,4% i 1993 til 6,6% i 2010. Den høyeste andelen som ble målt i 2010 var i Oslo på 20,9%. Den laveste andelen samme år var i Surnadal på 2,1%.

6 Analyse

I denne delen av oppgaven vil resultatene fra de ulike modellene bli fremlagt. Alle regresjonene er gjennomført i statistikkprogrammet STATA/MP 15.1. Først vil paneldatamodellen for privat sektor blir fremlagt, før det samme vil bli gjort for den offentlige sektor. Videre vil interessevariabelen i regresjonene bli endret, hvor hensikten er en robusthetssjekk av resultatene fra de foregående regresjonene. For å se på utviklingen til elastisitetene over tid følger videre en tverrsnittsdatamodell, med regresjoner fra alle årene i tidsperioden 1993 til 2010. Tilslutt vil det bli innført en instrumentvariabel for å kontrollere for eventuelle endogenitetsproblemer knyttet til lønn og befolkningsstørrelsen.

6.1 Paneldata

Denne delen benytter seg av data fra hele tidsperioden. For å se på utviklingen i de ulike sektorene har det blitt kjørt to separate regresjoner. Basert på mulige problemer knyttet til sortering, nevnt i kapittel 3.1.2, vil regresjonene for offentlig og privat sektor inneholde kontrollvariabler som kan antas å påvirke lønnen i de ulike regionene. Hvorfor disse blir sett på som relevante er beskrevet for hver variabel i kapittel fem.

6.1.1 Privat sektor

Tabell 6.1. Paneldata privat sektor.

	Lnwage_pri v (1)	Lnwage_pri v (2)	Lnwage_pri v (3)	Lnwage_pri v (4)	Lnwage_pri v (5)
Lnbefolk	0,0708*** (0,0022)	0,0451*** (0,0028)	0,036*** (0,0028)	0,0387*** (0,0027)	0,0411*** (0,0034)
Høyere utdanning		0,7974*** (0,0572)	0,5161*** (0,0623)	0,7116*** (0,0613)	0,7176*** (0,0615)
Finans				-1,9989*** (0,157)	-0,813 (1,039)
Finans*befolk					-0,109 (0,0944)
Kvinner,innvandrere	Nei Ja	Nei Ja	Ja Ja	Ja Ja	Ja Ja
Årsdummier	1602	1602	1602	1602	1602
Observasjoner	0,771	0,796	0,811	0,829	0,829
R ²					

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Fra tabellen kan det observeres at estimatet for agglomerasjonselastisiteten i den private sektoren ligger på 7,08%. Tolkningen av dette er at en fordobling av befolkningsstørrelsen vil øke lønningene i den private sektoren med 7,08%.

Grunnet potensielle problemer knyttet til sortering av arbeidere finner vi i kolonne 2 en regresjon hvor andelen med høy utdanning i regionen har blitt tatt med. Estimatet for agglomerasjonselastisiteten gikk da ned til 4,51%. Koeffisienten til høyere utdanning variabelen forteller at en 100%-poengs økning i andelen personer med høy utdanning estimert øker lønnen med 79,7%. Denne økningen er ikke veldig reell, slik at det er mer nærliggende å observere at en standardavviks økning, 0,053 eller 5,3%-poeng, fører til en estimert økning på 4,2% i lønnen gitt at alt annet holdes konstant.

Ved inkludering av andel kvinner og personer med innvandrerbakgrunn faller elastisiteten enda mer til 3,6%. Estimaten for både andel kvinner og innvandrere har en positiv koeffisient.

Andelen heltidsansatte innenfor næringen finansierings og forsikringsvirksomhet har ingen stor påvirkning på elastisiteten. Estimatet øker til 3,87%, noe som kan være en indikasjon på

at denne effekten allerede har blitt kontrollert for ved de andre forklaringsvariablene.

Koeffisienten foran variabelen er sterkt negativ samtidig som den er signifikant, noe som er motsatt av det som er forventet.

Modellen med et interaksjonsledd mellom andel innenfor finansiering- og forsikringsnæringen og befolkningsstørrelsen gir et estimat på agglomerasjonselastisiteten lik 3,94% ved bruk av gjennomsnittsverdien på finansvariabelen på 1,6%. Det kan bemerkes at dette interaksjonsleddet ikke er signifikant.

6.1.2 Offentlig sektor

Tabell 6.2. Paneldata offentlig sektor.

	Lnwage_pub (1)	Lnwage_pub (2)	Lnwage_pub (3)	Lnwage_pub (4)	Lnwage_pub (5)
Lnbefolk	0,0393*** (0,0012)	0,0227*** (0,0014)	0,0188*** (0,0015)	0,0191*** (0,0015)	0,0306*** (0,0047)
Høyere utdanning		0,5157*** (0,0293)	0,4605*** (0,032)	0,4546*** (0,0318)	0,4492*** (0,0318)
Offentlig adm				0,0632*** (0,0143)	0,5139*** (0,1747)
Off.adm*befolk					-0,045** (0,0174)
Kvinner,innvandrere	Nei	Nei	Ja	Ja	Ja
Årsdummier	Ja	Ja	Ja	Ja	Ja
Observasjoner	1602	1602	1602	1602	1602
R ²	0,909	0,924	0,929	0,93	0,93

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Agglomerasjonselastisiteten for den offentlige sektoren blir her estimert til 3,93% når det ikke har blitt inkludert noen kontrollvariabler. Dette er et betydelig lavere estimat enn det for privat sektor.

Når andelen personer med høy utdanning tas med, faller elastisiteten ned til 2,27%. På samme måte som for den private sektoren kan dette være en indikasjon på at de større regionene tiltrekker seg flere høyt utdannende personer. Den estimerte lønnsøkningen for personer med høy utdanning er lavere for heltidsansatte i offentlig sektor. En økning i andelen på 10%-poeng vil estimert øke lønnen med 5,1%.

Når andel kvinner og personer med innvandrerbakgrunn også tas med får vi et videre fall ned til 1,88%. Koeffisientene foran disse variablene er på lik linje med de for privat sektor positive.

Med inkludering av andelen i offentlig sektor som jobber innenfor næringen offentlig administrasjon økes elastisiteten til 1,91%. Koeffisienten foran variabelen viser at en 10% økning av heltidsansatte i offentlig sektor innenfor næringen for offentlig administrasjon estimert øker lønnen til offentlig ansatte med 0,6%. Dette er en signifikant effekt ned på ett 1%-nivå.

Interaksjonsleddet mellom *lnbefolkning* og andel offentlig administrasjon er signifikant ned på et 5%-nivå. Ved å benytte gjennomsnittsverdien til variabelen andel personer som jobber innenfor offentlig administrasjon og forsvar, 0,255, kommer en fram til en elastisitet på 1,91%. Dette er likt som før inkluderingen av interaksjonsleddet. Ettersom koeffisienten til interaksjonsleddet er negativ vil agglomerasjonselastisiteten estimert gå ned ved en økning på andelen offentlige arbeidstakere som jobber i offentlig administrasjon. Estimert vil den for eksempel være 1,1% for Brekstad som hadde den høyeste andelen av heltidsansatte i offentlig administrasjon i 2010.

Resultatene fra regresjonene gir både forventede og uforventede resultater. Det at den statiske agglomerasjonselastisiteten er høyere i privat enn i offentlig sektor er på lik linje med forventningene. Elastisiteten for privat sektor er også relativt lik estimatene fra tidligere forskning. Videre senkes elastisiteten for begge sektorene ved inkludering av forklaringsvariabler som kan forklare observerbare karakteristikk som er ulike for arbeidere i de ulike regionene. Dette kan tyde på et sorteringsproblem ved at arbeidsstyrken i de ulike regionene ikke er homogen.

