



Er lønn og produktivitet høyere i tett befolkede områder?

Tor Fabian Aaserud

Juni 2017

MASTEROPPGAVE

Institutt for samfunnsøkonomi

NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET

Veileder: Professor Hildegunn Ekroll Stokke

Innhold

| | |
|---|-----------|
| 1 Innledning | 1 |
| 1.1 Problemstilling | 2 |
| 2 Teori | 5 |
| 2.1 Agglomerasjonseffekter | 6 |
| 2.1.1 Deling | 6 |
| 2.1.2 Matching | 11 |
| 2.1.3 Læring | 12 |
| 3 Tidligere forskning | 17 |
| 3.1 Statistiske og dynamiske effekter | 17 |
| 3.2 Tidligere litteratur | 18 |
| 4 Empirisk tilrettelegging | 25 |
| 4.1 Avhengig variabel | 25 |
| 4.2 Forklaringsvariabler | 26 |
| 4.3 Data | 28 |
| 4.3.1 Økonometrisk tilrettelegging | 30 |
| 5 Økonometriske utfordringer | 33 |
| 5.1 Brudd på eksogenitetsutfordringen | 35 |
| 5.2 Håndtering av brudd på eksogenitetsbetingelsen | 38 |
| 6 Urbant lønnspremium | 41 |
| 6.1 MKM - produktivitet og lønn i økonomiske regioner | 42 |
| 6.2 IV - produktivitet og lønn i økonomiske regioner | 45 |
| 6.3 Urbant lønnspremium i ulike sektorer | 50 |
| 6.3.1 MKM | 50 |

| | | |
|----------|--|-----------|
| 6.3.2 | IV | 54 |
| 6.4 | Robusthetstesting | 57 |
| 7 | Konklusjon | 59 |
| 8 | Bibliografi | i |
| | Referanser | i |
| 9 | A Appendiks | v |
| 9.1 | Deling av fordelene fra bredere variasjon av tilbydere av inputs | v |
| 9.2 | Robusthetsresultater | xi |
| 9.3 | Liste over økonomiske regioner | xiii |

Tabeller

| | | |
|-----|---|-----|
| 4.1 | Deskriptiv statistikk | 29 |
| 4.2 | Deskriptiv statistikk sektorlønn | 30 |
| 6.1 | MKM - hovedanalyse | 42 |
| 6.2 | IV - hovedanalyse. | 45 |
| 6.3 | Førstesteget for spesifikasjon (1) | 47 |
| 6.4 | Førstesteget for spesifikasjon (2). | 47 |
| 6.5 | Førstesteget for spesifikasjon (3). | 48 |
| 6.6 | MKM - Sektor og næring | 51 |
| 6.7 | IV - Sektor og næring | 54 |
| 9.1 | Alternative instrument for spesifikasjon (2) og (3) | xi |
| 9.2 | Alternativ hovedforklaringsvariabel absoluttverdien av populasjon | xi |
| 9.3 | Sektor og næring - alternativt instrument | xii |

Forord

Denne masteroppgaven er avslutning på en mastergrad i samfunnsøkonomi ved Institutt for samfunnsøkonomi på NTNU. Jeg ønsker å rette en stor takk til min veileder Hildegunn Ekroll Stokke for verdifull og utrolig god tilbakemelding, og som alltid har hatt en åpen dør når jeg har hatt spørsmål underveis. Videre ønsker jeg å rette en stor takk til Luka Marcinko, Eirik Theie Ødegaard og Ulrik Brændengen for gjennomlesing av oppgaven. Jeg vil også takke Institutt for samfunnsøkonomi og Pareto fotballklubb for en fin studietid, og alle som har gjort mine år som student fantastisk.

Kapittel 1

Innledning

Sammenhengen mellom geografisk lokalisering og produktivitet fikk oppmerksomhet allerede av Smith (1776). Populasjonsendringer til urbane områder er et viktig aspekt innen økonomisk vekst. Høyere inntekt i byer tiltrekker flere individer. Rattsø og Stokke (2014) viser til store skift i populasjon mellom geografiske regioner, og da spesifikt til urbane områder over en tidsperiode fra 1972 til 2008.¹ Migrasjonen på tvers av regioner affekterer inntektsfordelingen på tvers av disse geografiske områdene. Forståelsen av hvorvidt dette er et faktum for Norge, og hvilke faktorer som eventuelt underbygger dette er viktig for forståelsen av sysselsetting, urbanisering, og ikke minst verdiskaping, som er sentrale faktorer i determinering av politikk både på mikro- og makronivå.

Norge er inndelt i 89 økonomiske regioner, hvorav regionene varierer i både innbyggertall og inntektsfordeling. Det er dermed nærliggende å tro at det er forskjeller i gjennomsnittlig timelønn, og dermed produktivitet, mellom disse regionene. Nylig har Carlsen, Rattsø og Stokke (2016) studert sammenhengen mellom urbant lønnspremium og utdanningsnivå, samt identifisering av agglomerasjonsmekanismer, og funnet resultat i favør av byregioner. I et land som fortsatt har en desentralisert bosettingsstruktur, men hvor tendensene har vært sterk urbanisering de siste 30 årene, virker forståelsen av hvorvidt de befolkede områdene besitter enkelte fordeler å være viktig i utformingen av optimal distriktpolitikk.

Med urbanisering er det også nærliggende å tro at bedrifters gevinster forsterkes ytterligere av bevisst geografisk lokalisering. Dette medfører at næringsliv svekkes i regionale strøk. Over tid vil dette i teorien styrke de urbane områdene, og svekke distriktene. Det kan tenkes at dette forsterker næringslivets insentiver til urban lokalisering. Dersom indi-

¹Som et kombinert resultat av innenlands migrasjon, internasjonal migrasjon, fertilitet og dødelighet

vider og bedrifter ser fordeler i urban geografisk lokalisering, vil dette som nevnt svekke de regionale strøk. Om dette faktisk er tilfellet, og hvorvidt dette varierer mellom sektorene er av interesse å undersøke

1.1 Problemstilling

Oppgaven tar til hensikt å analysere hvorvidt urbane områder i Norge har et høyere lønnspremium, og dermed produktivitet. Deretter om dette lønnspremiumet også varierer mellom ulike sektorer og næringer i økonomien. Utgangspunktet vil være regional tverrsnittsdata for 2010, fokuset vil dermed være på statiske agglomerasjonsmekanismer i Norge.

P.-P. Combes, Duranton, Gobillon og Roux (2010) finner i lignende studier at elastisiteten mellom lokal produktivitet og sysselsettingstetthet typisk ligger mellom 4 til 10 prosent ved dobling av sysselsettingstettheten når de tar for seg franske arbeidsmarkedsregioner. Mer nærliggende er Carlsen et al. (2013) som finner, når de tar for seg norske arbeidsmarkedsregioner, at gjennomsnittlig lønn ligger mellom 3,4 til 5,5 prosentpoeng høyere ved dobling av populasjonstetthet, avhengig av metode brukt.

I denne oppgaven varierer det urbane lønnspremiumet mellom 3,04 og 4,18 prosent avhengig av metode og spesifisering.

Videre nyter privat sektor mer av bevisst urban lokalisering enn hva offentlig sektor gjør. Forskjellen her ligger typisk mellom 2,22 og 1,76 prosentpoeng ved dobling av populasjonstetthet. Innen privat sektor nyter servicenæringen bedre av bevisst geografisk lokalisering mot urbane områder enn hva industri gjør, og denne forskjellen ligger typisk mellom 1,65 og 1,57 prosentpoeng høyere ved dobling av populasjon per kvadratkilometer.

Disposisjon

I kapittel 2 presenterer jeg det teoretiske fundamentet som underbygger produktivetsfordeler av tett befolkede områder, her er det primært tre faktorer som spiller en vesentlig rolle. I kapittel 3 presenterer jeg tidligere forskning, deriblant forskjellen mellom statiske og dynamiske effekter, og tidligere litteratur. I tillegg til dette en diskusjon av det som i litteraturen kalles det kvalitative og det kvantitative endogenitetsproblemet. Kapittel 4 omfatter den empiriske tilretteleggingen, mer spesifikt inngående variabler i metodedelen, databeskrivelse og den økonometriske tilretteleggingen. I kapittel 5 går jeg gjennom de

økonometriske utfordringene til dels diskutert omfanget av i kapittel 3. Kapittel 6 inneholder resultater oppnådd i analysen. Her har jeg primært brukt minste kvadraters metode og instrumentvariabelmetoden for ulike spesifikasjoner. Til slutt drar jeg i kapittel 7 noen konkluderende kommentarer basert på funn gjort i oppgaven.

Kapittel 2

Teori

Som nevnt innledningsvis er det hovedsakelig tre teoretiske mekanismer som forklarer høyere produktivitet; agglomerasjonseffekter, sortering og seleksjon. Agglomerasjonseffekter er effekter eksterne for bedrifter, og gir tiltakende avkastning i klynge- og byområder. Sortering av arbeidere innebærer at mer talentfulle individer søker seg til byområder, som fordrer høyere produktivitet i disse områdene. Mens seleksjon følger av at større byer faciliterer større markeder som i tur gjør seleksjon tøffere, og dermed overlever kun de mest produktive bedriftene. Dette omfanget vil naturlig nok være større i områder med mange bedrifter, derav er også disse områdene mer produktive. I teorien er de ulike mekanismene komplementære. Tøffere seleksjon krever mer talentfulle arbeidere. Dette fører til at mer talentfulle arbeidere ender opp i mer produktive bedrifter, som styrker agglomerasjonsgevinstene. Bartelsman og Doms (2000); Foster, Haltiwanger og Syverson (2008) viser at mindre produktive bedrifter forlater markeder med høyere frekvens. Syverson (2004) finner at nivået til produktivitet for å bli utkonkurrert av markedet er høyere i større markeder. Etter å ha korrigert for både sortering og agglomerasjon er det ingen bevis for at seleksjon faktisk er tøffere i byregioner.

Fokuset for teoridelen vil være det som er mest relevant for oppgaven, agglomerasjonseffekter. Som nevnt ovenfor er sortering og agglomerasjon komplementært, noe som gjør at jeg kommer tilbake til eventuelle problemer som kan oppstå som konsekvens av dette, og relaterte løsninger. Teoridelen er i hovedsak tatt fra Duranton og Puga (2004) sitt mikrogrunnlag for agglomerasjonsmekanismer: deling, matching og læring.

2.1 Agglomerasjonseffekter

Som nevnt tidligere er agglomerasjonseffekter eksterne for bedrifter, og gir tiltakende avkastning i klynge- og byområder. Det finnes et bredt spekter av ulike teoretiske modeller som forklarer urbane og lokale fordeler innen tettere befolkede områder. Jeg presenterer nedenunder flere modeller, men går kun i dybden på et utvalgt antall.

2.1.1 Deling

Det er flere mekanismer under deling som fremmer agglomerasjonsgevinster. Deling av udelelige fasiliteter, dele gevinst fra større variasjon i produksjon, urban struktur, gevinstene av individuell spesialisering samt deling av risiko.

Deling av udelelige fasiliteter er gjerne fasiliteter som krever en høy fast investeringskostnad, men lave variable kostnader. Med dette vil snittkostnaden avta med antallet arbeidere eller individer som benytter fasiliteten, som fører til tiltakende skalaavkastning i områder som er høyt eller tett befolket. Dette kan være alt fra fabrikker til ishaller.

Gevinst fra større variasjon i produksjon

Med gevinst fra større variasjon i produksjon er fokuset på koblinger fra mellom- til sluttprodukter. Hvor flere produsenter i mellomsektor fører til tiltakende avkastning i sluttproduksjon via at økt arbeidskraft til sluttproduksjon skyldes flere mellomprodusenter. Presenterer her modell fra Duranton og Puga (2004), som utleder en aggregert produktfunksjon som viser aggregert tiltakende skalaavkastning i produksjon av sluttprodukter i et perfekt konkurransemarked. Dette kommer hovedsakelig fra det faktum at sluttprodusenter nyter produksjonsfordeler av å dele en bredere variasjon av ulike innsatsvarer til produksjon.

Økonomien antas å bestå av m -sektorer indeksert på $j = 1, \dots, m$. I hver sektor produserer bedrifter konsumvarer med konstant skalautbytte under perfekt konkurranse. Disse sluttprodusentene benytter innsatsvarer fra mellomsektor til sluttproduksjon, som er spesifikke for hver enkelt sektor. Innsatsvarene brukt til sluttproduksjon inngår i fabrikkens teknologi med en konstant substitusjonselastisitet gitt ved:

$$\frac{1 + \epsilon_j}{\epsilon_j}, \epsilon_j > 0 \quad (2.1)$$

Dette er substitusjonselastisiteten mellom innsatsvarer. Desto høyere ϵ_j , jo lavere er elastisiteten. ϵ_j kan sees på som en parameter som måler graden av monopolistisk konkurran-

se i mellomsektor. En lav parameter-verdi representerer en høy elastisitet, som impliserer høyere konkurranse via flere tilbydere av innsatsgodet, og dermed mindre markeds makt i mellomsektor. Elastisiteten antas å være lik mellom alle variantene av en bestemt vare og kjent på forhånd.

Videre er aggregert produksjon i sluttsektor j gitt ved:

$$Y_j = \left(\int_0^{n_j} (X_j(h))^{\frac{1}{1+\epsilon_j}} \right)^{1+\epsilon_j} \quad (2.2)$$

Hvor $X_j(h)$ er samlet mengde av innsatsvare h_j , mens n_j er antall varianter av innsatsvarene. Aggregert produksjon av et sluttgode Y_j i gitt sektor kommer fra innsatsvarene gjennom en konstant substitusjonselastisitet.

Videre produseres innsatsvarer i mellomsektor av bedrifter i et marked med monopolistisk konkurranse:

$$X_j(h) = \beta_j l_j(h) - \alpha_j \quad (2.3)$$

Hvor $l_j(h)$ er den enkelte bedrifts bruk av arbeidskraft, β_j er marginalproduktet av arbeidskraft og α_j representerer faste kostnader i sektoren. I en slik situasjon vil det være tiltakende skalaavkastning i produksjonen av hvert innsatsgode. Dette sett i sammenheng med at det er en ubegrenset rekkevidde av mellomgoder, tilsier at hvert firma som produserer innsatsvarer har sitt unike innsatsgode som de produserer. Dette fordi det vil være kostnadsbesparende for den enkelte bedrift kun å produsere én variant av innsatsgodet. $X_j(h)$ øker med antallet bedrifter i innsatsproduksjon som øker aggregert produksjon Y , og dette gir tiltakende skalaavkastning i produksjon av innsatsgoder.

I produksjon av det ferdige godet er totale kostnader gitt ved:

$$C = \int_0^n q_j(h) X_j(h) dh \quad (2.4)$$

Hvor $q_j(h)$ er prisen på innsatsvaren h i sektor j . Ved minimering av totale kostnader i produksjon av det ferdige godet, gitt produktfunksjonen (2.2), gir dette den betingede etterspørselen etter innsatsvarer:¹

$$X_j(h) = \frac{q_j(h)^{-(1+\epsilon_j/\epsilon_j)} * Y_j}{\int_0^n (q_j(h)^{(-1/\epsilon_j)} dh)^{1+\epsilon_j}} \quad (2.5)$$

¹For mellomregninger se appendiks.

Dette uttrykket er determinert utifra produsenter i sluttgode-sektor sitt minimeringsproblem. Fra dette uttrykket ser jeg at hver produsent av innsatsfaktorer står overfor en etterspørsel elastisitet med hensyn på sin egen pris.² Denne illustrerer hvordan etterspørselen etter innsatsvarer endres når prisen endres, gitt ved uttrykket:

$$El_{q_j(h)}X_j(h) = \frac{-(1 + \epsilon_j)}{\epsilon_j} \quad (2.6)$$

Jo lettere det er å substituere mellom innsatsvarene, lav ϵ_j , desto større blir fallet i etterspørsel om bedriften setter opp prisen.

Videre vil den enkelte bedrifts profitt i sluttsektor være gitt ved:

$$\Pi = q_j * X_j - w_j * l_j \quad (2.7)$$

Hvor w_j er lønn per enhet arbeidskraft, q_j er pris på innsatsfaktorer til produksjon. Den profittmaksimerende prisen vil dermed for hvert innsatsgode være gitt ved en fast mark-up over marginalkostnadene.³

Førsteordensbetingelsen til profittmaksimerende bedrift vil være gitt ved $\frac{d\Pi}{dq} = 0$ som gir:⁴

$$q_j = \frac{1 + \epsilon_j}{\beta_j} * w_j \quad (2.8)$$

Prisen på innsatsvarer er en mark-up over de marginale kostnadene $\frac{w_j}{\beta_j}$. Differansen mellom pris satt og de marginale kostnadene vil være større jo lavere substitusjonselastisiteten mellom innsatsvarer er. En høy ϵ_j -verdi gjenspeiler en høy mark-up. På grunn av fri-entre og muligheter for å trekke seg fra markedet, vil profitten konkurreres til 0. Slik at:

$$\Pi = 0 \Rightarrow q_j * X_j - w_j * l_j = 0 \quad (2.9)$$

Ved å kombinere relasjon (2.3), (2.8) og (2.9) kan jeg utvide sistnevnte relasjon og løse for X_j . Det eneste produksjonsnivå til en bedrift som produserer innsatsfaktorer som også er konsistent med profitt lik null vil være gitt ved:⁵

$$X_j = \frac{\alpha_j}{\epsilon_j} \quad (2.10)$$

²Etterspørselens priselastisitet er et mål på hvor responsiv etterspurt mengde av et gode er for endringer i prisen

³Mark-up er pris satt over marginalkostnaden

⁴For mellomregninger se appendiks.

⁵For mellomregninger se appendiks.

Kombinerer relasjon (2.10) med (2.3) og finner mengden arbeidskraft mellomsektoren benytter til produksjon:

$$l_j = \frac{\alpha_j(1 + \epsilon_j)}{\beta_j * \epsilon_j} \quad (2.11)$$

Som representerer antall arbeidere. Finner likevektstallet ved å implementere (2.11) i $n = \frac{L}{l}$, som gir: ⁶

$$n_j = \frac{L_j}{l_j} = \frac{\beta_j * \epsilon_j}{\alpha_j(1 + \epsilon_j)} * L_j. \quad (2.12)$$

Husk at hver enkelt bedrift vil produsere ett innsatsgode da dette er kostnadsbesparende, derav vil antallet bedrifter representere antallet innsatsvarer produsert. Slik at n_j er det totale antallet innsatsvarianter til sluttgodeproduksjon.

Videre finner jeg aggregert produksjon i hver sektor ved å sette (2.10) og (2.12) inn i (2.2), får da:

$$Y = (n_j(x_j)^{\frac{1}{1+\epsilon_j}})^{1+\epsilon_j} = L^{1+\epsilon} \quad (2.13)$$

Økt arbeidsinnsats i sektor j assosieres med flere produsenter i mellomsektoren. Fra (2.12) og (2.2) observerer jeg at sluttprodusenter blir mer produktive med et økt antall mellomprodusenter. Med andre ord vil en økt sluttgodeproduksjon i kraft av deling av en større variasjon av innsatsvarer kreve en mindre enn proporsjonal økning i primærfaktorer som gjør sluttgodeprodusenter mer produktive. ⁷

Denne effekten er blitt kritisert for å være mekanisk. Det at høyere arbeidsstyrke fører til produksjon av flere typer innsatsvarer, og dette øker produksjon av sluttgoder mer enn proporsjonalt på grunn av konstant substitusjonselastisitetssaggregering hos produsenter av sluttgoder er et teoretisk argument. Enhver utvidelse av produksjon av innsatsvarer blir gitt av en økning i antall tilbydere av innsatsvarer, og ikke gitt av en økning i skala innafor produsenter av innsatsgoder. Med andre ord vil økt arbeidsstyrke endre marginen til produksjon.

Dele gevinster av individuell spesialisering

En alternativ fremstilling til denne teorien ble utviklet av Adam Smith (1776). Hvor utgangspunktet er likt som i mellom- og sluttkoblinger av produksjon, men fokuset er noe annet. Smith mener det handler om å dele gevinster av individuell spesialisering, hvor han heller tar for seg den intensive marginen av produksjon som grunnlag for gevinster. Flere arbeidere øker produksjon mer enn proporsjonalt fordi det frigjør arbeidere til å

⁶For mellomregninger se appendiks.

