

Wermundsen, Leif Henrik Solland
Geier, Kristian Schultz

Inntektsgapet mellom kvinner og menn

En empirisk analyse av hvilke variabler som påvirker inntektsgapet mellom kvinner og menn, med fokus på inntektsgapet i urbane og rurale områder

Masteroppgave i Samfunnsøkonomi
Veileder: Hildegunn Ekroll Stokke
Trondheim, juni 2018

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi

Forord

Denne oppgaven avslutter vårt kapittel som studenter ved NTNU. Det har vært noen fantastiske år vi aldri kommer til å glemme. En stor takk rettes til Hildegunn Ekroll Stokke for verdifull og utrolig gode tilbakemeldinger. Vi vil også benytte anledningen til å gratulere henne med 1. plass i instituttets Fantasy Premier League liga. En takk rettes også til moren til Kristian og broren til Henrik for korrekturlesing av oppgaven. Til slutt vil vi også takke Radioresepsjonen for at de leverer en humoristisk podcast som vi har kunnet nyte i de 3 månedene vi brukte på å lage datasettet. Kristian og Henrik out.

Trondheim, 1 Juni 2018.

Leif Henrik Solland Wermundsen & Kristian Schultz Geier.

Sammendrag

Denne oppgaven anvender paneldata for 160 bo- og arbeidsmarkedsregioner i Norge, for å undersøke inntektsgapet mellom kvinner og menn, med spesielt fokus på inntektsgapet mellom by- og rurale regioner. Resultatet fra hovedmodellen viser at inntektsgapet mellom kvinner og menn er 2.3% større i by regioner enn i rurale regioner og er robust med flere modellspesifikasjoner. I analysedelen av oppgaven blir det inkludert ulike variabler for å undersøke om det har en effekt på inntektsgapet. Analysen antyder at en større andel menn med høyere utdanning relativt til kvinner vil øke inntektsgapet. Videre viser vi i analysen at en økt andel menn ansatt i offentlig sektor relativt til kvinner vil redusere inntektsgapet. Vi har inkludert en variabel som tar hensyn til andelen barn i aldersgruppen 0-9 år og resultatene viser at en økt andel barn i denne aldersgruppen påvirker inntektsgapet negativt. Vi kan også konkludere med at en økning i mannlige deltidsarbeidere relativt til kvinner vil redusere inntektsgapet. En økning i andelen sysselsatte i aldersgruppen 20-25 år vil også ha en positiv innvirkning på inntektsgapet.

Abstract

This thesis uses panel data for 160 labor market regions in Norway to examine the gender income gap, with an extra focus on the gender income gap between urban and rural regions. Our results show that the gender income gap is 2.3%-points larger in urban regions compared to rural regions in our main model. This results is robust to different model specifications. The analysis suggest that a larger share of men with higher education relative to females will increase the gender income gap. Further, our analysis concludes that a greater share of males employed in the public sector relative to females results in a reduction in the gender income gap. We have also included different age groups. We have included a variable accounting for the share of children in the age group 0-9, and the results shows that an increased share of children in this age group effects the gender income gap negatively. We can conclude that an increase in part-time employed males relative to females have a positive effect on the gender income gap. Also, our thesis indicates that an increased share of the employed work force that are in the age group 20-25 years, has a positive effect on the gender income gap.

Innhold

1	Innledning	1
1.1	Oppgavens formål	1
1.2	Oppgavens oppsett	2
2	Tidligere litteratur	3
2.1	Lønnsgap mellom kvinner og menn	3
2.2	Regionale forskjeller i lønnsgapet	5
2.3	Agglomerasjonseffekter	7
2.4	Lønnsgapet mellom privat og offentlig sektor	7
2.5	Ulike mål på lønnsforskjeller	9
3	Data og deskriptiv statistikk	11
3.1	Introduksjon	11
3.2	160 Bo-og Arbeidsmarkedsregioner	11
3.2.1	Metode anvendt ved inndeling av BA-regioner	12
3.3	Datamaterialet	13
3.4	Oversikt over variabler og deskriptiv statistikk	14
3.4.1	By	14
3.4.2	Inntektsspreium	15
3.4.3	Høyere utdanning	16
3.4.4	Offentlig sektor	18
3.4.5	Deltid	18
3.4.6	Leder	19
3.4.7	Sysselsatte aldersgrupper	20
3.4.8	Barn	20
3.5	Næringer	21
3.5.1	Gini-Koeffisient	22
3.6	Hypoteser	23
3.6.1	Variabler som er inkludert i datasettet	26
4	Økonometrisk spesifisering og utfordringer	27
4.1	Introduksjon	27

4.2	Økonometriske utfordringer	27
4.2.1	Målefeil	29
4.2.2	Utelatt variabel	29
4.2.3	Simultanitet	30
4.2.4	Feilspesifikasjon	30
4.2.5	Heteroskedastisitet	30
4.2.6	Autokorrelerte feilledd	31
4.3	Økonometrisk spesifikasjon	33
4.4	Potensielle problemer med datamaterialet	34
5	Analyse	35
5.1	Enkel modell	35
5.2	Utdanning og offentlig sektor	35
5.3	Ledere, deltid og fertilitet	38
5.4	Avsluttende modell	41
5.5	Robusthetsanalyse	43
5.5.1	Inntektsgapet i tidlig fase av arbeidslivet	43
5.5.2	Sektor-variabler, Fixed effects og Gini	44
5.5.3	Ulike by definisjoner	46
5.5.4	Interaksjonsledd	48
6	Konklusjon	51
A	Vedlegg	57
A	Manglende observasjoner i datasettet	57
B	Konstruksjon av datasett	58
C	Korrelasjonsmatrise	58
D	160 bo- og arbeidsmarkedsregioner	60

1 Innledning

Norge har i lengre tid tilrettelagt for at kvinner skal få viktige posisjoner i arbeidslivet. En utgivelse gjort av "World Economic Forum" i 2017 (Forum 2007), plasserte Norge som nummer to over land med lavest lønnsforskjell mellom kvinner og menn, plassen etter Island. Resultatene fra rapporten viser at kvinner tjener 83 % av menns lønn, en konklusjon som tilsvarer Statistisk Sentralbyrå sine tall (Kristoffersen 2017). De øvrige nordiske landene Sverige, Danmark og Finland er på plassene bak Norge.

Antall kvinner i lederstillinger relativt til menn har lenge vært lav i Norge. I 2011 var 32 % av alle ledere kvinner. Problemet er størst i privat sektor, der syv av ti ledere er menn. Det har de siste årene blitt satt større fokus på å få kvinner inn i lederstillinger i Norge, og dette har resultert i at andelen kvinnelige ledere økte til 40 % i 2016. Forbedringen tilsvarer 8 %-poeng på 5 år. I dag er tre av fire ansatte i kommunal sektor kvinner og totalt sett i staten er denne dominansen mindre, der seks av ti er kvinner.

Andelen kvinner som tar høyere utdanning har gradvis økt i Norge. I 2008 var det flere kvinner enn menn som tok høyere utdanning. For ansatte med bakgrunn fra høyskole eller universitet på masternivå er lønnsforskjellene større enn for personer med et lavere utdanningsnivå. Noe av forklaringen på dette kan være at det er generelt større lønnsspredning i grupper som har høyere utdanning. En annen faktor er at på bachelor-nivå er det generelt sett flere ulike yrkesvalg kvinner og menn kan velge.

1.1 Oppgavens formål

Ved bruk av paneldata innhentet fra SSB, tar denne oppgaven for seg problemstillingen: *Hvilke variabler påvirker inntektsgapet mellom kvinner og menn, med fokus på inntektsgapet i urbane og rurale områder.* Inntektsforskjeller mellom kvinner og menn har i lengre tid vært ansett som et samfunnsproblem. Å kunne forstå hvordan inntektsgapet utvikler seg på mellom ulike regioner har vært motivasjonen bak valg av oppgave. Vi har av den grunn hatt et ekstra fokus på inntektsforskjellene mellom by og rurale områder i hver av

de ulike regresjonene presentert i kapittel 5.

1.2 Oppgavens oppsett

Oppgavens oppbygning er som følger:

Kapittel 2 presenterer et utvalg av tidligere litteratur knyttet til lønnsgapet mellom kvinner og menn.

Kapittel 3 presenterer deskriptiv statistikk og variabelbeskrivelse.

Kapittel 4 presenterer økonometriske utfordringer og spesifikasjoner.

Kapittel 5 presenterer resultatene fra analysen.

Kapittel 6 presenterer konklusjon og oppsummering.

2 Tidligere litteratur

Litteraturen har i lengre tid forsøkt å forklare lønnsgapet¹ mellom kvinner og menn. Denne oppgaven har tatt utgangspunkt i ulike forskningsartikler som inspirasjon for bruk av de ulike variablene oppgaven inkluderer.

2.1 Lønnsgap mellom kvinner og menn

Sammenlignet med etterkrigstiden har lønnsgapet mellom kvinner og menn blitt redusert. Dette blir forklart i artikkelen til Blau & Kahn (2000) som anvender data fra USA i tidsperioden 1950-2000 tallet. Det blir konkludert med at lønnsgapet ble redusert frem til 1990 og det har siden holdt seg relativt stabilt. Blau & Kahn (2017) antyder det samme som Blau & Kahn (2000), og forklarer at fra 1980-1989 ble lønnsgapet redusert grunnet kvinners økte utdanningsnivå og høyere deltakelse i arbeidslivet. I samme periode falt den uforklarlige delen² av lønnsgapet fra 21-29 % i 1980 til 8-18 % i 1989 og har holdt seg stabilt siden.

Det blir fastslått i artikkelen til Blau & Kahn (2017) at tradisjonelle variabler som utdanning og arbeidserfaring fremdeles er viktige for å tolke lønnsforskjeller mellom kvinner og menn, men at disse variablene har blitt mindre viktig med årene.

Begrunnelsen er at det i nyere tid er flere kvinner enn menn som utdanner seg, og dette har resultert i at «arbeidserfaringsgapet» har blitt redusert.

I artikkelen til Bayard et al. (2003) anvendes datamaterialet som dekker alle de viktige sektorene og yrkene på tvers av regioner i USA og dekomponerer lønnsgapet i en forklart og uforklart del for å identifisere sentrale faktorer bak lønnsgapet mellom kvinner og menn. Datamaterialet anvendt i artikkelen er "The New Worker-Establishment Characteristics Database" som omfatter arbeidstakere og bedrifter fra sektorer og regioner i

¹I økonomisk tolkning deles lønnsgap mellom kvinner og menn ofte inn i en forklart og en uforklart del. Det forklarte lønnsgapet kan observeres ved å se på ulikheter, mens den uforklarte delen er det resterende gapet etter at man har kontrollert for de observerte variablene.

²Diskriminering og lignende.

USA. Resultatene indikerer at selv om kvinner oftere enn menn velger å arbeide i sektorer som har en generelt lavere lønn, skyldes en vesentlig del av lønnsgapet fortsatt individets kjønn. Det vises også til at yrkesvalg utgjør over halvparten av lønnsgapet mellom kvinner og menn.

Å se på lønnsgapet i ulike deler av jobbkarrieren har vært viktig for å kunne forklare lønnsforskjeller mellom kvinner og menn. Dette undersøker Manning & Swaffield (2008) ved å studere data fra "British Household Panel Study" fra 1991-2002 og "Labour Force Survey". Manning & Swaffield (2008) konkluderer med at kvinner med lik gjennomsnittslønn som menn i tidlig del av arbeidslivet, har 8% lavere gjennomsnittslønn enn menn 10 år senere. Dette gjelder kvinner som i tidsperioden ikke har fått barn, men jobbet fulltid og har tilnærmet like personlige egenskaper som mennene de sammenlignes med. Det er hovedsakelig to grunner som forklarer dette: Job-Shopping teorier ³ og psykologiske forskjeller.

Stokke (2016) anvender data for arbeidsgivere og arbeidstakere i Norge og identifiserer at det ikke er et lønnsgap når individer med like ferdigheter ⁴ inntar arbeidsmarkedet for første gang. Dette samsvarer med resultatene til Manning & Swaffield (2008). Videre finner Stokke (2016) at lønnsgapet øker raskt i løpet av de første 10-15 årene i arbeidslivet før det stabiliserer seg. Under halvparten av lønnsgapet kan forklares ved observerte variabler, mens det resterende uforklarte lønnsgapet forklares gjennom mindre avkastning på arbeidslivskarakteristika og spesielt akkumulert arbeidserfaring for kvinner sammenlignet med menn. Kunze (2005) finner motstridene resultater for Tyskland. Ved å anvende paneldata bestående av unge arbeidstakere i traineestillinger fra gamle Vest-Tyskland, konkluderer artikkelen at det er et vedvarende lønnsgap mellom kvinner og menn de første 10 årene, men at det er avtagende med antall år med arbeidserfaring.

Artikkelen til Manning & Swaffield (2008) og Stokke (2016) er vår motivasjon for å kontrollere ulike sysselsatte arbeidsgrupper i analysen i kapittel 5.

³Periode med høy jobbmobilitet. Denne er ofte høyere hos menn enn hos kvinner.

⁴like observerte og uobserverte karakteristika.

2.2 Regionale forskjeller i lønnsgapet

For å forklare lønnsforskjeller mellom kvinner og menn har tidligere litteratur sett på det generelle lønnsgapet i de ulike landene. Huertas et al. (2017) utvider litteraturen og ser på de regionale lønnsforskjellene mellom kvinner og menn. Artikkelen studerer mikrodata for arbeidsgiver og arbeidstaker statistikk for årene 2002, 2006 og 2010⁵ i 17 ulike regioner i Spania. Huertas et al. (2017) konkluderer at det eksisterer regional heterogenitet i lønnsgapet mellom regionene og at lønnsgapet er relatert til personlige egenskaper, men også fertilitetsraten hos kvinner. Fremdeles er det en betydelig del av de observerte lønnsforskjellene som fortsatt ikke kan forklares ved observerbare data. Hirsch et al. (2013) anvender administrativt datasett fra områder vest i Tyskland for å kunne forklare regionale lønnsforskjeller mellom kvinner og menn i Tyskland. Det blir presentert fire ulike hypoteser:

1. De forventer at arbeidere tjener mer i urbane områder enn i rurale områder, fordi det er flere arbeidere i disse områdene og høyere konkurranse blant arbeiderne.
2. Kvinnelige arbeidere har lavere lønn, alt annet likt.
3. Siden det er forventet at reisekostnader er blitt mindre betydningsfulle, forventes det at lønnsgapet over tid skal reduseres.
4. Det er forventet at lønnsforskjellene er mindre i urbane områder enn i rurale områder.

Resultatene ble som følger:

1. Hypotese 1 beholdes: Gjennomsnittslønnen er høyere for begge kjønn i urbane områder.
2. Hypotese 2 beholdes. Unge jenter som jobber fulltid tjener mindre enn unge menn med like egenskaper. Hypotesen beholdes når man kontrollerer for erfaring, ansienitet og utdanning. Dette gjelder både for arbeidere i rurale og urbane områder.
3. Hypotese 3 beholdes: Det har med årene blitt færre mennesker i rurale områder, som igjen gir lavere reisekostnader inn til byene. Dette har vært med på å redusere

⁵Encuesta de Estructura Salarial - lønnsstrukturundersøkelse.

lønnsgapet mellom kvinner og menn.

4. Hypotese 4 beholdes: Det er større lønnsforskjeller mellom kvinner og menn på landsbygda enn i byer. Lønnsgapet er 10 %-poeng lavere i by enn i rurale områder, som er motstridende til det Huertas et al. (2017) konkluderer med.

2.3 Agglomerasjonseffekter

De La Roca & Puga (2017) behandler et rikt datasett fra Spania og estimerer en modell som baserer seg på å identifisere individer som flytter mellom byer av ulik størrelse. Det argumenteres for at den statiske elastisiteten til lønnsgapet mellom kvinner og menn øker i størrelsen på byene tilsvarende 0.0455, ved MKM estimering. Ved å kontrollere for faste effekter hos arbeidere, konkluderes det med at elastisiteten til lønnsgapet reduseres til 0.0241 og at dette korrigerer for deler av heterogeniteten i datasettet.