Det som derimot ikke er helt forventet er koeffisientene foran flere av forklaringsvariablene. Det at både kvinneandel og innvandrerandel har en positiv innvirkning på lønnen i begge sektorene er stikk i strid med lønnstallene fra SSB (2018a). En eventuell forklaring på dette kan være at både andelen kvinner og innvandrere er høyere for de sju mest befolkede regionene som beviselig har høyere lønn enn de mindre befolkede.¹⁸

¹⁸ Se appendiks B.1 for tabell.

6.1.3 Storregion-dummy

Her har befolkningsstørrelsen blitt erstattet med en storregion-dummy som innebefatter regioner som i år 2010 hadde over 150.000 innbyggere. Denne metoden er lignende den brukt av Carlsen et. al (2016) hvor de sammenlignet lønnen i Oslo med seks andre storbyer og resten av landet. Modellene vil bruke de samme kontrollvariablene som i standardmodellene.

Tabell 6.3. Paneldata privat sektor storregion-dummy.

	Lnwage_priv (1)	Lnwage_priv (2)	Lnwage_priv (3)	Lnwage_priv (4)
Storregion	0,1995*** (0,0087)	0,0864*** (0,0097)	0,0631*** (0,0097)	0,0733*** (0,0095)
Høyere utdanning		0,7974*** (0,0572)	0,7248*** (0,0619)	0,8567*** (0,0626)
Finans	Nei	Nei	Nei	Ja
Kvinner,innvandrere	Nei	Nei	Ja	Ja
Årsdummier	Ja	Ja	Ja	Ja
Observasjoner	1602	1602	1602	1602
R ²	0,72	0,796	0,797	0,806

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Tabell 6.4. Paneldata offentlig sektor storregion-dummy.

	Lnwage_pub (1)	Lnwage_pub (2)	Lnwage_pub (3)	Lnwage_pub (4)
Storregion	0,0844*** (0,005)	0,0004 (0,0051)	-0,0058 (0,005)	-0,0047 (0,005)
Høyere utdanning		0,817*** (0,0297)	0,6603*** (0,0323)	0,6551*** (0,0322)
Offentlig adm				0,0547*** (0,0143)
Off.adm*befolk				
Kvinner,innvandrere	Nei	Nei	Ja	Ja
Årsdummier	Ja	Ja	Ja	Ja
Observasjoner	1602	1602	1602	1602
R ²	0,869	0,912	0,921	0,922

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Av resultatene her kan det observeres at privat sektor har et lønnspremium for de sju største regionene, i forhold til resten av regionene. Dette er derimot ikke tilfellet for de samme sju regionene for den offentlige sektoren når det har blitt inkludert kontrollvariabler. Dette kan tyde på at det er forskjellene i befolkningsstørrelsen utenfor de sju største regionene som driver agglomerasjonseffekten i offentlig sektor.

6.1.4 Konklusjon paneldatamodell

Det har i denne delen blitt undersøkt om agglomerasjonselastisiteten er lik for de to sektorene. Videre har det blitt kontrollert for eventuelle observerbare karakteristikk som potensielt kan forklare ulikheter mellom regionens arbeidere. Det har i tillegg blitt inkludert årsdummier for å kontrollere for eventuelle endringer i den økonomiske situasjonen i de ulike årene. På bakgrunn av at dataen i oppgaven ikke er av typen individdata har det ikke blitt benyttet individfaste effekter. Dette medfører potensielle problemer knyttet til skjevhet i uobserverbare karakteristikk til arbeiderne i de ulike regionene. Det ble på et tidspunkt inkludert en dummy for regionene for å kontrollere for uobserverbare karakteristikk mellom regionene, men resultatene av dette var lave verdier på alle koeffisientene.¹⁹ Årsaken til dette kan være at man

¹⁹ Se appendiks B.2 for tabell.

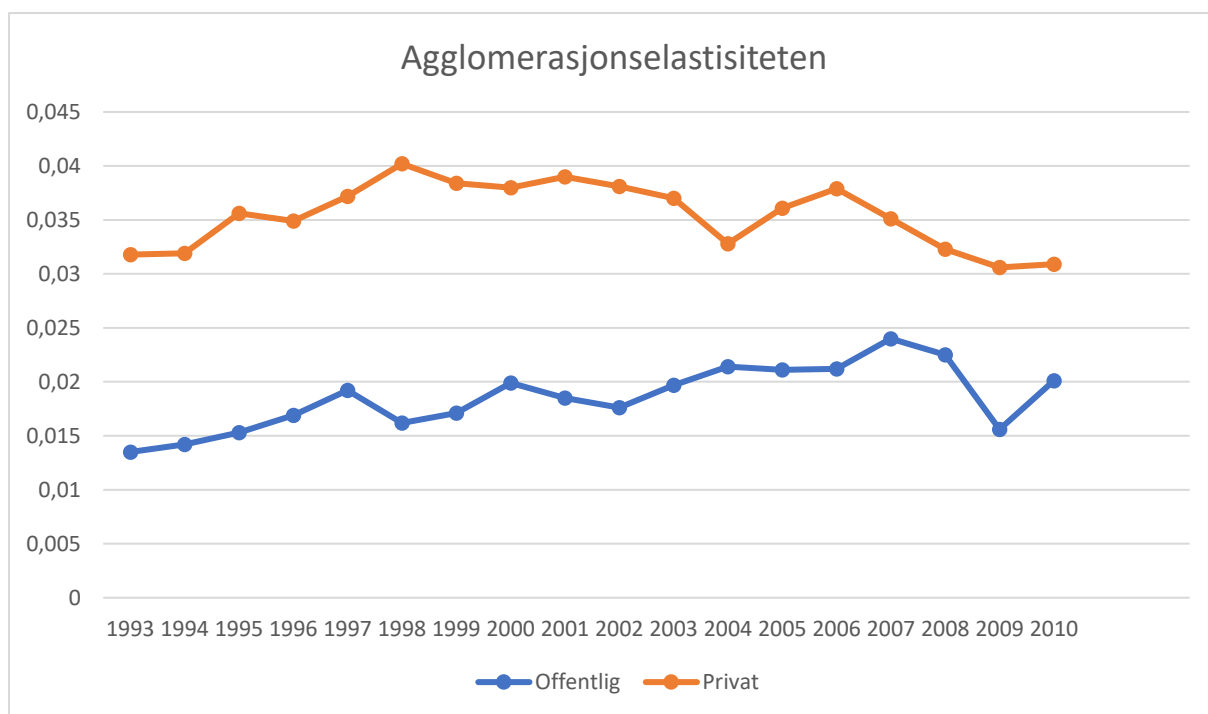
ved faste effekter kun ser på personer som migrerer mellom de geografiske områdene, og det kan være slik at det over tid er relativt lite av dette mellom de ulike regionene.

Instrumentvariabler har heller ikke blitt brukt da dette er en ganske komplisert prosess for paneldata.

Den viktigste konklusjonen fra paneldataen er at resultatene tilsier en større agglomerasjonseffekt i privat sektor enn i offentlig sektor. Det finnes også et lønnspremium for de sju storregionene i privat sektor, mens dette ikke er tilfellet for offentlig sektor.

6.2 Tverrsnittsdata

For å se på utviklingen til agglomerasjonselastisiteten over tid i de to sektorene vil det i denne delen bli benyttet tverrsnittsdata for hvert enkelt år i tidsperioden. Basert på regresjonene i avsnitt 6.1, vil modellene hvor andel høyere utdanning, kvinner og personer med innvandrerbakgrunn i regionene har blitt brukt som forklaringsvariabler i tillegg til befolkningsstørrelsen bli benyttet. Andel heltidsansatte i næringene med høy lønnsvekst, offentlig administrasjon for offentlig sektor, og finansiering og forsikring for privat sektor, er også tatt med sammen med et interaksjonsledd mellom disse og befolkningsstørrelsen i de ulike regionene det samme året. Da det er utviklingen av agglomerasjonselastisiteten i de to sektorene over tid som er problemstillingen til oppgaven, er det disse som vil bli sammenlignet.