⁷Primærfaktorer er land, arbeidskraft og kapitalgoder.

spesialisere seg til et mer snevert sett av arbeidsoppgaver, og ikke fordi nye arbeidere kan gjøre nye arbeidsoppgaver. Dette underbygges med hovedsakelig tre forskjellige årsaker. Arbeidere som utfører én arbeidsoppgave, blir gode på akkurat denne arbeidsoppgaven. Arbeidere slipper å omstille seg til andre arbeidsoppgaver, som gir lavere transaksjonskostnader, og til slutt gir en større arbeidsstyrke flere arbeidsbesparende innovasjoner fordi enklere oppgaver kan mekaniseres.

Her vil det være tiltakende skalaavkastning i produksjon av hver arbeidsoppgave, dermed er det i denne modellen gevinster på individnivå. Disse individuelle gevinstene kommer av hver enkelt arbeiders marginale produktivitet i en gitt arbeidsoppgave øker med spesialisering i den oppgaven som utføres. Graden av skalaavkastning er drevet av gevinsten fra arbeidspesialisering, og ikke av substitusjonseffekten som tidligere. Økt arbeidsstyrke gir dypere fordeling som gjør hver enkelt arbeider flinkere i utførelsen av gitt arbeidsoppgave, dette øker produktiviteten. På grunn av spesialisering øker produksjon mer enn proporsjonalt. Med fokus på læring i praksis vil dette gjøre den enkelte arbeider mer produktiv.

Deling av risiko

Til slutt har vi deling av risiko, som bygger på at lokalisert industri har konstant tilgang på faglært arbeidskraft, eller i det minste minimerer risikoen for at de ikke har tilgang på dette i tider hvor det trengs. Denne effekten er også kalt arbeidsstokk-effekten.⁸ Gevinsten av dette kommer via det faktum at en større arbeidsstokk via en samlet og stor industri har fordeler ved at den kontinuerlig vil kunne tilby et konstant marked for ferdigheter. Marshall(1890) påpekte denne gevinsten, mens Krugman (1991) formaliserer dette resonnetet ved å se på en rekke sektorer hvor bedrifter møter ideosynkratiske sjokk.⁹ Hver arbeider blir betalt sitt marginale produkt, og individuell profitt er konveks i disse bedriftsspesifikke sjokkene fordi bedriftene responderer sjokkene ved å justere produksjon og antall sysselsatte. Føventet lønn er tiltakende i antallet bedrifter, dette fordi en reduksjon i sysselsatte i hver bedrift gir høyere marginalprodukt av arbeidskraft, og dermed høyere lønn.

Endringer i bedriftens sysselsatte affekterer lokal lønn, og effekten er større desto lengre unna etablert bedrift er fra andre bedrifter i sektoren. Dersom lønn er høyere når bedrifter ønsker å ekspandere produksjon i respons til et positivt sjokk, og redusere i møte med negativt sjokk, vil bedriftens evne til å justere antall sysselsatte i gode og dårlige tider

⁸Arbeidsstokkeffekten er labor pooling effect.

⁹Ideosynratiske sjokk er sjokk som kommer uforventet for bedriften.

være geografisk betinget. Dette gjør at bedrifter som erfarer store endringer i sysselsatte relativt til andre bedrifter som benytter arbeidere med samme ferdighetssett, finner fordelene ved å lokalisere seg i tettere befolkede områder hvor de har tilgang på denne typen arbeidere. Her vil agglomerasjonsfordeler oppstå fordi en stor konsentrasjon av arbeidere jevner ut disse ideosynkratiske sjokkene, og fasiliterer overføring av arbeidskraft fra bedrifter med lav til høy produktivitet.

2.1.2 Matching

Matching baserer seg primært på mekanismer hvor agglomerasjon forbedrer forventet kvalitet av matcher, øker sannsynligheten for at matcher faktisk finner sted, og eliminerer eller minimerer problemer knyttet til forsinkelser.

Eksternaliteter knyttet til matching er knyttet til det teoretiske faktum at økt antall individer som prøver å finne en match, vil øke kvaliteten på matchen som finner sted.

Produksjon til inntekt per arbeider

Produksjon til inntekt per arbeider bygger på at individuell nytte øker med arbeidsstyrkens størrelse. Ikke bare på grunn av at økt konkurranse gir aggregert tiltakende skalaavkastning, men også på grunn av eksternaliteter fra matching. I en slik setting inkorporeres mis-match kostnader, som er kostnader påført det enkelte individ dersom dets ferdigheter ikke står i stil med jobben som skal gjøres. En økt arbeidsstyrke og økt antall bedrifter gjør at den gjennomsnittlige arbeider med høyere sannsynlighet finner en arbeidsgiver som står i stil med arbeiderens ferdigheter. Gitt at kostnaden betales av arbeideren selv, påføres individet en lavere mis-match kostnad som øker inntekten.

Økt sannsynlighet for matching

Økt sannsynlighet for matching betyr at et økt antall agenter som ønsker å matche, vil øke sannsynligheten for i det hele tatt å finne sitt motsvar.¹⁰ En proporsjonal økning i antallet jobbsøkere og ledige stillinger resulterer i en større enn proporsjonal økning i antallet jobbtreff. Flere individer i et område vil redusere funksjoner som kommer av søken etter arbeidere til ansettelse av ledige stillinger, samt gi fall i antall arbeidsledige og ledige stillinger. I en slik setting er det to kilder til agglomerasjonseffekter. Først og fremst vil det

¹⁰Forståelsen av å finne sitt motsvar er i denne settingen at bedriftens ferdighetskrav står i stil med arbeiderens ferdigheter.

være en matchingeksternalitet. Forståelsen av denne er at desto flere agenter som forsøker å matche, jo bedre blir matchene. Det andre aspektet er det at sterkere konkurranse vil gi sparing i form av reduserte faste kostnader. Disse representerer blant annet ressurser brukt til rekruttering og monitorering av arbeidere for å finne den best egnede kandidaten til ferdighetskravene. Grunnen til at dette skjer er fordi antall bedrifter øker mindre enn proporsjonalt med arbeidsstyrken. Her er kilden til tiltakende skalaavkastning konkurranse mellom bedrifter, da økt arbeidsstyrke gir økt arbeidsmarkedskonkurranse. Dette gjør at søken etter kandidater krever mindre ressurser siden markedet er fullt av potensielle arbeidere via en større arbeidsstyrke.

Eliminering av forsinkelser grunnet kontraktmessige ufullstendigheter

Eliminering av problemer knyttet til forsinkelser innebærer at kontraktmessige ufullstendigheter og forholdsmessige investeringer ofte er utsatt for ex-post¹¹reforhandlinger. Investorer gjør investeringer ex-ante,¹² hvor en slik reforhandlingsprosess i etterkant av en hendelse er lite gunstig for aktør som må ta en slik beslutning. Dersom denne aktøren kan ta ut sine investeringer før hendelse er inntruffet, og substituere seg til alternative partnere, vil problemer knyttet til slike forsinkelser være et mindre problem, og i beste fall eliminert. Det kan argumenteres for at urbane områder fasiliteter flere alternative partnere, og på denne måten besitter fordeler innen forretningsmessige avtaler.

2.1.3 Læring

Læringsmekanismer består primært av utvikling, diffusjon og spredning av kunnskap. Alt-omfattende utdanning, trening og forskning er en viktig aktivitet både i form av ressurser dedikert til det, og dets bidrag til økonomisk utvikling. Et velkjent trekk ved læring er at det sjelden eller aldri skjer i ensomhet, men mellom ansikt-til-ansikt interaksjoner. Derav kan også læring fasiliteres enkelt og med større kvalitet i tett befolkede områder. Mulighetene for læring i byområder er et sterkt argument for byens eksistens. Marshall (1890) argumenterte for at byer har fordeler i spredning av innovasjon og idéer. Jacobs (1969) vektlegger byens miljø for spredning av gode idéer. Det er naturlig å tro at byer fasiliteter kunnskap på en bedre måte nettopp på grunn av at samspillet mellom individer er en betingelse for at det i det hele tatt utfolder seg. Disse koblingene mellom mennesker gir nye impulser som potensielt kan fostre gode idéer. Læring klassifiseres som nevnt ovenfor

¹¹Ex-post er etter en hendelse har intruffet.

¹²Ex-ante er før en hendelse finner sted.

primært i tre kategorier: kunnskapsutvikling, spredning og akkumulering.

Kunnskapsutvikling

Kunnskapsutvikling handler i prinsippet om at nye bedrifter søker inspirasjon til sin ideale produksjonsprosess. Hvor mangfoldige urbane miljøer spiller en rolle i søken og eksperimentering av innovasjoner. Kombinasjonen av opphoppingskostnader¹³ og lokaliseringsekonomi gir statiske fordeler av urban spesialisering. Dette via at bedrifter prøver seg fram med ulike prosesser før de finner sin optimale produksjonsprosess. Den initielle investering for den enkelte bedrift er kostbar på grunn av disse opphoppingskostnadene, likevel anses det som en god investering da bedrifter må gjennom en rekke produksjonsprosesser for å finne den ideale. Bedrifter oppstår fordi andre bedrifter nedlegges, optimale investeringer vil da sørge for at de nye bedriftene produserer nye produkter.

Kunnskapsspredning

Kunnskapsspredning innebærer transmisjonen av kunnskap og ferdigheter, hvor nærhet til individer med høy kunnskap eller ferdigheter tilrettelegger for oppnåelse av disse egenskapene. Jovanovic, Rob (1989); Jovanovic, Nyarko (1995) og Glaeser (1999) presenterer en formell fremstilling av denne mekanismen.

Modellen bygger på at individer er ufaglærte initielt, men har muligheten til å oppnå kunnskap i løpet av sitt unge liv. Denne kunnskapen belæres av eldre individ. Her argumenteres det med at tett bebygde områder fasiliteter spredning av kunnskap bedre nettopp på grunn av betingelsen om ansikt-til-ansikt interaksjoner. Også i dette scenariet er investering i kunnskap en kostbar affære. Med basis i mange ulike kostnadsscenarioer, vil individene velge å søke etter kunnskap eller avstå fra det etter avkastningen av kunnskap kontra en kostnad som bæres av å bo i byregioner. Denne avkastningen er avgjørende for foredling av kunnskap over tidshorizonten fordi den tilrettelegger at kunnskapsrike mennesker forblir i byer, samtidig som den gir insentiver til oppnåelse av kunnskap blant unge individ.

Kunnskaps- og informasjonspredning

Kunnskaps- og informasjonsspredning har et litt annerledes fokus på spredning av informasjon enn tradisjonelle læringsmodeller. Denne effekten er ikke basert på individers

¹³Opphoppingskostnader er kostnader som kommer som konsekvens av at mange lokaliserer seg i et område.

ferdigheter, men heller på sosial læring. Dette kan eksempelvis være forplantning av rykter i byområder, agglomerasjon av kafeer i områder med mange restauranter, adaptering av gjødsel hos enkelte bønder samt læring av informasjonsspredning i nabolag.

Disse modellene har to avgjørende egenskaper. Først og fremst, ved å følge Banerjee (1992) og Bikhchandani, Hirshleifer og Welch (1992), finnes det muligheter for ineffektive investeringer. Med dette er forståelsen at dersom investeringer tas sekvensielt, er det sannsynlig at en dårlig investering hos en bedrift avler dårlige investeringer hos denne bedriftens følgere da de vil vekke den dårlige investeringen sterkt i sin beslutning.

Den andre egenskapen ved sosiale læringsmodeller er muligheten for strategiske forsinkelser. Dette bygger på at dersom timingen av beslutninger endogeniseres, vil ingen ønske å gjøre den første investeringen. I et slikt scenarie vil aktørene ha sterke insentiver til å avvente, da de kan få mye informasjon ut av å observere den første investeringen, men at det er for risikabelt å gjøre den selv da utfallet er ukjent.

Dette kan eksempelvis være utgangspunktet for et sentralisert forretningsområde. Fujita, Ogawa (1982) og Imai (1982) utleder et slikt endogent næringsområde, og gir en slik type spesifisering flere potensielle muligheter avhengig av kostnaden av å være langt borte fra sentralt næringsområde relativt til en opphopningskostnad. Mange forskjellige urbane konfigureringer er mulig, fra et komplett sentrum av bedrifter til stor spredning i regionen. Det presenteres likevel lite detaljert mikrogrunnlag for eksistensen av mekanismer på mindre arealmessig skala enn på bynivå.

Akkumuleringen av kunnskap

Akkumuleringen av kunnskap bygger, som de fleste vekstmodeller, på to relasjoner som beskriver produksjonen av ulike goder og akkumuleringen av produksjonsfaktorer. Hver arbeider har avtakende utbytte til akkumulering av human kapital, mens på by-nivå har vi tiltakende avkastning. Intuisjonen bak dette er at desto flere individer med kunnskapskapasitet, desto mindre vil verdien av den enkeltes human kapital være fordi flere tilbyr den samme kunnskapen. På aggregert nivå vil et større human kapitalt omfang medføre positive eksternaliteter via at flere kunnskapsrike individer gir positive eksternaliteter i form av ideer og innovasjon, og dermed er utbytte tiltakende. Dette medfører positive eksternaliteter, men det er ingenting som entydig tilsier at dette er grunnet kunnskapsoverføringer mellom arbeidere. Grunnlaget kan like gjerne være matching eller deling som læring.

I følge Lucas (1988) og Eaton, Eckstein (1997) er det mulig å måle urban vekst ved hjelp

av **dynamiske eksternaliteter**. Her er vekst drevet av eksternaliteter i akkumulasjon av human kapital i byregionen. Agglomerasjonsmekanisme via det faktum at byregioner har høyere grad av kapitalmengde enn mindre regioner.

Følgelig kan det argumenteres for at ulike mikroøkonomiske mekanismer kan rettferdiggjøre eksistensen av byregioner. Dette viser at agglomerasjonsmekanismer er robust i mange ulike spesifikasjoner, og for ulike mikroøkonomiske rammeverk. På den andre siden er dette en utfordring hva gjelder økonometrisk identifikasjon. Med dette mener Duranton og Puga (2004) at det er svært vanskelig å skille mekanismene fra hverandre økonometrisk, og derav konkret påpeke hvilke av mekanismene som spiller en dominant rolle i ulike geografiske områder. Dette er grunnen til at jeg har presentert et bredt omfang av teori som underbygger urbane fordeler. Siden den metodiske delen er lite spesifikk er det vanskelig å finne noe teorispesifikt grunnlag som underbygger de mer generelle resultatene.

En annen utfordring som oppstår er heterogenitet blant bedrifter og arbeidere. Ulike bedrifter besitter ulike forutsetninger for produksjon, og modellene er forskjellige i form av horisontale karakteristikk¹⁴ av godet de produserer. Slik horisontal differensiering er meget vanskelig å fange opp økonometrisk, det samme vil gjelde for heterogenitet i ferdigheter blant arbeidere.

Ufullstendig informasjon spiller også en avgjørende rolle. Byer gjør det enklere å finne ”inputs” som arbeidere, mellomgoder, kunder, fasiliteter eksperimenter og oppdage nye muligheter. Denne ideen er spesifikt transparent i læringsmodeller, men også i både matching og deling. Dette underbygger at gevinsten fra byer sannsynligvis er enda viktigere når de økonomiske omstendighetene er usikre og i rask endring.

Både deling og matching er velutviklede mikroøkonomiske grunnlag for agglomerasjonsgevinster. Læring er derimot ikke like velutviklet som for kunnskapsoverføringer. Gitt den viktigheten av slike overføringer for både utviklingen av byer, vekst og innovasjon, bør dette dedikeres et sterkere fokus i kommende forskning.

Har i dette kapitlet tatt for meg det teoretiske fundamentet som underbygger urbane fordeler. Mekanismene er mange og identifikasjon vanskelig, mer om dette samt relaterte løsninger i kommende kapittel.

¹⁴Horisontale karakteristikk er produktegenskaper som objektivt ikke kan skilles mellom hverandre

Kapittel 3

Tidligere forskning

Kapittelet tar til hensikt å skille mellom de typiske inndelingene av mekanismer, deretter diskutere de to største utfordringene knyttet til identifikasjon; det kvalitative og kvantitative endogenitetsproblemet. Til slutt hvordan tidligere litteratur har løst utfordringene knyttet dette, og egne løsninger.

3.1 Statiske og dynamiske effekter

Tidsaspektet av agglomerasjonsmekanismene er avgjørende. Aksept for at produktivitet og lønn justeres relativt raskt til variasjon i agglomerasjonsmekanismene gjennom priser og enhetskostnader på andre innsatsvarer enn arbeidskraft kan argumenteres for. Men de tilpasses typisk ikke like raskt til rene lokale eksternaliteter som affekterer produktivitet eller arbeidernes ferdigheter i et område. Dette gjør at faglitteraturen typisk har skilt mellom statiske og dynamiske agglomerasjonsmekanismer.

Ved statiske mekanismer vil den lokale sammensatte produktiviteten umiddelbart affekteres av lokale karakteristikk, men ikke av tidligere verdier. Dette betyr at en større bystørrelse i et gitt år påvirker lokal produktivitet kun i det året, og en fremtidig endring i bystørrelse vil umiddelbart overføres til en endring i lokal produktivitet.

Glaeser, Kallal, Scheinkman og Shleifer (1992) og V. Henderson, Kuncoro og Turner (1995) var noen av de første til å studere og poengtere dynamiske effekter av agglomerasjon. Til forskjell fra statiske mekanismer studerer de dynamiske mekanismene langsiktige effekter

av lokale karakteristikk. Disse mekanismene bygger på utvikling og konsistente effekter over tid. Dette kan eksempelvis være hvordan teknologiske overføringer over tidsperioder øker produktivitetsveksten, eller individers evne til å lære mer eller raskere i større byer. Det argumenteres til og med for at når individer flytter fra en større by til en mindre, vil en del av produktivitetsgevinsten oppnådd i urbant område kunne overføres til ny lokasjon. Dermed vil individet være mer produktiv i ny lokasjon relativt til individer som ikke har bodd i storbyregion. I slike tilfeller vil dynamiske effekter operere gjennom effekten av lokale karakteristikk på ulike mål for produktivitet og arbeidernes ferdigheter i lokalt område. Ekvivalent med statiske agglomerasjonsmekanismer kan også dynamiske mekanismer ha en effekt via pris og enhetskostnad på innsatsvarer annet enn arbeidskraft. Dette ved at mekanismene tilrettelegger for spredning av informasjon om kvaliteten på varer og inputs, som medfører prisvariasjon på tvers av tidsperioder.

De dynamiske effektene er mer omfattende og man evner å gå mer detaljert til verks når man studerer disse effektene. Denne oppgaven har som nevnt fokus på de umiddelbare effektene, det vil si statiske agglomerasjonsfordeler av urbane områder.

3.2 Tidligere litteratur

Litteraturen på sammenhengen mellom geografi og produktivitet er blitt avdekket grundig gjennom historien, det samme gjelder for forståelsen av hvorvidt byer besitter en høyere økonomisk aktivitet. Adam Smith var allerede oppmerksom på sammenhengen mellom økonomiske gevinster og geografisk beliggenhet i 1776. Mekanismene som underbygget disse fordelene ble først påpekt av Alfred Marshall (1920), som understreket teknologiske overføringer fra en bedrift til en annen i nærheten. Henderson (1974) formaliserte Marshall sine ideer og demonstrerte at i likevekt vil ulemper som kommer av agglomerasjon for husholdninger kunne motvirke produktivitetsgevinstene bedriftene opplever av det. Krugman (1991) demonstrerte at agglomerasjon oppstår selv når transportkostnadene er relativt lave og arbeiderne er mobile. Ciccone og Hall (1993) argumenterte for at agglomerasjon forsterkes når teknologi fastsettes endogen. Glaeser og Mare (2001) finner ved å sammenlikne lønn mellom områder et mer direkte bevis på eksistensen og størrelsen på agglomerasjonsmekanismer. De finner at lønna er høyere i byer, og at dette urbane lønnspremiumet er høyere desto lenger arbeideren har vært i dette området. En felles begrunnelse for denne metoden er at i konkurranseutsatte arbeidsmarkeder er arbeidskraften betalt sitt marginale produkt. Selv uten konkurranseutsatte arbeidsmarkeder kan varia-

sjon i lønn representere høyere produktivitet, fordi høyere lønninger har en kompenserende effekt for høyere bolig- og pendlekostnader i tettere befolkede områder. Her diskuterer de i tillegg utfordringer knyttet til tolkning av det urbane lønnspremium som bevis for agglomerasjonsmekanismer. Problemet relatert til dette er at arbeidere med større ferdigheter kan være systematisk lokalisert i favør av enkelte geografiske områder. Dekle og Eaton (1999) ser på sammenhengen mellom leie betalt i ulike områder, hvor tanken er at dersom bedrifter finner lokalisering i befolkede områder, der leieprisen er høyere, må dette være grunnet noen underliggende produktivitetsfordeler dette området representerer.