De La Roca & Puga (2017) trekker frem at ved å bytte arbeidsplass fra rurale områder til en storby, gir en umiddelbar statisk lønnspremium. Ved å binde arbeidserfaring og bosted sammen undersøker De La Roca & Puga (2017) for hvordan urbanisering påvirker læringseffekter og det dynamiske lønnsgapet mellom kvinner og menn. Det blir også argumentert for at arbeidere i større byer akkumulerer mer verdifull arbeidserfaring, som i størst grad gjelder for personer med høyere utdanning.

Carlsen et al. (2016) finner tilsvarende trender som De La Roca & Puga (2017), at personer med høy utdanning i større grad tenderer til å bo i byer og at det er statiske lønnsforskjeller mellom rurale og urbane områder i Norge. Videre indikerer Carlsen et al. (2016) at det også er dynamiske lønnsforskjeller mellom by og rurale områder og at denne effekten øker med utdanningsgrad. Høyt utdannede personer i høytlønnede sektorer har størst fordel av å bo i tettbebygde områder. Konsentrasjonen av høyt utdannede i byer forklarer en stor del av det observerte lønnsgapet mellom byer og rurale områder. Carlsen et al. (2016) finner at det statiske og dynamiske lønnsgapet for Oslo⁶ er samlet på 17%, hvor 7% forklares som et statisk lønnsgap, mens de resterende 10% skyldes effekten av erfaring.⁷

2.4 Lønnsgapet mellom privat og offentlig sektor

Tidligere litteratur konkluderer med at mye av lønnsgapet mellom kvinner og menn er forklart ved kvinners yrkesvalg. Kvinner er i større grad representert i offentlig sektor enn

⁶Hovedstaden i Norge.

⁷Gjennomsnittlig erfaring er 8 år, og er beregnet utifra dette.

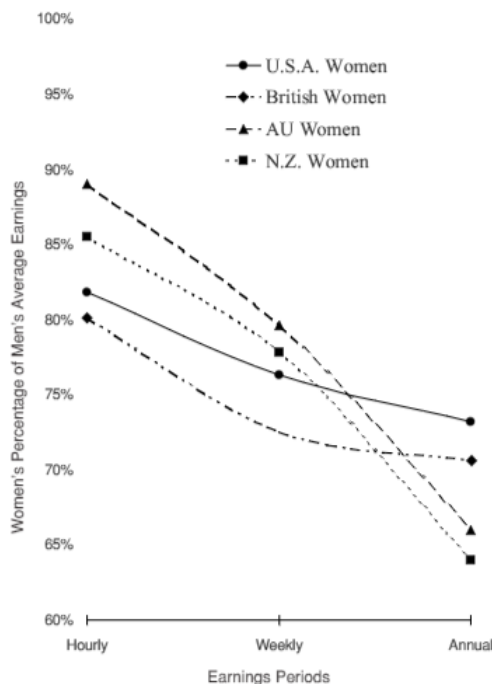
menn, og dette kan være en medvirkende årsak til forskjeller i lønnsgapet. Denne problemstillingen studerer Rattsø & Stokke (2018) ved å anvende seg av 89 arbeidsmarkedsregioner. Det blir benyttet fixed effects⁸ for å kontrollere for uobserverte individ karakteristika hos arbeiderene. Rattsø & Stokke (2018) konkluderer med at lønnsgapet mellom privat og offentlig sektor er 15% for personer med høyere utdanning. Det blir argumentert for at det er en forskjell i sektorene med hensyn på hvilken utdanningsgrad individene har. Det er 77 % av de ansatte i privat sektor som har lavere utdanning mot 45 % i offentlig sektor. Resultatene viser også at den statiske lønnsforskjellen mellom privat og offentlig sektor er 3,9%-poeng høyere i byer enn i rurale områder. Det blir også tatt hensyn til at avkastningen av erfaring kan variere mellom offentlig- og privat sektor. Når dette kontrolleres for øker lønnsgapet mellom privat og offentlig sektor utenfor byer fra 2.5% til 3.8%. I byer er akkumulert arbeidserfaring mer verdifull, som vil si at det aggregerte lønnsgapet mellom privat og offentlig sektor øker fra 6.4 % til 12.5 %. Dette er vår motivasjon for å inkludere en variabel for offentlig sektor i analysen i kapittel 5.

⁸Fixed effects er en estimeringsmetode hvor man kontrollerer for individ faste effekter som ikke varierer over tid Wooldridge (2015).

2.5 Ulike mål på lønnsforskjeller

Lønnsgapet mellom kvinner og menn kan bli predikert ulikt ut ifra hvilket mål man anvender på lønn. Lips (2003) tar for seg problemer ved rapportering av inntektsforskjeller mellom kvinner og menn. Det argumenteres at resultater blir ulike avhengig av hvilket mål på inntekt man bruker. Lønnsgapet mellom kvinner og menn er forskjellig avhengig av om man ser på timelønn, ukelønn eller årslønn. Menn har større sannsynlighet for å få ”fordeler og premiering”⁹ enn det kvinner har (Rubery et al. 2005), og ifølge Lips (2003) blir ekstra inntekter ekskludert når man beregner for kortere lønnsintervaller som timelønn.

Figur 1 viser tydelig forskjellen mellom timelønn, ukelønn og årslønn målt som kvinners inntekt relativt til menn.¹⁰



Figur 1: Måleenhet
(Lips 2003)

I alle de fire landene er det en lik utvikling i rapportert lønnsforskjell, men ulik helning. I USA og Storbritannia er det en avtagende konveks kurve, mens det i Australia og New Zealand er avtagende konkav kurve. Fellestrekket er at lønnsgapet øker når man bruker årslønn i stedet for timelønn eller ukelønn.

⁹Lønnsstlegg, bonus og andre inntektsrelaterte ordninger.

¹⁰Artikkelen baserer seg på inntektsdata fra U.S. Bureau of Labor Statistics, Statistics New Zealand, Britain's Equal Opportunities Commission, and the Australian Bureau of Statistics.

3 Data og deskriptiv statistikk

3.1 Introduksjon

I dette kapitlet presenteres datamaterialet og regionstrukturen vi benytter for å gjennomføre den økonometriske analysen, samt en detaljert beskrivelse av de ulike variablene som er inkludert i datasettet vi har produsert. I datasettet mangler informasjon for 2015 for variabelen ”M/K Offentlig sektor” og tallene for 2016 er ikke sammenlignbare med tidligere år (2008-2014). Dette kan forårsake systematiske målefeil som det tas hensyn til ved å kun se på tidsperioden 2008-2014 i analysen i kapittel 5.

3.2 160 Bo-og Arbeidsmarkedsregioner

Vi har valgt å benytte 160 Bo- og Arbeidsmarkedsregioner¹¹ (BA-regioner), definert av Norsk institutt for by- og regionforskning. Inndelingen i BA-regioner skal være et funksjonelt uttrykk for den geografiske tilknyttingen mellom arbeidsliv og hushold, og skal fungere som en geografisk byggestein i ulike analyser. NIBR definerer en region som, *et avgrenset området med en eller annen form for indre enhet som skiller det fra de geografiske områdene rundt* (Gundersen & Juvkam 2013). Inndelingen tar hensyn til grad av arbeidspendling, arbeidsmarkedets størrelse, tettstedets størrelse, tilgjengelige tjenester og offentlige institusjoner. Inndelingsarbeidet er basert på ulike kriterier, som i hovedsak er ulike kombinasjoner av tre kjennetegn med kommunen: pendling mellom de ulike kommunene, hvordan kommunen er plassert i senterstrukturen¹² og reisetid mellom kommunesentra.

¹¹By- og regionforskningsinstituttet NIBR, som er underlagt kunnskapsdepartementet

¹²Kommuner med bosetning i et tettstedsområde tilhører samme senter

3.2.1 Metode anvendt ved inndeling av BA-regioner

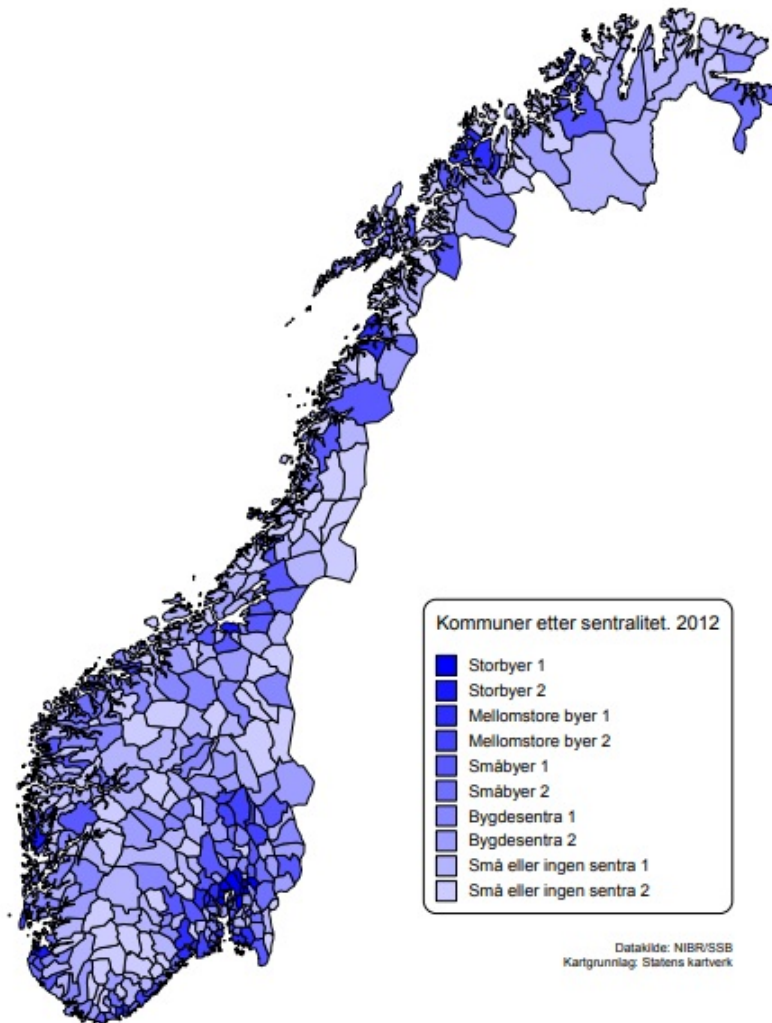
Tettsteder¹³ skal være byggesteinen for inndelingen av BA-regionene. Drammen tettsted vil eksempelvis bestå av Drammen, Lier og Nedre Eiker kommune. Et høyt pendlingsnivå mellom ulike tettsteder vil kunne medføre at tettsteder tilhører samme region. Det kan også være flere tettsteder i samme kommune og da er kommunen tilknyttet til det største tettstedet kommunen har innbyggere i. Følgende variabler er blitt benyttet for å utarbeide 160 BA-regioner:

- Antall personer bosatt i de ulike tettstedene.
- Antall personer bosatt i tettstedet i den enkelte kommune.
- Antall arbeidsplasser i kommunen.
- Antall personer som pendler inn til kommunen.
- Hvilke tjenester som er tilgjengelig i kommunen.
- Hvilke offentlige institusjoner som finnes i de ulike kommunene.

I NIBR-rapporten er byer grovinn delt i fem kategorier etter folketall. By med innbyggertall større en 500.000 er definert i hovedstadsområdet. Videre for 2, 3, 4 og 5 er innbyggertallene henholdsvis 150-500.000, 35-150.000, 19-43.000 og 5-25.000 innbyggere.

Figur 2 illustrerer fininndelingen av kommuner etter kjennemerker for tettstedsstørrelse og funksjonalitet utført av NIBR.

¹³Av SSB er tettsted definert som *en hussamling skal registreres som et tettsted dersom det bor minst 200 personer der*



Figur 2: Fininndeling av kommunene etter kjennemerker for tettstedsstørrelse og funksjonalitet (Gundersen & Juvkam 2013). Denne oppgaven har brukt en annen by-definisjon en det NIBR har gjort, dette blir beskrevet senere i dette kapitlet.

3.3 Datamaterialet

Datamaterialet vi har hentet fra SSB¹⁴ er på kommunenivå i perioden 2008-2014. Vi har gjennom et grundig arbeid omgjort datamaterialet til 160 bo-og arbeidsmarkedsregioner, for å kunne ta hensyn til personer som pendler mellom kommuner. Ved å befolkningsvekter variablene i datasettet tas det hensyn til kommunens folketall ved sammenslåing til bo og arbeidsmarkedsregioner. De ulike variablene brukt i denne oppgaven er beskrevet i de kommende delkapitlene.

¹⁴SSB er en faglig uavhengig institusjon ansvarlig for å samle inn, produsere og publisere offisiell statistikk relatert til økonomi, befolkning og samfunnet på nasjonalt, regionalt og lokalt nivå.

3.4 Oversikt over variabler og deskriptiv statistikk

3.4.1 By

I datasettet er ti av regionene definert som by-regioner med befolkning ≥ 100.000 i 2014. Vi har valgt denne grensen med bakgrunn i tidligere litteratur (Rattsø & Stokke 2018). De største byene som ikke har blitt definert som by-region i datasettet er Tromsø, Ålesund og Molde. Disse inngår i gruppen med folketall < 100.000 . Omlag 60% av befolkningen i Norge bor i en by-region og de resterende 40% bor i rurale regioner, etter vår definisjon.

I denne oppgaven ønsker vi spesielt å se nærmere på inntektsforskjellene mellom kvinner og menn i by-og rurale regioner. Derfor er det vesentlig å inkludere en dummyvariabel som tar verdien = 1 for by-regioner og = 0 for rurale regioner.

Regionene definert som by-regioner i datasettet er:

Oslo = 31 kommuner som inngår i regionen.

Bergen = 13 kommuner som inngår i regionen.

Stavanger = 14 kommuner som inngår i regionen.

Trondheim = 10 kommuner som inngår i regionen.

Drammen = 8 kommuner som inngår i regionen.

Fred/Sarp = 5 kommuner som inngår i regionen.

Kristiansand = 7 kommuner som inngår i regionen.

Grenland = 7 kommuner som inngår i regionen.

Tønsberg = 9 kommuner som inngår i regionen.

Haugesund = 7 kommuner som inngår i regionen.