Figur 6.1. Agglomerasjonselastisiteten i offentlig og privat sektor 1993-2010.

Figuren sammenligner agglomerasjonselastisiteten i offentlig og privat sektor over tidsperioden 1993 til 2010. Resultatet tilsier en høyere agglomerasjonselastisitet i den private sektoren, som sammenfaller med resultatene fra paneldatamodellene. Utviklingen ser ut til å være positiv for begge sektorene fram til finanskrisen i 2008.

For å undersøke det mulige identifiseringsproblemet vil det videre bli brukt en instrumentvariabel på elastisiteten for år 2010.

6.2.1 IV-estimering

Metoden til Ciccone og Hall (1996), ved bruk av historisk befolkningsdata som instrument for nyere befolkningsdata, vil bli benyttet i denne delen. Dette vil kunne finne et eventuelt identifiseringsproblem mellom lønn og befolkningsstørrelse, og dermed om de er simultant bestemt. Historisk befolkningsdata har blitt hentet inn fra NSD sin kommunedatabase.

Combes et al. (2010) sin idé om å benytte seg av geologiske faktorer som kvalitet på jord og meter over havet vil ikke bli benyttet da dette er noe det er vanskelig å finne gode tall på.

Årstallet som har blitt valgt som instrument er 1875, da dette kan være langt nok tilbake i tid til å ikke påvirke lønnen i år 2010 samtidig som det kan antas å ha en sammenheng med befolkningsnivået det samme året. Årsaken til at det ikke har blitt benyttet befolkningstall for årene før 1875 er at folketellingene i disse er nokså usikre, noe som vil kunne resultere i potensielle målefeil. Videre vil det også være usikkerhet knyttet til om det er for langt tilbake i tid til å ha sammenheng med befolkningsnivået i 2010. For årene etter 1875 vil problemet kunne være at de kan påvirke lønnsnivået i 2010, som vil føre at de ikke er et gyldig instrument.

Først vil det første steget i IV-prosessen for å undersøke hvor godt instrumentet er.

Tabell 6.5. Første-steg IV-estimering 2010.

	Lnbefolk (1)	Lnbefolk (2)	Lnbefolk (3)
Lnbefolk1875	1,1201*** (0,0798)	0,908*** (0,061)	0,8769*** (0,0618)
Høyere utdanning		8,2217*** (0,882)	6,9645*** (1,0415)
Andel kvinner og innvandrere	Nei	Nei	Ja
Observasjoner	89	89	89
R ²	0,694	0,848	0,857
F	196,98	239,18	126,14

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Den første-stegs IV-estimeringen viser oss hvor sterkt instrumenter er. Basert på Stock and Yogo (2005) anbefales det at F-statistikken ligger på minst 10 for instrumentvariablene. Tabell 6.5 viser at dette er tilfellet for alle første-steg i regresjonen, noe som indikerer at instrumentet er relevant.

Gitt at instrumentet virker å holde de gitte kriteriene vil det nå undersøkes hvilken effekt inkluderingen av dette har på estimatene for agglomerasjonselastisiteten for år 2010.

Fremgangsmåten her vil være å først presentere resultatene fra tverrsnittsdatamodellene for år 2010 for henholdsvis privat og offentlig sektor uten instrumentvariabelen, for deretter å inkludere denne. Estimaten fra disse regresjonene vil så bli sammenlignet.

Tabell 6.6. Tverrsnittsdata 2010 privat sektor.

	Lnwage_priv (1)	Lnwage_pub (2)	Lnwage_pub (3)
Lnbefolk	0,0663*** (0,0091)	0,0408*** (0,0118)	0,0309** (0,0122)
Høyere utdanning		0,7351*** (0,2297)	0,5312*** (0,2403)
Andel kvinner og innvandrere	Nei	Nei	Ja
Observasjoner	89	89	89
R ²	0,375	0,442	0,483

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Tabell 6.7. Tverrsnittsdata 2010 offentlig sektor.

	Lnwage_pub (1)	Lnwage_pub (2)	Lnwage_pub (3)
Lnbefolk	0,0389*** (0,0044)	0,0227*** (0,0055)	0,0201*** (0,0058)
Høyere utdanning		0,465*** (0,1064)	0,4028*** (0,1136)
Andel kvinner og innvandrere	Nei	Nei	Ja
Observasjoner	89	89	89
R ²	0,468	0,565	0,581

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Tabell 6.8. Andre-steg IV-estimering privat sektor.

1875 IV	Lnwage_priv (1)	Lnwage_priv (2)	Lnwage_priv (3)
Lnbefolk	0,0556*** (0,011)	0,0378*** (0,0137)	0,0296*** (0,0141)
Høyere utdanning		0,774*** (0,245)	0,5432*** (0,2441)
Andel kvinner og innvandrere	Nei	Nei	Ja
Observasjoner	89	89	89
R ²	0,366	0,441	0,483

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Tabell 6.9. Andre-steg IV-estimering offentlig sektor.

1875 IV	Lnwage_pub (1)	Lnwage_pub (2)	Lnwage_pub (3)
Lnbefolk	0,0381*** (0,0053)	0,0294*** (0,0064)	0,0274*** (0,0067)
Høyere utdanning		0,3773*** (0,1145)	0,3353*** (0,1164)
Andel kvinner og innvandrere	Nei	Nei	Ja
Observasjoner	89	89	89
R ²	0,468	0,557	0,573

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

Ved å sammenligne estimatene for agglomerasjonselastisiteten i privat sektor for år 2010 med og uten instrumentvariabler viser tabell 6.6 og 6.8 at effekten ved kun inkludering av historisk befolkningsdata for 1875 at estimatet minker fra 6,6 til 5,6%. Dette sammenfaller med effekten Combes et al. (2010) fant med sin franske data, hvor elastisiteten ble redusert fra 4,8 til 4%. Sammenligning av kolonne (2) og (3) i de samme tabellene viser en mindre reduksjon ved inkludering av flere instrumentvariabler, hvor kun å inkludere høyere utdanning gir en nedgang fra 4,1 til 3,8% mens å tillegg inkludere andel kvinner og innvandrere gir et fall fra 3,1 til 3%. Disse resultatene kan indikere på et større problem knyttet til sortering av arbeidere enn det knyttet til identifisering. Dette samsvarer også med resultatene til Combes et al. (2010).

Tabell 6.7 og 6.9 viser estimatene for offentlig sektor henholdsvis uten og med instrumentvariabel og indikerer at den estimerte effekten av agglomerasjon ikke minker nevneverdig ved inkludering av kun historisk befolkningsdata for 1875 som instrumentvariabel. Ved å inkludere flere instrumentvariabler viser det seg at agglomerasjonselastisiteten øker, noe som er motsatt av det som ble funnet for privat sektor. Årsaken til dette er uviss, men det kan indikere på at modellen mangler relevante variabler.

6.2.2 Konklusjon tverrsnittsdatamodell

Tverrsnittsdatamodellene ble inkludert for å undersøke utviklingen til agglomerasjonselastisiteten over tid i de to sektorene. Resultatene tilsier at effekten har historisk sett vært høyere i privat enn offentlig sektor.

Det viser seg derimot at en inkludering av instrumentvariabel påvirker elastisitetene i de to sektorene ulikt. For privat sektor minker estimatene i langt mindre grad ved inkludering av instrumentvariablene enn de gjør for kontroll av observerbare karaktrestikker. Det at inkludering av instrumentvariabler i offentlig sektor gir et økt estimat for agglomerasjonselastisiteten tyder på at modellen ikke er fullkommen. Uansett tilsier resultatene at sortering av arbeidere er et større problem i estimeringen av agglomerasjonselastisiteten i begge sektorer.