Det teoretiske fundamentet som underbygger agglomerasjonsfordeler er velutviklet og står sterkt innen faglitteraturen. Innen de empiriske studiene har utfordringen vært å skille mellom effektene som leder til agglomerasjonsfordeler. Her har fokuset vært på by- og industristørrelse som determinanter for produktivitet og teknologiske overføringer. Sveikauskas (1975), Segal (1976) og Moomaw (1981) estimerte effekt av bystørrelse mot produktivitet. J.V. Henderson (1986) fant at produktiviteten i bedrifter øker med størrelsen på industrien, målt i industriell sysselsetting. De empiriske bevisene antyder at større markeder har flere potensielle entreprenører, hvor kun et lite antall av disse som entrer markedet vil overleve. De som overlever markedet vil prestere signifikant bedre, noe som insinuerer at tettere markeder har større grad av ulikhet. Syverson (2004, 2007); Campbell og Hopenhayn (2005) styrker beviset for seleksjon ved å studere større markeder i USA. Combes et al (2012) finner ingen bevis for seleksjonseffekter når de sammenlikner små og større byregioner i Frankrike. Her i Norge har Carlsen et al. (2016) studert sammenhengen mellom urbant lønnspremium og utdanningsnivå samt identifiseringen av agglomerasjonsmekanismer, og funnet klare resultat i favør byregioner.

P.-P. Combes et al. (2010) finner at elastisiteten mellom lokal produktivitet og sysselsettingstetthet for franske regioner ligger mellom 0,04 og 0,10.¹ Carlsen, Rattsø og Stokke (2013) finner at den rene elastisiteten av gjennomsnittlig lønn mot populasjon ligger på 0,076, videre finner de ved korrigerer av observerte arbeidseffekter et estimat på 0,045, ikke ulikt Combes et al.(2010). Når de korrigerer for endogenitetsproblemet med historisk befolkningsdata får de et estimat på 0,054, og til slutt faller koeffisienten til 0,033 ved kontroll for både endogenitet og sortering. Ved bruk av populasjonstetthet får de et rent elasisiteseestimert på 0,055 som faller ved korrigerer av sortering til 0,04, og til slutt 0,034 når de korrigerer for både endogenitet og sortering.

¹En elastisitet mellom 0,04 og 0,1 betyr at ved en dobling av sysselsetting i et område vil produktivitet øke med mellom 4 og 10 prosent

Det kan med andre ord tyde på at både det teoretiske- og empiriske fundamentet trekker i retning av at befolkede områder opplever enkelte fordeler. Hvor de empiriske resultatene, til tross for utfordringer relatert til identifisering av de reelle agglomerasjonsfordelene, indikerer at det teoretiske fundamentet kan underbygges statistisk sett.

Har ovenfor skilt mellom statiske og dynamiske mekanismer, avdekket tidligere faglitteratur på området samt gjort meg kjent med de klassiske økonometriske utfordringene ved estimering av agglomerasjon. For denne oppgaven vil de økonometriske utfordringene primært være sortering av arbeidere, og simultan bestemmelse av mål på urbanitet og produktivitet. Belyser nedenunder hvorfor dette er utfordringer, og hvordan jeg potensielt kan løse for dette.

Diskusjon av sortering

Sorteringsproblemet er i faglitteraturen omtalt som det kvalitative endogenitetsproblemet. Forståelsen av dette er at bedrifter og arbeidere ikke er homogene i sine karakteristikk, men ofte, i større eller mindre grad, er forskjellige fra hverandre.

Videre skilles slike karakteristikk mellom det som typisk er enkelt observerbart og det som er uobserverbart. For individers del kan dette være utdanningsnivå, som er observerbar. Eller eksempelvis personlighetstrekk, antall barn i familien og mobilitet i barndommen som er uobserverbare eller vanskelig å fange opp. Dersom disse ferdighetene skulle være geografisk betinget, det vil si at de er oppnådd i det regionale området, vil det ikke være nødvendig å korrigere for slike effekter i den økonometriske tilnærmingen fordi fordelene kommer av området i seg selv.

J.V. Henderson (1974) illustrerer en komplementaritet mellom talent og agglomerasjonsmekanismer. Hvor en større by gir en disproporsjonal gevinst mot mer talentfulle individer i byregionen. Videre vises det til at tøffere seleksjon virker frastøtende på individer, som gjør at de mest talentfulle blir igjen. Det er også et omvendt kausalt forhold mellom disse komplementære faktorene da komposisjonen av talent i en by varierer med bystørrelse, dette er seleksjon. Behrens, Duranton og Robert-Nicoud (2014) drar resonnementet videre og studerer kombinasjonen av sortering, seleksjon og agglomerasjon, og relaterte komplementariteter. Slike komplementariteter er vanskelig å skille økonometrisk. Helsley og Strange (2014) viser at diversifiserte og spesialiserte byer eksisterer side om side. Duranton og Puga (2005) viser at byer kan være forskjellige både horisontalt i form av industrier de besitter og vertikalt i form av hvilke funksjoner de utspiller. P.-P. Combes et al. (2010)

jobber med fransk data, og finner at elastisiteten av lønnspremiumet assosiert med hver byregion omtrentlig halveres når faste effekter er introdusert som metode. Dette tolker de som at effekten av sortering potensielt er stor.

Carlsen et al. (2016) adresserer sortering i Norge ved å sammenlikne arbeideres utdanningsnivå med individspesifikk data over en tidsperiode på 7 år.² Her undersøker de agglomerasjonseffekter og sortering basert på observerte og uobserverte karakteristikk på tvers av utdanningsgrupper. Utdanningsgruppene er grovt inndelt i fullført grunnskole, fullført minst ett år videregående, og fullført minst ett år høyere utdanning. De estimerer både effekten i Oslo og seks andre storbyregioner.³ Her finner de, fra ren elastisitet til korreksjon av både observerte og uobserverte karakteristikk, at endringen typisk er på 12,2 prosentpoeng for Oslo⁴, og 8,2 prosentpoeng for de andre storbyregionene.⁵ Dette indikerer at sortering, både av observerte og uobserverte karakteristikk, er et faktum, og desto større i befolkede områder. I tillegg til dette benytter de en kontinuerlig tetthetsvariabel som måler populasjon per kvadratkilometer. Kort oppsummert finner de at sortering basert på uobserverte karakteristikk er av betydning for universitetsutdannede arbeidere, selv når de tar hensyn til dynamiske læringseffekter. Dette er også tilsynelatende viktigere tidligere i arbeidskarrieren. Avkastningen av høyere utdannede individer er både på kort og mellomlang sikt høyere enn for andre utdanningsgrupper.

Tidligere forskning gir klare indikasjoner på at sortering er reelt, og at dette skjer i favør av urbane områder. Det er naturlig nok vanskelig å påpeke én faktor som determinerer sorteringen til befolkede områder, men det er ofte i faglitteraturen vist til at slike individer verdsetter bekvemmeligheter som kulturliv⁶ mer enn andre. Eller som Wheeler (2001); Glaeser og Resseger (2010) viser, at agglomerasjonsgevinster er sterkere blant talentfulle individer, og at disse individene trekker til områder hvor denne gevinsten oppleves å være størst.⁷ Carlsen et al. (2013) viser at norsk data er konsistent med disse funnene, graden av statiske agglomerasjonsgevinster arbeideren opplever øker med utdanningsnivået.

Grunnen til at dette er en utfordring er at ferdigheter og talent korresponderer med en sammensatt effekt av lokal arbeidsstøkk, og ikke økt produktivitet grunnet lokal in-

²Fra 2003 til 2010.

³Regioner med over 150.000 innbyggere.

⁴Fra 18,7 prosent i ren elastisitet, til 10,8 prosent ved kun observerte faktorer, og ytterligere til 6,5 prosent ved korreksjon.

⁵Fra 12,7 prosent i ren elastisitet til 6,5 prosent ved kun observerte faktorer, og ytterligere til 4,5 prosent ved korreksjon.

⁶Bekvemmeligheter er på fagspråket kalt "amenities"

⁷Definisjonen på talentfull er vid, begrenser meg her til karakteristikk som omfatter evnerik og flink

teraksjon blant arbeidere med høyere ferdigheter. Ergo gjenspeiler slike karakteristikk fordeler, men de er ikke nødvendigvis relatert til agglomerasjonsmekanismer. Rollen til ferdighetssettet til lokale arbeidere med ferdigheter som ikke er determinert av området de representerer, men oppnådd via arv eller utdanning, kan medføre positive skjevheter av agglomerasjon fordi det fanges opp av koeffisienten foran forklaringsvariabelen som representerer urbanitet.

Resonnementene ovenfor underbygger at slik heterogenitet lenge har vært et anerkjent faktum innen faglitteraturen på området, og kan potensielt skape en stor skjevhet i estimatene. Carlsen et al. (2016) beviser at dette også gjelder for Norge ved å studere individers avkastning av urbane områder for ulike utdanningsnivå.

Norge har historisk hatt et kvantitativt fokus på utdanning, noe som betyr at spredningen av kunnskap blant befolkningen har vært vektet tungt. Analysen videre har ikke til hensikt å skille mellom hvor omfattende skjevheten kan være, slik at en aksept og bevissthet for at den faktisk er tilstedeværende er tilstrekkelig for videre resonnement.

Tradisjonell håndtering av sorteringsproblemet

For korrigerende effekter relatert til sortering er det hensiktsmessig å benytte detaljert individdata. Sortering innebærer både observerte og uobserverte karakteristikk. Carlsen et al. (2016) bruker utdanningsnivå hos individer som kontroll for de observerte karakteristikkene for individer. For de uobserverte karakteristikkene som typisk er vanskelig å måle har en utbredt økonometrisk metode for korrigerende vært bruken av individdata og faste effekter. En slik metode var først foreslått av Glaeser og Mare (2001). Dette muliggjør kontrollering for individkarakteristikk relatert til individets evne til å adaptere ferdigheter som ikke varierer over tid. Dette inkluderer blant annet karakteristikk som utdanningsnivå til foreldre og besteforeldre, antall barn i familien, mobilitet i barndommen og personkarakteristikk.

En annen fordel med bruken av individdata er at de faste effektene kontrollerer for individenes utdanningsnivå, slik at en kan implementere den lokale andelen av utdanningsnivå for å vurdere hvorvidt arbeidere med høyere ferdigheter utøver en positiv eksternalitet i spredning av human kapital til andre arbeidere i regionen. I et slikt tilfelle vil målet på urbanitet reflektere alle andre mekanismer enn kunnskapsoverføringer. Litteraturen har tradisjonelt sett benyttet detaljert data på arbeideres karakteristikk og faste effekter som økonometrisk metode for å korrigere for sortering (P.-P. Combes, Duranton & Gobillon, 2008) og (Matano, Naticchioni et al., 2009; Matano & Naticchioni, 2011).

Håndtering av sorteringsproblemet

En forutsetning for å kunne bruke den tradisjonelt beste tilnærmingen til korrigerings av sorteringsproblemet krever detaljert individdata over tid.⁸ I min oppgave vil denne tilnærmingen være vanskelig fordi jeg benytter meg av regional data for 2010.

Observerte faktorer korrigerer jeg for i analysen ved å inkludere total andel med høyere utdanning og videregående skole. Videre finnes det også andre uobserverbare karakteristikk som mobilitet i barndom, antall søsken, utdanningsnivå til foreldre og personkarakteristikk. En ulempe ved å bruke regional data er at dette blir noe vanskeligere å kontrollere for på regionalt nivå. Et alternativ kunne vært å benytte en proxy for disse uobserverte karakteristikkene, men dette er svært utfordrende å finne på regional basis.⁹ Resultatene jeg oppnår i analysen vil dermed etter all sannsynlighet inneha en grad av skjevhet. Dette fordi jeg ikke kontrollerer for de vanskelig målbare karakteristikkene.

Det kvantitative endogenitetsproblemet

Det kvantitative endogenitetsproblemet bygger på det faktum at en høyere lønn i enkelte områder, tiltrekker flere individer til dette området. Dette betyr at lønn og populasjon potensielt er simultant bestemt. Høyere lønn trekker populasjon til tett befolkede områder som igjen øker lønna som tiltrekker flere individer til området. En tradisjonell økonometrisk metode til håndtering av endogenitetsproblemet er å benytte et historisk befolkningsinstrument på populasjonsvariabelen. Det var Ciccone og Hall (1996) som var de første til å behandle endogenitetsproblemet ved å benytte lange laggs på populasjon som instrument. P.-P. Combes et al. (2008) argumenterer for at historisk populasjon kan være affektet av permanente karakteristikk, og dermed nåværende produktivitet. Her foreslår de et alternativt instrument i geologiske karakteristikk, og at dette kan være et potensielt godt instrument.

Til tross for potensielle problemer ved bruk av historisk befolkningsdata som instrument på populasjonsestimatet, kommer jeg i denne oppgaven til å bruke populasjon per kvadratkilometer i 1910 som instrument. Det er naturlig å tro at populasjon per kvadratkilometer i 1910 påvirker bosetning i 2010, men ikke nødvendigvis lønn i 2010.

I dette kapittelet har jeg gjort meg kjent med de ulike mekanismene, tidligere faglitteratur og de økonometriske utfordringene. Går i de kommende kapitlene inn på den formelle

⁸Dette er ekvivalent med paneldata.

⁹En proxy er en fungerende representant som forklarer og er representativ for det jeg ønsker å forklare.

metoden, og de økonometriske utfordringene nevnt ovenfor samt relaterte metodiske løsninger.

Kapittel 4

Empirisk tilrettelegging

Presenterer her inngående variable i analysen, datamaterialet, og det empiriske grunnlaget i form av tilrettelegging og utfordringer relatert til analysen.

De regionale inndelingene er hentet fra SSB sitt rammeverk *Standard for økonomiske regioner*¹ og er delt inn i totalt 89 økonomiske regioner.² Økonomisk region er en regional inndeling på nivået mellom fylke og kommune. Inndelingen bygger på kriterier som handel og arbeidsmarked og skal representere et hensiktsmessig publiseringsnivå for statistikk samtidig som den skal tilsvare det regionale nivået som EU har definert som sin NUTS4-inndeling. En konsekvens av dette, er at de økonomiske regionene ikke kan krysse fylkesgrensene.

4.1 Avhengig variabel

Som avhengig variabel har det i litteraturen primært vært enten lønn eller total faktorproduksjon blitt brukt som tilnærming til produktivitet. Lønn er det enkleste målet på produktivitet da definisjonene på total faktorproduksjon kan være mange. Den enkleste måten å implementere relasjonen gjennomsnittlig lønn er fra P.-P. Combes et al. (2010) definert ved å ta snittet til hvert enkelt individ i ulike kommuner og aggregere til økonomisk region.

I denne oppgaven er data hentet fra Carlsen et al. (2016), er aggregert opp fra individnivå, og tilsvarende gjennomsnittlig timelønn i aktuell region for alle fulltidarbeidere i aldersgruppen 20-65 år, både menn og kvinner. Arbeidere i primærnæringene jordbruk, skogbruk og

¹<http://www.ssb.no/offentlig-sektor/artikler-og-publikasjoner/standard-for-okonomiske-regioner>

²Nordre Sunnhordland endret navn til Sunnhordland og nummer til 1296. Den økonomiske regionen 1292 Søndre Sunnhordland utgikk som økonomisk region.

fiske er ekskludert, siden lønningene her er spesiell og har lite med urbanisering å gjøre. Lønns målet er timelønn inkludert arbeidsgiveravgift. Grunnen til at arbeidsgiveravgift inkluderes er at jeg ønsker et mål på produktivitet, og da er lønnskostnaden for bedriftene det nærmeste jeg kommer.

4.2 Forklaringsvariabler

Disse variablene har til hensikt å fange opp både områdekarakteristikker og individ- eller bedriftskarakteristikker. Siden jeg i oppgaven tar for meg regionale forskjeller, vil individkarakteristikker være vanskelig å fange opp.

Hovedforklaringsvariabelen i oppgaven er ulike mål på urbanitet. Her har jeg potensielt flere alternativer, blant annet logaritmisk tilnærming til populasjonen per kvadratkilometer samt absoluttverdien av populasjon. Jeg har valgt å bruke populasjon per kvadratkilometer som hovedforklaringsvariabel, og absoluttverdien av populasjon i en robusthetsanalyse. Det er fordeler og ulemper ved å bruke populasjonstetthet som hovedforklaringsvariabel. Det kan tenkes at tettheten i større grad reflekterer bosettingsstrukturen innad i de økonomiske regionene. Samtidig vil kommunene innad i de økonomiske regionene variere i areal og populasjon, slik at for enkelte, tettere befolkede områder, kan det tenkes at andre kommuner i utvalget trekker deres reelle populasjonstetthet ned.³

Hovedforklaringsvariabelen er formelt gitt ved:

$$\ln\left(\frac{pop_i}{kvkm_i}\right) \quad (4.1)$$

Som er befolkningstetthet i hver enkelt økonomisk region. Hvor pop_i representerer populasjonen i økonomisk region og $kvkm_i$ er antall kvadratkilometer i regionen. Dette er populasjon per kvadratkilometer.

I tillegg til dette inkluderer jeg også dummy for byregioner som et siste alternativ. Disse tverrsnittsdummiene er definert for henholdsvis storbyregioner, som er byer over 150.000 innbyggere, og mindre byer med innbyggertall over 100.000. Dummy-resultatene er både interessante og nyttige med tanke på urbant gjennomsnittlig lønnspremium, og det potensielle endogenitetsproblemet ved korrelasjon mellom høyere utdanning og uobserverte karakteristikker. Derfor inkluderer jeg en liten kommentar på disse i hovedanalysen. Storbydummyen er definert av Oslo, Bergen, Trondheim, Stavanger, Drammen, Bærum og Lillestrøm. I byregion er Fredrikstad, Follo, Tønsberg, Skien, Kristiansand og Haugesund

³Et eksempel på dette er Trondheim, som er et relativt tett befolket område, men er omringet av mindre befolkede kommuner. Dette trekker deres snittverdi ned på overordnet økonomisk regionnivå.

inkludert.

Siden jeg ønsker å måle hvorvidt produktiviteten er høyere i tettere befolkede områder, og deretter hvorvidt det urbane lønnspremiumet varierer mellom sektorer, så er hensikten her å korrigere for eventuelle andre faktorer enn produktivitet med basis i agglomerasjonsmekanismer som påvirker produktivitet på regionalt nivå.

Utdanning på videregående- og universitetsnivå inkluderes, her benytter jeg andel av befolkning med videregående utdanning, og andelen av befolkning med høyere utdanning som impliserer fullførte 3-år eller mer på høyskole- eller universitetsnivå.⁴ Med befolkning mener jeg da de som har fullført graden i 2010 fordelt på total populasjon i økonomisk region, fremstilt ved:

$$\frac{vgs_i}{Populasjon_i} \quad (4.2)$$

$$\frac{Universitet_i}{Populasjon_i} \quad (4.3)$$

Hvor (4.2) er prosentvis andel av befolkning med fullført videregående utdanning, og (4.3) representerer prosentvis andel med høyere utdanning.

Ciccone og Hall (1996) argumenterer for at klima virker produktivitetsfremmende for arbeidere, og som en av faktorene for at individer flytter til byene. Bruker gjennomsnittlig januartemperatur for perioden 1994-2002 basert på data fra Meteorologisk Institutt, eklimate databasen:

$$\frac{\sum_{1996}^{2002} C_i}{T}, T = 7. \quad (4.4)$$

Grunnen til at denne perioden benyttes er at eklimate databasen besitter begrenset data på geografiske områder som er aktuelle for min oppgave.

Andelen sysselsatte i offentlig sektor er interessant og en nyttig indikator for min sekundære analyse. Prosentvis andel arbeidende i offentlig sektor i en gitt økonomisk region vil dermed være gitt ved:

$$\frac{L_{i,offentlig}}{(L_{i,Offentlig} + L_{i,Privat})} \quad (4.5)$$

Hvor L er sysselsatte og i representerer de respektive økonomiske regionene. Dataen er hentet fra SSB sin statistikkbank, og er aggregert opp fra kommunenivå til økonomiske regioner. Denne variabelen kan være interessant fordi det er naturlig å tro at agglomera-

⁴I en slik setting vil individer med fullført ett eller to år høyere utdanning kategoriseres i andelen fullført videregående skole.

sjonsmekansimene potensielt er sterkere for private næringer, og det offentlig til dels har en større grad av desentralisering enn hva privat sektor har.