3.4.2 Inntektsspremium

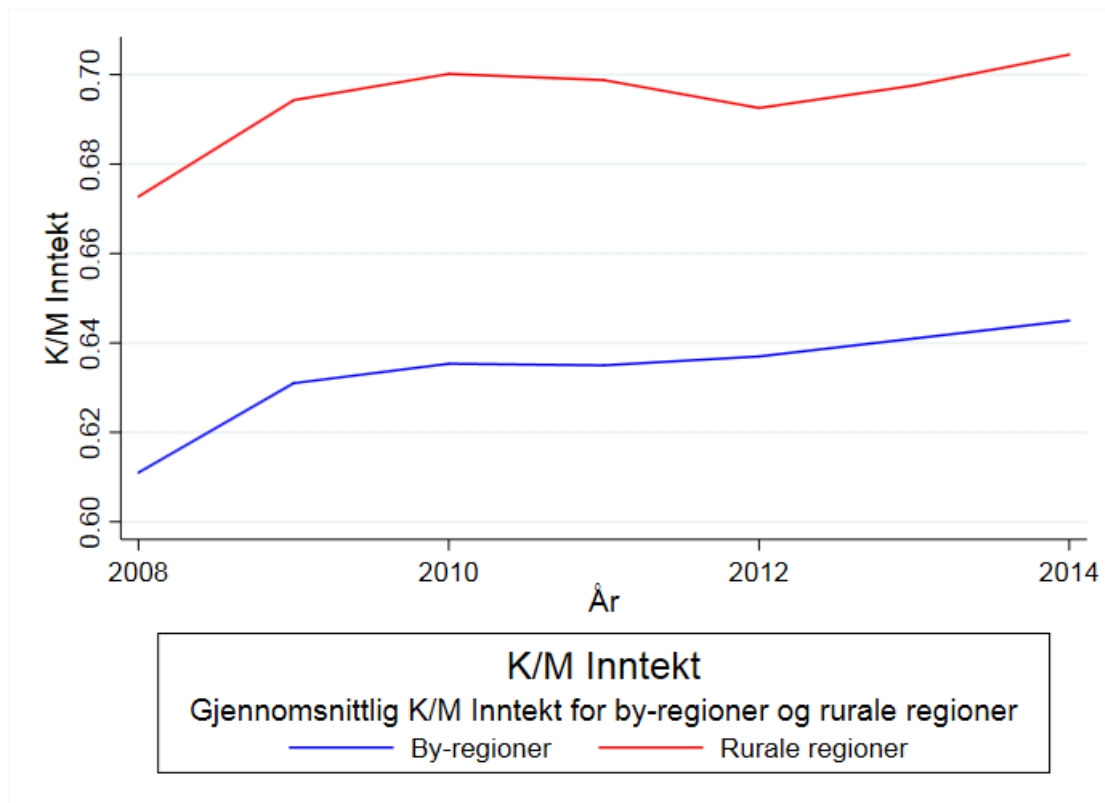
”K/M Inntekt”¹⁵ er den avhengige variabelen i analysen og er definert som: *Årlig gjennomsnittlig bruttoinntekt for kvinner dividert på årlig gjennomsnittlig bruttoinntekt for menn.*

$$\text{K/M Inntekt} = \frac{\text{Årlig gjennomsnittlig bruttoinntekt for kvinner}}{\text{Årlig gjennomsnittlig bruttoinntekt for menn}} \quad (3.1)$$

Gjennomsnittsverdien til ”K/M Inntekt” er 0.69 som vil si at kvinner i snitt tjener 69 % av det en mann tjener. Som forklart i innledningen til denne oppgaven konkluderer (Kristoffersen 2017) at kvinner i Norge tjener 86% av menns lønn. Resultatene er dermed ulike som skyldes at Kristoffersen (2017) ser på den gjennomsnittlige månedslønnen mens vi ser på årlig bruttoinntekt.

Tabell 1 illustrerer deskriptiv statistikk for de ulike variablene. Minimumsverdien til ”K/M Inntekt” er relativt lav og dette skyldes at menn i regionen Kvæningen hadde svært høy gjennomsnittsinntekt i 2011 etter salget av Jøkelfjord Laks i 2010. Verdien til ”K/M Inntekt” i samme region, falt følgelig til 0.26 i 2011. Kvæningen har, med unntak av 2011 en verdi for ”K/M Inntekt” på 0.71 til 0.79. Fra samme tabell kan vi også observere at maksimumsverdien for ”K/M Inntekt” er 1.00. Dette kommer av at det i regionen Karasjok var tilnærmet lik inntekt mellom kvinner og menn i 2009. Det er gjennomgående små inntektsforskjeller i denne regionen. Fra tabell 2 kan vi se at gjennomsnittsverdien til ”K/M Inntekt” er lavere i by-regioner enn i rurale. Dette kan indikere at det er et større inntektsgap mellom kvinner og menn i by-regioner enn i rurale-regioner, noe som er tydeligere forklart i kapittel 5.

¹⁵Inntekt omfatter lønnsinntekter, næringsinntekter, pensjoner og kapitalinntekter (Statistikkbanken 2018b).



Figur 3: K/M Inntekt

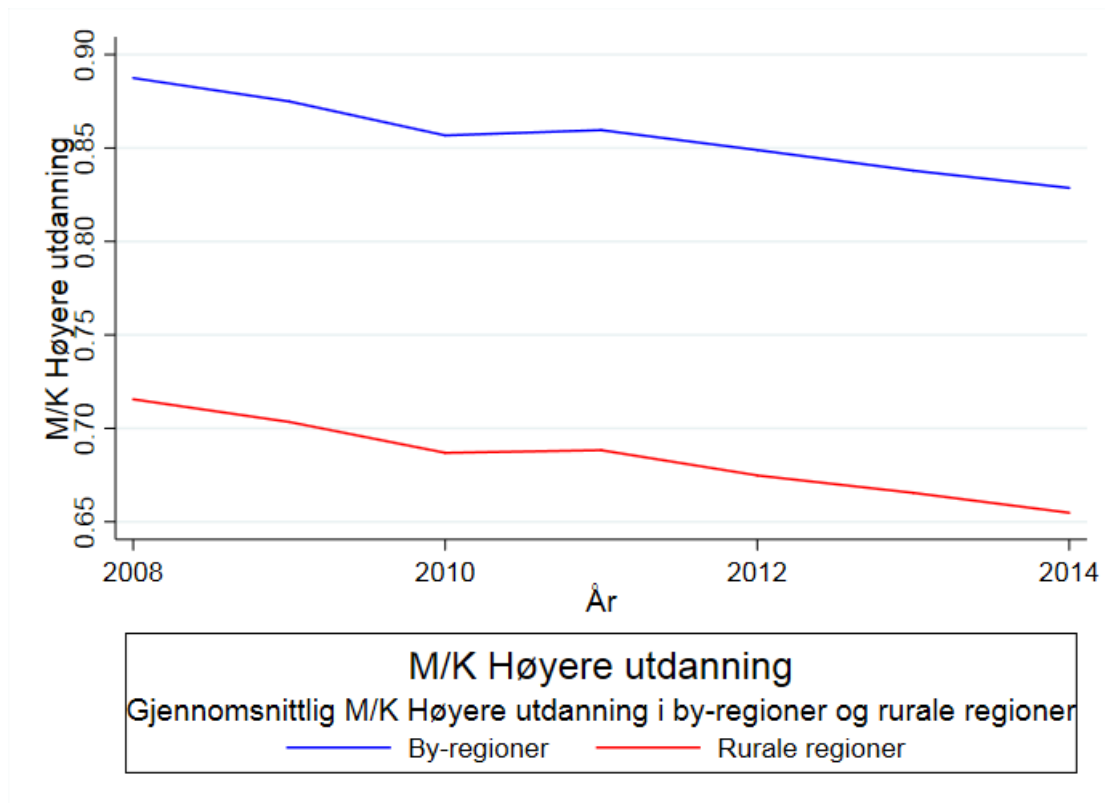
Deskriptivt gjennomsnittlig inntektsgap mellom by-regioner og rurale regioner i perioden 2008-2014

3.4.3 Høyere utdanning

Variabelen "M/K Høyere utdanning" er definert som: *Antall menn med høyere utdanning relativt til befolkningen, dividert på antall kvinner med høyere utdanning relativt til befolkningen.*

$$\text{M/K Høyere utdanning} = \frac{\text{Andel menn med høyere utdanning}}{\text{Andel kvinner med høyere utdanning}} \quad (3.2)$$

Man kan lese fra tabell 1 at gjennomsnittet til "M/K Høyere utdanning" er på 0.69 som vil si at andel menn med høyere utdanning utgjør 69 % av andelen kvinner med høyere utdanning. Effekten av utdanningsnivå er beskrevet i tidligere litteratur og kan ha stor påvirkning på lønn. Minimum- og maksimumsverdien til variabelen varierer mellom 0.38 - 1.1. Den høye verdien skyldes Kongsberg som er den eneste regionen i datasettet som har verdi høyere enn én gjennom hele perioden. Det vil si at andelen menn med høyere utdanning er større enn andelen kvinner med høyere utdanning i denne regionen.



Figur 4: M/K Høyere utdanning
 Deskriptiv gjennomsnittlig forskjell i utdanningsvariabelen mellom by-regioner og rurale regioner

Fra tabell 2 kan vi se at gjennomsnittsverdien til "M/K Høyere utdanning" i by-og rurale regioner er henholdsvis 0.85 og 0.68. Dette betyr at det bor en større andel menn med høyere utdanning relativt til kvinner i by-regionene, sammenlignet med de rurale regionene. En av grunnen kan være at bedriftene som etterspør arbeidskraft med høyere utdanning tenderer å tilhøre byene.

Vi ser en tydelig trend i figur 4 som illustrerer samlet gjennomsnitt for de ti by-regionene sammen med gjennomsnittet for de 150 rurale regionene. Trenden indikerer at kvinner fra 2008 til 2014 i større grad enn menn tar høyere utdanning. Avtakende kurver indikerer at variabelen som er definert i 3.2 reduseres når nevneren, andelen kvinner med høyere utdanning, øker relativt til andelen menn med høyere utdanning. Det er så godt som lik utvikling mellom by-regionene og de rurale regionene. Dette samsvarer med resultatene til Tuhus (2010).

3.4.4 Offentlig sektor

”M/K Offentlig sektor”¹⁶ er definert som: *antall menn sysselsatt i offentlig sektor relativt til alle sysselsatte menn, dividert på antall kvinner sysselsatt i offentlig sektor relativt til alle sysselsatte kvinner.*

$$\text{M/K Offentlig sektor} = \frac{\text{Andel menn som jobber i offentlig sektor}}{\text{Andel kvinner som jobber i offentlig sektor}} \quad (3.3)$$

Variabelens verdi er mellom 0 og 1 som skyldes at kvinner i større grad enn menn jobber i offentlig sektor (Tuhus 2010). Fra tabell 1 har variabelen en maksimumsverdi nært én som skyldes regionen Måsely, der kjønnsbalansen i offentlig sektor er tilnærmet lik gjennom hele perioden. Størst ulikhet finner vi i regionene Åseral, Moskenes og Beiarn, hvor sistnevnte har lavest verdi gjennom perioden 2008-2014. Fra tabell 1 kan vi se at gjennomsnittsverdien til ”M/K offentlig sektor” er 0.32, som vil si at andelen menn relativt til kvinner i offentlig sektor utgjør 32%. Videre forteller tabell 2 at det er en høyere andel menn relativt til kvinner som jobber i offentlig sektor i by-regionene sammenlignet med de rurale regionene. Rattsø & Stokke (2014) konkluderer med at det statiske lønnsgapet mellom offentlig og privat sektor er 3,8 %-poeng. Det er derfor rimelig å tro at noe av forklaringen på inntektsgapet mellom kvinner og menn skyldes lønnsgapet mellom privat og offentlig sektor.

3.4.5 Deltid

”M/K Deltid”¹⁷ er definert som: *antall menn som jobber i deltidsstilling relativt til alle sysselsatte menn, dividert på antall kvinner som jobber i deltidsstilling relativt til alle sysselsatte kvinner*¹⁸

¹⁶Det mangler 2 observasjoner for denne variabelen - se tabell 9 i Vedlegg.

¹⁷Det mangler 22 observasjoner for denne variabelen - se tabell 9 i Vedlegg.

¹⁸Sysselsatte menn og kvinner er i aldersgruppen 20-66 år.

$$\text{M/K Deltid} = \frac{\text{Andel menn som jobber deltid}}{\text{Andel kvinner som jobber deltid}} \quad (3.4)$$

Gjennomsnittlig andel kvinner ansatt i deltidsstillinger i datasettet er 42 %, mot 15 % for menn. Dette kan være en påvirkende faktor til inntektsgapet mellom kvinner og menn, som er grunnen til at vi inkluderer variabelen i analysen. Tabell 1 illustrerer at gjennomsnittet til variabelen er 0.375, som tilsvarer at andelen menn som jobber deltid utgjør 37,5 % av andelen kvinner som jobber deltid. I regionene Alta og Berlevåg er det flere menn relativt til kvinner som jobber deltid sammenlignet med gjennomsnittet. Regionene ligger gjennomgående opp mot maksimumsverdien. Stord, Sauda, Vanylven og Sandøy har lave verdier gjennom hele perioden, som vil si at det er færre menn relativt til kvinner som jobber deltid i disse regionene. Fra tabell 2 kan vi se at gjennomsnittsverdien til ”M/K deltid” i by-og rurale regioner er relativt like.

3.4.6 Leder

”K/M Leder”¹⁹ er definert som: *Antall kvinnelige ledere relativt til andel mannlige ledere.*

$$\text{K/M Leder} = \frac{\text{Antall kvinnelige ledere}}{\text{Antall mannlige ledere}} \quad (3.5)$$

Gjennomsnittsverdien til variabelen har verdi lik 0.6, som vil si at menn i større grad enn kvinner er representert i lederstillinger. Fra tabell 1 observerer vi at maksimumsverdien til variabelen er 1.75 som skyldes at det i noen regioner er flere kvinner i lederstillinger relativt til menn. I perioden 2008-2014 har eksempelvis regionen Leka flere kvinner i lederstillinger ²⁰ enn menn. Ved å inkludere variabelen ”K/M Leder” kontrollerer vi for forholdet mellom kvinner og menn i lederstillinger, da personer i lederstillinger tradisjonelt sett har høyere inntekt. Fra tabell 2 kan vi se at gjennomsnittsverdien til ”K/M Leder” i by-og rurale regioner er henholdsvis 0.49 og 0.61. Dette indikerer at det er en større andel kvinner relativt til menn som er ledere i by-regioner enn i rurale-regioner.

¹⁹Det mangler 8 observasjoner for denne variabelen - se tabell 9 i Vedlegg.

²⁰Ny yrkesstandard (STYRK-08) for årene 2011 og 2012. Indikatoren er allikevel sammenlignbar over tid. (SSB 2018)

3.4.7 Sysselsatte aldersgrupper

Aldersgruppene "Andelen 20-24 år", "Andelen 20-39 år" og "Andelen 40-66 år" er definert som: *Antall sysselsatte personer i den aktuelle aldersgruppen dividert på hele arbeidsstyrken*.²¹

$$20 - 24 \text{ år} = \frac{\text{Antall i aldersgruppen 20-24 år som er i arbeidsstyrken}}{\text{Alle i arbeidsstyrken}} \quad (3.6)$$

Fra tabell 1 kan vi se at en større del av de sysselsatte i datasettet er i aldersgruppen 40-66 år. Korrelasjonen mellom variablene "Andelen 20-39 år" og "Andelen 40-66 år" er lik 1, som betyr at vi ikke kan inkludere disse i samme regresjon grunnet brudd på forutsetning 3 ved estimering med MKM ²². Tabell 2 illustrerer at det er en større andel sysselsatte i aldersgruppen 20-39 år i by-regionene sammenlignet med de rurale. Som følge av dette er det en større andel i aldersgruppen 40-66 i de rurale regionene. Andelen sysselsatte i aldersgruppen 20-24 år er relativt lik mellom begge regionstypene. Variablene er inkludert for å ta høyde for at inntektsgapet mellom menn og kvinner varierer i løpet av karrieren.

3.4.8 Barn

Variabelen er definert som: *Antall barn i aldersgruppen 0-9 år dividert på hele befolkningen*.

$$\text{Andelen 0-9 år} = \frac{\text{Antall i befolkningen som er i aldersgruppen 0-9 år}}{\text{Hele befolkningen}} \quad (3.7)$$

"Andelen 0-9 år" er inkludert som en indikator på fertilitet. Gjennomsnittsverdien til variabelen er 11,1 %, og har tilhørende minimum -og maksimumsverdi tilsvarende 5-16 % av

²¹Arbeidsstyrken er definert som sysselsatte i aldersgruppen 20-66 år.

²²Minste kvadraters metode.

befolkningen. Om andel av befolkningen i denne aldersgruppen er høy, vil det kunne tenkes at kvinner i større grad enn menn er hjemmeværende. Dette kan resultere i at kvinner i større grad blir mindre delaktig i arbeidslivet. Fra tabell 2 kan vi observere at det er 1.5%-poeng høyere andel av befolkningen som er i aldersgruppen 0-9 år i by-regionene, sammenlignet med de 150 øvrige regionene.