7 Konklusjon

Formålet med denne oppgaven har vært å bidra til den relativt tynne forskningen på den statiske agglomerasjonselastisiteten i offentlig sektor. Dette har blitt gjort ved å sammenligne denne med den tilsvarende i privat sektor. Videre har det blitt studert om forskjellen mellom sektorene har endret seg over tid.

For å besvare problemstillingen knyttet til sammenhengen mellom den statiske agglomerasjonselastisiteten i privat og offentlig sektor, har data fra SSB sin statistikkbank og NSD sin kommunedata blitt innhentet. I tillegg til dette har det blitt benyttet lønnsdata for 89 økonomiske regioner i Norge.

Resultatene fra paneldatamodellene tyder på at agglomerasjonseffektene også er tilstede i begge sektorer, men at den er sterkere i den private. Dette underbygges ved en undersøkelse av lønnsnivået i sju storregioner mot resten, som viser en signifikant forskjell for privat sektor mens lønnsnivået tilsynelatende er likt i offentlig sektor. utfordringer knyttet til sortering og identifisering av arbeidere virker å begge være til stede for privat sektor, hvor effekten av sortering virker å være størst. Offentlig sektor virker å ha en lik sorteringseffekt, mens effekten knyttet til identifisering virker å være fraværende.

Siden dataen er på et regionsnivå har det ikke vært mulig å inkludere individfaste effekter for å kontrollere for eventuelle uobserverbare karaktrestikker knyttet til arbeiderne i de ulike regionene. Dette svekker troverdigheten til resultatene som følge av at dette historisk beviselig har vært tilfellet ved flere anledninger.

7.1 Videre forskning

På bakgrunn av manglende forskning på agglomerasjonselastisiteten i offentlig sektor har denne oppgaven basert seg på forskning med data fra privat sektor. Det vil derfor være ønskelig at det videre forskes på agglomerasjonseffektene på offentlig sektor, samtidig som man finner. Videre kan det benyttes individdata ettersom dette øker mulighetene for forskningen.

8 Referanser

- Bhuller, M., S. (2009). Inndeling av Norge i arbeidsmarkedsregioner. (2009/24). SSB: Oslo, Norge.
- Carlsen, F., Rattsø, J. & Stokke, H., E. (2016). Education, experience, and urban wage premium. *Regional Science and Urban Economics*, 60: 39–49.
- Ciccone, A., og Hall, R. (1996). Productivity and the Density of Economic Activity. *American Economic Review*, 86(1): 54-70
- Combes, P.-P., Duranton, G., Gobillon, og Roux, S., (2010). Estimating Agglomeration Economies with History, Geology, and Worker Effects. *Agglomeration Economics*, ed. Glaeser., E.L. Chicago, US: University of Chicago Press
- De la Roca, J., og Puga D., (2017). Learning by Working in Big Cities. *Review of Economic Studies*, 84: 106-142
- Dekle, R. & Eaton, J. (1999). Agglomeration and land rents: evidence from the prefectures. *Journal of Urban Economics*, 46(2): 200–214.
- Duranton, G. & Puga, D. (2004): «Micro-foundations of Urban Agglomeration», *Handbook of Urban and Regional Economics*, vol. 4 ed. V. Henderson and J.-F. Thisse, 2063-117. Amsterdam
- Glaeser, E. L. (2011). *Triumph of the City: How Our Greatest Invention Makes Us Richer, Smarter, Greener, Healthier, and Happier* (London: MacMillan)
- Glaser, E.L. & Maré, D.C. (2001). Cities and Skills. *Journal of Labour Economics*, 19(2): 316- 342.
- Helsley, R., W. and Strange, C., W. (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics* 20(2):189–212.
- Henderson, J., V. (1986). Efficiency of Resource Usage and City Size. *Journal of Urban Economics* 19(1), pp. 47-70.
- Jovanovic, B. (1997). Learning and growth. In David M. Keps and Kenneth F. Wallis (eds.) *Advances in Economics and Econometrics: Theory and applications*, volume 2. Cambridge: Cambridge University Press, 318–339.
- Kim, S. (1989). Labor specialization and the extent of the market. *Journal of Political Economy* 97(3):692–705.

- Krugman, P., R. (1991). *Geography and Trade*. Cambridge, MA: MIT Press
- Marshall, A. (1920). *Principles of Economics*, 8th Ed. London: MacMillan
- Melo, P.C., Graham D., J., og Noland R., B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics* 39: 332-342.
- Moomaw, R. L. (1981). Productivity and citysize: a critique of the evidence. *The Quarterly Journal of Economics*, 96(4), 675–688.
- Nakamura, R. (1985). Agglomeration Economies in Urban Manufacturing Industries: A Case of Japanese Cities. *Journal of Urban Economics*, 17(1):108-124
- Smith, A. (1776). *An inquiry into the nature and causes of the wealth of nations: Volume one*.
- SSB (2000). *Standard for økonomiske regioner*. SSB: Oslo, Norge.
- SSB (2018a). Gjennomsnittlig månedslønn for kvinner og menn i ulike sektorer. Tilgjengelig fra: <https://www.ssb.no/lonnansatt>
- SSB (2018b). Folkemengde og areal i tettsteder 1. januar. Tilgjengelig fra: <https://www.ssb.no/befolkning/statistikker/bef tett/aar>
- Stock, J., Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. *Identification and Inference for Econometric Models: Essay in honor of Thomas Rothenberg*, 80-108. Cambridge: Cambridge University Press.
- Sveikauskas, L. (1975). The productivity of cities. *The Quarterly Journal of Economics*, 89(3), 393–413.
- Venables, A., J. (2002). *Labour sorting by cities: Partnerships, self-selection, and agglomeration*. Processed, London School of Economics.
- Wooldridge, J., M. (2015). *Introductory Econometrics: A Modern Approach* (6 utg.). Boston: Cengage Learning.

Appendiks

A Utledninger teoridel

A.1 Etterspørselen etter mellomliggende goder

Ser her på et tilfelle med to mellomliggende goder.

$$q_1 X_1 + q_2 X_2 \quad (1)$$

Gitt produktfunksjonen

$$Y = \left(x_1^{\frac{1}{1+\varepsilon}} + x_2^{\frac{1}{1+\varepsilon}} \right)^{1+\varepsilon} \rightarrow Y^{\frac{1}{1+\varepsilon}} = x_1^{\frac{1}{1+\varepsilon}} + x_2^{\frac{1}{1+\varepsilon}} \quad (2)$$

Vil bruke Lagrange for å minimere uttrykket:

$$L = q_1 x_1 + q_2 x_2 - \lambda \left(Y^{\frac{1}{1+\varepsilon}} - x_1^{\frac{1}{1+\varepsilon}} - x_2^{\frac{1}{1+\varepsilon}} \right) \quad (3)$$

Dette gir følgende førsteordensbetingelser:

$$\frac{dL}{dx} = q_1 + \frac{1}{1+\varepsilon} * \lambda x_1^{\left(\frac{1}{1+\varepsilon}\right)-1} = 0 \quad (4)$$

Ved å løse for λ fås:

$$\lambda = -q_1(1+\varepsilon)x_1^{\frac{\varepsilon}{1+\varepsilon}} \quad (5)$$

Gitt at de to er identiske blir det også:

$$\lambda = -q_2(1+\varepsilon)x_2^{\frac{\varepsilon}{1+\varepsilon}} \quad (6)$$

Kan ta sette $\lambda = \lambda$ å få:

$$-q_1(1+\varepsilon)x_1^{\frac{\varepsilon}{1+\varepsilon}} = -q_2(1+\varepsilon)x_2^{\frac{\varepsilon}{1+\varepsilon}} \rightarrow x_1 = \left(\frac{q_2}{q_1}\right)^{\frac{1+\varepsilon}{\varepsilon}} * x_2 \quad (7)$$