Kultur er en bekvemmelighet for individer. Combes, Duranton, Gobillon og Roux (2010) argumenterer for at kulturtilbudet er større i byregioner, og at dette er et motiv for individer til å flytte til urbane områder. Her benytter jeg netto driftsutgifter til kultursektoren av den enkelte økonomiske regions totale driftsutgifter.⁵

4.3 Data

Dataen er primært hentet fra kommunedatabasen, statistisk sentralbyrå og eKlima. Kommunedatabasen NSD er et nasjonalt arkiv for forskningsdata. NSD arkiverer, tilrettelegger og formidler data til forskningsmiljøer, nasjonalt og internasjonalt, og utarbeider teknologiske løsninger.⁶

SSB sørger for at innbyggerne i Norge kan debattere, planlegge og ta beslutninger på grunnlag av pålitelig statistisk informasjon. Det meste av SSBs statistikk utarbeides med utgangspunkt i data fra administrative registre og spørreundersøkelser, og stadig mer informasjon hentes direkte fra datasystemene til næringsliv og kommuner. Hvis dataene ikke finnes i et administrativt register, kan informasjonen samles ved elektronisk rapportering. I tillegg gjennomfører SSB intervjuer, enten via telefon eller ved å oppsøke folk hjemme.⁷ eKlima er portalen til Meteorologisk institutts klimadatabase som er åpen for alle og gratis å bruke. eKlima inneholder data fra alle værstasjoner som Meteorologisk institutt har i drift i dag, og har drevet tidligere, samt data som andre lar de distribuere.⁸

Verken NSD, SSB eller eKlima er ansvarlig for analysen av data eller de tolkninger som er gjort her.

Deskriptiv statistikk

Presenterer i tabell 4.1 deskriptiv statistikk på alle inngående variable i hovedanalysen, mål på henholdsvis produktivitet og urbanitet, gitt ved lønn og populasjon per kvadrat-

⁵Driftsutgiftene er summen av lønnsutgifter inkludert sosiale utgifter, inventar og utstyr til driftsformål, andre kostnader til driftsformål, stønader, subsidier og overføringer til stat/andre kommuner/andre fylkeskommuner.

⁶<http://www.nsd.uib.no/om/>

⁷<https://www.ssb.no/omssb/om-oss/vaar-virksomhet>

⁸www.eklima.met.no

kilometer, utdanningsnivå, klima og driftsutgifter til kultur.

Tabell 4.1: Deskriptiv statistikk

| Variabel | Gjennomsnitt | Standardavvik | Minimum | Maksimum |
|--|--------------|---------------|---------|----------|
| Timelønn | 267,92 | 22,33 | 228,78 | 352,76 |
| Populasjon per kvadratkilometer | 45,08 | 149,62 | 0,735 | 1292,38 |
| Populasjon | 54651,71 | 83908,29 | 5219 | 586861 |
| Andel med universitet | 0,22 | 0,05 | 0,144 | 0,447 |
| Andel med videregående | 0,44 | 0,038 | 0,31 | 0,52 |
| Andel arbeidere i offentlig sektor | 0,33 | 0,05 | 0,219 | 0,54 |
| Gjennomsnittlig januartemperatur | -1,99 | 3,4 | -9,72 | 4,04 |
| Netto driftsutgifter kultur av totale driftsutgifter | 0,04 | 0,018 | 1,85 | 7,32 |

Ser at gjennomsnittlig timelønn for alle sektorer ligger på 267,92 kr timen, hvor Bærum har høyest gjennomsnittlig timelønn på 352,76 kr, mens Nord-Troms har lavest gjennomsnittlige timelønn på 228,78 kr.

Når det gjelder populasjonsparametrene presenterer jeg her populasjon per kvadratkilometer og absoluttverdien til populasjon. Ved populasjon per kvadratkilometer ligger snittet på 45 innbyggere per kvadratkilometer, hvor det tettest befolkede området er Oslo med 1292,38 innbyggere per kvadratkilometer. Minimum er på 0,73 innbyggere per kvadratkilometer, og er tilhørende Grong-regionen.

Gjennomsnittlig populasjon i de ulike regionene ligger på 54651,71, hvor Oslo har høyest innbyggertall på 581642 innbyggere, og Grong har det laveste innbyggertallet på 5219.

Andelen med høyere utdanning, gitt ved universitet, har et gjennomsnitt på 22 prosent. Frøya har den laveste andelen med høyere utdanning på 14,4 prosent, og den høyest utdannede regionen er Bærum med 44,7 prosent høyere utdannede.

For andelen fullført videregående er snittet over alle regioner på 44 prosent, hvor Oslo har lavest andel fullført videregående opplæring på 31 prosent, mens Odda har den høyeste andelen på 52,27 prosent.

Andelen arbeidere i offentlig sektor er i gjennomsnitt på 33 prosent av sysselsatte over alle økonomiske regioner. Ulsteinvik har lavest andel arbeidende i offentlig sektor på 21,9 prosent, mens Andselv har flest med sine 54 prosent av sysselsatte.

Den gjennomsnittlige januartemperaturen over 7 år ligger for hele Norge på -1,99 grader. Den kaldeste regionen over tidsperioden var, ikke overraskende, Kirkenes med sine -9,72 grader i gjennomsnitt. Det varmeste strøket over perioden var Ulsteinvik med 4,04 grader. Netto driftsutgifter til kultur av de totale utgiftene i den økonomiske regionen lå gjennomsnittlig på 4,4 prosent av budsjettene for alle økonomiske regioner. Hvor Mo i Rana dedikerte minst av budsjettet til kultur med sine 1,85 prosent av totale driftsutgifter og Setesdal med den høyeste andelen på 7,32 prosent.

Presenterer i tabell 4.2 deskriptiv statistikk relatert til gjennomsnittlig timelønn i de ulike sektorene.

Tabell 4.2: Deskriptiv statistikk sektorlønn

| Timelønn | Gjennomsnitt | Standardavvik | Minimum | Maksimum |
|-----------|--------------|---------------|---------|----------|
| Offentlig | 262,37 | 13,93 | 222,35 | 302,61 |
| Privat | 269,7 | 27,44 | 225,17 | 376,95 |
| Industri | 279,35 | 32,106 | 216,94 | 390,72 |
| Service | 259,93 | 26,36 | 216,84 | 373,41 |

Offentlig sektor har snittlønn på 262,37 kr med høyest snittlønn i Bærum på 302,61 kr og lavest i Nord-Troms på 222,35 kr i timen.

Timelønna i privat sektor har et snitt på 269,7 kr. Her er det Bærum som har høyest gjennomsnittlig timelønn på 376,95 kr, mens Nord-gudbrandsdalen besitter det laveste snittet på 225,17 kr i timen.

I industrien ligger gjennomsnittlig timelønn for alle økonomiske regioner på 279,35kr, her har Bærum høyest snittlønn på 390,72 kr, mens Vadsø ligger lavest med 216,94kr.

Servicenæringen har en snittlig timelønn på 259,93kr med høyest snittlønn i Bærum på 373,41 kr og lavest i Nord-Gudbrandsdal med snittlønn på 216,84kr i timen.

4.3.1 Økonometrisk tilrettelegging

Metodene for å kartlegge produktivitetsforskjeller har grovt sett vært inndelt i tre kategorier. Den første går ut på å bevise at produktive aktiviteter er mye mer klynget enn det som ville forventes dersom plasseringen var et tilfeldig utfall, eller bare reflekterte underliggende forskjeller på tvers av geografiske områder som fører til komparative fortrinn. Den andre tilnærmingen studerer forskjeller i områdekaraktéristikker i lønn og boligpriser. Dersom bedrifter og arbeidere er mobile, og lønn samt boligpriser varierer mellom områder, bør høyere lønn og boligpriser i enkelte områder reflektere produktivetsfordeler. Den siste metoden går på å analysere produktivitetsforskjeller på tvers av ulike områder ved bruk av detaljert data på individ- og bedriftsnivå.

Den metodiske basisen brukt i denne oppgaven vil være ekvivalent med å studere lønn mot områdekaraktéristikker, med et mål på urbanitet som hovedforklaringsvariabel. Den økonometriske delen er primært hentet fra Woolridge (2009). Basisen for den økonometriske

metoden er minste kvadraters metode.⁹ Denne metoden bygger på å minimere summen av kvadrerte avvik (Jeffrey et al., 2009).

I valget av kontrollvariable i modellen, er lønn og ulike mål på populasjon de viktigste variablene. Deretter inkluderer jeg enkelte forklaringsvariable der hensikten er å kontrollere ulikheter på regionalt nivå.¹⁰ Dermed vil basismodellen være gitt ved gjennomsnittlig lønn for ulike økonomiske regioner modellert på ulike mål på populasjon i de ulike regionene:

$$\ln W_i = \beta_0 + \beta_1 \ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right) + U_i \quad (4.6)$$

Hvor $i = 1, 2, \dots, 89$ økonomiske regionene i Norge. $\ln W_i$ representerer en logaritmisk tilnærming til gjennomsnittlig timelønn i de ulike økonomiske regionene, $\ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right)$ representerer logaritmisk tilnærming til populasjonstetthet, og U_i er et stokastisk restledd som fanger opp uobserverte variabler. Estimater foran hovedforklaringsvariablen er det som er av størst interesse i oppgaven, gitt ved β_1 .

Som nevnt i teoridelen vil tøffere seleksjon i byer kreve bedre arbeidere, som impliserer at tøffere seleksjon forsterker sortering av arbeidere på tvers av geografiske områder. Wheeler (2001); Glaeser og Resseger (2010) finner at agglomerasjonsgevinster er sterkere blant talentfulle individer, og at disse individene vil trekke til områder hvor gevinsten er størst. Dette gir sortering av arbeidere. Utvider her basismodellen med et mål på utdanning, slik at:

$$\ln W_i = \beta_0 + \beta_1 \ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right) + \beta_2 \text{vgs}_i + \beta_3 \text{universitet}_i + U_i \quad (4.7)$$

Videre inkluderer jeg regionspesifikke variable som andelen arbeidere i offentlig sektor, klima og kultur:

$$\ln W_i = \beta_0 + \beta_1 \ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right) + \beta_2 \text{vgs}_i + \beta_3 \text{universitet}_i + \beta_4 \text{offentlig}_i + \beta_5 \text{klima}_i + \beta_6 \text{kultur}_i + U_i \quad (4.8)$$

Minste kvadraters metode bygger da på å minimere summen av kvadrerte avvik fra sann, teoretisk modell:

$$\min \sum_{i=1}^{n=89} = (\ln W_i - (\hat{\beta}_0 + \hat{\beta}_1 \ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right) + \hat{\beta}_2 \text{vgs}_i + \hat{\beta}_3 \text{universitet}_i + \hat{\beta}_4 \text{offentlig}_i + \hat{\beta}_5 \text{klima}_i + \hat{\beta}_6 \text{kultur}_i))^2 \quad (4.9)$$

Hvor $\ln W_i$ er den sanne, teoretiske verdien.

⁹MKM er en estimeringsmetode for å finne sammenhengen mellom en eller flere forklaringsvariable og en responsvariabel. MKM estimerer ved å finne en sammenheng mellom variablene som minimerer variansen.

¹⁰I min oppgave er nordre og søndre sunnhordaland slått sammen, noe de ikke er i denne artikkelen.

Kapittel 5

Økonometriske utfordringer

Utgangspunktet for den metodiske delen av oppgaven er som tidligere nevnt MKM. Denne metoden bygger på å minimere summen av kvadrerte avvik. For at estimatene skal være ”BLUE”¹, best linear unbiased estimators, og dermed at estimatene har variasjon mindre enn alternative estimat, så må en rekke forutsetninger være oppfylt (Jeffrey et al., 2009).

Det må være linearitet i parametrene, mer spesifikt må avhengig variabel være en lineær funksjon av et sett med uavhengige forklaringsvariable. I dette tilfellet $\ln(\frac{pop_i}{kvkm_i})$, vgs_i , $universitet_i$, $klima_i$, $offentlig_i$, $kultur_i$ og et stokastisk restledd, U_i . Dersom jeg ikke har linearitet vil estimatene og determinantene være feil fordi sammenhengene ikke kan fanges opp av en lineær modell.

Jeg må i tillegg ha et tilfeldig utvalg av observasjoner som følger populasjonsmodellen i relasjon (4.8).

Den forventede verdien til det stokastiske restleddet må være lik 0:²

$$E(U_i|X_i) = E(U_i) = 0 \quad (5.1)$$

For alle observasjoner i utvalget. Av denne følger det at snittverdien for residualen er uavhengig av forklaringsvariablene i modellen.

$$cov(U_i, X_{ij}) = 0; j = 1, 2, \dots, 89. \quad (5.2)$$

¹Med forståelsen av beste estimat er det som gir lavest varians sammenlignet med andre potensielle forventningsrette, lineære estimatorer.

²Hvor $X_i = (X_{i1}, X_{i2}, \dots, X_{i89})$

Dersom dette ikke er tilfellet har jeg utelatt viktig informasjon fra predikert modell, som dermed fanges opp av den tilfeldige variasjonen som gjør at stokastikken ikke lenger er preget av tilfeldigheter, men systematikk.

Den betingede variasjonen til restleddet er konstant i alle forklaringsvariable. Restleddsvariansen måler usikkerheten i modellen og denne antagelsen impliserer homoskedastiske restledd. Dette betyr at variasjonen er konstant, som igjen underbygger at modellusikkerheten er lik på tvers av alle observasjoner i modellen. Dersom jeg skulle ha heteroskedastiske restledd, som betyr at modellens usikkerhet vil variere fra observasjon til observasjon, må jeg korrigere for dette. Det behøver ikke nødvendigvis være et problem for estimering, men tradisjonelle inferensmetoder vil ikke være gyldige i slike tilfeller. Selv om det ikke nødvendigvis trenger å være et problem, kan det indikere at jeg har et utelatt variabel problem, som vil gi inkonsistente- og feilestimat. Formelt er fremstillingen gitt ved:

$$VAR(U_i) = E(U_i^2) = \sigma_U^2 = konstant \quad (5.3)$$

Videre må restleddet være uavhengig distribuert og det er ingen korrelasjon mellom observasjoner av uavhengige variable på tvers av økonomiske regioner, slik at:

$$COV(U_i, U_j) = E(U_i, U_j) = 0; i \neq j \quad (5.4)$$

Ingen av de uavhengige variablene i utvalget kan være konstante. Det er ingen eksakte lineære sammenhenger mellom de ulike uavhengige variablene. Forutsetning om at forklaringsvariablene skal være helt uavhengige av hverandre er lite realistiske, det viktigste i slike tilfeller er at de ikke forklarer hverandre perfekt eller for sterkt. Da kan det være at estimatene blir noe utydelige på grunn av høy samvariasjon, slik at variablene potensielt forklarer hverandre. I slike tilfeller er det vanskelig å fange opp den reelle, kausale effekten.

Forventningsverdien til residualen betinget på forklaringsvariablene er lik 0. Forklaringsvariablene i modellen er deterministiske og ikke stokastiske. Med dette er forståelsen at inngående forklaringsvariable ikke korrelerer med residualen, U_i :

$$E(\ln(\frac{pop_i}{kvkm_i}), vgs_i, universitet_i, klima_i, kultur_i, Offentlig_i, klima_i, kultur_i) * E(U_i) \quad (5.5)$$

forenkler videre ved å sette X_i lik alle inkluderte forklaringsvariable, $\ln pop_i$, vgs_i , $universitet_i$, $klima_i$, $of fentlig_i$ og $kultur_i$:

$$E(X_i * U_i) - E(X_i) * E(U_i) \quad (5.6)$$

som gir:

$$X_i * E(U_i) - X_i * E(U_i) = 0 \quad (5.7)$$

Dersom denne ikke er oppfylt kan dette tyde på målefeil, multikollinearitet, utelatt variabel problem, endogenitet eller simultanitet.

Gitt at forutsetningene ovenfor holder vil jeg som nevnt innledningsvis ha forventningsrette og konsistente estimat.

5.1 Brudd på eksogenitetsutfordringen

Foruten sortering har det i faglitteraturen primært vært simultanitet som har vært den utpregede økonometriske utfordringen. I tillegg til dette er målefeil eller utelatt variabel problem potensielle utfordringer jeg bør være observant på ved utforming av en økonometrisk analyse.

Eksogenitetsbetingelsen, $E(U_i | X_1, X_2 \dots X_{89}) = 0$, og brudd på denne kan skyldes flere faktorer som jeg diskuterer nærmere under.

Dersom jeg skulle ha endogene forklaringsvariable vil dette implisere at forklaringsvariablene ikke tilstrekkelig er forklarlige for modellen. Det kan skyldes at jeg har utelatt en viktig variabel som korrelerer med allerede inngåtte variable, at jeg har gjort tilstrekkelige store målefeil slik at det er systematikk i utelatte observasjoner eller simultanitet som betyr at to variable er simultant bestemt.

Presenterer i det kommende utfordringer relatert til dette, og potensielle løsninger.

Utelatt variabel problem

Et utelatt variabel problem oppstår dersom jeg har utelatt en relevant forklaringsvariabel som korrelerer med inngående forklaringsvariable i predikert modell. Ser av forutsetningene for forventningsrette og konsistente estimat at dette vil gi brudd på $E(U_i) = 0$. Dette medfører at forklaringsvariablen i predikert modell vil korrelere med stokastisk restledd som gjør at modellen ikke er sann og estimat ikke lenger er "BLUE."

Skal med utgangspunkt i basismodellen gitt ved relasjon (4.8) vise at dette gir skjevhet og inkonsistens i estimatet.

Antar at den sanne modellen er gitt ved:

$$\ln W_i = \beta_0 + \beta_1 \ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right) + \beta_2 \text{vgs}_i + \beta_3 \text{universitet}_i + \beta_4 \text{klima}_i + \beta_5 \text{offentlig} + \beta_6 \text{kultur} + U_i \quad (5.8)$$

Hvor $E(U_i | \ln(\frac{\text{pop}_i}{\text{kvkm}_i}), \text{vgs}_i, \text{universitet}_i, \text{klima}_i, \text{offentlig}_i, \text{kultur}_i) = 0$.

Jeg estimerer den underspesifiserte modellen:

$$\ln W_i = \beta_0 + \beta_1 \ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right) + \beta_2 \text{vgs}_i + \beta_3 \text{universitet}_i + \beta_4 \text{klima}_i + \beta_5 \text{Offentlig} + V_i. \quad (5.9)$$

Der:

$$V_i = \beta_6 \text{kultur} + \epsilon_i \quad (5.10)$$

Jeg ser av relasjon (5.8) og (5.9) at sistnevnte relasjon vil være underspesifisert, hvor jeg har utelatt en relevant forklaringsvariabel. Denne fanges da opp av det stokastiske restleddet i (5.9), vist ved (5.10).

Forventningsverdien til β_1 -koeffisienten i den underspesifiserte modellen vil dermed være lik:

$$E(\beta_1) = \beta_1 \ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right) + \beta_6 * \delta_1 \text{kultur}_i \quad (5.11)$$

Her er δ_1 samvariasjonen mellom inkludert og ekskludert forklaringsvariabel fordelt på standardavvik til inkludert forklaringsvariabel. I en slik situasjon vil tolkning av estimatet foran målet på urbanitet være lite troverdig da det reflekterer en tilleggseffekt gjennom den utelatte variabelen. Graden av dette bestemmes av størrelsen på $\beta_6 \delta_1$.

I denne oppgaven kan uobserverte karakteristikk potensielt korrelere med universitetsvariabelen inngående i modellen. Dette er det mest åpenbare utelatte variabel problemet i oppgaven.

Målefeil

I de aller fleste økonometriske analyser vil variablene ofte være affektet, i større eller mindre grad, av målefeil (Jeffrey et al., 2009, s.307-313). Med dette er forståelsen et avvik mellom observert og teoretisk verdi av variablene. Det skilles mellom tre ulike scenarie for beheftet målefeil.

Dersom jeg skulle ha et tilfelle av målefeil i venstresidevariabelen, hvor observert verdi for $\ln W_i$ er forskjellig fra sin teoretiske verdi med et avvik tilsvarende ϵ , men denne ikke korrelerer med inngående forklaringsvariable, vil ikke en slik målefeil gi skjevhet i estimatene. Det vil generelt gi høyere både varians og standardavvik som impliserer mer upresise estimat.

Skulle jeg ha et tilfelle av målefeil i forklaringsvariabel, men målefeilen er uavhengig av observert verdi vil det, ekvivalent som i forrige målefeilsproblem, være en beheftet målefeil tilhørende en eller flere av de inkluderte forklaringsvariablene $\ln(\frac{pop_i}{kvkm_i})$, vgs_i , $universitet_i$, $offentlig_i$ eller $klima_i$. Målefeilen vil som i forrige tilfellet øke restleddsvariansen, men ikke gi skjevhet i estimatene.