3.5 Næringer

Jordbruk, skogbruk og fiske er definert som:²³ *Antall sysselsatte som jobber i en av de aktuelle næringene, relativt til hele arbeidsstyrken i en region.*

$$\text{Jordbruk, skogbruk og fiske} = \frac{\text{Antall som jobber i jordbruk, skogbruk og fiske}}{\text{arbeidsstyrken i regionen}} \quad (3.8)$$

Tilsvarende som 3.8 er variabelen "Bergverksdrift og utvinning" definert som:²⁴ *Antall sysselsatte som jobber i en av de aktuelle næringene, relativt til hele arbeidsstyrken i en region*

$$\text{Bergverksdrift og utvinning} = \frac{\text{Antall som jobber i bergverksdrift og utvinning}}{\text{arbeidsstyrken i regionen}} \quad (3.9)$$

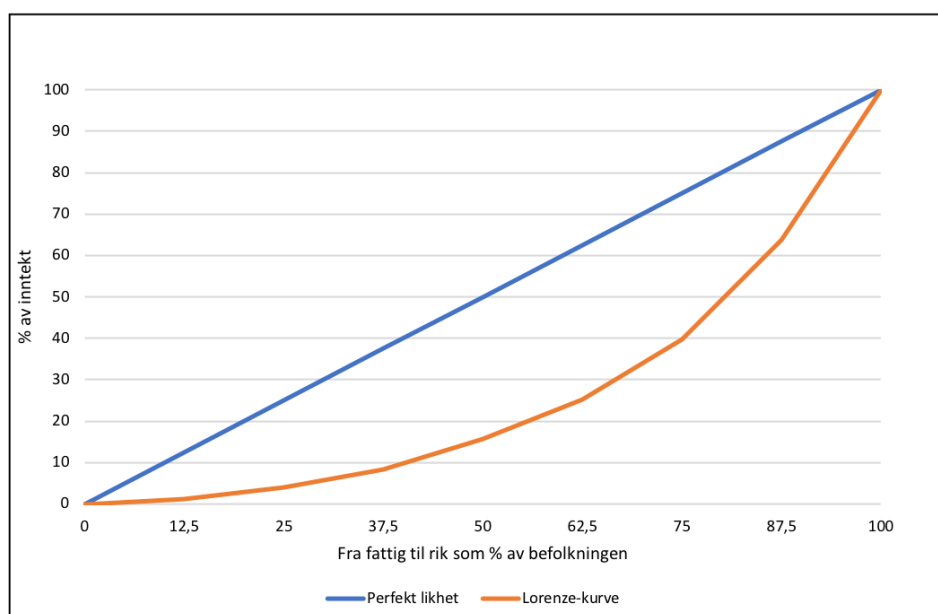
Variablene er inkludert ettersom næringene ofte er mannsdominerte og kjent for å ha relativt høye lønninger. Gjennomsnittet til "Jordbruk, skogbruk og fiske" er høyere i rurale regioner sammenlignet med by-regioner, som kan skyldes at tilgangen til nødvendige ressurser er større i disse områdene. Dette er ikke tilfellet for "Bergverksdrift og utvinning" der det i snitt er flere personer som jobber i denne næringen i by-regioner enn i rurale. En grunn til dette kan være at bedrifter relatert oljevirkosomhet er stasjonert i de store byene som Stavanger, Bergen og Oslo.

²³Det er tatt hensyn til personvern i SSB sin statistikk "Kommuner som kun har 1-2 arbeidere i en sektoren er endret til 0 eller 3 for å ivareta personvernet".(Statistikkbanken 2017) som medfører at det er bevisste målefeil i variabelen.

²⁴Se fotnote 23.

3.5.1 Gini-Koeffisient

Corrado Gini (1912) presenterte i boken *Variabilitá e Mutabilitá* gini-indeksen og brukes til å beregne inntekts og formue ulikheter i forskjellige land. Koeffisienten angis som en tallverdi mellom 0 og 1, der 1 indikerer at én person eier all inntekt og formue, og 0 indikerer at alle innbyggere har lik inntekt og formue. ”Gini” er befolkningsvektet likt som de øvrige variablene.



Figur 5: Gini

Fra tabell 1 kan vi se at den gjennomsnittlige verdien til ”Gini” er 0.20. Videre er minimum- og maksimumsverdien henholdsvis 0.098 og 0.65. Minimumsverdien kommer av at regionen Ørland har gjennomgående lav gini-koeffisient. Maksimumsverdien skyldes regionen Kvæningen etter salg av Jøkelfjord Laks i 2010, som er tydeligere beskrevet i delkapittel 3.4.2.

Tabellen 5 illustrerer gini-indeksen sammen med Lorenze-kurven. På den horisontale X-aksen avsettes befolkningen proporsjonalt bortover, der de fattigste innbyggerne til venstre, og de rikeste til høyre. På Y-aksen avsettes inntekten til befolkningen. Gini-indeksen er differansen mellom den blå og oransje linje, dividert på hele trekanten under den blå linjen. Desto mer den oransje linjen buer utover, desto skjevare er inntekten. Gini-koeffisient representert med 45 graders vertikal linje, indikerer full likhet i befolkningen. (Ultsch & Lötsch 2017).

3.6 Hypoteser

Som en avslutningsdel på kapitel 3 ønsker vi å presentere ulike hypoteser. Vi baserer de ulike hypotesene på hva som er konkludert med i tidligere litteratur og indikasjoner fra deskriptiv statistikk.

1. Øker variabelen "Offentlig sektor" så vil dette, alt annet likt, redusere inntekts-gapet mellom kvinner og menn.
2. En økning i andelen utdannede menn relativt til kvinner, alt annet likt, vil øke inntekts-gapet mellom kvinner og menn.
3. En økning i kvinnelige ledere relativt til menn, alt annet likt, vil redusere inntekts-gapet.
4. Det er større inntekts-gap mellom menn og kvinner i by-region sammenlignet med rurale regioner, alt annet likt.
5. Inntektsforskjellene er mindre ved et tidlig stadige av arbeidslivet.
6. En økning i variabelen "0-9 år" vil, alt annet likt, øke inntektsforskjellene mellom kvinner og menn.

Tabell 1: 160 regioner

Variabler	Obs.	Mean	Std. Dev.	Min	Max
By	1,200	.04861	.2151	0	1
K/M Inntekt	1,120	.6905	.0741	.26	1.00
M/K Høyere utdanning	1,120	.6904	.1129	.3821	1.10
M/K Offentlig sektor	1,118	.3211	.1025	.1112	.7619
K/M Leder	1,112	.6004	.1882	.0416	1.74
M/K Deltid	1,098	.3751	.0922	.1341	.8323
Andelen 20-24 år	1,200	9.42	1.81	2.05	18.26
Andelen 20-39 år	1,200	38.23	6.19	10.26	58.62
Andelen 40-66 år	1,200	61.76	6.19	41.37	89.73
Andelen 0-9 år	1,120	11.10	1.50	5.55	15.34
Jordbruk, skogbruk og fiske	1,200	3.57	4.37	0	50
Bergverksdrift og utvinning	1,200	.6994	1.72	0	14.28
Gini	640	0.2036	.0398	.098	.65

Oppsummering av variablene i alle regionene

I Tabell 1 presenteres deskriptiv statistikk for de ulike variablene. Antall observasjoner for de ulike variablene endrer seg noe grunnet manglende observasjoner for de forskjellige kommunene. Variabelen "Gini" har lavere antall observasjoner enn resterende fordi vi kun ser på årene (2011-2014).

Tabell 2: By-regioner og rurale regioner

Variabler	By-regioner			Rurale-regioner		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
By	70	1	0	1050	0	0
K/M Inntekt	70	.6336	.0369	1050	.6944	.0744
M/K Høyere utdanning	70	.8505	.0563	1050	.6799	.1077
M/K Offentlig sektor	69	.3567	.0672	1049	.3188	.1040
K/M Leder	68	.4923	.0656	1044	.6110	.1887
M/K Deltid	70	.3689	.0618	1028	.3755	.0939
Andelen 20-24 år	70	9.79	.9244	1050	9.40	1.85
Andelen 20-39 år	70	45.25	4.08	1,050	37.76	6.031
Andelen 40-66 år	70	54.74	4.08	1,050	62.23	6.03
Andelen 0-9 år	69	12.69	.8431	1050	11.00	1.47
Jordbruk, skogbruk og fiske	70	.7331	.4233	1050	3.76	4.45
Bergverksdrift og utvinning	70	2.14	3.72	1050	.6028	1.45
Gini	40	.1895	.0984	600	.2045	.0323

Oppsummering av variabler for by-regionene og de resterende regionene som ikke er definert som en by-region.

Tabell 2 er delt opp i to seksjoner, by-regioner og rurale regioner. Dette er grunnen til at antall observasjoner er lavere for by-region, enn for rurale regioner. Følgelig er ekstremverdiene for seksjonene forskjellige. Å kunne tyde og forklare gjennomsnitt og ekstremverdiene til de ulike variablene er viktig for å kunne utføre en grundig analyse av inntektsgapet mellom kvinner og menn.

3.6.1 Variabler som er inkludert i datasettet

K/M Inntekt – Kvinners inntekt relativt til menns inntekt.

M/K Høyere utdanning – Andel menn med høyere utdanning relativt til kvinner.

By – Dummy variabel som er lik 1 hvis befolkningen er større eller lik 100.000 i 2014.

M/K Offentlig sektor – Andel kvinner i off. sektor relativt til menn.

K/M Leder – Andel kvinner i lederstilling relativt til menn.

M/K Deltid – Andel kvinner som jobber deltid relativt til menn.

Andelen 20-24 år – Andel av arbeidsstyrken som er i aldersgruppen 20-24 år

Andelen 20-39 år – Andel av arbeidsstyrken som er i aldersgruppen 20-39 år.

Andelen 39-66 år – Andel av arbeidsstyrken som er i aldersgruppen 39-66 år.

Andelen 0-9 år – Andel av befolkningen som er i aldersgruppen 0-9 år.

Jordbruk, skogbruk og fiske – Andel av arbeidsstyrken som jobber innen

jordbruk, skogbruk eller fiske

Bergverksdrift og utvinning. – Andel av arbeidsstyrken som jobber innen

bergverksdrift eller utvinning.

Gini – Koeffisient mellom 0 og 1 som forklarer inntektsforskjeller i ulike land.

4 Økonometrisk spesifikasjon og utfordringer

4.1 Introduksjon

Vi skal i dette kapitlet se nærmere på økonometrisk spesifikasjon og utfordringer som kan oppstå ved estimering av modellen. Problemene vi har tatt for oss i dette kapitlet er problemer vi særlig ser for oss kan forekomme ved bruk av paneldata.

4.2 Økonometriske utfordringer

Forholdet mellom en forklaringsvariabel og den avhengige variabelen har et kausalt forhold om en endring i forklaringsvariabelen, alt annet likt, forårsaker en endring i den avhengige variabelen (Sucarrat 2016). Den forventede verdien til den avhengige variabelen kan formuleres på følgende måte:

$$E(y|x, Z) \tag{4.1}$$

I ligning 4.1 er y den avhengige variabelen, x er forklaringsvariabelen og Z representerer et sett med ytterligere forklaringsvariabler.²⁵ Ved å analysere et datasett definert som paneldata ser man på flere observasjoner for samme enhet over tid. Dette medfører at man har to dimensjoner, enhet og tid, som i analysen er henholdsvis 160 bo- og arbeidsmarkedsregioner og antall år som analyseres (2008-2014). For å forklare hvordan vi estimerer modellen, kan man ta utgangspunkt i følgende ligning:

$$Y_{it} = b_0 + b_1 X_{it} + u_{it} \quad \begin{array}{l} i = 1, 2, \dots, n \\ t = 1, 2, \dots, T \end{array} \tag{4.2}$$

der:

Y_{it} - Den avhengige variabelen, som i oppgaven vil være K/M Inntekt.

²⁵Vi har i vår analyse inkludert flere forklaringsvariabler for å kontrollere for ulike forhold og se på utviklingen til modellen.

b_0 - Konstantledd.

X_{it} - er et sett med forklaringsvariabler med tilhørende sett av β_1 koeffisienter.²⁶

u_{it} - Et stokastisk feilledd.

Ved estimering av modellen anvendes minste kvadraters metode (MKM). Det kvadrerte avviket mellom den observerte og den estimerte, er variansen, som er grunnen til navnet minste kvadraters metode. Ved bruk av MKM må fem forutsetninger oppfylles for å få konsistente estimater. Disse forutsetningene er:

1. Modellen er lineær i sine parametre.
2. Tilfeldig utvalg av observasjoner.
3. Ingen perfekt multikollinearitet. Hvis en avhengig variabel har en eksakt lineær kombinasjon av de andre avhengige variablene, kan vi si at modellen er perfekt kollinert og MKM kan ikke anvendes.
4. Eksogene forklaringsvariabler, som vil si at ingen av forklaringsvariablene kan være konstante. $E(U_{it}|X_{it})=0$, som betyr at forklaringsvariablene ikke er korrelert med feilleddet.
5. Feilleddet er normalfordelt, $\text{Var}(U_{it}|x_1, x_2, \dots, x_k) = \sigma^2$

I ligning 4.2 beskriver fotskrift "i" hvilken region vi analyserer og fotskrift "t" beskriver en bestemt periode, "N" er antall tverrsnittsenheter og "T" er antall perioder. Ved balansert data har man opplysninger om alle enheter på alle tidspunkter og totalt antall observasjoner blir da $n * T$. For at vi skal kunne få korrekte og konsistente estimeringer ved bruk av forklaringsvariabler, må forutsetning 4 holde. Hvis denne ikke er oppfylt er det potensielt fire hovedproblemer som kan oppstå, disse er målefeil, utelatt variabel skjevhet, simultanitet eller feilspesifikasjon (Wooldridge 2015).

²⁶ $X_{it} = [X_{1it}, X_{2it}, \dots, X_{jit}]$ er et sett med forklaringsvariabler med dimensjon $(1 \times j)$, og $\beta_1 = \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1j} \end{bmatrix}$.

4.2.1 Målefeil

Målefeil deles inn i ”systematisk”- og ”tilfeldig” og oppstår når den observerte variabelen avviker fra sin faktiske verdi. Tilfeldige målefeil er generelt ikke et problem så lenge det er relativt få feil og at størrelsen er relativt lav. Systematisk målefeil i forklaringsvariablene kan forårsake skjevhet i estimatene i den avhengige variabelen (Wooldridge 2015). Som beskrevet i kapittel 3 har vi korrigert for eventuelle systematiske målefeil ved å ekskludere observasjoner fra årene 2015-2016.

4.2.2 Utelatt variabel

Dersom en utelatt variabel er korrelert med en eller flere av de uavhengige variablene, vil dette forårsake brudd på forutsetning 4 og gi skjeve estimater ved estimering (Wooldridge 2015). Problemet med utelatt variabel skjevhet kan forekomme ved manglende data. For å gardere analysen for utelatt variabel problemet må vi i teorien inkludere alle variabler som har forklaringskraft på den avhengige variabelen.

Ta utgangspunkt i modellen:

$$Y = a_1 + a_2X_1 + a_3X_2 + u \quad (4.3)$$

Anta at vi beregner:

$$Y = b_1 + b_2X_1 + v \quad (4.4)$$

X_2 er utelatt i modellen 4.4. Er variablene X_1 og X_2 korrelerte, vil det si at b_2 potensielt ikke er forventningsrett eller konsistent beregning av a_2 , som er effekten X_2 har på Y . Vi kan endre problemet til at X_1 og X_2 ikke er korrelerte. Da vil beregningen til b_2 være en riktig beregning ettersom forutsetningene til MKM er oppfylt. Størrelsen på skjevheten vil avhenge av variabelens påvirkningskraft på Y . I denne oppgaven vil det være vanskelig å unngå å få et utelatt variabel problem, ettersom det er flere faktorer som kan påvirke inntektsgapet mellom kvinner og menn.

4.2.3 Simultanitet

Simultanitet oppstår ved at en eller flere forklaringsvariabler bestemmes simultant med den avhengige variabelen. En konsekvens av simultanitet vil være at den avhengige variabelen vil være korrelert med feilledet. Dette fører til skjeve og inkonsistente estimater. Et eksempel kan være at "K/M Inntekt" påvirker antall kvinner som tar høyere utdanning og ved å anvende variabelen "M/K Høyere utdanning" til å estimere "K/M Inntekt" vil det potensielt oppstå et simultanitetsproblem. Hvis den uobserverte variabelen er uavhengig av tid og kun varierer i tverrsnittet så kan vi kontrollere for dette ved bruk av fixed effects metoden.