Uttrykket for X_1 settes inn i (2):

$$Y^{\frac{1}{1+\varepsilon}} = \left(\left(\frac{q_2}{q_1}\right)^{\frac{1+\varepsilon}{\varepsilon}} * x_2 \right)^{\frac{1}{1+\varepsilon}} + x_2^{\frac{1}{1+\varepsilon}} \quad (8)$$

Kan sette $X_2^{\frac{1}{1+\varepsilon}}$ utenfor en parentes:

$$Y^{\frac{1}{1+\varepsilon}} = x_2^{\frac{1}{1+\varepsilon}} \left(1 + \left(\frac{q_2}{q_1}\right)^{\frac{1}{\varepsilon}} \right) \quad (9)$$

Løser for x_2 :

$$x_2^{\frac{1}{1+\varepsilon}} = \frac{Y^{\frac{1}{1+\varepsilon}}}{1 + \left(\frac{q_2}{q_1}\right)^{\frac{1}{\varepsilon}}} \quad (10)$$

Ved å gange teller og nevner med $\frac{q_2^{\frac{1}{\varepsilon}}}{q_2^{\frac{1}{\varepsilon}}}$ blir det:

$$x_2^{\frac{1}{1+\varepsilon}} = \frac{Y^{\frac{1}{1+\varepsilon}} * q_2^{-\frac{1}{\varepsilon}}}{q_2^{\frac{1}{\varepsilon}} + q_1^{\frac{1}{\varepsilon}}}$$

$$x_2 = \frac{Y * q_2^{-\frac{1+\varepsilon}{\varepsilon}}}{\left(q_1^{\frac{1}{\varepsilon}} + q_2^{\frac{1}{\varepsilon}}\right)^{1+\varepsilon}} \quad (11)$$

Ved å sette (11) inn i (7) blir uttrykket:

$$x_1 = \left(\frac{q_2}{q_1}\right)^{\frac{1+\varepsilon}{\varepsilon}} * \frac{Y * q_2^{-\frac{1+\varepsilon}{\varepsilon}}}{\left(q_1^{\frac{1}{\varepsilon}} + q_2^{\frac{1}{\varepsilon}}\right)^{1+\varepsilon}} \quad (12)$$

Hvor en strykning av q_2 i leddet foran med den i telleren gir:

$$x_1 = \frac{Y * q_1^{-\frac{1+\varepsilon}{\varepsilon}}}{\left(q_1^{\frac{1}{\varepsilon}} + q_2^{\frac{1}{\varepsilon}}\right)^{1+\varepsilon}} \quad (13)$$

En generell formalisering av denne gir uttrykk (4) i teoridelen:

$$x^j(h) = \frac{\left([q^j(h)]^{-\frac{1+\varepsilon^j}{\varepsilon^j}} Y^j\right)}{\left\{\int_0^{n^j} [q^j(h')]^{-\frac{1}{\varepsilon^j}} dh'\right\}^{1+\varepsilon^j}}$$

A.2 Utregning av pris på innsatsfaktorer

Skal nå bruke (2) og (5) fra teoridelen til å vise hvordan vi kommer fram til (6) (dropper subskrift j):

$$x^j = \beta^j l^j - a^j \rightarrow l^j = \frac{x^j + a^j}{\beta^j} = \frac{x^j}{\beta^j} + \frac{a^j}{\beta^j}$$

Setter dette inn for l i (5):

$$\pi^j = q^j * x^j - w^j \left(\frac{x^j}{\beta^j} + \frac{a^j}{\beta^j} \right)$$

Setter x utenfor parentesen:

$$\pi^j = x^j \left(q^j - \frac{w^j}{\beta^j} \right) - \frac{w^j * a^j}{\beta^j}$$

Finner videre profittmaksimerende pris å sette den deriverte av denne lik 0:

$$\frac{d\pi^j}{dq^j} = 0 \rightarrow x^j + \left(q^j - \frac{w^j}{\beta^j} * \frac{dx^j}{dq^j} \right) = 0$$

Vi har tidligere at elastisiteten av x^j med hensyn på q^j er $-\left(\frac{1+\varepsilon^j}{\varepsilon^j}\right)$ slik at $\frac{dx^j}{dq^j} = -\left(\frac{1+\varepsilon^j}{\varepsilon^j}\right) * \frac{x^j}{q^j}$ som gir:

$$x^j + q^j - \frac{w^j}{\beta^j} * (-) \frac{(1 + \varepsilon^j)}{\varepsilon^j} * \frac{x^j}{q^j} = 0$$

Løser så alt for q^j :

$$q^j = \frac{\left(\frac{w^j}{\beta^j}\right)}{\frac{x^j * \varepsilon^j * q^j}{(1 + \varepsilon^j)x^j}}$$

Stryker x^j – ene og får:

$$q^j = \frac{w^j}{\beta^j} + \frac{\varepsilon^j * q^j}{1 + \varepsilon^j}$$

Ved å omorganisere får man:

$$q^j \left(1 - \frac{\varepsilon^j}{1 + \varepsilon^j} \right) = \frac{w^j}{\beta^j}$$

Ved å løse denne parentesen og flytte over får vi:

$$q^j = \frac{(1 + \varepsilon^j)}{\beta^j} w^j$$

Som er (6) i teoridelen.

A.3 Produksjonsnivå ved nullprofitt

Ved å benytte oss av (2) og (6) samt anta nullprofitt skal det utledes hvordan man kommer fram til (7) i teoridelen:

$$0 = \left(\frac{1+\varepsilon^j}{\beta^j}\right) w^j * (\beta^j l^j - a^j) - w^j l^j.$$

Ved å løse parentesene samt flytte $w^j l^j$ over på den andre siden fås:

$$w^j l^j = \frac{(1 + \varepsilon^j)\beta^j l^j}{\beta^j} w^j - \frac{1 + \varepsilon^j}{\beta^j} a^j w^j$$

Ved å dele begge sider med w^j blir det:

$$l^j = l^j + l^j \varepsilon^j - \frac{a^j + a^j \varepsilon^j}{\beta^j} \rightarrow \beta^j l^j = \beta^j (l^j + l^j \varepsilon^j) - a^j - a^j \varepsilon^j$$

Med dette kan $\beta^j l^j$ på hver side strykes mot hverandre:

$$-\beta^j l^j \varepsilon^j = -a^j - a^j \varepsilon^j$$

Deler på ε^j og flytter litt rundt:

$$\beta^j l^j - a^j = \frac{a^j}{\varepsilon^j}$$

Fra teoridelen benyttes (2) som gir:

$$x^j = \frac{a^j}{\varepsilon^j}$$

Som er (7) i teoridelen.

A.4 Antall bedrifter i likevekt

Skal nå komme fram til (8) i teoridelen ved hjelp av (2) og (7):

$$\frac{a^j}{\varepsilon^j} = \beta^j l^j - a^j \rightarrow a^j = \beta^j l^j \varepsilon^j - a^j \varepsilon^j \rightarrow a^j(1 + \varepsilon^j) = \beta^j l^j \varepsilon^j$$

Hvor vi kan løse for l^j å få:

$$l^j = \frac{a^j(1 + \varepsilon^j)}{\beta^j \varepsilon^j}$$

Som er (8) i teoridelen.