Gitt at jeg har målefeil i forklaringsvariabel, ukorrelert med uobserverbar variabel, vil restleddet korrelere med inkludert forklaringsvariabel som gir estimatskjevhet. Dette impliserer altså at OLS vil gi skjeve og inkonsistente estimatorer, og disse er dermed ikke lenger "BLUE."

Simultanitet

Potensielle simultanitetsproblemer ved estimering av agglomerasjon var først poengtert av Moomaw(1981). Dersom regioner med høy produktivitet tiltrekker flere individer og dermed øker befolkningsstørrelsen, kan jeg potensielt ha med et simultanitetsproblem å gjøre. Dette kan være et grunnleggende estimeringsproblem da høyere produktivitet i regionen gir flere tilflyttere som øker befolkningsstørrelsen i denne regionen. Dette er et tradisjonelt simultanitetsproblem ved estimering av agglomerasjonsmekanismer, poengtert av blant annet P. Combes, Duranton, Gobillon, Puga og Roux (2007). Typisk vil dette gi en overestimering av effekten i urbane områder.

Simultane ligningssett vil i dette tilfellet være gitt ved:

$$\ln W_i = \beta_0 + \beta_1 \ln\left(\frac{pop_i}{kvkm_i}\right) + \beta_2 vgs_i + \beta_3 universitet_i + \beta_4 offentlig + \beta_5 klima + U_i \quad (5.12)$$

$$\ln\left(\frac{pop_i}{kvkm_i}\right) = \alpha_0 + \alpha_1 \ln W_i + \alpha_2 \rho_1 + V_i \quad (5.13)$$

Her vil estimering av (5.12) være et problem fordi $\ln(\frac{pop_i}{kvkm_i})$ vil korrelere med restleddet U_i , og dermed vil ikke predikert modell være en tilnærming til sann modell. I en slik situasjon

vil både $\ln(\frac{pop_i}{kvkm_i})$ og $\ln W_i$ være endogene variable, mens ρ_1 er en eksogen variabel.

5.2 Håndtering av brudd på eksogenitetsbetingelsen

Presenterer under potensielle løsninger på utfordringene relatert til økonometrien. I analysen instrumenterer jeg populasjon per kvadratkilometer, grunnet et potensielt simultanitetsproblem presentert tidligere.

Instrumentvariabelmetoden

I en simultan ligningsmodell definert tidligere, vil eksogene variable som er utelatt fra modellen jeg ønsker å estimere, men inngår i andre ligninger, være gyldige instrument under enkelte forutsetninger. I følge Jeffrey et al. (2009, s.492) er det primært to kriterier som stilles til et valid instrument, relevans- og eksklusjonskriteriet.

Eksklusjonskriteriet er gitt ved:

$$Cov(\rho_1, U_i) = 0 \quad (5.14)$$

Kriteriet for relevans er gitt ved:

$$Cov(\rho_1, \ln(\frac{pop_i}{kvkm_i})) \neq 0 \quad (5.15)$$

Eksklusjonskriteriet sier at instrumentet jeg benytter ikke korrelerer med det stokastiske restleddet i strukturligningen. Dersom dette er tilfellet er ikke instrumentet eksogent, og derav vil ikke instrumentet gjenspeile systematisk variasjon i variabel jeg ønsker å instrumentere. Dette fordi en del av variasjonen i instrumentvariabel vil være determinert av stokastikk, som representerer tilfeldigheter, og gir ingen systematisk informasjon om verken instrumentet eller endogen variabel.

Relevanskriteriet krever at instrumentet korrelerer med endogen variabel i strukturligning. Dette fordi hensikten med instrumentet er å forklare mest mulig, og dermed identifisere endogen variabel nærmest sin sanne verdi. Følgelig vil graden av korrelasjon mellom disse variablene gjenspeile graden av variasjon i endogen variabel instrumentet identifiserer. Dersom det finnes flere instrument til en variabel, vil variabelen være overidentifisert. Forutsatt at jeg har ett instrument er endogen variabel eksakt identifisert.

IV-metoden løser da problemet ved å inkludere instrumentvariabelen og alle andre forklaringsvariabler, for deretter å gjøre en MKM regresjon på den endogene variabelen (Jeffrey et al., 2009, s.506).³ Dette kalles for førstesteget. Fra problemet knyttet til simultanitet, må jeg først finne redusert form likningen for populasjon per kvadratkilometer. Dette gjør jeg ved å implementere strukturlikningen (5.12) i relasjonen for populasjon per kvadratkilometer, (5.13), og løse for målet på urbanitet:⁴

$$\ln\left(\frac{pop_i}{kvkm_i}\right) = \alpha_0 + \alpha_1(\beta_0 + \beta_1 \ln\left(\frac{pop_i}{kvkm_i}\right) + \beta_2 vgs_i + \beta_3 universitet_i + \beta_4 of\ fentlig + \beta_5 klima + U_i) + \alpha_2 \rho_1 + V_i \quad (5.16)$$

Flytter over $\ln\left(\frac{pop_i}{kvkm_i}\right)$ på venstre side og får:

$$(1 - \alpha_1 \beta_1) \ln\left(\frac{pop_i}{kvkm_i}\right) = \alpha_0 + \alpha_1 \beta_0 + \alpha_1 \beta_2 vgs_i + \alpha_1 \beta_3 universitet_i + \alpha_1 \beta_4 of\ fentlig_i + \alpha_1 \beta_5 klima_i + \alpha_2 \rho_1 + V_i + \alpha_1 U_i \quad (5.17)$$

Slik at:

$$\ln\left(\frac{pop_i}{kvkm_i}\right) = \frac{\alpha_0 + \alpha_1 \beta_0}{(1 - \alpha_1 \beta_1)} + \frac{\alpha_1 \beta_2}{(1 - \alpha_1 \beta_1)} vgs_i + \frac{\alpha_1 \beta_3}{(1 - \alpha_1 \beta_1)} universitet_i + \frac{\alpha_1 \beta_4}{(1 - \alpha_1 \beta_1)} of\ fentlig_i + \frac{\alpha_1 \beta_5}{(1 - \alpha_1 \beta_1)} klima_i + \frac{\alpha_2}{(1 - \alpha_1 \beta_1)} \rho_1 + \frac{V_i + \alpha_1 U_i}{(1 - \alpha_1 \beta_1)} \quad (5.18)$$

Forenkler videre ved å sette:

$$\pi_o = \frac{\alpha_0 + \alpha_1 \beta_0}{(1 - \alpha_1 \beta_1)}, \pi_2 = \frac{\alpha_1 \beta_2}{(1 - \alpha_1 \beta_1)}, \pi_3 = \frac{\alpha_1 \beta_3}{(1 - \alpha_1 \beta_1)}, \pi_4 = \frac{\alpha_1 \beta_4}{(1 - \alpha_1 \beta_1)}, \pi_5 = \frac{\alpha_1 \beta_5}{(1 - \alpha_1 \beta_1)}, \pi_6 = \frac{\alpha_2}{(1 - \alpha_1 \beta_1)}, \epsilon_i = \frac{V_i + \alpha_1 U_i}{(1 - \alpha_1 \beta_1)}$$

Dette gir redusert form likningen for populasjon per kvadratkilometer:

$$\ln\left(\frac{pop_i}{kvkm_i}\right) = \pi_o + \pi_2 vgs_i + \pi_3 universitet_i + \pi_4 of\ fentlig_i + \pi_5 klima_i + \pi_6 \rho_1 + \epsilon_i \quad (5.19)$$

En betingelse for identifisering er at $\pi_6 \neq 0$, som nevnt tidligere kan relevanskriteriet testes for empirisk, hvor tommelfingerregelen er at F-verdien til førstesteget er over 10. En slik empirisk tilnærming er ikke mulig for eksklusjonskriteriet.

Videre antas det i en slik situasjon at $\pi_2 vgs_i + \pi_3 universitet_i + \pi_4 klima_i + \pi_5 of\ fentlig_i$

³IV-metoden er instrumentvariabelmetoden.

⁴Ved å sette uttrykket for $\ln W_i$ direkte inn i (34) og løse for populasjon per kvadratkilometer.

ikke har noe særlig effekt på populasjon per kvadratkilometer når jeg har kontrollert for instrumentvariabelen ρ_1 .

Ved å benytte minste kvadraters metode på relasjon (5.19) får jeg et estimat for $\ln(\frac{pop_i}{kvkm_i})$, som jeg implementerer i strukturligningen.

Benytter estimatet til populasjon per kvadratkilometer fra førsteregresjonen på strukturligningen:

$$\ln W_i = \beta_0 + \beta_1 \ln\left(\frac{\hat{pop}_i}{kvkm_i}\right) + \beta_2 vgs_i + \beta_3 universitet_i + \beta_4 of fentlig + \beta_5 klima_i + U_i \quad (5.20)$$

Bruker nok en gang MKM på (5.20) til å estimere andresteget. Estimatorene jeg får vil kalles for 2SLS-estimer, og er konsistente dersom relevans og eksklusjonskriteriet holder. I praksis vil dermed populasjon per kvadratkilometer i dette tilfellet estimeres to ganger.

Som nevnt tidligere benytter jeg historisk befolkningsdata for instrumentering av populasjonsestimatet. Her har jeg flere alternativer, men benytter primært befolkningsdata fra starten av 1900-tallet. Relatert til IV-metoden ovenfor vil dette være representert av ρ_1 for populasjon per kvadratkilometer. Validitet og gyldighet av instrument kommer jeg tilbake til i førstestegsregresjonen senere.

Kapittel 6

Urbant lønnspremium

I analysen har jeg primært delt inn i tre ulike modellspesifikasjoner. I hver spesifisering utvider jeg modellene i tre steg, hvor hvert steg vil være inkludert nye forklaringsvariable i henhold til relasjonene (4.6), (4.7) og (4.8).

Først presenterer jeg enkel MKM, deretter instrumenterer jeg hovedforklaringsvariabelen, populasjon per kvadratkilometer. I analysens resultater kommenterer jeg kun estimatene foran målet for urbanitet, og forskjeller i denne ved ulike spesifikasjoner og økonometriske metoder. Grunnen til det er at det er denne som er av interesse for mine problemstillinger.

For å fastslå en kausal sammenheng mellom avhengig- og en kontrollvariabel er begrepet *ceteris paribus* sentralt. I følge Jeffrey et al. (2009, s.12-14) betyr dette hvordan en endring i relevant forklaringsvariabel affekterer avhengig variabel, alt annet likt. Gjennomgående i analysen vil tolkningen av hovedforklaringsvariabelen være i fokus. I alle spesifikasjonene vil denne *ceteris paribus* betingelsen være gjeldende, slik at alle andre variabler holdes konstant ved tolkningen.

Som nevnt tidligere finner P.-P. Combes et al. (2010) at elastisiteten mellom lokal produktivitet og sysselsettingstetthet for franske regioner ligger mellom 0,04 og 0,10.¹

Carlsen et al. (2013) finner at den rene elastisiteten av gjennomsnittlig lønn mot populasjon ligger på 0,076, videre finner de ved korrigering av observerte arbeidseffekter et estimat på 0,048, ikke ulikt Combes et al.(2010). Når de korrigerer for endogenitetsproblemet med historisk befolkningsdata får de et estimat på 0,054, og til slutt faller koeffisienten til 0,033 ved kontroll for både endogenitet og sortering. Når de bruker populasjonstetthet får de et rent elasisitetsestimert på 0,055 som faller ved korrigering av sortering til 0,04 og

¹En dobling av sysselsettingstetthet gir økt produktivitet på mellom 4-10 prosent i regionen.

til slutt til 0,034 når de korrigerer for både endogenitet og sortering.²

Alle regresjoner og tester relatert til analysen er gjort i Stata.

6.1 MKM - produktivitet og lønn i økonomiske regioner

Tabell 6.1: MKM - hovedanalyse

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|-----------------------|-----------------------|-------------------------|
| | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn |
| Lnpopkvkm | 0.0428*** (0.00405) | 0.0358*** (0.00396) | 0.0332*** (0.00513) | | | | | | |
| vgs | | 0.851*** (0.146) | 0.671*** (0.160) | | 0.932*** (0.187) | 0.560*** (0.190) | | 0.760*** (0.192) | 0.449** (0.190) |
| universitet | | 0.816*** (0.125) | 0.769*** (0.125) | | 0.915*** (0.163) | 0.884*** (0.153) | | 1.001*** (0.171) | 0.960*** (0.153) |
| klima | | | 0.00309** (0.00145) | | | 0.00435** (0.00167) | | | 0.00463*** (0.00171) |
| Offentlig | | | -0.0598 (0.102) | | | -0.332*** (0.104) | | | -0.377*** (0.105) |
| kultur | | | 0.00956** (0.00445) | | | 0.00722 (0.00517) | | | 0.00298 (0.00536) |
| Lnpop | | | | 0.0507*** (0.00702) | 0.0391*** (0.00819) | 0.0250*** (0.00841) | | | |
| storby | | | | | | | 0.158*** (0.0265) | 0.0912*** (0.0290) | 0.0492* (0.0277) |
| byregion | | | | | | | 0.101*** (0.0284) | 0.0650** (0.0251) | 0.0441* (0.0227) |
| _cons | 5.480*** (0.0117) | 4.938*** (0.0818) | 5.023*** (0.103) | 5.063*** (0.0728) | 4.566*** (0.134) | 4.975*** (0.158) | 5.568*** (0.00769) | 5.016*** (0.107) | 5.291*** (0.113) |
| N | 89 | 89 | 89 | 89 | 89 | 89 | 89 | 89 | 89 |
| adj. R ² | 0.557 | 0.721 | 0.737 | 0.368 | 0.568 | 0.642 | 0.329 | 0.523 | 0.624 |

Standard errors in parentheses

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

Fra hovedanalysen i tabell 6.1 finner jeg i resultatene fra modell (1) at den rene elastisiteten av populasjon per kvadratkilometer er gitt ved:

$$\frac{d \ln W_i}{d \ln \left(\frac{\text{pop}_i}{\text{kvkm}_i} \right)} = 0,0428. \quad (6.1)$$

Dette impliserer at en dobling av populasjonen per kvadratkilometer gir 4,28 prosent høyere gjennomsnittlig timelønn, og dermed produktivitet. P.-P. Combes et al. (2010)

²Dobling av populasjonstetthet gir 5,5 prosent høyere produktivitet i regionen. Korrigering av sortering gir 4 prosent høyere produktivitet i regionen og 3,4 prosent når de korrigerer for både kvantitativ og kvalitativ endogenitet.

finner at den rene elastisiteten av sysselsettingstetthet til lønn typisk ligger på 0,05 for franske regioner når de bruker paneldata. Sveikauskas (1975) finner, når han regresserer produksjon per arbeider mot populasjon, at en dobling av populasjon gir 6 prosent høyere produksjon per arbeider. Moonaw(1983) kritiserte bruken av produksjon per arbeider da høyere akkumulasjon av kapitalmengde i byer etter all sannsynlighet vil medføre en positiv skjevhet i estimatet. Rosenthal og Strange (2004) finner at en dobling av populasjon i byregionen bør gi økt produktivitet på mellom 3-8 prosent.

Det er lite sannsynlig at resultatet i modell (1) gjenspeiler de sanne agglomerasjonsfordelene i urbane områder.

Utvider derfor videre til modellspesifikasjon (2), hvor jeg både har inkludert andelen med fullført videregående utdannelse samt høyere utdanning i de respektive økonomiske regionene, dette gir et estimat på:

$$\frac{dlnW_i}{dln(\frac{pop_i}{kvkm_i})} = 0,0358 \quad (6.2)$$

Som impliserer at en dobling av populasjonen per kvadratkilometer i regionen vil gi 3,58 prosent høyere gjennomsnittlig timelønn. Dette er et fall fra modellspesifikasjon (1) på 0,7 prosentpoeng i gjennomsnittlig timelønn ved dobling av populasjon per kvadratkilometer i regionen. P.-P. Combes et al. (2010) finner, når de kontrollerer for individuelle karakteristikk, et estimat på 0,033, ikke ulikt egne resultater. Sammenlignet med Carlsen et al. (2013) som finner en endring på rundt 1,5 prosentpoeng mellom ren elastisitet og korrigering for sortering, er estimatendringen i mitt tilfelle relativt lite. Carlsen et al. benytter individspesifikk data, slik at deres estimat høyst antagelig gjenspeiler de reelle endringene i større grad enn hva mitt estimat gjør. Dette er en indikasjon på en svakhet ved egen analyse.

I modellspesifikasjon (3) korrigerer jeg i tillegg for andre regionale variable som driftsutgifter til kultur av totalt driftsbudsjett, gjennomsnittlig januarstemperatur fra 1994-2002 og andelen arbeidere i offentlig sektor. I denne modellen blir populasjon per kvadratkilometer:

$$\frac{dlnW_i}{dln(\frac{pop_i}{kvkm_i})} = 0,0332 \quad (6.3)$$

Den rene elastisiteten er på 0,0332 hvorav en dobling av populasjonen per kvadratkilometer gir en økning i gjennomsnittlig timelønn på 3,32 prosent. Dette er en nedgang på 0,26 prosentpoeng fra (2). Carlsen et al. (2013) finner ved bruk av til sammenlikning, ved bruk av populasjonstetthet, et estimat på 0,04 når de korrigerer for sortering av arbeidere.

Alle resultatene for hovedforklaringsvariabelen er signifikante på 1 prosent nivå, dette impliserer at estimatene i 2010 var reelle, med en sannsynlighet på 99 prosent.

Forklaringskraften til de ulike modellene representert ved den justerte R^2 .³ Denne sier noe om hvor nære mine observasjoner er den tilpassede regresjonslinjen. I multiple regresjonsmodeller kan dette være et godt verktøy fordi den straffer bruken av unødig inkluderte kontrollvariable, slik at en enkelt kan eliminere ulike modellspekifikasjoner utifra verdien av denne. Ser at denne øker med modellene, noe som kan bety at spesifikasjon (3) er den modellen som er mest nærliggende sin teoretiske, sanne modell.

Dummy

I tillegg til ulike mål på populasjon har jeg inkludert dummyer for storbyer samt byregioner. Kategoriseringen av disse ble nevnt innledningsvis i kapitlet for databeskrivelse. Resultatene her er presentert i tabell 6.1 modell (7),(8) og (9).

I modell (7) ser jeg at ved kun å inkludere dummyer er gjennomsnittlig timelønn i storbyregioner 15,8 prosent høyere, og 10,1 prosent høyere i byregion enn resten av Norge, alt annet likt.

Ved å korrigere for observerte karakteristikk faller estimatet foran storbyregioner med hele 6,68 prosentpoeng, fra 15,8 til 9,12 prosent. Estimaten foran byregion faller med 3,6 prosentpoeng fra 10,1 til 6,5 prosent. Dette representert ved overgangen fra modell (7) til (8). Sortering av individ basert på utdanningsnivå var altså nesten dobbelt så stor i storbyregioner kontra byregioner i 2010. Disse resultatene gir en indikasjon på hvor stort omfang sorteringsproblemet medfører til større og mindre byer i Norge. Det er åpenbart at dette er omfattende i byregioner, noe som også underbygger de sterke sentraliseringsprosessene de siste 30 årene. Begge dummyene i modell (8) signifikante på ett prosent nivå.

Som et siste alternativ inkluderer jeg i (9) regionale karakteristikk hvor estimatene faller til henholdsvis 0,0492 for storby og 0,0411 for byregion. Her er estimatene signifikante på et 10 prosent nivå.

Til sammenligning studerer Carlsen et al. (2013) effekten av arbeidserfaring og fast ansettelse og finner ved statiske agglomerasjonsmekanismer en ren elastisitet i byregion⁴ på 14,1 prosent, ikke ulikt Overman (2014). Kontroll for observert heterogenitet faller esimatet til 7,5 prosent. Til slutt benytter de faste effekter hvor estimatet faller til 5,1 prosent, korrigert for både observert og uobserverte karakteristikk.

³Justert R^2 sier noe om forklaringskraften til estimert modell, dette er en statistisk tilnærming til hvor nære dataen er den tilpassede regresjonslinjen.

⁴Definert ved områder med over 150k innbyggere.

6.2 IV - produktivitet og lønn i økonomiske regioner

I tabell 6.2 presenteres resultatene hvor jeg har instrumentert populasjon per kvadratkilometer med populasjon per kvadratkilometer i 1910 for de respektive spesifikasjonene (1)-(3).