4.2.4 Feilspesifikasjon

Feilspesifikasjon oppstår når man estimerer en modell som for eksempel har utelatt en variabel med forklaringskraft på den avhengige variabelen i modellen. Feilspesifikasjon kan oppstå om man tolker effekten av en forklaringsvariabel som en lineær variabel, hvor den egentlig er en avtakende funksjon. Ved å utelate den kvadrert forklaringsvariabelen vil man estimere en feilspesifisert modell.

4.2.5 Heteroskedastisitet

En antagelse ved bruk av MKM er at variansen til feilledet er konstant som kalles homoskedastisitet. Hvis denne antakelsen ikke er oppfylt har man å gjøre med et heteroskedastisitetsproblem. Anta:

$$y_i = b_0 + b_1x_{1i} + u_i \quad (4.5)$$

Vi forutsetter at de første fire forutsetningene i Gauss-Markov er oppfylt. Anta modellen:

$$y_i = b_0 + b_2x_{2i} + b_3Dkvinne_i + u_i \quad (4.6)$$

Hvor y_i i dette tilfellet er lønnen til en person, x_2 er antall år med arbeidserfaring, og D_{kvinne} er en dummyvariabel lik 1 for kvinner og 0 for menn. Presisjonen til denne modellen er målt ved $\sigma^2 = E(u_i^2)$. Desto høyere denne er, jo dårligere presisjon. Kjentegnet til heteroskedastisitet er at modellen i gjennomsnitt ikke er like presis for forskjellige verdier på X-ene. En slik variasjon er nødvendig å unngå, siden presisjonen til venstresidevariabelen avhenger av de ulike x verdiene. Om disse verdiene ikke er konsistente, blir heller ikke \hat{y} riktig predikert. Den mest alvorlige konsekvensen av heteroskedastisitet er at t- og f-tester blir ugyldige. I praksis betyr dette at man kan forkaste en nullhypotese som aldri skulle vært forkastet. En indikator på om det er heteroskedastisitet i datasettet, er å se om enkelte variabler har ekstreme verdier som ligger langt ifra gjennomsnittet.

Den mest anvendte testen for heteroskedastisitet er en Breusch-pagan test. Anta

$$u^2 = A_1 + A_2x_2 + \dots + A_kx_k + w_i \quad (4.7)$$

der:

- a_i = konstantledd
- $A_2 \dots A_k$ = er stigningstall
- w_i = er feilleddet.

En nullhypotese her vil være at stigningstallene er = 0 og alternativhypotesen, dvs feilleddet er heteroskedastisk, at en eller flere av stigningstallene er $\neq 0$.

4.2.6 Autokorrelerte feilledd

En viktig forutsetning ved estimering av OLS er at feilleddet u_{it} ikke er autokorrelert.

$$Corr(u_t, u_{t-1}) = 0 \quad (4.8)$$

$$Corr(u_t, u_{t-2}) = 0 \quad (4.9)$$

Er det slik at en av disse feilleddene er ulik fra null vil det si at feilleddet er autokorrelert. En konsekvens av at feilleddet er autokorrelert er at MKM ikke lengre oppfyller de 5 forutsetningene, og at F-tester og T-tester ikke lengre er gyldige.

Anta feilleddet i modellen er autokorrelert:

$$Y_t = \beta_1 + \beta_2 X_t + u_t \quad (4.10)$$

En måte å fjerne autokorrelasjons problemet på, er å inkludere Y_{t-1} som ny forklaringsvariabel, slik at modellen vil se slik ut:

$$Y_t = \beta_1 + \beta_2 X_t + \beta_3 Y_{t-1} + u_t \quad (4.11)$$

som vil være en modell som ikke har variabler som er autokorrelerte med feilleddet. To andre årsaker til autokorrelerte feilledd er at man har utelatt andre viktige høyresidevariabler, da en lagget verdi av Y_t . Den siste er strukturelle brudd, som vil si at sammenhenger mellom de ulike variablene endrer seg over tid.

Den enkleste måten å kvitte seg med et autokorrelert feilledd er å inkludere utelatte høyresidevariabler. En annen måte å unngå dette problemet på er å transformere den avhengige variabelen til endringsform. Den vanligste måten å teste for autokorrelasjon er ved en Breusch-Godfrey test.

4.3 Økonometrisk spesifikasjon

Heterogenitet er et potensielt problem som er vanskelig å unngå når vi analyserer paneldata. MKM benytter variasjonen mellom regioner og tid, men ikke til variabler som er unike for hver region. Utifra vår problemstilling har vi valgt å benytte MKM som estimeringsmetode i vår hovedmodell. Ved å anvende fixed effects metode tar man hensyn til variabler som er unike for hver region, noe vi har utført som en robusthetssjekk i kapittel 5.

Modellen vi vil analysere er følgende:

$$\text{K/M Inntekt}_{it} = \alpha + \beta_1 \text{By}_i + \beta_j X_{it} + \eta_t + u_{it} \quad (4.12)$$

Der:

1. α er konstantleddet.
2. β_1 er koeffisienten foran dummyvariabelen. By_i
3. X_{it} - er et sett med forklaringsvariabler med tilhørende sett av β_j koeffisienter.²⁷
4. η_t er sett med årsummer.
5. u_{it} er restleddet.

Fotskrift "i" og "t" benevner henholdsvis region og tid. $j = 1, 2, \dots, j$. β_1 illustrerer effekten variabelen har på inntektsgapet, av å bo i en by-region. Restleddet u_{it} fanger opp alle de utelatte variablene som ikke er inkludert i modellen, men som har påvirkningskraft på inntektsgapet.

Tidsdummyer inkluderes for å kontrollere for årlige effekter. Ved bruk av paneldata som varierer i tverrsnittet og over tid kan vi ved bruk av tidsdummyer kontrollere for aggregerte variabler som økonomisk vekst og inflasjon. I datasettet er "individet", 160 bo- og arbeidsmarkedsregioner, som har noen lovpålagte fellestrekk. Eksempelvis er ledighetstrygd felles for alle "individene", men vil variere over tid. Ved å ikke inkludere tidsdummyer vil effektene som beskrevet over kunne påvirke analysen, og vi vil ikke få gode estimater på

²⁷ $X_{it} = [X_{1it}, X_{2it}, \dots, X_{jit}]$ er et sett med forklaringsvariabler med dimensjon $(1 \times j)$, og $\beta_1 = \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1j} \end{bmatrix}$

forholdene mellom variablene.

4.4 Potensielle problemer med datamaterialet

Vi har konsekvent innhentet data fra statistikkbanken til SSB (Statistikkbanken 2018*c*). Dataene er konstruert utifra befolkning- og likestillingsstatistikk fra Statistikkbanken (2018*c*). Det kan ha oppstått tilfeldig målefeil når vi omgjorde kommunedata til regionsdata, men sannsynligheten er liten.

5 Analyse

5.1 Enkel modell

I dette kapittelet tar vi utgangspunkt i modellen i ligning 4.12, og videre inkludere flere variabler. Dette vil tydeliggjøre hvordan effekten av dummyvariabelen "By" endrer seg ved å kontrollere for flere variabler. Som forklart i kapittel 3 er variabelen "K/M Inntekt" det relative inntektsforholdet mellom kvinner og menn. Positive koeffisienter vil resultere i et redusert inntektsgap mellom kvinner og menn, motsatt for negative koeffisienter. Tabell 3 illustrerer regresjon (1), hvor "by" er den eneste variabelen inkludert. I denne regresjonen er rurale regioner referanse kategorien ("By"=0), og resultatene er signifikant på en prosentnivå. Ved å ikke kontrollere for noen andre variabler har variabelen "By" en negativ koeffisient med tilhørende verdi på 0.0608. Det vil si at inntektsgapet mellom menn og kvinner er 6.08 %-poeng større i by-regioner enn i rurale-regioner. Det er trolig flere faktorer som påvirker inntektsgapet mellom kvinner og menn, vi velger derfor å kontrollere for flere variabler i de neste delkapitlene.

5.2 Utdanning og offentlig sektor

Blau og Kahn (2016) argumenterte for at utdanning fremdeles er viktig for å kunne forklare inntektsforskjellene mellom kvinner og menn. Dette er årsaken til at vi inkluderer variabelen "M/K Høyere utdanning" i regresjon (2). Fra tabell 3 ser vi at koeffisienten til "By" reduseres med omtrent 1/3 når det kontrolleres for høyere utdanning. Dette kan tyde på at regresjon (1) overestimerer effekten av å bo i en by-region, som kan skyldes at det er større andel menn med høyere utdanning relativt til kvinner i by-regionene. Modellen estimert i regresjon (2) kan skrives som:

$$\text{K/M Inntekt}_{it} = \alpha - 0.0413\text{By} - 0.114\text{Høyere utdanning}_{it} + \eta_i + u_{it} \quad (5.1)$$

Ved å derivere med hensyn på "M/K Høyere utdanning" ser vi på én enhets endring i variabelen:

$$\frac{\partial \text{K/M Inntekt}}{\partial \text{Høyere utdanning}} = -0.114 \quad (5.2)$$

Variablene inkludert i de ulike regresjonene har verdi mellom 0 og 1, med noen unntak. Vi vil derfor ved tolkning se på 0.1 enhetsøkning, som vil være en realistisk økning i forhold til variabelens verdi. En økning med 0,1 enhet i variabelen "M/K Høyere utdanning" vil øke inntektsgapet med: $\frac{-0.114}{10} = -0.0114$ tilsvarende -1.14%-poeng. Dette er i tråd med oppgavens hypotese om at "M/K Høyere utdanning" og "K/M Inntekt" er negativt korrelert.

I regresjon (3) kontrolleres det for variabelen "M/K Offentlig sektor". Dette resulterer i at variabelen "By" reduseres med 0,0069 til -0.0344 fra regresjon (2) til (3). At det er lavere lønn i offentlig sektor sammenlignet med privat, er beskrevet i (SSB 2016), som kan være en medvirkende faktor til at forskjellen i inntektsgapet mellom by-og rurale regioner blir redusert når det kontrolleres for "M/K offentlig sektor". En økning med 0,1 enheter i samme variabel vil, alt annet likt, redusere inntektsgapet mellom kvinner og menn med 3,86 %-poeng. For at "M/K Offentlig sektor" skal øke i verdi, må antall menn relativt til kvinner i offentlig sektor øke. Flere menn i offentlig sektor vil resultere i et redusert inntektsgap. Ved å kontrollere for "M/K Høyere utdanning" og "M/K Offentlig sektor" har effekten på inntektsgapet av å bo i by blitt redusert med 2.64 %-poeng. Fremdeles forklarer regresjon (3) kun 33.2 % av "K/M Inntekt". Tabell 2 beskriver at det er en høyere andel menn ansatt i offentlig sektor relativt til kvinner i by-regionene, sammenlignet med de rurale regionene. Dette trekker i retning av at effekten av "By" reduseres når vi kontrollerer for "M/K Offentlig sektor". I neste delkapittel inkluderes flere av variablene som er beskrevet i kapittel 3.

Tabell 3: Kontroll for høyere utdanning og offentlig sektor

Variabler	(1) K/M lønn	(2) K/M lønn	(3) K/M lønn
By dummy	-0.0608*** (0.00891)	-0.0413*** (0.00947)	-0.0344*** (0.00814)
M/K Høyere utdanning		-0.114*** (0.0206)	-0.241*** (0.0187)
M/K Offentlig sektor			0.386*** (0.0189)
Konstant	0.673*** (0.00573)	0.754*** (0.0157)	0.724*** (0.0135)
Års dummyer	Ja	Ja	Ja
Observasjoner	1,120	1,120	1,118
Periode	2008-2014	2008-2014	2008-2014
Metode	MKM	MKM	MKM
Observations	1,120	1,120	1,118
R^2	0.056	0.081	0.332

Signifikansnivå, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

K/M lønn-Kvinner årlige, gjennomsnittlige bruttoinntekt relativt til menn sin årlige, gjennomsnittlige bruttoinntekt. M/K Høyere utdanning-Andelen menn med høyere utdanning relativt til andelen kvinner med høyere utdanning. M/K Offentlig sektor-Andelen menn i som jobber i offentlig sektor relativt til andelen kvinner som jobber i offentlig sektor.

5.3 Ledere, deltid og fertilitet

Tabell 4 illustrerer fire nye regresjoner. I regresjon (4) er det kontrollert for antall kvinnelige ledere relativt til mannlige. Koeffisienten foran variabelen "K/M leder" er positiv med tilhørende verdi på 0.0256. En økning med 0,1 enheter i denne variabelen medfører at inntektsgapet reduseres med 0.256 %-poeng. For at "K/M leder" skal øke i verdi, må flere kvinner relativt til menn bli ansatt som ledere. Ettersom lederstillinger har relativt høye lønninger vil det resultere i redusert inntektsgap (Statistikbanken 2018a). Fra tabell 2 ser vi at det er færre kvinnelige ledere relativt til mannlige i by-regionene sammenlignet med de rurale. Dette kan være en av årsakene til at inntektsgapet er høyere i by-regionene enn i urbane. "By" holder seg relativt stabil fra regresjon (4) til (5) noe R^2 , modellens forklaringskraft på venstresidevariabelen, også gjør. En potensiell årsak til variabelens lave betydning kan skyldes at det er relativt få personer som er i lederstillinger sammenlignet med resten av arbeidsstyrken.

I regresjon (5) er variabelen "M/K Deltid" inkludert. Ved 0.1 enhets økning i variabelen, alt annet likt, vil inntektsgapet reduseres med 2.5 %-poeng. Koeffisienten foran "By" reduseres, som vil si at forskjellen i inntektsgapet mellom by-og rurale regioner øker med 0.0063 fra regresjonen (4) til (5). Dette kan skyldes at det i by-regionene er en større andel kvinner som jobber deltid relativt til menn, sammenlignet med de rurale regionene. En økning i andelen kvinner relativt til menn i deltidsstillinger vil, alt annet likt, føre til økt inntektsgap.

Ved å kontrollere for variabelen "M/K Deltid" reduseres koeffisienten foran "M/K Offentlig sektor" med 0.093. Dette kan forklares med at deltidsansatte tenderer til å jobbe i offentlig sektor (Kristoffersen 2017). Variabelen "M/K Offentlig sektor" fanget opp effekten av "M/K Deltid" og følgelig ble koeffisienten overestimert i regresjon (3). Videre reduseres effekten av høyere utdanning med 0.058, som kan være et resultat av at det er en lavere andel med høyere utdanning som jobber i deltidsstillinger.

I regresjonen (6) er variabelen "Andelen 0-9 år" inkludert for å kontrollere for om størrelsen på denne aldersgruppen påvirker inntektsgapet. En økning på én enhet i variabelen

²⁸ vil øke inntektsgapet med 1.148 %-poeng. Effekten av "By" blir mindre negativ når vi kontrollerer for "Andelen 0-9 år" og variabelen reduseres med 0.0166 fra regresjon (5) til (6). Dette kan skyldes at det i by-regionene er en større andel av befolkningen som er i aldersgruppen 0-9 år, sammenlignet med de rurale regionene. Når vi kontrollerer for størrelsen på denne gruppen reduseres effekten av "By" fordi variabelen muligens fanger opp denne effekten i regresjon (5). "K/M Leder" er ikke lengre signifikant i regresjon (5) og (6) på noen av de tradisjonelle nivåene som er grunnen til at vi utelater denne variabelen i regresjon (7). Dette gir ikke store utslag på noen av koeffisientene, som støtter påstanden om at variabelen har lite å si for inntektsforskjellene mellom kvinner og menn.

²⁸tilsvarende at andelen av befolkningen i aldersgruppen 0-9 år øker med 1%-poeng.