A.5 Skalautbytte sektornivå

Starter med en forenkling av (1) fra teoridelen:

$$Y^j = \left(n^j * x^{j \frac{1}{1+\varepsilon}} \right) \rightarrow Y^j = n^{j^{1+\varepsilon}} * x^j$$

Setter så inn (7) og (9) fra teoridelen:

$$Y^j = \left(\frac{\beta^j \varepsilon^j}{a^j(1 + \varepsilon^j)} L^j \right)^{1+\varepsilon^j} \left(\frac{a^j}{\varepsilon^j} \right) \rightarrow Y^j = \left(\beta^{j^{1+\varepsilon^j}} \varepsilon^{j^{1+\varepsilon^j}} L^{j^{1+\varepsilon^j}} \right) a^{j^{1+\varepsilon}} \left(\frac{a^j}{\varepsilon^j} \right)$$

Ved å gange inn $\frac{a^j}{\varepsilon^j}$ fås:

$$Y^j = \frac{\left(\beta^{j^{1+\varepsilon}} \varepsilon^{j^\varepsilon} \right)}{(1 + \varepsilon^j)^{1+\varepsilon^j} a^{j^\varepsilon}} L^{j^{1+\varepsilon}}$$

Ved å anta at $\beta^j = (1 + \varepsilon^j) \left(\frac{a^j}{\varepsilon^j} \right)^{\frac{\varepsilon}{1+\varepsilon}}$ vil den første delen være lik 1 slik at:

$$Y^j = (L^j)^{1+\varepsilon^j}$$

Som er (10) fra teoridelen.

B Tabeller

B.1 Andel kvinner og innvandrere i storregioner sammenlignet med andre regioner

	Storregioner	Andre regioner
Andelen kvinner	51,2%	50,4%
Andelen innvandrere	7,7%	3,6%

B.2 Paneldata med region faste effekter privat sektor

	Lnwage_priv (1)	Lnwage_priv (2)	Lnwage_priv (3)	Lnwage_priv (4)
Lnbefolk	-0,0129 (0,0105)	-0,0259** (0,012)	-0,0262** (0,0132)	-0,0282** (0,0136)
Høyere utdanning		0,1905** (0,0845)	0,1342 (0,0859)	0,1416 (0,0867)
Finans	Nei	Nei	Nei	Ja
Kvinner,innvandrere	Nei	Nei	Ja	Ja
Årsdummier	Ja	Ja	Ja	Ja
Observasjoner	1602	1602	1602	1602
R ²	0,994	0,994	0,994	0,994

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

B.3 Paneldata med region faste effekter offentlig sektor

	Lnwage_pub (1)	Lnwage_pub (2)	Lnwage_pub (3)	Lnwage_pub (4)
Lnbefolk	-0,02** (0,01)	-0,0545*** (0,0112)	-0,0424*** (0,0124)	-0,0434*** (0,0129)
Høyere utdanning		0,5053*** (0,0792)	0,4707*** (0,0806)	0,474*** (0,0814)
Off.adm	Nei	Nei	Nei	Ja
Kvinner,innvandrere	Nei	Nei	Ja	Ja
Årsdummier	Ja	Ja	Ja	Ja
Observasjoner	1602	1602	1602	1602
R ²	0,993	0,993	0,993	0,993

Estimerte standardavvik i parentes. Signifikansnivå: * 10%, ** 5%, *** 1%

C Liste over økonomiske regioner SSB 2000

Liste over 90 økonomiske regioner med tilhørende kommuner og folkekemengde i år 2000. Det er disse som legger grunnlaget for dataen i denne oppgaven.

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Østfold	Halden 0191	0101 Halden	26 733
		0118 Aremark	1 451
		I alt	28 184
	Moss 0192	0104 Moss	26 633
		0135 Råde	6 217
		0136 Rygge	13 288
		0137 Våler	4 059
		I alt	50 197
	Fredrikstad/Sarpsborg 0193	0106 Fredrikstad	67 761
		0105 Sarpsborg	47 447
		0111 Hvaler	3 494
		0128 Rakkestad	7 076
		I alt	125 778
	Askim/Mysen 0194	0124 Askim	13 521
		0125 Eidsberg	9 703
		0119 Marker	3 300
		0121 Rømskog	674
		0122 Trøgstad	4 868
		0123 Spydeberg	4 486
		0127 Skiptvet	3 140
0138 Hobøl		4 366	
I alt	44 058		

Fylke	Økonomisk region/kode	Kommune	Folkemengde
Akershus	Follo 0291	0213 Ski	25 394
		0211 Vestby	11 815
		0214 Ås	13 568
		0215 Frogn	12 972
		0216 Nesodden	15 448
		0217 Oppegård	22 953
		I alt	102 150
	Bærum/Asker 0292	0219 Bærum	101 494
		0220 Asker	49 284
		I alt	150 778
	Lillestrøm 0293	0231 Skedsmo	38 701
		0221 Aurskog-Høland	12 561
		0226 Sørums	12 133
		0227 Fet	9 271
		0228 Rælingen	14 493
		0229 Enebakk	8 680
		0230 Lørenskog	29 505
		0233 Nittedal	18 639
		0234 Gjerdrum	4 554
		0236 Nes	16 288
	I alt	164 825	
	Ullensaker/Eidsvoll 0294	0235 Ullensaker	20 160
		0237 Eidsvoll	17 524
		0238 Nannestad	8 967
		0239 Hurdal	2 648
		I alt	49 299

Fylke	Økonomisk region/Kode	Bydel	Folkemengde
Oslo	Oslo 0391	01 Bygdøy-Frogner	20 326
		02 Majorstua-Uranienborg	23 809
		03 St.Hanshaugen-Ullevål	28 259
		04 Sagene-Torshov	27 888
		05 Grünerløkka-Sofienberg	26 683
		06 Gamle Oslo	25 433
		07 Ekeberg-Bekkelaget	16 587
		08 Nordstrand	17 349
		09 Søndre Nordstrand	31 380
		10 Lambertseter	10 230
		11 Bøler	13 170
		12 Manglerud	12 309
		13 Østensjø	15 433
		14 Helsefyr-Sinsen	21 100
		15 Hellerud	15 691
		16 Furuset	29 290
		17 Stovner	20 968
		18 Romsås	6 822
		19 Grorud	17 075
		20 Bjerke	22 821
		21 Grefsen-Kjelsås	17 765
		22 Sogn	15 823
		23 Vinderen	19 612
		24 Røa	21 310
		25 Ullern	26 607
		26 Sentrum	1 194
		27 Marka	1 647
		Uoppgitt	886
	I alt	507 467	

Fylke	Økonomisk region/Kode	Kommune	Folkemengde
Hedmark	Kongsvinger 0491	0402 Kongsvinger	17 349
		0418 Nord-Odal	5 089
		0419 Sør-Odal	7 349
		0420 Eidskog	6 409
		0423 Grue	5 442
		0425 Åsnes	8 112
		I alt	49 750
	Hamar 0492	0403 Hamar	26 545
		0412 Ringsaker	31 622
		0415 Løten	7 188
		0417 Stange	17 928
		I alt	83 283
	Elverum 0493	0427 Elverum	18 046
		0426 Våler	4 063
		0428 Trysil	7 069
		0429 Åmot	4 379
		0430 Stor-Elvdal	3 012
		0434 Engerdal	1 580
		I alt	38 149
	Tynset 0494	0437 Tynset	5 473
		0432 Rendalen	2 257
		0436 Tolga	1 812
		0438 Alvdal	2 417
		0439 Folldal	1 814
		0441 Os	2 148
		I alt	15 921

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Oppland	Lillehammer 0591	0501 Lillehammer	24 724
		0521 Øyer	4 859
		0522 Gausdal	6 186
		I alt	35 769
	Gjøvik 0592	0502 Gjøvik	27 013
		0528 Østre Toten	14 103
		0529 Vestre Toten	13 065
		0536 Søndre Land	6 073
		0538 Nordre Land	6 950
		I alt	67 204
	Midt-Gudbrandsdalen 0593	0516 Nord-Fron	5 953
		0519 Sør-Fron	3 322
		0520 Ringebu	4 752
		I alt	14 027
	Nord-Gudbrandsdalen 0594	0517 Sel	6 273
		0511 Dovre	2 851
		0512 Lesja	2 303
		0513 Skjåk	2 386
		0514 Lom	2 567
		0515 Vågå	3 818
		I alt	20 198
	Hadeland 0595	0534 Gran	12 877
		0532 Jevnaker	5 995
		0533 Lunner	8 264
		I alt	27 136
	Valdres 0596	0542 Nord-Aurdal	6 560
		0540 Sør-Aurdal	3 389
		0541 Etnedal	1 401
		0543 Vestre Slidre	2 282
		0544 Øystre Slidre	3 060
0545 Vang		1 675	
I alt		18 367	