Tabell 6.2: IV - hovedanalyse.

| | (1) | (2) | (3) |
|----------------------------|------------------------|------------------------|------------------------|
| | Lntimelonn | Lntimelonn | Lntimelonn |
| Lnpopkvkm | 0.0404*** (0.00420) | 0.0342*** (0.00405) | 0.0304*** (0.00531) |
| universitet | | 0.839*** (0.123) | 0.797*** (0.122) |
| vgs | | 0.844*** (0.143) | 0.645*** (0.155) |
| klima | | | 0.00325** (0.00140) |
| Offentlig | | | -0.0920 (0.101) |
| kultur | | | 0.00919** (0.00428) |
| _cons | 5.486*** (0.0120) | 4.940*** (0.0800) | 5.048*** (0.101) |
| <i>N</i> | 89 | 89 | 89 |
| adj. <i>R</i> ² | 0.555 | 0.720 | 0.737 |

Standard errors in parentheses

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

Fra (1) ser jeg at den rene elastisiteten av populasjon per kvadratkilometer vil reduseres til:

$$\frac{d\ln W_i}{d\ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right)} = 0,0404 \quad (6.4)$$

Dette betyr at en dobling av populasjon per kvadratkilometer i regionen gir høyere gjennomsnittlig timelønn på 4,04 prosent, som tilsvarer en nedgang på 0,24 prosentpoeng fra MKM-tilnærmingen. Dette tilsier at det kvantitative endogenitetsproblemet isolert sett er tilstedeværende, men tilsynelatende mindre i omfang enn hva sorteringsproblemet er. Videre, ved å inkludere utdanningsvariabler, ser jeg fra (2) at:

$$\frac{d\ln W_i}{d\ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right)} = 0,0342 \quad (6.5)$$

En dobling av populasjonen gir 3,42 prosent høyere gjennomsnittlig timelønn. Dette er et fall på 0,16 prosentpoeng fra tilsvarende spesifisering med MKM som metode. Alle variable inkludert i spesifisering (2) er signifikante på ett prosent nivå.

Ved å kontrollere for andre regionale variable som klima, andel sysselsatte i offentlig sektor og netto driftsutgifter til kultur, reduseres estimatet til:

$$\frac{d\ln W_i}{d\ln\left(\frac{\text{pop}_i}{\text{kvkm}_i}\right)} = 0,0304 \quad (6.6)$$

Som betyr at en dobling av populasjonen per kvadratkilometer gir økt snittlig timelønn på 3,04 prosent i regionen. Dette er en forskjell på 0,28 prosentpoeng ved dobling av populasjon per kvadratkilometer når jeg bruker ekvivalent modellspesifisering, men med MKM som metode. Her er populasjonskoeffisienten og de respektive utdanningsvariablene signifikante på ett prosent nivå.

P.-P. Combes et al. (2010) finner, når de korrigerer for kvalitativ og kvantitativ endogenitet, en populasjonskoeffisient på 0,027. En mer nærliggende sammenligning vil være Carlsen et al. (2013) som finner et estimat foran populasjon per kvadratkilometer på 0,034 ved korrigering av både det kvantitative og kvalitative endogenitetsproblemet. Sammenlignet med egne resultater er ikke forskjellen altfor stor. Dette viser at en god andel av målte agglomerasjonseffekter må tilskrives uobserverte karakteristikk av populasjonen i aktuell region. Av de tre spesifiseringene er det modell (3) som har den høyeste justerte forklaringskraften. Dette kan tyde på at denne modellen er estimatene nærest sin sanne, teoretiske verdi.

Førstesteget

Hvorvidt instrumentene oppfyller kravene for relevans kan vi som nevnt tidligere empirisk teste. I denne seksjonen presenterer jeg førstesteget for validitet av instrumentet benyttet. Tabell 6.3 viser instrumentet til den rene elastisiteten, tabell 6.4 er ekvivalent med spesifikkasjon (2), mens 6.5 er for instrumentet i modell (3).

Tabell 6.3: Førstesteget for spesifikkasjon (1)

| |
|------------------------|
| Number of obs = 89 |
| F(1,87)=904,59 |
| Prob >F = 0.0000 |
| R-squared = 0.9123 |
| Adj R-squared = 0.9113 |
| Root MSE = 0.4269 |

| Lnpopkvkm | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|---|----------|-----------|-------|-------|----------------------|
| $\ln\left(\frac{pop_i}{kvkm_i}\right)_{1910}$ | 1.149643 | .038224 | 30.08 | 0.000 | 1.073668 1.225617 |
| _cons | .2020287 | .0890202 | 2.27 | 0.026 | .0250913 .378966 |

Tommelfingerregelen for et valid instrument er at F-verdien til instrumentet er over 10. Ser av tabellen ovenfor at befolkning per kvadratkilometer har en F-verdi på 904,59 som impliserer at relevanskriteriet er oppfylt. Et annet poeng, som indikerer at instrumentet er godt, er at koeffisienten foran historisk populasjonstetthet er positiv. Dette gir mening, og kan tyde på at regresjonen er riktig.

Tabell 6.4: Førstesteget for spesifikkasjon (2).

| |
|------------------------|
| Number of obs = 89 |
| F(3,85) = 482.26 |
| Prob >F = 0.0000 |
| R-squared = 0.9445 |
| Adj R-squared = 0.9426 |
| Root MSE = 0.3435 |

| Lnpopkvkm | Coef. Std. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------------|------------|-----------|-------|-------|-----------------------|
| vgs | -1.844373 | 1.156729 | -1.59 | 0.115 | (-4.14426, .4555129) |
| universitet | 4.692169 | .9364014 | 5.01 | 0.000 | (2.830352, 6.553986) |
| ln1910 | 1.043599 | .0342689 | 30.45 | 0.000 | .9754635 1.111735 |
| _cons | .1965855 | .6489525 | 0.30 | 0.763 | (-1.093706, 1.486877) |

Her er populasjon per kvadratkilometer estimert mot de andre eksogene variablene,

andel fullført videregående skole og universitet, samt instrumentet jeg benytter for identifisering. Ekvivalent som ved instrumentering av (1), kan det tilsynelatende virke som om relevanskriteriet er gyldig her og, med en F-verdi på 482,26.

Tabell 6.5: Førstesteget for spesifikasjon (3).

| |
|-----------------------|
| Number of obs=89 |
| F(6, 82) =255.43 |
| Prob >F =0.0000 |
| R-squared =0.9492 |
| Adj R-squared =0.9455 |
| Root MSE =0.3346 |

| Lnpopkvkm | Coef. | Std. Err. | t | P> t | [95% Conf.Interval] |
|-------------|-----------|-----------|-------|-------|------------------------|
| vgs | -2.880663 | 1.251429 | -2.30 | 0.024 | (-5.370155, -.3911718) |
| universitet | 4.499346 | .9420671 | 4.78 | 0.000 | (2.625274, 6.373417) |
| Offentlig | -2.151611 | .7826026 | -2.75 | 0.007 | (-3.708457, -.5947656) |
| klima | -.0014449 | .0116087 | -0.12 | 0.901 | (-.0245383, .0216485) |
| kultur | .0079272 | .0355781 | 0.22 | 0.824 | (-.062849, .0787035) |
| ln1910 | .9812314 | .0432455 | 22.69 | 0.000 | (.8952024, 1.06726) |
| _cons | 1.510141 | .8089283 | 1.87 | 0.065 | (-.0990754, 3.119357) |

Førstesteget etter spesifikasjon (3) i instrumentvariabelmetoden. Inkludert alle eksogene variable og instrumentet for å sjekke om dette er relevant. Fra tidligere førstesteg faller F-verdien, men relevanskriteriet holder fortsatt med en F-verdi på 255.43.

Oppsummering

Fra hovedanalysen og MKM får jeg ved de ulike estimeringsstrategiene koeffisienter fra 0,0428 i ren populasjonselastisitet, til 0,0358 ved inklusive ulike former for utdanning, og til slutt utvides spesifikasjonen hvor estimatet faller til 0,0332. Til sammenligning får jeg ved IV-metoden 0,0404 i ren populasjonselastisitet til 0,0342 ved inklusive utdanning før jeg til slutt etter en siste utvidelse får 0,0304 ved å instrumentere for populasjonsvariabelen.

Endogenitets- og sorteringsproblemet er tilstedeværende, og medfører en positiv skjevhet i estimatene. Dette ser jeg av resultatene ved bruk av forskjellig metode, og ved de ulike utvidelsene av modellene. Det kan virke som sorteringsproblemet gir en større skjevhet enn hva simultanitetsproblemet gjør, det er nettopp endringene i estimatene og sammenligning ved utvidelser og metode brukt som underbygger dette faktum.

Populasjon per kvadratkilometer er signifikant til 0,01 prosent nivå på nærmest alle mo-

dellspesifikasjoner i hovedanalysen.

Noen konkrete svakheter ved analysen er at jeg ikke korrigerer for uobserverte karakteristikk. Dette vil etter all sannsynlighet medføre en grad av skjevhet i resultatene. En direkte indikator av dette er sammenlikning med Carlsen et al. (2016, 2013) som også studerer norsk data, men får noe annerledes resultater enn undertegnede. Et annet aspekt er P.-P. Combes et al. (2008) sin bevisstgjøring av historisk befolkningsdata som potensielt dårlige instrument, og at dette kan korrelere med produktivitet i senere år. Det vil isåfall kunne gi noe svake resultater ved instrumentvariabelmetoden. På tross av disse svakhetene, kan resultatene indikere at agglomerasjonsmekanismer er et faktum, og disse medfører større gevinster i tett befolkede områder.

6.3 Urbant lønnspremium i ulike sektorer

Hensikten her er å se hvorvidt det urbane lønnspremiumet varierer mellom ulike sektorer og næring i økonomien. Mer spesifikt tar jeg her for meg privat og offentlig sektor, samt service- og industrinæringene.

Privat sektor er grovt inndelt i service og industri, slik at jeg først presenterer forskjellen mellom offentlig og privat, for deretter å gå inn på privat sektor og se på forskjellen mellom service og industri.

Presenterer i resultatene 4 modellspesifikasjoner, hvor hver og en representerer en av sektorene. Oppbyggingen av resultatene er ellers ekvivalente som i hovedanalysen. Først presenterer jeg MKM, deretter IV hvor jeg initielt kun instrumenterer populasjon for deretter å utvide til instrumentering av populasjon per kvadratkilometer.

6.3.1 MKM

En betingelse for å se på variasjonen mellom sektorer er at jeg må benytte samme datagrunnlag for alle regresjonene. Her bruker jeg tilnærmet det samme datagrunnlag som i hovedanalysen, men fjerner andelen arbeidere i offentlig sektor. Lønn forhandles i alle deler av norsk arbeids- og næringsliv. Måten dette gjøres på varierer fra sektor til sektor og fra sted til sted. Derfor er det også naturlig å tro at lønnspremiumet varierer. Resultatene presenteres i tabell 6.6, modell (1)-(4).

Tabell 6.6: MKM - Sektor og næring

| | (1) | (2) | (3) | (4) |
|----------------------------|------------------------|-------------------------|-------------------------|------------------------|
| | LNOffentliglønn | LNPrivatlønn | LNIndustrilønn | LNServicelønn |
| Lnpopkvkm | 0.0204*** (0.00294) | 0.0426*** (0.00532) | 0.0348*** (0.00678) | 0.0513*** (0.00463) |
| universitet | 0.616*** (0.0882) | 0.805*** (0.160) | 1.101*** (0.203) | 0.545*** (0.139) |
| vgs | 0.496*** (0.107) | 0.793*** (0.193) | 1.000*** (0.246) | 0.405** (0.168) |
| klima | -0.000939 (0.00102) | 0.00513*** (0.00185) | 0.00819*** (0.00235) | 0.00162 (0.00161) |
| KULTUR | -0.000224 (0.00314) | 0.0140** (0.00567) | 0.0229*** (0.00723) | 0.00235 (0.00494) |
| _cons | 5.159*** (0.0580) | 4.908*** (0.105) | 4.774*** (0.134) | 5.120*** (0.0914) |
| <i>N</i> | 89 | 89 | 89 | 89 |
| adj. <i>R</i> ² | 0.694 | 0.709 | 0.631 | 0.772 |

Standard errors in parentheses

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

Offentlig

Offentlig sektor består av følgende tre store forhandlingsområder; staten, kommunal sektorområdet og Oslo kommune. Avtalene innenfor disse områdene er bygd opp forskjellig. I staten har man fortsatt et regulativlønnssystem med innplassering av den enkelte stillingstype i lønnsramme. I kommunal sektor er lønnssystemet bygd opp som et minstelønnssystem, men med garanterte minstelønnsatser for ulike hovedgrupper ansatte. Både i staten og kommunene settes det vanligvis av midler til lokale forhandlinger. I motsetning til privat sektor, skjer slike forhandlinger innenfor en økonomisk ramme som avtales i de sentrale forhandlingene. Størrelsen på denne ramma varierer fra år til år. I de lokale forhandlingene er det ofte de lokale tillitsvalgte og den enkelte kommunale virksomheten som forhandler om lønn.⁵ Den rene elastisiteten er i dette tilfellet:

$$\frac{d \ln W_{\text{Offentlig}}}{d \ln \left(\frac{\text{pop}_i}{\text{kvkm}_i} \right)} = 0,0204 \quad (6.7)$$

⁵<http://www.arbeidslivet.no/Lonn/Tariffavtaler/Sentrale-og-lokale-lonnsforhandlinger/>

Som impliserer at en dobling av populasjon per kvadratkilometer gir 2,04 prosent høyere gjennomsnittlig timelønn i offentlig sektor.

Privat

Mange av bedriftene i privat sektor er organisert i en arbeidsgiverforening. Ofte vil virksomheten være bundet av en tariffavtale, som arbeidsgiverforeningen har forhandlet frem med aktuelle motparter på arbeidstakersiden. Den eller disse avtalene vil i utgangspunktet gjelde for bestemte grupper av ansatte, det vil si de som har medlemskap i arbeidstakerorganisasjonen som det er inngått avtale med.

Elastisiteten er i dette tilfellet:

$$\frac{d \ln W_{privat}}{d \ln \left(\frac{pop_i}{kvkm_i} \right)} = 0,0426 \quad (6.8)$$

Den rene elastisiteten er i privat sektor 0,0222 høyere enn i offentlig sektor. Dette betyr at en dobling av populasjonen gir 4,26 prosent høyere gjennomsnittlig timelønn i privat sektor, som er 2,22 prosentpoeng høyere enn i offentlig.

Det kan tenkes at lokaliseringsaspektet er viktigere for private bedrifter enn det vil være for det offentlige. Begge sektorer vil naturlig nok nyte godt av lokalisering i et befolkningstett område, men private bedrifter er i større grad utsatt for konkurranse. Dette gjør at privat sektor får en annen dimensjon i gevinster av sentralisering. I et konkurranseperspektiv vil en kunne observere og monitorere avgjørelser til konkurrenter, og ta beslutninger på grunnlag av dette. I tillegg til dette ha en nærere tilgang til et større marked, som i teorien også avler en raskere responstid på akkurat disse beslutningene. På denne måten vil aktørene kunne lære av andre og bruke dette som basis for egen utvikling. Det kan også tenkes at privat sektor i større grad nyter sentraliseringsgevinster på grunn av sortering av arbeidere og er flinkere til å tilpasse lønn i henhold til slike arbeidere.

Industri

Industrinæringen er bestående av verksted-, tekstil-, prosess-, næringsmiddel- og kjemisk industri. Noen store norske industribedrifter er Hydro, Aker Kværner, Elkem, Orkla, Veidekke og Kongsberg Gruppen. I industrien er det som regel verkstedindustrien som er frontfaget innen lønnsforhandlinger. Dette impliserer at organisasjoner innen denne næringen først går til forhandlinger, og deretter følger de andre næringene etter med ramene lagt av de initielle forhandlingene. Dette har primært med at industrien er en del

av konkurranseutsatt sektor. Derav vil sterk lønnsvekst medføre svekket konkurransekraft hos de ulike industriaktørene, som i verste fall kan føre til nedleggelse av bedrifter eller outsourcing til lavkostnadsland. En annen aktuell faktor er det faktum at industrien er preget av en desentralisert struktur. I en del distriktsregioner er industrien en viktig del av næringslivet, og ofte en hjørnesteinsbedrift som legger grunnlaget for lokalt næringsliv. I et fortsatt relativt desentralisert land som Norge, men hvor den historiske tendensen de siste 30 årene har vært sterk urbanisering, spiller disse hjørnesteinsbedriftene naturlig nok en viktig rolle i opprettholdelsen av bosetting i regionale områder, og derav en brikke i politikken. Dette vil i sum trekke i retning av at en skulle tro det urbane lønnspremiu- met er relativt lavt i denne sektoren. Fra modell (3) tabell 6.6 finner jeg resultatene for industrinæringen. Den rene elastisiteten er gitt ved:

$$\frac{d \ln W_{\text{industri}}}{d \ln \left(\frac{\text{pop}_i}{\text{kvkm}_i} \right)} = 0,0348 \quad (6.9)$$

Dette impliserer at en dobling av populasjonen per kvadratkilometer gir høyere gjennomsnittlig timelønn i industrien på 3,48 prosent høyere lønnspremium i denne næringen.

Service

Service næringen er tjenesteytende virksomhet, denne næringen omfatter mange ulike bransjer. Varehandel, transport, maritim næring, reiseliv, IKT, finans, privat tjenesteyting er blant næringene servicesektor består av. Alle disse næringene er besittende av en, i større eller mindre grad, desentralisert struktur. For eksempel vil bank, varehandel, reiseliv alle være lokalisert i distriktsregioner. Samtidig er det naturlig å tro at disse også besitter fordeler av sentralisering. Tjenesteyting krever interaksjoner mellom mennesker. Det er naturlig å tro at dette tilrettelegges bedre, og at dette er av et større i omfang i tett befolkede områder. Derav vil det være naturlig at gevinsten er større i disse områdene. Den rene elastisiteten er i dette tilfellet:

$$\frac{dW_{\text{Service}}}{d \ln \left(\frac{\text{pop}_i}{\text{kvkm}_i} \right)} = 0,0513 \quad (6.10)$$

Som impliserer at en dobling av populasjonen per kvadratkilometer vil øke gjennomsnittlig timelønn i servicesektoren med 5,13 prosent. Noe som er 1,65 prosentpoeng høyere enn i industriell sektor. Med utgangspunktet i resonnementene ovenfor er ikke forskjellene i disse resultatene nevneverdig overraskende.

Alle resultatene er signifikante på 1 prosent nivå. Vi kan altså si at estimatene i 2010 var reelle, men med en usikkerhet på 1 prosent.

6.3.2 IV

Instrumentet benyttet i dette tilfellet er ekvivalent som i hovedanalysen populasjon per kvadratkilometer i 1910 i de respektive økonomiske regionene.

Førstestegsregresjonen blir ikke presentert i denne delen, da verdiene vil være ekvivalente som i hovedanalysens førsteregresjon.

Tabell 6.7: IV - Sektor og næring

| | (1) | (2) | (3) | (4) |
|----------------------------|------------------------|-------------------------|-------------------------|------------------------|
| | LNOffentliglønn | LNPrivatlønn | LNIndustrilønn | LNServicelønn |
| Lnpopkvkm | 0.0219*** (0.00300) | 0.0395*** (0.00542) | 0.0320*** (0.00690) | 0.0477*** (0.00473) |
| universitet | 0.593*** (0.0864) | 0.851*** (0.156) | 1.140*** (0.199) | 0.596*** (0.136) |
| vgs | 0.504*** (0.103) | 0.776*** (0.187) | 0.986*** (0.238) | 0.387** (0.163) |
| klima | -0.00110 (0.000992) | 0.00546*** (0.00179) | 0.00848*** (0.00229) | 0.00200 (0.00157) |
| KULTUR | 0.000106 (0.00304) | 0.0133** (0.00550) | 0.0223*** (0.00700) | 0.00159 (0.00480) |
| _cons | 5.155*** (0.0562) | 4.917*** (0.102) | 4.782*** (0.129) | 5.129*** (0.0887) |
| <i>N</i> | 89 | 89 | 89 | 89 |
| adj. <i>R</i> ² | 0.693 | 0.707 | 0.630 | 0.770 |

Standard errors in parentheses

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

Offentlig

Når jeg kun instrumenterer populasjon per kvadratkilometer fra tabell 6.7 modell (1), blir elastisiteten på:

$$\frac{d \ln W_{\text{offentlig}}}{d \ln \left(\frac{\text{pop}_i}{\text{kvkm}_i} \right)} = 0,0219 \quad (6.11)$$

Dette tilsier at en dobling av populasjon gir 2,19 prosent høyere gjennomsnittlig offentlig timelønn. Fra MKM-estimatet på 2,04 prosent gir dette en forskjell på 0,15 prosentpoeng gjennomsnittlig timelønn ved dobling av populasjon. Dette betyr, overraskende nok, at det kvantitative endogenitetsproblemet tilsynelatende trekker estimatet ned.