Tabell 4: Kontroll for ledere, deltid og fertilitet

Variabler	(4)	(5)	(6)	(7)
	K/M Lønn	K/M Lønn	K/M Lønn	K/M Lønn
By dummy	-0.0337*** (0.00820)	-0.0400*** (0.00775)	-0.0234*** (0.00739)	-0.0228*** (0.00734)
M/K Høyere utdanning	-0.229*** (0.0195)	-0.171*** (0.0195)	-0.138*** (0.0185)	-0.136*** (0.0180)
M/K Offentlig sektor	0.382*** (0.0190)	0.289*** (0.0206)	0.295*** (0.0193)	0.294*** (0.0193)
K/M Leder	0.0256** (0.0104)	0.0130 (0.00981)	-0.00843 (0.00937)	
M/K Deltid		0.250*** (0.0217)	0.244*** (0.0204)	0.242*** (0.0202)
Andelen 0-9 år			-0.0148*** (0.00121)	-0.0145*** (0.00118)
Konstant	0.702*** (0.0164)	0.607*** (0.0177)	0.767*** (0.0211)	0.758*** (0.0187)
Års dummyer	Ja	Ja	Ja	Ja
Observasjoner	1,110	1,093	1,093	1,096
Periode	2008-2014	2008-2014	2008-2014	2008-2014
Metode	MKM	MKM	MKM	MKM
R^2	0.337	0.421	0.492	0.491

Signifikansnivå, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

K/M lønn-Kvinner årlige, gjennomsnittlige bruttoinntekt relativt til menn sin årlige, gjennomsnittlige bruttoinntekt. M/K Høyere utdanning-Andelen menn med høyere utdanning relativt til andelen kvinner med høyere utdanning. M/K Offentlig sektor-Andelen menn i som jobber i offentlig sektor relativt til andelen kvinner som jobber i offentlig sektor. K/M Leder-Antall kvinnelige ledere relativt til antall mannlige ledere. M/K Deltid-Andelen menn som jobber i deltidsstilling relativt til andelen kvinner som jobber i deltidsstillinger. Andelen 0-9 år-Andelen av befolkningen i aldersgruppen 0-9 år

5.4 Avsluttende modell

Tabell 5 illustrerer to ulike regresjoner. I regresjon (8) er den demografiske variabelen ”Andelen 20-24 år” inkludert for å illustrere hvordan inntektsgapet mellom kvinner og menn utvikler seg i en tidlig fase av arbeidskarrieren. koeffisienten til ”Andelen 20-24 år” er positiv og en økning med én enhet²⁹ i denne variabelen vil, alt annet likt, føre til at inntektsgapet reduseres med 0.41 %-poeng. Vi kan dermed konkludere med at en økning i antall sysselsatte personer i denne aldersgruppen vil redusere inntektsgapet.

Effekten av å bo i en by-region reduseres med 0.002 fra regresjon (7) i tabell 4 til regresjon (8) i tabell 5. Den lave reduksjonen i koeffisienten til ”By” kan skyldes at andelen av befolkningen som er i aldersgruppen 0-9 år er relativt lik mellom regionstypene. Dette er vår avsluttende modell, hvor vi har inkludert de variablene som vi mener har en relevant innvirkning på inntektsgapet mellom kvinner og menn, samt inntektsgapet mellom by- og rurale regioner. I det neste delkapittel illustrere vi ulike robusthetsanalyser for å undersøke om resultatene er konsistente med tidligere litteratur.

²⁹Som tilsvarer 1%-poeng

Tabell 5: Kontroll for aldersgrupper

Variabler	(8)	(9)
	K/M Lønn	K/M Lønn
By dummy	-0.0230*** (0.00727)	-0.0272*** (0.00739)
M/K Høyere utdanning	-0.139*** (0.0178)	-0.143*** (0.0179)
M/K Offentlig sektor	0.291*** (0.0191)	0.291*** (0.0192)
M/K Deltid	0.248*** (0.0201)	0.243*** (0.0201)
Andelen 0-9 år	-0.0150*** (0.00118)	-0.0162*** (0.00126)
Andelen 20-24 år	0.00410*** (0.000896)	
Andelen 20-39 år		0.00112*** (0.000297)
Konstant	0.726*** (0.0198)	0.738*** (0.0193)
Års dummyer	Ja	Ja
Observasjoner	1,095	1,095
Periode	2008-2014	2008-2014
Metode	MKM	MKM
R^2	0.500	0.497

Signifikansnivå, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

K/M lønn-Kvinnens årlige, gjennomsnittlige bruttolønn relativt til menn sin årlige, gjennomsnittlige bruttolønn. M/K Høyere utdanning-Andelen menn med høyere utdanning relativt til andelen kvinner med høyere utdanning. M/K Offentlig sektor-Andelen menn i som jobber i offentlig sektor relativt til andelen kvinner som jobber i offentlig sektor. K/M Leder-Antall kvinnelige ledere relativt til antall mannlige ledere. M/K Deltid-Andelen menn som jobber i deltidsstilling relativt til andelen kvinner som jobber i deltidsstillinger. Andelen 0-9 år-Andelen av befolkningen i aldersgruppen 0-9 år. Andelen 20-24 år-Antall sysselsatte personer i arbeidsstyrken i aldersgruppen 20-24 år relativt til hele arbeidsstyrken (20-66 år). Andelen 20-39 år-Antall sysselsatte personer i arbeidsstyrken i aldersgruppen 20-39 år relativt til hele arbeidsstyrken. Andelen 40-66 år-Antall sysselsatte personer i arbeidsstyrken i aldersgruppen 40-66 år relativt til hele arbeidsstyrken

5.5 Robusthetsanalyse

5.5.1 Inntektsgapet i tidlig fase av arbeidslivet

Fra regresjon (8) til (9) er aldersgruppen 25-39 år slått sammen med 20-24 år. Dette medfører at koeffisienten reduseres, som tyder på at det tidlig i arbeidslivet (20-24 år) er mindre inntektsforskjeller mellom kvinner og menn sammenlignet med aldersgruppen 25-39 år. I regresjon (9) har vi kontrollert for variabelen "Andelen 20-39 år". Variabelen har en tilhørende positiv koeffisient, som vil si at inntektsgapet reduseres om andelen personer i denne aldersgruppen øker, et resultat ulikt med hva Manning & Swaffield (2008) konkluderte. Fra tabell (8) ser vi at koeffisienten til "Andelen 20-24 år" er større enn koeffisienten til "Andelen 20-39 år" i regresjon (9). Dette tyder på at det er mindre inntektsforskjeller i aldersgruppen 20-24 år og at inntektsgapet øker fra aldersgruppen 20-24 år til 25-39 år. Motstykket til "Andelen 20-39 år" er "Andelen 40-66 år", en økning i denne variabelen vil trekke i retning av økt inntektsgap mellom kvinner og menn. Grunnen til dette er at variabelen er perfekt negativt korrelerte. Variablene det er kontrollert for i tabell 5 er alle signifikante på et prosentnivå.

5.5.2 Sektor-variabler, Fixed effects og Gini

Ved å inkludere "Jordbruk, skogbruk og fiske" i regresjon (10) endrer koeffisienten til "By" seg i liten grad fra regresjon (8). Videre fra (10) til (11) reduseres effekten av å bo i by med 0.0016 når vi kontrollerer for "Bergverksdrift og utvinning". Ved én enhets³⁰ økning i variabelen "Jordbruk, skogbruk og fiske", øker inntektsgapet mellom kvinner og menn med 0.068 %-poeng. Denne variabelen er ikke signifikant på hverken, en, fem eller ti prosentnivå. Vi kan dermed konkludere at variabelen har liten eller ingen effekt på inntektsgapet mellom kvinner og menn.

Regresjon (12) er estimert med fixed effects som eliminerer alle regionsfaste effekter. Dette er begrunnelsen for å erstatte variabelen "by" med logaritmen til befolkningen. Modellen predikerer at ved befolkningsvekst, vil inntektsgapet mellom kvinner og menn reduseres. Variablene inkludert i regresjon (11) har en lavere effekt på inntektsgapet, ved fixed effects estimering i regresjon (12), og er heller ikke signifikant på noen av de tradisjonelle nivåene. Dette kan tyde på at tverrsnittsvariasjonene påvirker inntektsgapet, men at tidsvariasjonene ikke affiserer inntektsgapet i nevneverdig grad.

I regresjon (13) er variabelen "Gini" inkludert. Den tilhørende koeffisienten har en negativ verdi lik -0.227, som vil si at en økning med 0,1 enheter i variabelen, øker inntektsgapet med 2,27 %-poeng. En økning i variabelen "Gini", fra 0,3 til 0,4 vil medføre at Lorenzekurven buer mer utover. Dette illustrerer økte inntektsforskjeller i befolkninger og færre personer som deler den totale formuen i Norge. Fra tabell 2 i kapittel 3 kan vi se at "Gini" har en gjennomsnittsverdi tilsvarende 0.20. Brunvoll (2014) publiserte i 2014 en rapport som tok for seg inntektsutviklingen i Norge fra 1986 til 2012. I samme periode endret gini-koeffisienten seg fra 0,21 til 0,25. Ved å sammenligne SSB sine data med resultatene fra denne oppgaven, ser vi at gjennomsnittet til gini-koeffisient er 0.05 enheter lavere enn det SSB konkluderte i 2012 (Brunvoll 2014).

³⁰som tilsier at andelen av befolkningen som jobber med jordbruk, skogbruk eller fiske øker med 1%-poeng.

Tabell 6: Kontroll for sektor og Fixed Effects og Gini

Variabler	(10)	(11)	(12)	(13)
	K/M Lønn	K/M Lønn	K/M Lønn	K/M Lønn
By	-0.0233*** (0.00734)	-0.0217*** (0.00753)		-0.0239** (0.00986)
M/K Høyere utdanning	-0.149*** (0.0187)	-0.151*** (0.0188)	0.01985 (0.02553)	-0.135*** (0.0250)
M/K Offentlig sektor	0.285*** (0.0194)	0.285*** (0.0194)	0.006554 (0.037249)	0.297*** (0.0258)
M/K Deltid	0.250*** (0.0210)	0.249*** (0.0210)	0.034512 (0.020065)	0.254*** (0.0275)
Andelen 0-9 år	-0.0152*** (0.00118)	-0.0151*** (0.00119)	-0.00036 (0.001448)	-0.0151*** (0.00157)
Andelen 20-24 år	0.00435*** (0.000903)	0.00436*** (0.000903)	-0.0008345 (0.000657)	0.00427*** (0.00121)
Jordbruk, skogbruk og fiske	-0.000797** (0.000405)	-0.000799** (0.000405)		
Bergverksdrift og utvinning		-0.000921 (0.000953)		
Log Befolkning			0.0038616 (0.0165176)	
Gini				-0.227*** (0.0539)
Konstant	0.737*** (0.0212)	0.738*** (0.0212)	0.617*** (.1495)	0.770*** (0.0273)
Års dummyer	Ja	Ja	Ja	Ja
Observasjoner	1,095	1,095	1,095	624
Periode	2008-2014	2008-2014	2008-2014	2011-2014
Metode	MKM	MKM	FE	MKM
R^2 Within			0.2134	
R^2	0.500	0.501		0.507
Antall regioner			160	

Signifikansnivå, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

K/M lønn-Kvinneres årlige, gjennomsnittlige bruttoinntekt relativt til menn sin årlige, gjennomsnittlige bruttoinntekt. M/K Høyere utdanning-Andelen menn med høyere utdanning relativt til andelen kvinner med høyere utdanning. M/K Offentlig sektor-Andelen menn som jobber i offentlig sektor relativt til andelen kvinner som jobber i offentlig sektor. K/M Leder-Antall kvinnelige ledere relativt til antall mannlige ledere. M/K Deltid-Andelen menn som jobber i deltidsstilling relativt til andelen kvinner som jobber i deltidsstillinger. Andelen 0-9 år-Andelen av befolkningen i aldersgruppen 0-9 år. Jordbruk, Skogbruk og fiske-Antall som jobber i sektoren relativt til hele arbeidsstyrken. Bergverksdrift og utvinning-Antall som jobber i sektoren relativt til hele arbeidsstyrken. Log Befolkning-Logaritmen til befolkningen. Gini-indeks mellom 0 og 1, hvor 0 er helt lik, og 1 er helt ulik.

5.5.3 Ulike by definisjoner

I dette delkapittelet defineres variabelen "By" på følgende tre måter for å se om det gir en endring i inntektsgapet:

- "By" - Definert som de ti regionene som har folketall ≥ 100.000 i 2014
- "By fire" - Definert som de fire regionene som har folketall ≥ 250.000 i 2014
- "By to" - Definert som de to regionene som har folketall ≥ 350.000 i 2014

Regresjon (14) i tabell 7 er identisk med regresjon (8) i tabell 5. I regresjon (15) er de fire største regionene inkludert i variabelen "By fire". Det er 3.16 %-poeng større forskjell i inntektsgapet i regresjon (15), dette er en økning på 0.84 %-poeng fra regresjon (14). Koeffisienten foran by dummyen reduseres fra regresjon (14) til (16) som vil si at inntektsgapet er større når det kontrolleres for færre, men større byregioner.

Effekten av "By" er sterkest når vi kontrollerer for den strengeste definisjonen av variabelen vist i regresjon (16). I denne regresjonen er forskjellen i inntektsgapet mellom kvinner og menn 3.47 %-poeng større i by-regioner sammenlignet med de 158 rurale regionene som er referansegruppen. Trenden indikerer at folketallet i by-regionene har en negativ innvirkning på inntektsgapet.

Tabell 7: Kontroll for ulike bydefinisjoner

Variabler	(14) K/M Lønn	(15) K/M Lønn	(16) K/M Lønn
By	-0.0232*** (0.00735)		
By fire		-0.0316*** (0.0109)	
By to			-0.0347** (0.0149)
M/K Høyere utdanning	-0.138*** (0.0179)	-0.145*** (0.0175)	-0.149*** (0.0173)
M/K Offentlig sektor	0.290*** (0.0193)	0.293*** (0.0192)	0.295*** (0.0192)
M/K Deltid	0.254*** (0.0209)	0.256*** (0.0210)	0.254*** (0.0210)
Andelen 0-9 år	-0.0150*** (0.00118)	-0.0151*** (0.00118)	-0.0154*** (0.00117)
Andelen 20-24 år	0.00415*** (0.000898)	0.00423*** (0.000900)	0.00413*** (0.000900)
Konstant	0.723*** (0.0200)	0.726*** (0.0199)	0.733*** (0.0195)
Års dummyer	Ja	Ja	Ja
Observasjoner	1,095	1,095	1,095
Periode	2008-2014	2008-2014	2008-2014
Metode	MKM	MKM	MKM
R^2	0.499	0.498	0.496

Signifikansnivå, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

K/M lønn-Kvinnens årlige, gjennomsnittlige bruttoinntekt relativt til menn sin årlige, gjennomsnittlige bruttoinntekt. M/K Høyere utdanning-Andelen menn med høyere utdanning relativt til andelen kvinner med høyere utdanning. M/K Offentlig sektor-Andelen menn i som jobber i offentlig sektor relativt til andelen kvinner som jobber i offentlig sektor. K/M Leder-Antall kvinnelige ledere relativt til antall mannlige ledere. M/K Deltid-Andelen menn som jobber i deltidsstilling relativt til andelen kvinner som jobber i deltidsstillinger. Andelen 0-9 år-Andelen av befolkningen i aldersgruppen 0-9 år. Andelen 20-24 år-Antall sysselsatte personer i arbeidsstyrken i aldersgruppen 20-24 år relativt til hele arbeidsstyrken (20-66 år). By to-Kun de to mest folkerike regionene (Oslo og Bergen) som er definert som by-region. By fire- Kun de fire mest folkerike regionene (Oslo, Bergen, Stavanger og Trondheim) som er definert som by-region.