Fylke	Økonomisk region/Kode	Kommune	Folkemengde
Buskerud	Drammen 0691	0602 Drammen	54 816
		0621 Sigdal	3 556
		0623 Modum	12 366
		0624 Øvre Eiker	15 058
		0625 Nedre Eiker	20 502
		0626 Lier	21 308
		0627 Røyken	16 245
		0628 Hurum	8 363
		I alt	152 214
		Kongsberg	0692
0631 Flesberg	2 491		
0632 Rollag	1 492		
0633 Nore og Uvdal	2 764		
I alt	29 040		
Hønefoss	0693	0605 Ringerike	27 917
		0612 Hole	4 977
		0622 Krødsherad	2 254
		I alt	35 148
Hallingdal	0694	0617 Gol	4 390
		0615 Flå	1 102
		0616 Nes	3 528
		0618 Hemsedal	1 958
		0619 Ål	4 789
		0620 Hol	4 642
		I alt	20 409

Fylke	Økonomisk region/Kode	Kommune	Folkemengde
Vestfold	Tønsberg/Horten 0791	0704 Tønsberg	34 716
		0701 Borre	23 764
		0716 Våle	4 219
		0718 Ramnes	3 717
		0719 Andebu	4 678
		0720 Stokke	9 557
		0722 Nøtterøy	19 601
		0723 Tjøme	4 505
		I alt	104 757
	Holmestrand 0792	0702 Holmestrand	9 384
		0714 Hof	2 891
		I alt	12 275
	Sandefjord/Larvik 0793	0706 Sandefjord	39 317
		0709 Larvik	40 386
		0728 Lardal	2 379
		I alt	82 082
	Sande/Svelvik 0794	0713 Sande	7 377
0711 Svelvik		6 284	
I alt		13 661	

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Telemark	Skien/Porsgrunn 0891	0806 Skien	49 592
		0805 Porsgrunn	32 892
		0811 Siljan	2 269
		0814 Bamble	14 142
		0819 Nome	6 648
		I alt	105 543
	Notodden/Bø 0892	0807 Notodden	12 272
		0821 Bø	4 964
		0822 Sauherad	4 411
		0827 Hjartdal	1 687
		I alt	23 334
	Kragereø 0893	0815 Kragereø	10 656
		0817 Drangedal	4 175
		I alt	14 831
	Rjukan 0894	0826 Tinn	6 560
	Vest-Telemark 0895	0828 Seljord	2 928
		0829 Kviteseid	2 667
		0830 Nissedal	1 447
		0831 Fyresdal	1 353
		0833 Tokke	2 506
		0834 Vinje	3 869
I alt		14 770	

Fylke	Økonomisk region/Kode	Kommune	Folkemengde
Aust-Agder	Risør 0991	0901 Risør	7 000
		0911 Gjerstad	2 509
		I alt	9 509
	Arendal 0992	0906 Arendal	39 446
		0904 Grimstad	17 821
		0912 Vegårshei	1 838
		0914 Tvedestrand	5 967
		0919 Froland	4 497
		0929 Åmli	1 862
		I alt	71 431
	Lillesand 0993	0926 Lillesand	8 816
		0928 Birkenes	4 290
		I alt	13 106
	Setesdal 0994	0937 Evje og Hornnes	3 346
		0935 Iveland	1 128
0938 Bygland		1 351	
0940 Valle		1 439	
0941 Bykle		868	
I alt		8 132	

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Vest-Agder	Kristiansand 1091	1001 Kristiansand	72 395
		1014 Vennesla	12 141
		1017 Songdalen	5 455
		1018 Søgne	8 929
		I alt	98 920
	Mandal 1092	1002 Mandal	13 316
		1021 Marnardal	2 198
		1026 Åseral	878
		1027 Audnedal	1 530
		1029 Lindesnes	4 395
	I alt	22 317	
	Lyngdal/Farsund 1093	1003 Farsund	9 630
		1032 Lyngdal	7 064
		1034 Hægebostad	1 621
		I alt	18 315
	Flekkefjord 1094	1004 Flekkefjord	8 851
		1037 Kvinesdal	5 554
		1046 Sirdal	1 734
		I alt	16 139

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Rogaland	Egersund 1191	1101 Eigersund	13 288
		1111 Sokndal	3 358
		1112 Lund	3 096
		1114 Bjerkeim	2 456
		I alt	22 198
	Stavanger/Sandnes 1192	1103 Stavanger	108 818
		1102 Sandnes	52 998
		1122 Gjesdal	8 911
		1124 Sola	18 915
		1127 Randaberg	8 773
		1129 Forsand	1 018
		1130 Strand	10 122
		1133 Hjelmeland	2 756
		1141 Finnøy	2 865
		1142 Rennesøy	3 111
		1144 Kvitsøy	523
		I alt	218 810
		Haugesund 1193	1106 Haugesund
	1134 Suldal		4 043
	1135 Sauda		5 081
	1145 Bokn		786
	1146 Tysvær		8 828
	1149 Karmøy		36 971
	1151 Utsira		256
	1154 Vindafjord		4 848
	I alt		91 175
	Jæren 1194	1121 Time	13 317
		1119 Hå	13 921
		1120 Klepp	13 789
		I alt	41 027

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Hordaland	Bergen 1291	1201 Bergen	229 496
		1238 Kvern	8 592
		1241 Fusa	3 684
		1242 Samnanger	2 282
		1243 Os	13 896
		1244 Austevoll	4 406
		1245 Sund	5 160
		1246 Fjell	18 178
		1247 Askøy	19 727
		1251 Vaksdal	4 192
		1252 Modalen	354
		1253 Osterøy	7 006
		1256 Meland	5 353
		1259 Øygarden	3 623
		1260 Radøy	4 585
		1263 Lindås	12 492
		1264 Austrheim	2 527
		1265 Fedje	682
		1266 Masfjorden	1 774
			I alt
	Søndre Sunnhordland		
	1292	1214 Ølen	3 287
		1211 Etne	3 917
		1216 Sveio	4 623
		I alt	11 827
	Nordre Sunnhordland		
	1293	1221 Stord	16 144
		1219 Bømlo	10 739
		1222 Fitjar	2 992
		1223 Tysnes	2 843
		1224 Kvinnherad	13 196
		I alt	45 914
	Odda		
	1294	1228 Odda	7 727
		1227 Jondal	1 151
		1231 Ullensvang	3 562
		1232 Eidfjord	1 037
		I alt	13 477
	Voss		
	1295	1235 Voss	13 726
		1233 Ulvik	1 222
		1234 Granvin	1 044
		I alt	15 992

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Sogn og Fjordane	Florø 1491	1401 Flora	11 226
		1438 Bremanger	4 170
		I alt	15 396
	Høyanger 1492	1416 Høyanger	4 677
		1411 Gulen	2 489
		1412 Solund	959
		1418 Balestrand	1 513
		I alt	9 638
	Sogndal/Årdal 1493	1420 Sogndal	6 666
		1424 Årdal	5 797
		1417 Vik	2 965
		1419 Leikanger	2 183
		1421 Aurland	1 833
		1422 Lærdal	2 202
		1426 Luster	5 003
	I alt	26 649	
	Førde 1494	1432 Førde	10 473
		1413 Hyllestad	1 554
		1428 Askvoll	3 349
		1429 Fjaler	2 964
		1430 Gaular	2 886
		1431 Jølster	2 957
		1433 Naustdal	2 736
	I alt	26 919	
	Nordfjord 1495	1439 Vågsøy	6 479
		1443 Eid	5 780
		1445 Gloppen	5 769
		1449 Stryn	6 666
		1441 Selje	3 066
		1444 Hornindal	1 227
	I alt	28 987	