Privat

Privat sektor har et urbant lønnspremium, når instrumentert for populasjon per kvadratkilometer, på:

$$\frac{d\ln W_{privat}}{d\ln\left(\frac{pop_i}{kvkm_i}\right)} = 0,0395 \quad (6.12)$$

Som tilsvarer at en dobling i populasjon gir 3,95 prosent høyere gjennomsnittlig lønn i privat sektor. Dette er en nedgang i estimatet på 0,31 prosentpoeng fra MKM som metode. Simultanitetsproblemet medfører en positiv skjevhet i estimatene for privat sektor.

Industri

Fra tabell 6.7 modell (3) ser jeg at når jeg kun instrumenterer for populasjon vil estimatet ligge på:

$$\frac{d\ln W_{industri}}{d\ln\left(\frac{pop_i}{kvkm_i}\right)} = 0,0320 \quad (6.13)$$

Dette er en forskjell på 0,0028 i nedgang av ren elasticitet fra enkel OLS-tilnærming. Denne forskjellen blir ved en dobling av populasjonen per kvadratkilometer på 0,28 prosentpoeng høyere gjennomsnittlig timelønn i industrisektoren.

Service

Fra modell (4), når jeg kun instrumenterer populasjon ligger elasticiteten på:

$$\frac{d\ln W_{service}}{d\ln\left(\frac{pop_i}{kvkm_i}\right)} = 0,0477 \quad (6.14)$$

Hvor da en dobling av populasjonen per kvadratkilometer gir en økning i gjennomsnittlig lønnspremium i servicesektor på 4,77 prosent. Dette er en nedgang på 0,36 prosentpoeng fra MKM. I henhold til dette resultatet, kan det virke som den simultane effekten er størst i servicenæringen.

Oppsummering

I den andre analysen har jeg estimert det urbane premiumet for ulike sektorer i økonomien. Her har jeg inkludert offentlig, privat, industri og service.

Her ligger estimatene for offentlig på henholdsvis 0,0204 ved MKM og 0,0219 ved instrumentering av mål på urbanitet. For privat sektor er estimatet noe høyere, på 0,0426 ved MKM og 0,0395 ved instrumentering. Estimatene er relativt kompakte.

Industri har en ren elastisitet på 0,0348 ved MKM, når jeg instrumenterer får jeg 0,0320. Service-estimatet ligger noe høyere, mer spesifikt på 0,0513 ved MKM og på 0,0477 når jeg instrumenterer populasjon per kvadratkilometer.

Siden modellen ikke bygges opp i samme grad som i hovedanalysen, er det vanskelig å si noe om omfanget av sortering i de ulike sektorene. Simultanitetsproblemet virker tilsynelatende å være størst i servicenæringen med en endring på 0,36 prosentpoeng fra MKM til IV ved dobling av populasjon per kvadratkilometer. Overraskende nok medfører korreksjon av simultanitet en positiv effekt på offentlig sektor. Ekvivalent med hovedanalysen, vil også denne delen inneha en rekke svakheter. Det er likevel naturlig å tro at privat sektor har større fordeler av bevisst geografisk lokalisering mot urbane områder, og videre at service besitter de samme fordelene målt opp mot industrinæringen.

6.4 Robusthetstesting

I denne delen ønsker jeg å se om estimatene foran populasjon er konsistente dersom jeg endrer til absoluttverdien av populasjon samt benytter andre instrumenter. Alle resultatene er presentert i appendiks.

Hovedanalyse

I oppgaven har jeg inkludert en rekke modellspesifikasjoner, og ulike varianter av instrumentvariabelmetoden.

Ved bruk av alternativ forklaringsvariabel, absolutt populasjon, og MKM som metode får jeg estimat presentert i tabell 6.1 spesifikasjon (4)-(6) varierende fra 0,0507 i ren elastisitet, til 0,0391 ved inklusive utdanning, og til slutt 0,0250 ved korrigerende for alle områdekaraktistikker. Dette er en forskjell på henholdsvis 0,79, 0,33 og 0,59 prosentpoeng for de ulike utvidelsene sammenliknet med bruken av populasjon per kvadratkilometer. For alle spesifikasjonene tilskrives agglomerasjonsmekanismene en større effekt ved absoluttverdien av populasjon kontra populasjon per kvadratkilometer. Estimaten er i alle spesifikasjoner signifikante på ett prosent nivå. Forklaringskraften, gitt ved justert R^2 , er lavere ved absoluttverdien av populasjon enn ved bruken av populasjonstetthet.

Videre har jeg benyttet ulike instrument for populasjonsparameteren. I tabell 7.1 i appendiks har jeg instrumentert populasjon med ulike historiske befolkningsdata. Her er følgende instrument brukt: modell (1) befolkning 1845, (2) stemmeberettigede i 1898, (3) bebodde hus i 1900, (4) befolkning i 1769, (5) populasjon 1835. Spesifikasjonene her korresponderer med (2) og (3) fra hovedanalysen. Fra (1)-(5) er modelloppsettet fra (3) i tabell 6.2 brukt, mens (2) er benyttet for (6)-(10). Resultatene i modellene (1)-(5) i tabell 7.1 er konsistente. Alle resultatene er utenom (2)⁶ signifikante på ett prosent. Koeffisienten foran populasjon per kvadratkilometer varierer fra 0,0334 med stemmeberettigede i 1898 som instrument, til 0,0357 med populasjon per kvadratkilometer i 1835 som instrument. Fra (6)-(10) er estimatene også konsistente og rimelig kompakte. Alle koeffisientene er signifikante på ett prosent nivå og estimatet varierer fra 0,0359 til 0,0403.

⁶Her hadde ikke kvinner stemmerett enda, så instrumentet antas å være svakt.

Privat, offentlig, industri og service

Fremgangsmåten er her ekvivalent som i hovedanalysen. Modellerer med alternativ hovedforklaringsvariabel, instrumenterer denne med historisk data og deretter instrumenteres både populasjon og høyere utdanning. Premissene er altså like som tidligere.

Presenterer resultatene i tabell 7.2 og 7.3. I tabell 7.2 er det alternativ tilnærming til mål på urbanitet, absoluttverdien av populasjon. Metoden benyttet er både enkel MKM, i modell (1)-(4), og ved instrumentvariabelmetoden i (5)-(8), for de respektive sektorene. Instrumentene brukt er befolkning i 1910 for spesifikasjonene (5)-(8). Ved MKM er koeffisientene ikke helt ulike populasjon per kvadratkilometer. Estimater varierer fra 0,0255 i offentlig sektor til 0,0488 i servicenæringen. Rangeringen av hvem som nyter best av befolkede områder er lik som ved bruk av populasjon per kvadratkilometer. Alle koeffisientene i MKM-spesifikasjonene er signifikante på 0,1 prosent nivå.

I (5)-(8) er det alternative målet på urbanitet instrumentert med populasjonen i 1910, her er hovedkoeffisientene kompakte, og varierer mellom 0,0184 i industrinæringen til 0,278 i servicenæringen. Offentlig er signifikant på ett prosent nivå, service er signifikant på 5 prosent, og privat på 10 prosent. Industrikoeffisienten er insignifikant.

I tabell 7.3 er resultatene ved bruk av alternativt instrument, populasjon per kvadratkilometer i 1801, presentert. Her er estimatene signifikante på 0,1 prosent nivå. Det er fortsatt service som nyter best av urbane områder, på 0,0576 i ren elasticitet. Offentlig har minst nytte av befolkede områder med en elasticitet på 0,0276. Privat og industri ligger på henholdsvis 0,0498 og 0,0405. Fra bruken av populasjon per kvadratkilometer i 1910, så er estimatene gjennomgående noe høyere, men fortsatt kompakte.

Det finnes en rekke svakheter nevnt tidligere, til tross for dette er resultatene likevel interessante og gir en god indikasjon på tingenes tilstand selv om en grad av skjevhet er tilstede.

Kapittel 7

Konklusjon

I denne oppgaven har intensjonen vært å undersøke hvorvidt tett befolkede områder besitter enkelte fordeler som gjør de mer produktive. Deretter hvorvidt enkelte sektorer og næringer drar fordeler av å lokalisere seg i disse områdene, og om disse gevinstene varierer mellom de.

Det er vanskelig å finne noe teorispesifikt grunnlag som underbygger resultatene. Dette kan skyldes at de ulike teoretiske momentene typisk ikke er uavhengige av hverandre, dermed blir de vanskelige å skille fra hverandre, og konkret påpeke hvilket grunnlag som underbygger resultatene. Jeg kan derfor ikke konkludere med hva det er som underbygger disse resultatene annet enn at det bygger på mikrogrunnlaget drøftet i teoridelen av oppgaven. Det er enkelte økonometriske svakheter ved analysen og. Bruken av regional data begrenser metodene for å løse de klassiske økonometriske problemene av identifisering av agglomerasjon. I denne oppgaven har dette vært relatert til uobservert heterogenitet blant individer. Videre argumenterer P.-P. Combes et al. (2008) for at bruken av historisk populasjon som instrument kan skape et nytt potensielt problem, da dette kan ha en effekt på produktivitet i senere år.

Til tross for disse svakhetene, indikerer resultatene oppnådd i analysen at det urbane lønnspremiet typisk var høyere i tett befolkede områder. Ved MKM som metode er den rene elastisiteten er typisk på 4,28 prosent høyere ved en dobling av populasjon per kvadratkilometer, og 3,32 prosent ved korrigerings av observerte karakteristikk. Dette er ikke helt ulikt Carlsen et al. (2013). Ved instrumentvariabelmetoden faller de rene gevinstene, reflektert av elastisiteten, til 4,04 prosent ved en dobling av populasjon per kvadratkilometer. For korreksjon av observerte karakteristikk og eliminering av simultaneitet faller fordelene ytterligere til 3,04 prosent ved dobling av populasjon per kvadratkilometer.

I den andre analysen var fokuset på variasjon i dette premiumet mellom sektorene og næringer i økonomien. Mer spesifikt mellom privat og offentlig sektor, for deretter å se på service- og industrinæringen innen privat sektor. Innen privat og offentlig sektor ligger forskjellen på 2,22 prosentpoeng i favør privat sektor ved MKM som metode, når jeg doubler populasjonen per kvadratkilometer. Ved instrumentvariabelmetoden faller denne differansen til 1,76 prosentpoeng ved dobling av populasjon per kvadratkilometer. Gitt konkurranseaspektet og tilpasningsdyktigheten ovenfor dyktige individer mange av de private aktørene besitter hva gjelder lønnsforhandlinger, diskutert nærmere i kapittel 6.3, er ikke resultatene urovekkende.

Til slutt deler jeg privat sektor inn i service- og industrinæringene. Her er det service-næringen som har gevinster av å lokalisere seg i tett befolkede områder. Ved MKM som metode ligger differansen mellom de to på 1,65 prosentpoeng i favør service, ved en dobling av populasjon per kvadratkilometer. Ved utvidelse til instrumentvariabelmetoden faller denne differansen til 1,57 prosentpoeng ved dobling av populasjon per kvadratkilometer.

Analysen inneholder enkelte svakheter relatert til blant annet uobserverte karakteristikk blant individer, og potensielle problemer relatert til instrumentet benyttet. Til tross for dette viser resultatene fra hovedanalysen at, uavhengig av hvilken modellspesifikasjon eller metode jeg benytter, så var høyere gjennomsnittlig timelønn i 2010 i urbane strøk et faktum. Dette gjenspeiler også en høyere produktivitet i urbane regioner. Omfanget av dette varierer med spesifikasjon og metode, men er gjennomgående konsistente.

Videre kan det fra den sekundære analysen argumenteres for at det urbane lønnspremiumet også varierte mellom ulike sektorer og næringer i økonomien i 2010.

For videre forskning kan det være interessant å se mer kvalitativt på produktivitetsforskjeller i ulike sektorer over lengre tidsperioder.

Kapittel 8

Bibliografi

Referanser

- Bartelsman, E.J. & Doms, M. (2000). Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic literature*, 38(3), 569–594.
- Behrens, K., Duranton, G. & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3), 507–553.
- Campbell, J.R. & Hopenhayn, H.A. (2005). Market size matters. *The Journal of Industrial Economics*, 53(1), 1–25.
- Carlsen, F., Rattsø, J. & Stokke, H.E. (2013). Education, experience and dynamic urban wage premium. *Norwegian University of Science and Technology working paper*, 14, 35.
- Carlsen, F., Rattsø, J. & Stokke, H.E. (2016). Education, experience, and urban wage premium. *Regional Science and Urban Economics*, 60, 39–49.
- Ciccone, A. & Hall, R.E. (1993). Productivity and the density of economic activity.
- Combes, P., Duranton, G., Gobillon, L., Puga, D. & Roux, S. (2007). The productivity advantage of large markets: Distinguishing agglomeration from firm selection. *University of Toronto, mimeograph*.
- Combes, P.-P., Duranton, G. & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2), 723–742.
- Combes, P.-P., Duranton, G., Gobillon, L. & Roux, S. (2010). Estimating agglomeration economies with history, geology, and worker effects. *Agglomeration Economics*, 15–66.
- 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 economies. *Handbook of regional and urban economics*, 4, 2063–2117.
- Duranton, G. & Puga, D. (2005). From sectoral to functional urban specialisation. *Journal of urban Economics*, 57(2), 343–370.
- Foster, L., Haltiwanger, J. & Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? *The American economic review*, 98(1), 394–425.
- Glaeser, E.L., Kallal, H.D., Scheinkman, J.A. & Shleifer, A. (1992). Growth in cities. *Journal of political economy*, 100(6), 1126–1152.
- Glaeser, E.L. & Mare, D.C. (2001). Cities and skills. *Journal of labor economics*, 19(2), 316–342.
- Glaeser, E.L. & Resseger, M.G. (2010). The complementarity between cities and skills. *Journal of Regional Science*, 50(1), 221–244.
- Helsley, R.W. & Strange, W.C. (2014). Coagglomeration, clusters, and the scale and composition of cities. *Journal of Political Economy*, 122(5), 1064–1093.
- Henderson, J.V. (1974). The sizes and types of cities. *The American Economic Review*, 64(4), 640–656.
- Henderson, J.V. (1986). Efficiency of resource usage and city size. *Journal of Urban economics*, 19(1), 47–70.
- Henderson, V., Kuncoro, A. & Turner, M. (1995). Industrial development in cities. *Journal of political economy*, 103(5), 1067–1090.
- Jeffrey, M.W. et al. (2009). Introductory econometrics: A modern approach. *Canada: South-Western Cengage Learning*.
- Krugman, P.R. (1991). Geography and trade.
- Matano, A. & Naticchioni, P. (2011). Wage distribution and the spatial sorting of workers. *Journal of Economic Geography*, lbr013.
- Matano, A., Naticchioni, P. et al. (2009). Wage distribution and the spatial sorting of workers and firms. *DE-Isfol Discussion Paper*(8).
- Moomaw, R.L. (1981). Productivity and city size: a critique of the evidence. *The Quarterly Journal of Economics*, 96(4), 675–688.
- Rattsø, J. & Stokke, H.E. (2014). Population divergence and income convergence: Regional distribution dynamics for Norway. *Regional Studies*, 48(11), 1884–1895.
- Rosenthal, S.S. & Strange, W.C. (2004). Evidence on the nature and sources of agglomeration economies. *Handbook of regional and urban economics*, 4, 2119–2171.
- Segal, D. (1976). Are there returns to scale in city size? *The Review of Economics and*

- Statistics*, 339–350.
- Smith, A. (1776). An inquiry into the nature and causes of the wealth of nations: Volume one.
- Sveikauskas, L. (1975). The productivity of cities. *The Quarterly Journal of Economics*, 89(3), 393–413.
- Syverson, C. (2004). Market structure and productivity: A concrete example. *Journal of Political Economy*, 112(6), 1181–1222.
- Syverson, C. (2007). Prices, spatial competition and heterogeneous producers: an empirical test. *The Journal of Industrial Economics*, 55(2), 197–222.
- Wheeler, C.H. (2001). Search, sorting, and urban agglomeration. *Journal of Labor Economics*, 19(4), 879–899.

Kapittel 9

A Appendiks

9.1 Deling av fordelene fra bredere variasjon av tilbydere av inputs

Deling av fordelene fra bredere variasjon av tilbydere av inputs

Utregning i et tilfelle med to innsatsvarer, ønsker å minimere:

$$q_1 X_1 + q_2 X_2 \tag{9.1}$$

gitt:

$$(X_1^{(1/1+\epsilon)} + X_2^{(1/1+\epsilon)})^{(1+\epsilon)} \tag{9.2}$$

Skriver om bibetingelsen i (9.2)

$$Y^{1/1+\epsilon} = X_1^{1/1+\epsilon} + X_2^{1/1+\epsilon} \tag{9.3}$$

Benytter Lagrange metode for å minimere uttrykket

$$L = q_1 X_1 + q_2 X_2 - \lambda(Y^{1/1+\epsilon} - X_1^{1/1+\epsilon} - X_2^{1/1+\epsilon}) \tag{9.4}$$

Minimering av uttrykket mhp X_1 , X_2 og λ gir følgende førsteordensbetingelser

$$\frac{dL}{dX_1} = q_1 + \frac{1}{1+\epsilon} * \lambda X_1^{(1/1+\epsilon)-1} = 0 \tag{9.5}$$

Løser for λ og får

$$\lambda = -q_1(1 + \epsilon)X_1^{(\epsilon/1+\epsilon)} \quad (9.6)$$

Ekvivalent fremgangsmåte for X_2

$$\frac{dL}{dX_2} = q_2 + \frac{1}{1 + \epsilon} * \lambda X_2^{(1/1+\epsilon)-1} = 0 \quad (9.7)$$

Løser for λ

$$\lambda = -q_2(1 + \epsilon)X_2^{(\epsilon/1+\epsilon)} \quad (9.8)$$

Videre blir

$$\frac{dL}{d\lambda} = Y^{1/1+\epsilon} = X_1^{1/1+\epsilon} + X_2^{1/1+\epsilon} \quad (9.9)$$

Kombinerer relasjon (9.6) og (9.8) for å finne et uttrykk for X_1

$$q_1(1 + \epsilon)X_1^{(\epsilon/1+\epsilon)} = q_2(1 + \epsilon)X_2^{(\epsilon/1+\epsilon)} \quad (9.10)$$

$$X_1 = \left(\frac{q_2}{q_1}\right)^{(1+\epsilon)/\epsilon} * X_2 \quad (9.11)$$

Benytter relasjon (9.9) og setter inn for X_1

$$Y^{1/1+\epsilon} = \left(\frac{q_2}{q_1}\right)^{(1+\epsilon)/\epsilon} * X_2^{(1/1+\epsilon)} + X_2^{1/1+\epsilon} \quad (9.12)$$

$$Y^{1/1+\epsilon} = X_2^{(1/1+\epsilon)} \left(1 + \left(\frac{q_2}{q_1}\right)^{(1/\epsilon)}\right) \quad (9.13)$$

Løser for X_2

$$X_2^{1/1+\epsilon} = \frac{Y^{1/1+\epsilon}}{1 + (q_2/q_1)^{1/\epsilon}} \quad (9.14)$$

ved å løse parentesen $1 + (q_2/q_1)^{1/\epsilon}$ får jeg:

$$X_2^{1/1+\epsilon} = \frac{Y^{1/1+\epsilon}}{1 + q_2^{1/\epsilon} * q_1^{-1/\epsilon}} \quad (9.15)$$

Multipliserer teller og nevner med $\frac{q_2^{-1/\epsilon}}{q_2^{-1/\epsilon}}$ og får:

$$X_2^{1/1+\epsilon} = \frac{Y^{1/1+\epsilon} * q_2^{-1/\epsilon}}{q_2^{-1/\epsilon} * q_1^{-1/\epsilon} * 1} \quad (9.16)$$

9.1. DELING AV FORDELENE FRA BREDERE VARIASJON AV TILBYDERE AV INPUTS

Flytter til slutt over uttrykket $1/1 + \epsilon$ fra venstre siden over på høyreside og får et uttrykk for:

$$X_2 = \frac{Y * q_2^{-(1+\epsilon)/\epsilon}}{(q_2^{-1/\epsilon} + q_1^{-1/\epsilon})^{1+\epsilon}} \quad (9.17)$$

Løsningen for X_1 blir dermed:

$$X_1 = \left(\frac{q_2}{q_1}\right)^{(1+\epsilon)/\epsilon} * X_2 \quad (9.18)$$

Setter inn for X_2 fra relasjon (9.17)

$$X_1 = \left(\frac{q_2}{q_1}\right)^{(1+\epsilon)/\epsilon} * \frac{Y * q_2^{-(1+\epsilon)/\epsilon}}{(q_2^{-1/\epsilon} + q_1^{-1/\epsilon})^{1+\epsilon}} \quad (9.19)$$