5.5.4 Interaksjonsledd

I tabell 8 er det inkludert interaksjonsledd mellom "By" og henholdsvis, "M/K Offentlig sektor", "M/K Deltid" og "M/K Høyere utdanning". I Regresjon (17) kontrolleres det for interaksjonsleddet "By*M/K Offentlig sektor" som resulterer i at effekten av variabelen "by" reduseres kraftig. Ved å derivere (17) med hensyn på "M/K Offentlig sektor" får vi følgende:

$$\frac{\partial K/M \text{ Inntekt}}{\partial M/K \text{ Offentlig sektor}} = 0.289 + (-0.0475) * \text{By} \quad (5.3)$$

Det vil si at ved 0.1 enhets økning i variabelen "M/K Offentlig sektor" vil inntektsgapet i de rurale regionene reduseres med 2.89 %-poeng. En tilsvarende økning i by-regionene vil inntektsgapet reduseres med 2.42 %-poeng. En økning på landsbasis i "M/K Offentlig sektor" med 0.1 enhet vil inntektsgapet i by-regioner reduseres mindre sammenlignet med de rurale regionene.³¹

I regresjon (18) har vi kontrollert for interaksjonsledd mellom "By" og "M/K Deltid". Regresjonen illustrerer at "By" reduseres mindre sammenlignet med regresjon (17). En økning i variabelen "M/K Deltid" med 0.1 enhet i alle regionene fører til en større reduksjon i inntektsgapet for de rurale regionene, sammenlignet med by-regionene.

Regresjon (19) illustrerer hovedmodellen med interaksjonsleddet "By*M/K Høyere utdanning". Ved 0.1 enhets økning i "M/K Høyere utdanning" øker inntektsgapet i by-regionene med 1.49 %-poeng, mens det i de rurale regionene øker med 1.39 %-poeng. En økning i andelen menn med høyere utdanning relativt til kvinner, resulterer i at inntektsgapet i by-regionene øker mer enn i de rurale regionene. Ingen av resultatene over er signifikante på noen av de tradisjonelle nivåene, som vil si at vi ikke kan trekke noen konklusjon fra denne delen av analysen.

³¹inntektsgapet vil bli redusert mindre i by-regionene på grunn av den negative koeffisienten i interaksjonsleddet

Tabell 8: Kontroll for interaksjonsledd

Variabler	(17)	(18)	(19)
	K/M Lønn	K/M Lønn	K/M Lønn
By	-0.00593 (0.0354)	-0.00938 (0.0395)	-0.0145 (0.0994)
M/K Høyere utdanning	-0.139*** (0.0179)	-0.139*** (0.0179)	-0.140*** (0.0179)
M/K Offentlig sektor	0.289*** (0.0192)	0.289*** (0.0191)	0.289*** (0.0191)
M/K Deltid	0.255*** (0.0204)	0.255*** (0.0205)	0.254*** (0.0204)
Andelen 0-9 år	-0.0150*** (0.00117)	-0.0150*** (0.00117)	-0.0150*** (0.00118)
Andelen 20-24 år	0.00408*** (0.000895)	0.00408*** (0.000895)	0.00409*** (0.000895)
By*M/K Offentlig sektor	-0.0475 (0.0979)		
By*M/K Deltid		-0.0365 (0.106)	
By*M/K Høyere utdanning			-0.00975 (0.117)
Konstant	0.723*** (0.0199)	0.723*** (0.0200)	0.724*** (0.0199)
Års dummyer	Ja	Ja	Ja
Observasjoner	1,095	1,095	1,095
Periode	2008-2014	2008-2014	2008-2014
Metode	MKM	MKM	MKM
R^2	0.502	0.502	0.502

Signifikansnivå, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

K/M lønn-Kvinnens årlige, gjennomsnittlige bruttoinntekt relativt til menn sin årlige, gjennomsnittlige bruttoinntekt. M/K Høyere utdanning-Andelen menn med høyere utdanning relativt til andelen kvinner med høyere utdanning. M/K Offentlig sektor-Andelen menn som jobber i offentlig sektor relativt til andelen kvinner som jobber i offentlig sektor. K/M Leder-Antall kvinnelige ledere relativt til antall mannlige ledere. M/K Deltid-Andelen menn som jobber i deltidsstilling relativt til andelen kvinner som jobber i deltidsstillinger. Andelen 0-9 år-Andelen av befolkningen i aldersgruppen 0-9 år. Jordbruk, Skogbruk og fiske-Antall som jobber i sektoren relativt til hele arbeidsstyrken. Bergverksdrift og utvinning-Antall som jobber i sektoren relativt til hele arbeidsstyrken. By*M/K Offentlig sektor-Interaksjonsledd mellom By dummy og M/K Offentlig sektor. By*M/K Deltid-Interaksjonsledd mellom By dummy og M/K Deltid. By*M/K Høyere utdanning-Interaksjonsledd mellom By dummy og M/K Høyere utdanning.

I kapittel 3.6 ble seks hypoteser presentert. Fra analysen i dette kapitlet, har vi fått svar på om de ulike hypotesene kan beholdes eller forkastes.

1. En økning i andelen menn som jobber i offentlig sektor relativt til andelen kvinner vil, alt annet likt, redusere inntektsgapet mellom kvinner og menn. Hypotesen beholdes.
2. En økning i andelen utdannede menn relativt til kvinner vil, alt annet likt, øke inntektsgapet mellom kvinner og menn. Hypotesen beholdes.
3. En økning i antall kvinnelige ledere relativt til mannlige vil, alt annet likt, redusere inntektsgapet mellom kvinner og menn. Hypotesen beholdes.
4. Hovedmodellen konkluderer med at det er større inntektsgap i by-regioner, sammenlignet med rurale. Hypotesen beholdes.
5. Ifølge resultatene fra analysen, viser den at inntektsgapet mellom kvinner og menn reduseres når antall sysselsatte i denne aldersgruppen øker, dette gjelder også for aldersgruppen 25-39 år. Hypotesen beholdes.
6. Hypotese 5: En økning i antall barn mellom "0-9" år vil, alt annet likt, øke inntektsforskjellene mellom kvinner og menn. Hypotesen beholdes.

6 Konklusjon

Denne oppgaven har som formål å undersøke hvilke variabler som påvirker inntektsgapet mellom kvinner og menn med spesielt fokus på inntektsgapet mellom by- og rurale regioner. Oppgaven baserer seg på paneldata for Norge, inndelt i 160 bo og arbeidsmarkedsregioner. Datamaterialet er innhentet fra SSB, og er omgjort fra kommunenivå til regionsnivå for å kunne ta høyde for pendlere mellom de ulike kommunene. Hovedmodellen er estimert med MKM, og følgelig utnyttes all variasjon i data. Oppgavens hovedmål er å undersøke om det er inntektsforskjeller mellom by- og rurale regioner. Vi har håndtert dette ved å definere by-regionene som en dummy variabel som resulterer at vi ikke kontrollerer for regionsfaste effekter i vår hovedmodell. Vi har likevel gjennomført en robusthetsanalyse hvor vi kontrollerer for regionsfaste effekter, som resulterer i at ingen av variablene vi anser som grunnleggende for å forklare inntektsgapet er signifikante på noen av de tradisjonelle nivåene. Vi har inkludert tidsdummyer for å kontrollere for nasjonale endringer og svingninger over tid. Analysen viser til et rått inntektsgap på 6,08 %-poeng mellom by-regioner og rurale regioner. Ved å kontrollere for ulike variabler reduseres forskjellen i inntektsgapet til 2,3 %-poeng mellom by- og rurale regioner.

Analysen indikerer at dersom andelen kvinner med høyere utdanning fortsetter å øke relativt til menn, vil forskjellen i inntektsgapet mellom by- og rurale-regioner reduseres. Andelen menn relativt til kvinner i offentlig sektor har stor påvirkning på forskjellen i inntektsgapet mellom by- og rurale regioner. Det er en mindre andel menn relativt til kvinner som jobber i offentlig sektor i rurale-regioner. Dette betyr at om det i by-regionene er en mindre andel kvinner relativt til menn som jobber i offentlig sektor, vil forskjellen i inntektsgapet mellom kvinner og menn i by- og rurale regioner reduseres.

Vi finner at en reduksjon i andelen kvinner relativt til menn som er ansatt i deltidsstillinger vil påvirke inntektsgapet positivt.

Ved å kontrollere for andelen barn i aldersgruppen 0-9 år reduseres forskjellen i inntektsgapet mellom kvinner og menn i by- og rurale regioner. Grunnen til dette er at det er en høyere andel barn i aldersgruppen 0-9 år i by-regionene sammenlignet med de rurale-regionene.

Det er fremdeles en uforklarlig del av inntektsgapet mellom kvinner og menn vi ikke kan observere. Dette tyder sannsynligvis i retning av at kvinner blir diskriminert i arbeidslivet. Dette er omtalt i tidligere litteratur som en viktig faktor for å forklare inntektsgapet mellom kvinner og menn. Videre forskning på dette tema bør ta for seg en lengre tidsperiode og om mulig se på individdata med sektor og individspesifikke karakteristika.

Referanser

- Bayard, K., Hellerstein, J., Neumark, D. & Troske, K. (2003), 'New evidence on sex segregation and sex differences in wages from matched employee-employer data', *Journal of Labor Economics* **21**(4), 887–922.
- Blau, F. D. & Kahn, L. M. (2000), 'Gender differences in pay', *Journal of Economic Perspectives* **14**(4), 75–99.
- Blau, F. D. & Kahn, L. M. (2017), 'The gender wage gap: Extent, trends, and explanations', *Journal of Economic Literature* **55**(3), 789–865.
- Brunvoll, F. (2014), 'Utvikling i inntektsulikhet'. Hentet: 1.03.2018.
URL: <http://https://www.ssb.no/natur-og-miljo/barekraft/utvikling-i-inntektsulikhet>
- Carlsen, F., Rattsø, J. & Stokke, H. E. (2016), 'Education, experience and dynamic urban wage premium', *Regional Science and Urban-Economics* **60**,39-49 .
- De La Roca, J. & Puga, D. (2017), 'Learning by working in big cities', *The Review of Economic Studies* **84**(1), 106–142.
- Forum, W. E. (2007), The global gender gap report, World Economic Forum, Geneva, Switzerland.
- Gundersen, F. & Juvkam, D. (2013), 'Inndelinger i senterstruktur, sentralitet og barregioner', *NIBR-rapport* **1**, 15.
- Hirsch, B., König, M. & Möller, J. (2013), 'Is there a gap in the gap? regional differences in the gender pay gap', *Scottish Journal of Political Economy* **60**(4), 412–439.
- Huertas, I. P. M., Ramos, R. & Simon, H. (2017), 'Regional differences in the gender wage gap in Spain', *Social Indicators Research* **134**(3), 981–1008.
- Kristoffersen, S. (2017), 'Lønnsforskjellene mellom kvinner og menn vedvarer'. Hentet: 14.02.2018.
URL: <https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/lonnsforskjellene-mellom-kvinner-og-menn-vedvarer>
- Lips, H. M. (2003), 'The gender pay gap: Concrete indicator of women's progress toward equality', *Analyses of Social Issues and Public Policy* **3**(1), 87–109.

- Manning, A. & Swaffield, J. (2008), 'The gender gap in early-career wage growth', *The Economic Journal* **118**(530), 983–1024.
- Rattsø, J. & Stokke, H. E. (2014), 'Population divergence and income convergence: Regional distribution dynamics for Norway', *Regional Studies* **48**(11), 1884–1895.
- Rattsø, J. & Stokke, H. E. (2018), 'Dynamic private-public wage gap: Estimates based on sector-specific experience data', *Norwegian University of Science and Technology working paper* .
- Rubery, J., Grimshaw, D. & Figueiredo, H. (2005), 'How to close the gender pay gap in Europe: towards the gender mainstreaming of pay policy', *Industrial Relations Journal* **36**(3), 184–213.
- SSB (2016), 'Faktaside'. Hentet: 20.03.2018.
URL: <https://www.ssb.no/befolkning/faktaside/likestilling>
- SSB (2018), 'Indikatorer for kjønnslikestilling i kommunene'. Hentet: 20.01.2018.
URL: <https://www.ssb.no/statbank/table/09293/?rxid=e86dea03-0a77-4ce9-87f8-3204480814ad>
- Statistikkbanken (2017), 'Sysselsetting, registerbasert'. Hentet: 06.04.2018.
URL: <https://www.ssb.no/statbank/table/11365/?rxid=d57389e5-99d2-49ae-84ef-69a11afeac5a>
- Statistikkbanken (2018a), 'Lønn for ansatte, etter yrke, sektor, kjønn og arbeidstid 2015 - 2017'. Hentet: 04.04.2018.
URL: <https://www.ssb.no/statbank/table/11418/?rxid=8936b74d-8601-45e3-be1b-7c5a79a847b7>
- Statistikkbanken (2018b), 'Variabeldefinisjon bruttoinntekt'. Hentet: 04.05.2018.
URL: <https://www.ssb.no/a/metadata/conceptvariable/vardok/16/nb>
- Statistikkbanken, S. (2018c), 'Statistikkbanken'. Hentet: 27.04.2018.
URL: <https://www.ssb.no/statbank>
- Stokke, H. E. (2016), 'The gender wage gap and the early-career effect', *Department of Economics, Norwegian University of Science and Technology* .
- Sucarrat, G. (2016), *Metode og økonometri - en moderne innføring*, Fagbokforlaget.

Tuhus, P. T. (2010), 'En av tre har høyere utdanning'. Hentet: 05.02.2018.

URL: <https://www.ssb.no/utdanning/artikler-og-publikasjoner/en-av-tre-har-hoyere-utdanning>

Ultsch, A. & Lötsch, J. (2017), 'A data science based standardized gini index as a Lorenz dominance preserving measure of the inequality of distributions', *PloS one* **12**(8), e0181572.

Wooldridge, J. M. (2015), *Introductory econometrics: A modern approach*, Nelson Education.

A Vedlegg

A Manglende observasjoner i datasettet

Tabell 9: Manglende observasjoner

M/K Offentlig sektor							
Region	2008	2009	2010	2011	2012	2013	2014
Sunddal		X					
Tønsberg		X					
K/M Leder							
Region	2008	2009	2010	2011	2012	2013	2014
Fred/Sarp		X					
Tønsberg		X					
Harstad						X	X
Steinkjær						X	X
Ibestad						X	X
K/M Deltid							
Region	2008	2009	2010	2011	2012	2013	2014
Harstad						X	X
Steinkjær				X	X	X	X
Sortland	X						
Narvik	X	X	X	X			
Måselv					X	X	X
Ibestad	X	X	X	X	X		
Lierne					X	X	X

Tabellen viser for hvilke regioner og årstall det mangler observasjoner for i vårt datasett

B Konstruksjon av datasett

Datasett

Vi har konstruert datasettet som er anvendt i denne oppgaven utifra kommunestatistikk fra SSB.

Vi har befolkningsvektet variablene når vi har konvertert kommunene til 160 BA regioner. Dette har vi gjort slik at data blir så presist som mulig og at det ikke fremkommer noen systematiske eller tilfeldige målefeil.

For variabelen "M/K Høyere utdanning" så hentet vi ut andelen menn og kvinner med høyere utdanning i samtlige kommuner i Norge. Under har vi illustrert hvordan vi har vektet variabelen "M/K Høyere utdanning" for én region.

"M/K Høyere utdanning for regionene Halden:

$$\frac{(Y * X) + (Z * C)}{X + C} = M \quad (\text{A.1})$$

1. Y - Andelen menn med høyere utdanning i Aremark
2. X - Antall personer i Aremark
3. Z - Andelen menn med høyere utdanning i Halden
4. C - Antall personer i Halden
5. M - Andelen menn med høyere utdanning
6. K - Andelen kvinner med høyere utdanning

Tilsvarende er Andelen kvinner med høyere utdanning beregnet. Videre har vi konstruert variabelen "M/K Høyere utdanning" som M/K.

Tilsvarende har vi definert de fleste variablene i vårt datasett, med små variasjoner utifra hvordan variabelen er definert fra SSB sin kommunedata.

C Korrelasjonsmatrise

Tabell 10: Korrelasjon mellom variablene

(1)

	Lønn	By	Utdanning	Offentlig	Leder	Deltid	0-9	20-24	20-39	40-66
K/M Lønn	1									
By dummi	-0.199***	1								
M/K Høyere utdanning	-0.236***	0.365***	1							
M/K Offentlig sektor	0.403***	0.0888*	0.334***	1						
K/M Leder	0.194***	-0.145***	-0.322***	-0.00373	1					
M/K Deltid	0.512***	-0.0148	-0.177***	0.374***	0.187***	1				
Andelen 0-9 år	-0.367***	0.270***	0.312***	0.0933**	-0.284***	-0.0927**	1			
Andelen 20-24 år	0.0288	0.0525	0.108***	0.0619*	-0.0295	-0.0757*	0.127***	1		
Andelen 20-39 år	-0.114***	0.293***	0.293***	0.130***	-0.183***	-0.0568	0.449***	0.692***	1	
Andelen 40-66 år	0.114***	-0.293***	-0.293***	-0.130***	0.183***	0.0568	-0.449***	-0.692***	-1	1

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

K/M Lønn(Lønn)-Kvinnens årlige, gjennomsnittlige bruttoinntekt relativt til menn sin årlige, gjennomsnittlige bruttoinntekt.