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Møre og Romsdal	Molde 1591	1502 Molde	23 710
		1535 Vestnes	6 530
		1539 Rauma	7 415
		1543 Nesset	3 289
		1545 Midsund	1 978
		1547 Aukra	2 978
		1548 Fræna	9 005
		1551 Eide	3 193
		1557 Gjemnes	2 683
		I alt	60 781
	Kristiansund 1592	1503 Kristiansund	16 925
		1554 Averøy	5 409
		1556 Frei	5 199
		1569 Aure	2 782
		1572 Tustna	1 041
		1573 Smøla	2 432
		I alt	33 788
	Ålesund 1593	1504 Ålesund	38 855
		1523 Ørskog	2 076
		1524 Norddal	1 960
		1525 Stranda	4 677
		1526 Stordal	1 052
		1528 Sykkylven	7 280
		1529 Skodje	3 513
		1531 Sula	7 070
		1532 Giske	6 344
		1534 Haram	8 774
		1546 Sandøy	1 332
		I alt	82 933

Ulsteinvik	1516 Ulstein	6 541
1594	1511 Vanylven	3 584
	1514 Sande	3 081
	1515 Herøy	8 390
	1517 Hareid	4 748
	I alt	26 344
Ørsta/Volda	1520 Ørsta	10 276
1595	1519 Volda	8 322
	I alt	18 598
Sunnalsøra	1563 Sunndal	7 368
1596	1560 Tingvoll	3 148
	I alt	10 516
Surnadal	1566 Surnadal	6 252
1597	1567 Rindal	2 138
	1571 Halså	1 808
	I alt	10 198

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Sør-Trøndelag	Trondheim 1691	1601 Trondheim	148 859
		1624 Rissa	6 503
		1648 Midtre Gauldal	5 779
		1653 Melhus	13 169
		1657 Skaun	5 843
		1662 Klæbu	4 875
		1663 Malvik	11 132
		1664 Selbu	3 926
		1665 Tydal	949
		I alt	201 035
Frøya/Hitra	1692	1620 Frøya	4 115
		1617 Hitra	4 038
		I alt	8 153
Brekstad	1693	1621 Ørland	5 037
		1627 Bjugn	4 696
		1630 Ålfjord	3 403
		1632 Roan	1 121
		1633 Osen	1 194
		I alt	15 451
Oppdal	1694	1634 Oppdal	6 288
		1635 Rennebu	2 700
		I alt	8 988
Orkanger	1695	1638 Orkdal	10 250
		1612 Hemne	4 324
		1613 Snillfjord	1 093
		1622 Agdenes	1 783
		1636 Meldal	3 991
		I alt	21 441
Røros	1696	1640 Røros	5 545
		1644 Holtålen	2 239
		I alt	7 784

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Nord-Trøndelag	Steinkjer 1791	1702 Steinkjer	20 459
		1718 Leksvik	3 512
		1723 Mosvik	925
		1724 Verran	2 733
		1725 Namdalseid	1 831
		1729 Inderøy	5 802
		1736 Snåsa	2 397
		I alt	37 659
	Namsos 1792	1703 Namsos	12 325
		1743 Høylandet	1 335
		1744 Overhalla	3 659
		1748 Fosnes	794
		1749 Flatanger	1 238
		I alt	19 351
	Stjørdalshalsen 1793	1714 Stjørdal	18 238
		1711 Meråker	2 637
		I alt	20 875
	Levanger/Verdalsøra 1794	1719 Levanger	17 501
		1721 Verdal	13 644
		1717 Frostå	2 430
		I alt	33 575
	Grong 1795	1742 Grong	2 565
		1738 Lierne	1 565
		1739 Røyrvik	595
		1740 Namsskogan	989
		I alt	5 714
	Rørвик 1796	1750 Vikna	3 867
1751 Nærøy		5 353	
1755 Leka		714	
I alt		9 934	

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Nordland	Bodø 1891	1804 Bodø	41 367
		1836 Rødøy	1 570
		1837 Meløy	6 796
		1838 Gildeskål	2 351
		1839 Beiarn	1 311
		1840 Saltdal	4 887
		1841 Fauske	9 632
		1842 Skjerstad	1 098
		1845 Sørfold	2 367
		1848 Steigen	2 977
		1849 Hamarøy	2 008
		I alt	76 364
		Narvik 1892	1805 Narvik
	1850 Tysfjord		2 302
	1851 Lødingen		2 471
	1852 Tjeldsund		1 527
	1853 Evenes		1 523
	1854 Ballangen		2 775
	I alt		29 198
	Brønnøysund 1893	1813 Brønnøy	7 433
		1811 Bindal	1 921
		1812 Sømna	2 116
		1815 Vega	1 414
		1816 Vevelstad	592
		I alt	13 476
	Sandnessjøen 1894	1820 Alstahaug	7 440
		1818 Herøy	1 881
		1822 Leirfjord	2 242
		1827 Dønna	1 582
		1834 Lurøy	2 107
		1835 Træna	466
	I alt	15 718	

Mosjøen	1824 Vefsn	13 553
1895	1825 Grane	1 652
	1826 Hattfjelldal	1 634
	I alt	16 839
Mo i Rana	1833 Rana	25 255
1896	1828 Nesna	1 882
	1832 Hemnes	4 689
	I alt	31 826
Lofoten	1865 Vågan	9 229
1897	1856 Røst	666
	1857 Værøy	775
	1859 Flakstad	1 575
	1860 Vestvågøy	10 750
	1874 Moskenes	1 352
	I alt	24 347
Vesterålen	1870 Sortland	9 230
1898	1866 Hadsel	8 321
	1867 Bø	3 288
	1868 Øksnes	4 758
	1871 Andøy	5 744
	I alt	31 341

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Troms	Harstad 1991	1901 Harstad	23 025
		1911 Kvæfjord	3 287
		1913 Skånland	3 109
		1915 Bjarkøy	602
		1917 Ibestad	1 763
		I alt	31 786
	Tromsø 1992	1902 Tromsø	59 145
		1933 Balsfjord	5 749
		1936 Karlsøy	2 496
		1938 Lyngen	3 225
		1939 Storfjord	1 872
		I alt	72 487
	Andselv 1993	1924 Målselv	7 054
		1919 Gratangen	1 345
		1920 Lavangen	1 052
		1922 Bardu	3 889
		1923 Salangen	2 346
		I alt	15 686
	Finnsnes 1994	1931 Lenvik	11 039
		1925 Sørreisa	3 294
		1926 Dyrøy	1 337
		1927 Tranøy	1 695
		1928 Torsken	1 166
		1929 Berg	1 111
		I alt	19 642
	Nord-Troms 1995	1942 Nordreisa	4 821
		1941 Skjervøy	2 934
1943 Kvænangen		1 435	
1940 Gáivuotna-Kåfjord		2 369	
I alt		11 559	

Fylke	Økonomiske region/Kode	Kommune	Folkemengde
Finnmark	Vadsø 2091	2003 Vadsø	6 130
		2002 Vardø	2 705
		2024 Berlevåg	1 236
		2025 Deatnu-Tana	3 074
		2027 Unjárga-Nesseby	965
		2028 Båtsfjord	2 470
		I alt	16 580
		Hammerfest 2092	2004 Hammerfest
	2017 Kvalsund		1 106
	2018 Måsøy		1 477
	2019 Nordkapp		3 517
	2020 Porsanger		4 451
	2021 Kárášjohka-Karasjok		2 901
	2022 Lebesby		1 463
	2023 Gamvik		1 288
	I alt	25 416	
	Alta 2093	2012 Alta	16 837
		2011 Guovdageaidnu-Kautokeino	3 068
		2014 Loppa	1 426
		2015 Hasvik	1 200
		I alt	22 531
	Kirkenes 2094	2030 Sør-Varanger	9 532