Løser parentesene i det første uttrykket i (9.19), og får:

$$X_1 = q_2^{(1+\epsilon)/\epsilon} * q_1^{-(1+\epsilon)/\epsilon} * \frac{Y * q_2^{-(1+\epsilon)/\epsilon}}{(q_2^{-1/\epsilon} + q_1^{-1/\epsilon})^{1+\epsilon}} \quad (9.20)$$

$q_2^{(1+\epsilon)/\epsilon} * q_2^{-(1+\epsilon)/\epsilon} = 1$ som gir et uttrykk for X_1

$$X_1 = \frac{Y * q_1^{-(1+\epsilon)/\epsilon}}{(q_2^{-1/\epsilon} + q_1^{-1/\epsilon})^{1+\epsilon}} \quad (9.21)$$

Utregning av prisen på innsatsfaktorer, q . Arbeidskraft som trengs per produserte enhet, X , er gitt ved

$$X = \beta * l - \alpha \quad (9.22)$$

løser for l og får

$$l = \frac{X + \alpha}{\beta} = \frac{X}{\beta} + \frac{\alpha}{\beta} \quad (9.23)$$

Bedriftens profitt er gitt ved

$$\Pi = q * X - W * l \quad (9.24)$$

setter inn for l fra (9.23) og får

$$\Pi = q * X - W \left(\frac{X}{\beta} + \frac{\alpha}{\beta} \right) \quad (9.25)$$

Setter X utafør parentes og får uttrykket

$$\Pi = X\left(q - \frac{W}{\beta}\right) - \frac{W * \alpha}{\beta} \quad (9.26)$$

Etterspørselen etter innsatsvaren, X , avhenger av prisen, q . Slik at profittmaksimerende pris er determinert av $\frac{d\Pi}{dq} = 0$ som gir:

$$X + \left(q - \frac{W}{\beta} * \frac{dX}{dq}\right) = 0 \quad (9.27)$$

Løser videre for innsatsprisen, q :

$$q * \frac{W}{\beta} = \frac{-X}{dX/dq} \Rightarrow q = \frac{W}{\beta} \div \frac{X}{dX/dq} \quad (9.28)$$

Fra tidligere fant vi at elastisiteten av X med hensyn på q er $\frac{-X}{q} = \frac{-\epsilon}{1+\epsilon}$, slik at $\frac{dX}{dq} = \frac{-\epsilon}{1+\epsilon} * \frac{X}{q}$. Dette gjør at vi kan skrive om førsteordensbetingelsen

$$q = \frac{W}{\beta} \div \frac{X}{\frac{-\epsilon}{1+\epsilon} * \frac{X}{q}} \Rightarrow q = \frac{W}{\beta} \div \frac{\epsilon * X * q}{-(1+\epsilon) * X} \quad (9.29)$$

Stryker X -ene mot hverandre og får

$$q = \frac{W}{\beta} - \frac{\epsilon * q}{-(1+\epsilon)} \Rightarrow q = \frac{W}{\beta} + \frac{\epsilon * q}{1+\epsilon} \quad (9.30)$$

Omorganiserer q og $\frac{W}{\beta}$ og får $\frac{W}{\beta} = q * \left(1 - \frac{\epsilon}{1+\epsilon}\right)$. Dette gir

$$\frac{W}{\beta} = \frac{1}{1+\epsilon} * q \Rightarrow q = \frac{(1+\epsilon) * W}{\beta} \quad (9.31)$$

Utregning av marginalkostnad $\frac{W}{\beta}$ fra $\frac{W}{\beta} * X + \frac{W * \alpha}{\beta}$ blir

$$\frac{dkostnad}{dX} = \frac{W}{\beta} \quad (9.32)$$

Utregning av konsistent produksjonsnivå ved null profitt. Kombinerer (2.2), (2.9) og (2.10) for å finne X . Setter inn for henholdsvis l og q i profittfunksjonen og får

$$\frac{1+\epsilon}{\beta} * W * X - \frac{X + \alpha}{\beta} * W = 0 \quad (9.33)$$

9.1. DELING AV FORDELENE FRA BREDERE VARIASJON AV TILBYDERE AV INPUTS_{ix}

Multipliserer ut parentesene og forenkler

$$\frac{W * X}{\beta} + \frac{W * X * \epsilon}{\beta} - \frac{W * X}{\beta} - \frac{W * \alpha}{\beta} = 0 \Rightarrow X * \frac{W * \epsilon}{\beta} = \alpha * \frac{W}{\beta} \quad (9.34)$$

Løser for X, og får

$$X = \frac{\alpha * W * \beta}{\beta * W * \epsilon} \Rightarrow X = \frac{\alpha}{\epsilon} \quad (9.35)$$

Utregning av likevektsantallet bedrifter som produserer innsatsgoder.

$$\frac{\alpha}{\epsilon} = \beta * l - \alpha \Rightarrow \beta * l = \frac{\alpha}{\epsilon} + \alpha \quad (9.36)$$

Multipliserer uttrykket med ϵ og løser for l

$$\beta * l * \epsilon = \alpha(1 - \epsilon) \Rightarrow l = \frac{\alpha(1 + \epsilon)}{\beta * \epsilon} \quad (9.37)$$

Finner likevektstallet ved å implementere (9.37) i $n = \frac{L}{l}$ som gir

$$n = \frac{\beta * \epsilon}{\alpha(1 + \epsilon)} * L. \quad (9.38)$$

Videre finner jeg aggregert produksjon i hver sektor ved å sette (9.35) og (9.38) inn i produktfunksjonen. Ser at hver produsent av innsatsvarer produserer samme volum, kan da skrive:

$$Y = (n * X^{1/(1+\epsilon)}) \Rightarrow Y = n^{1+\epsilon} * X \quad (9.39)$$

Setter inn for uttrykkene for n og X fra henholdsvis (9.35) og (9.38)

$$Y = \left(\frac{\beta * \epsilon}{\alpha(1 + \epsilon)} * L\right)^{1+\epsilon} * \frac{\alpha}{\epsilon} \Rightarrow Y = \frac{\beta^{1+\epsilon} * \epsilon^{1+\epsilon} * L^{1+\epsilon}}{(1 + \epsilon)^{1+\epsilon}} * (\alpha^{1+\epsilon}) * \frac{\alpha}{\epsilon} \quad (9.40)$$

Multipliserer inn $\frac{\alpha}{\epsilon}$ og får

$$Y = \frac{\beta^{1+\epsilon} * \epsilon^{\epsilon}}{(1 + \epsilon)^{1+\epsilon} * \alpha^{\epsilon}} * L^{1+\epsilon} \quad (9.41)$$

Her er uttrykket i parentes en konstant, k. Slik at

$$Y = k * L^{1+\epsilon} \quad (9.42)$$

Dersom vi antar at $\beta = (1 + \epsilon) * \left(\frac{\alpha}{\epsilon}\right)^{(\epsilon/1+\epsilon)}$ får vi at k=1. Setter parentes=1 og løser for β . Setter inn for β inn igjen i parentes og får k=1. Dette gir da aggregert produksjon

$$Y = L^{1+\epsilon} \tag{9.43}$$

9.2 Robusthetsresultater

Hensikten med robusthetsanalysen er å se om estimatene varierer mye mellom ulike hovedforklaringsvariable og instrumenter.

Hovedanalyse

Tabell 9.1: Alternative instrument for spesifikasjon (2) og (3)

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---------------------|------------------------|-----------------------|----------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn | Lntimelonn |
| Lnpopkvkm | 0.0355*** (0.00555) | 0.0334* (0.0165) | 0.0380* (0.0173) | 0.0351*** (0.00565) | 0.0357*** (0.00557) | 0.0363*** (0.00414) | 0.0358*** (0.00816) | 0.0361*** (0.00893) | 0.0359*** (0.00418) | 0.0403*** (0.00909) |
| universitet | 0.745*** (0.123) | 0.766*** (0.201) | 0.719*** (0.209) | 0.749*** (0.124) | 0.743*** (0.123) | 0.809*** (0.124) | 0.816*** (0.159) | 0.811*** (0.167) | 0.815*** (0.124) | 0.752*** (0.169) |
| vgs | 0.693*** (0.156) | 0.674** (0.212) | 0.716** (0.219) | 0.689*** (0.156) | 0.695*** (0.156) | 0.854*** (0.143) | 0.851*** (0.147) | 0.853*** (0.147) | 0.852*** (0.143) | 0.871*** (0.149) |
| klima | 0.00295* (0.00140) | 0.00307 (0.00167) | 0.00280 (0.00171) | 0.00297* (0.00140) | 0.00294* (0.00140) | | | | | |
| Offentlig | -0.0324 (0.103) | -0.0564 (0.207) | -0.00349 (0.217) | -0.0369 (0.103) | -0.0301 (0.103) | | | | | |
| kultur | 0.00987* (0.00429) | 0.00960* (0.00475) | 0.0102* (0.00483) | 0.00982* (0.00429) | 0.00990* (0.00429) | | | | | |
| _cons | 5.002*** (0.102) | 5.020*** (0.172) | 4.980*** (0.179) | 5.005*** (0.102) | 5.000*** (0.102) | 4.937*** (0.0800) | 4.938*** (0.0805) | 4.937*** (0.0807) | 4.938*** (0.0800) | 4.932*** (0.0813) |
| N | 89 | 89 | 89 | 89 | 89 | 89 | 89 | 89 | 89 | 89 |
| adj. R ² | 0.737 | 0.737 | 0.735 | 0.737 | 0.737 | 0.721 | 0.721 | 0.721 | 0.721 | 0.716 |

Standard errors in parentheses

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

2.analyse

Tabell 9.2: Alternativ hovedforklaringsvariabel absoluttverdien av populasjon

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---------------------|------------------------|-------------------------|------------------------|------------------------|------------------------|-------------------------|-------------------------|-------------------------|
| | LNOffentliglonn | LNPrivatlonn | LNIndustrilonn | LNServicelonn | LNIndustrilonn | LNServicelonn | LNOffentliglonn | LNPrivatlonn |
| Lnpop | 0.0255*** (0.00524) | 0.0397*** (0.0105) | 0.0335*** (0.0120) | 0.0488*** (0.0104) | 0.0184 (0.0135) | 0.0278** (0.0120) | 0.0268*** (0.00587) | 0.0215* (0.0120) |
| universitet | 0.635*** (0.102) | 0.992*** (0.206) | 1.241*** (0.234) | 0.758*** (0.204) | 1.401*** (0.239) | 0.979*** (0.211) | 0.622*** (0.104) | 1.184*** (0.212) |
| vgs | 0.552*** (0.122) | 0.824*** (0.244) | 1.032*** (0.278) | 0.450* (0.242) | 0.937*** (0.274) | 0.318 (0.242) | 0.560*** (0.119) | 0.710*** (0.243) |
| klima | 0.0000216 (0.00110) | 0.00777*** (0.00222) | 0.0103*** (0.00252) | 0.00475** (0.00220) | 0.0110*** (0.00248) | 0.00572*** (0.00219) | -0.0000379 (0.00108) | 0.00861*** (0.00220) |
| KULTUR | -0.00163 (0.00346) | 0.00947 (0.00695) | 0.0194** (0.00789) | -0.00296 (0.00688) | 0.0176** (0.00773) | -0.00535 (0.00684) | -0.00149 (0.00336) | 0.00740 (0.00686) |
| _cons | 4.925*** (0.0871) | 4.573*** (0.175) | 4.488*** (0.199) | 4.705*** (0.173) | 4.660*** (0.209) | 4.943*** (0.185) | 4.910*** (0.0906) | 4.779*** (0.185) |
| N | 89 | 89 | 89 | 89 | 89 | 89 | 89 | 89 |
| adj. R ² | 0.625 | 0.559 | 0.556 | 0.553 | 0.548 | 0.531 | 0.624 | 0.543 |

Standard errors in parentheses

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

Tabell 9.3: Sektor og næring - alternativt instrument

| | (1) | (2) | (3) | (4) |
|----------------------------|-------------------------|------------------------|------------------------|------------------------|
| | LNIndustrilonn | LNServicelonn | LNOffentliglonn | LNPrivatlonn |
| Lnpopkvkm | 0.0405*** (0.00918) | 0.0576*** (0.00643) | 0.0276*** (0.00429) | 0.0498*** (0.00735) |
| universitet | 0.731 (0.450) | 0.162 (0.315) | 0.255 (0.210) | 0.391 (0.360) |
| vgs | 0.792* (0.338) | 0.198 (0.236) | 0.321* (0.158) | 0.573* (0.270) |
| klima | 0.00808*** (0.00234) | 0.00145 (0.00164) | -0.00125 (0.00109) | 0.00491** (0.00187) |
| KULTUR | 0.0262*** (0.00796) | 0.00581 (0.00557) | 0.00317 (0.00372) | 0.0178** (0.00637) |
| _cons | 4.920*** (0.211) | 5.267*** (0.148) | 5.284*** (0.0986) | 5.064*** (0.169) |
| <i>N</i> | 89 | 89 | 89 | 89 |
| adj. <i>R</i> ² | 0.617 | 0.751 | 0.631 | 0.685 |

Standard errors in parentheses

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

9.3 Liste over økonomiske regioner

Listen viser 90 økonomiske regioner og tilhørende kommuner. I min oppgave er nordre- og sunnhordland slått sammen til én region. Folkemengden er ikke ekvivalent med folke- mengde brukt i oppgaven da denne refererer til et annet år enn hva jeg bruker.

| Fylke | Økonomiske region/Kode | Kommune | Folkemengde |
|---------------|-------------------------------|------------------|----------------|
| Østfold | Halden 0191 | 0101 Halden | 26 733 |
| | | 0188 Aremark | 1451 |
| | | 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 0913 | 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 | Økonomiske 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 | 18 629 |
| | | 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 | |
| Oslo | Oslo 0391 | 0391 Oslo | 507 467 |
| | | I alt | 507 467 |

| Fylke | Økonomiske 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 Norde Land | 6 950 |
| | | I alt | 67 204 |
| | Midt-Gudbrandsdalen 0593 | 0516 Nord-Fron | 5 963 |
| | | 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 | 2 302 |
| | | 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 | Økonomiske 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 Ulvdal | 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 Hol | 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 | Økonomiske region/Kode | Kommune | Folkemengde |
|----------|----------------------------------|------------------|----------------|
| Vestfold | Tønsberg/Horten 0791 | 0704 Tønsberg | 34 716 |
| | | 0701 Borre | 23 764 |
| | | 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 | 4 694 |
| | | 0822 Sauherad | 4 411 |
| | | 0827 Hjartdal | 1 687 |
| | | I alt | 23 334 |
| | Kragerø 0893 | 0815 Kragerø | 10 656 |
| | | 0817 Drangedal | 4 175 |
| | | I alt | 14 831 |
| | Rjukan 0894 | 0826 Tinn | 6 560 |
| | | I alt | 6 560 |
| | Vest-Telemark 0895 | 0828 Selfjord | 2 928 |
| | | 0829 Kviteseid | 2 667 |
| | | 0830 Nissedal | 1 477 |
| | | 0831 Fyresdal | 1 353 |
| | | 0833 Tokke | 2 506 |
| 0834 Vinje | | 3 869 | |
| I alt | | 14 770 | |

| Fylke | Økonomiske region/Kode | Kommune | Folkemengde |
|------------|------------------------|---------------------|---------------|
| Aust-Adger | 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 Hornes | 3 346 |
| | | 0935 Iveland | 1 128 |
| | | 0938 Bygland | 1 351 |
| | | 0940 Valle | 1 439 |
| | | 0941 Bykle | 836 |
| | | 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 |
|----------|---------------------------|-------------------|----------------|
| Rogoland | Egersund 1191 | 1101 Eigersund | 13 288 |
| | | 1111 Sokndal | 3 358 |
| | | 1112 Lund | 3 096 |
| | | 1114 Bjerkreim | 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 | | 4848 |
| | 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 Kvam | 8 592 |
| | | 1241 Fusa | 3 864 |
| | | 1242 Samnager | 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 |

| Fylke | Økonomiske region/Kode | Kommune | Folkemengde |
|-----------|------------------------------|-----------------|---------------|
| Hordaland | Nordre Sunnhordoland 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 | 4 667 |
| | | 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 Vagsøy | 6 479 | |
| | 1443 Eid | 5 780 | |
| | 1445 Gloppen | 5 769 | |
| | 1449 Stryn | 6 666 | |
| | 1441 Selje | 3 066 | |
| | 1444 Hornidal | 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 |
| | 1523 Ørskog | | 2 076 |
| | 1524 Norddal | | 1 960 |
| | 1525 Strand | | 4 677 |
| | 1526 Stordal | | 1 052 |
| | 1528 Stykkylven | | 7 280 |
| | 1529 Skodje | | 3 513 |
| | 1531 Sula | | 7 070 |
| | 1532 Giske | | 6 344 |
| | 1534 Haram | | 8 744 |
| | 1546 Sandøy | | 1 332 |
| | I alt | | 82 933 |

| Fylke | Økonomiske region/Kode | Kommune | Folkemengde |
|-----------------|------------------------|---------------|---------------|
| Møre og Romsdal | Ulsteinvik 1594 | 1516 Ulstein | 6 541 |
| | | 1511 Vanylven | 3 854 |
| | | 1514 Sande | 3 081 |
| | | 1515 Herøy | 8 390 |
| | | 1517 Hareid | 4 748 |
| | | I alt | 26 344 |
| | Ørsta/Volda 1595 | 1520 Ørsta | 10 276 |
| | | 1519 Volda | 8 322 |
| | | I alt | 18 598 |
| | Sunndalsøra 1596 | 1563 Sunndal | 7 368 |
| | | 1560 Tingvoll | 3 148 |
| | | I alt | 10 516 |
| | Surnadal 1597 | 1566 Surnadal | 6 252 |
| | | 1567 Rindal | 2 183 |
| | | 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 Midte 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 |
| | | | 1632 Roan | 1 121 |
| | | | 1633 Osen | 1 194 |
| | | | I alt | 15 451 |
| | | Oppdal 1694 | 1638 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-Tøndelag | 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 Frosta | 2 430 |
| | | I alt | 33 575 |
| | Grong 1795 | 1742 Grong | 2 565 |
| | | 1738 Lierne | 1 565 |
| | | 1739 Røyrvik | 595 |
| | | 1740 Namskogan | 989 |
| | | I alt | 5 714 |
| | Rørвик 1796 | 1750 Vikna | 3867 |
| | | 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 | 9 653 |
| | | 1842 Skjerstad | 1 098 |
| | | 1845 Sørfold | 2 367 |
| | | 1848 Steigen | 2 977 |
| | | 1849 Hamarøy | 2 008 |
| | | I alt | 76 364 |
| | | | Narvik 1892 |
| 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 |

| Fylke | Økonomiske region/Kode | Kommune | Folkemengde |
|---------------------------|-----------------------------|-------------------|---------------|
| Nordland | 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 1895 | 1824 Vefsn | 13 553 |
| | | 1825 Grane | 1 652 |
| | | 1826 Hattfjelldal | 1 634 |
| | | I alt | 16 839 |
| | Mo i Rana 1896 | 1833 Rana | 25 255 |
| | | 1828 Nesna | 1 882 |
| | | 1832 Hemnes | 4 869 |
| | | I alt | 31 826 |
| | Lofoten 1897 | 1865 Vågan | 9 229 |
| | | 1856 Røst | 666 |
| | | 1857 Værøy | 775 |
| | | 1859 Flakstad | 1 575 |
| | | 1860 Vestvågøy | 10 750 |
| | | 1874 Moskenes | 1 352 |
| I alt | | 24 347 | |
| Vesterålen 1898 | 1870 Sortland | 9 230 | |
| | 1866 Hadsel | 8 321 | |
| | 1857 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 |
| | | 1919 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 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 Tana | 3 074 |
| | | 2027 Nesseby | 965 |
| | | 2028 Båtsfjord | 2 470 |
| | | I alt | 16 580 |
| | Hammerfest 2092 | 2004 Hammerfest | 9 213 |
| | | 2017 Kvalsund | 1 106 |
| | | 2018 Måsøy | 1 477 |
| | | 2019 Nordkapp | 3 517 |
| | | 2020 Porsanger | 4 451 |
| | | 2021 Karasjok | 2 901 |
| | | 2022 Lebesby | 1 463 |
| | | 2023 Gamvik | 1 288 |
| | | I alt | 25 416 |
| | Alta 2093 | 2012 Alta | 16 837 |
| | | 2011 Kautokeino | 3 068 |
| | | 2014 Loppa | 1 426 |
| | | 2015 Hasvik | 1 200 |
| | | I alt | 22 531 |
| | Kirkenes 2094 | Sør-Varanger | 9 532 |
| | | I alt | 9 532 |