M/K Høyere utdanning(Utdanning)-Andelen menn med høyere utdanning relativt til andelen kvinner med høyere utdanning.

M/K Offentlig sektor(Offentlig)-Andelen menn som jobber i offentlig sektor relativt til andelen kvinner som jobber i offentlig sektor.

K/M Leder(Leder)-Antall kvinnelige ledere relativt til antall mannlige ledere.

M/K Deltid(Deltid)-Andelen menn som jobber i deltidstilling relativt til andelen kvinner som jobber i deltidstillinger.

Andelen 0-9 år(0-9)-Andelen av befolkningen i aldersgruppen 0-9 år

Andelen 20-24 år(20-24)-Antall sysselsatte personer i arbeidsstyrken i aldersgruppen 20-24 år relativt til hele arbeidsstyrken (20-66 år).

Andelen 20-39 år(20-39)-Antall sysselsatte personer i arbeidsstyrken i aldersgruppen 20-39 år relativt til hele arbeidsstyrken.

Andelen 40-66 år(40-66)-Antall sysselsatte personer i arbeidsstyrken i aldersgruppen 40-66 år relativt til hele arbeidsstyrken

D 160 bo- og arbeidsmarkedsregioner

Liste over hvilke kommuner som inngår i de forskjellige, 160 bo- og arbeidsmarkedsregionene. (Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
1	Halden	0101	Halden
1	Halden	0118	Aremark
2	Moss	0104	Moss
2	Moss	0136	Rygge
2	Moss	0137	Våler
3	Fredrikstad/Sarpsborg	0105	Sarpsborg
3	Fredrikstad/Sarpsborg	0106	Fredrikstad
3	Fredrikstad/Sarpsborg	0111	Hvaler
3	Fredrikstad/Sarpsborg	0128	Rakkestad
3	Fredrikstad/Sarpsborg	0135	Råde
4	Askim/Eidsberg	0119	Marker
4	Askim/Eidsberg	0122	Trogstad
4	Askim/Eidsberg	0124	Askim
4	Askim/Eidsberg	0125	Eidsberg
4	Askim/Eidsberg	0127	Skiptvet
5	Oslo	0121	Romskog
5	Oslo	0123	Spydeberg
5	Oslo	0138	Hobøl
5	Oslo	0211	Vestby
5	Oslo	0213	Ski
5	Oslo	0214	Ås
5	Oslo	0215	Frogn
5	Oslo	0216	Nesodden
5	Oslo	0217	Oppegård
5	Oslo	0219	Barnum
5	Oslo	0220	Asker
5	Oslo	0221	Aurskog-Holand
5	Oslo	0226	Sorum
5	Oslo	0227	Fet
5	Oslo	0228	Rælingen
5	Oslo	0229	Enebakk
5	Oslo	0230	Lørenskog
5	Oslo	0231	Skedsmo
5	Oslo	0233	Nittedal
5	Oslo	0234	Gjerdrum
5	Oslo	0235	Ullensaker
5	Oslo	0236	Nes
5	Oslo	0237	Eidsvoll

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
5	Oslo	0238	Nannestad
5	Oslo	0239	Hurdal
5	Oslo	0301	Oslo
5	Oslo	0533	Lunner
5	Oslo	0534	Gran
5	Oslo	0627	Røyken
5	Oslo	0628	Hurum
6	Kongsvinger	0402	Kongsvinger
6	Kongsvinger	0418	Nord-Odal
6	Kongsvinger	0419	Sør-Odal
6	Kongsvinger	0420	Edsbyg
6	Kongsvinger	0423	Grue
6	Kongsvinger	0425	Åsnes
7	Hamar	0403	Hamar
7	Hamar	0412	Ringsaker
7	Hamar	0415	Loten
7	Hamar	0417	Stange
8	Elverum	0426	Våler
8	Elverum	0427	Elverum
8	Elverum	0429	Åmot
9	Trysil/Engerdal	0428	Trysil
9	Trysil/Engerdal	0434	Engerdal
10	Sør-Elvdal	0430	Stor-Elvdal
11	Tynset	0432	Rendalen
11	Tynset	0436	Tolga
11	Tynset	0437	Tynset
11	Tynset	0438	Alvdal
11	Tynset	0439	Folldal
12	Lillehammer	0501	Lillehammer
12	Lillehammer	0521	Øyer
12	Lillehammer	0522	Gausdal
13	Gjøvik	0502	Gjøvik
13	Gjøvik	0528	Østre Toten
13	Gjøvik	0529	Vestre Toten
13	Gjøvik	0536	Søndre Land
13	Gjøvik	0538	Nordre Land
14	Dovre	0511	Dovre
14	Dovre	0512	Lesja
15	Skjåk/Lom	0513	Skjåk
15	Skjåk/Lom	0514	Lom
16	Midt-Gudbrandsdal	0516	Nord-Fron
16	Midt-Gudbrandsdal	0519	Sør-Fron

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
16	Midt-Gudbrandsdal	0520	Ringsbu
17	Sel	0515	Vågå
17	Sel	0517	Sel
18	Fagernes	0540	Sør-Aurdal
18	Fagernes	0541	Etneidal
18	Fagernes	0542	Nord-Aurdal
18	Fagernes	0543	Vestre Slidre
18	Fagernes	0544	Øystre Slidre
18	Fagernes	0545	Vang
19	Drammen	0602	Drammen
19	Drammen	0621	Sigdal
19	Drammen	0623	Modum
19	Drammen	0624	Øvre Eiker
19	Drammen	0625	Nedre Eiker
19	Drammen	0626	Lier
19	Drammen	0711	Svelvik
19	Drammen	0713	Sande
20	Kongsberg	0604	Kongsberg
20	Kongsberg	0631	Flesberg
20	Kongsberg	0632	Rollag
21	Ringerike	0532	Jevnaker
21	Ringerike	0605	Ringerike
21	Ringerike	0612	Hole
21	Ringerike	0622	Krodsherad
22	Hallingdal	0615	Få
22	Hallingdal	0616	Nes
22	Hallingdal	0617	Gol
22	Hallingdal	0618	Hemsedal
22	Hallingdal	0619	Ål
22	Hallingdal	0620	Hol
23	Nore og Uvdal	0633	Nore og Uvdal
24	Tonsberg	0701	Horten
24	Tonsberg	0702	Holmestrand
24	Tonsberg	0704	Tonsberg
24	Tonsberg	0714	Hof
24	Tonsberg	0716	Re
24	Tonsberg	0719	Andebu
24	Tonsberg	0720	Stokke
24	Tonsberg	0722	Nøtterøy
24	Tonsberg	0723	Tjøme
25	Larvik/Sandefjord	0706	Sandefjord
25	Larvik/Sandefjord	0709	Larvik

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
25	Larvik/Sandefjord	0728	Lardal
26	Grenland	0805	Porsgrunn
26	Grenland	0806	Skien
26	Grenland	0811	Siljan
26	Grenland	0814	Bamble
26	Grenland	0815	Kragerø
26	Grenland	0817	Drangedal
26	Grenland	0819	Nome
27	Notodden	0807	Notodden
27	Notodden	0821	Bø
27	Notodden	0822	Sauherad
27	Notodden	0827	Hjørtdal
28	Tinn	0826	Tinn
29	Seljord/Kviteseid	0828	Seljord
29	Seljord/Kviteseid	0829	Kviteseid
30	Nissedal	0830	Nissedal
31	Fyresdal	0831	Fyresdal
32	Vinje/Tokke/Nissedal	0833	Tokke
32	Vinje/Tokke/Nissedal	0834	Vinje
33	Risør	0901	Risør
33	Risør	0911	Gjerstad
34	Arendal	0904	Grimstad
34	Arendal	0906	Arendal
34	Arendal	0912	Vegårshei
34	Arendal	0914	Tvedestrand
34	Arendal	0919	Froland
34	Arendal	0929	Åmli
35	Evje/Bygland	0937	Evje og Hornnes
35	Evje/Bygland	0938	Bygland
36	Valle/Bykle	0940	Valle
36	Valle/Bykle	0941	Bykle
37	Kristiansand	0926	Lillesand
37	Kristiansand	0928	Birkenes
37	Kristiansand	0935	Iveland
37	Kristiansand	1001	Kristiansand
37	Kristiansand	1014	Vennesla
37	Kristiansand	1017	Songdalen
37	Kristiansand	1018	Søgne
38	Mandal	1002	Mandal
38	Mandal	1021	Marnardal
38	Mandal	1027	Audnedal
38	Mandal	1029	Lindesnes

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
25	Larvik/Sandefjord	0728	Lardal
26	Grenland	0805	Poesgrunn
26	Grenland	0806	Skien
26	Grenland	0811	Siljan
26	Grenland	0814	Bamble
26	Grenland	0815	Kragerø
26	Grenland	0817	Drangedal
26	Grenland	0819	Nome
27	Notodden	0807	Notodden
27	Notodden	0821	Bø
27	Notodden	0822	Sauherad
27	Notodden	0827	Hjartdal
28	Tinn	0826	Tinn
29	Seljord/Kviteseid	0828	Seljord
29	Seljord/Kviteseid	0829	Kviteseid
30	Nissedal	0830	Nissedal
31	Fyresdal	0831	Fyresdal
32	Vinje/Tokke/Nissedal	0833	Tokke
32	Vinje/Tokke/Nissedal	0834	Vinje
33	Risør	0901	Risør
33	Risør	0911	Gjerstad
34	Arendal	0904	Grimstad
34	Arendal	0906	Arendal
34	Arendal	0912	Vegårshei
34	Arendal	0914	Tvedestrand
34	Arendal	0919	Froland
34	Arendal	0929	Åmli
35	Evje/Bygland	0937	Evje og Hornnes
35	Evje/Bygland	0938	Bygland
36	Valle/Bykle	0940	Valle
36	Valle/Bykle	0941	Bykle
37	Kristiansand	0926	Lillesand
37	Kristiansand	0928	Birkenes
37	Kristiansand	0935	Iveland
37	Kristiansand	1001	Kristiansand
37	Kristiansand	1014	Vennesla
37	Kristiansand	1017	Songdalen
37	Kristiansand	1018	Søgne
38	Mandal	1002	Mandal
38	Mandal	1021	Marnardal
38	Mandal	1027	Audnedal
38	Mandal	1029	Lindesnes

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunene.	Kommunenavn
39	Farsund/Lyngdal	1003	Farsund
39	Farsund/Lyngdal	1032	Lyngdal
39	Farsund/Lyngdal	1034	Hagebostad
40	Flekkefjord	1004	Flekkefjord
40	Flekkefjord	1037	Kvinesdal
40	Flekkefjord	1112	Lund
41	Åseral	1026	Åseral
42	Sirdal	1046	Sirdal
43	Eigersund	1101	Eigersund
43	Eigersund	1111	Sokndal
44	Stavanger/Sandnes	1102	Sandnes
44	Stavanger/Sandnes	1103	Stavanger
44	Stavanger/Sandnes	1114	Bjerkreim
44	Stavanger/Sandnes	1119	Hå
44	Stavanger/Sandnes	1120	Klepp
44	Stavanger/Sandnes	1121	Time
44	Stavanger/Sandnes	1122	Gjesdal
44	Stavanger/Sandnes	1124	Sola
44	Stavanger/Sandnes	1127	Randsberg
44	Stavanger/Sandnes	1129	Forsand
44	Stavanger/Sandnes	1130	Strand
44	Stavanger/Sandnes	1141	Finnøy
44	Stavanger/Sandnes	1142	Rennesøy
44	Stavanger/Sandnes	1144	Kvitøy
45	Haugesund	1106	Haugesund
45	Haugesund	1145	Bokn
45	Haugesund	1146	Tysvær
45	Haugesund	1149	Karmøy
45	Haugesund	1160	Vindafjord
45	Haugesund	1211	Etnø
45	Haugesund	1216	Sveio
46	Hjelmeland	1133	Hjelmeland
47	Suldal	1134	Suldal
48	Sanda	1135	Sanda
49	Utsira	1151	Utsira
50	Bergen	1201	Bergen
50	Bergen	1241	Fusa
50	Bergen	1242	Samnanger
50	Bergen	1243	Os
50	Bergen	1245	Sund
50	Bergen	1246	Fjell
50	Bergen	1247	Askøy

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
50	Bergen	1251	Vaksdal
50	Bergen	1253	Osterøy
50	Bergen	1256	Meland
50	Bergen	1259	Øygarden
50	Bergen	1260	Radøy
50	Bergen	1263	Lindås
50	Bergen	1264	Austrheim
51	Stord	1219	Bømlo
51	Stord	1221	Stord
51	Stord	1222	Fitjar
51	Stord	1223	Tysnes
52	Jondal/Kvam	1227	Jondal
52	Jondal/Kvam	1238	Kvam
53	Kvinherad	1224	Kvinherad
54	Odda	1228	Odda
54	Odda	1231	Ullensvang
54	Odda	1232	Eidfjord
55	Voss	1233	Ulvik
55	Voss	1234	Granvin
55	Voss	1235	Voss
56	Austevoll	1244	Austevoll
57	Modalen	1252	Modalen
58	Fedje	1265	Fedje
59	Masfjorden/Gulen	1266	Masfjorden
59	Masfjorden/Gulen	1411	Gulen
60	Flora	1401	Flora
61	Solund	1412	Solund
62	Høyanger	1416	Høyanger
62	Høyanger	1418	Balestrand
63	Vik	1417	Vik
64	Sogndal	1419	Leikanger
64	Sogndal	1420	Sogndal
64	Sogndal	1426	Luster
65	Aurland	1421	Aurland
66	Lærdal/Årdal	1422	Lærdal
66	Lærdal/Årdal	1424	Årdal
67	Hyllestad	1413	Hyllestad
68	Forde	1428	Askvoll
68	Forde	1429	Fjaler
68	Forde	1430	Gaular
68	Forde	1431	Jølster
68	Forde	1432	Forde

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
68	Førde	1433	Naustdal
69	Bremanger	1438	Bremanger
70	Vågsøy	1439	Vågsøy
70	Vågsøy	1441	Selje
71	Gloppen	1445	Gloppen
72	Stryn/Eid	1443	Eid
72	Stryn/Eid	1444	Hornindal
72	Stryn/Eid	1449	Stryn
73	Molde	1502	Molde
73	Molde	1535	Vestnes
73	Molde	1543	Nesset
73	Molde	1545	Midsund
73	Molde	1547	Aukra
73	Molde	1548	Fræna
73	Molde	1551	Eide
73	Molde	1557	Gjemnes
74	Kristiansund	1505	Kristiansund
74	Kristiansund	1554	Averøy
74	Kristiansund	1560	Tingvoll
75	Ålesund	1504	Ålesund
75	Ålesund	1523	Ørskog
75	Ålesund	1526	Stordal
75	Ålesund	1528	Sykkylven
75	Ålesund	1529	Skodje
75	Ålesund	1531	Sula
75	Ålesund	1532	Giske
75	Ålesund	1534	Haram
76	Vanylven	1511	Vanylven
77	Ulstein	1514	Sande
77	Ulstein	1515	Herøy
77	Ulstein	1516	Ulstein
77	Ulstein	1517	Hareid
78	Ørsta/Volda	1519	Volda
78	Ørsta/Volda	1520	Ørsta
79	Noeddal/Stranda	1524	Noeddal
79	Noeddal/Stranda	1525	Stranda
80	Rauma	1539	Rauma
81	Sandøy	1546	Sandøy
82	Sunnadal	1563	Sunnadal
83	Surnadal	1566	Surnadal
83	Surnadal	1567	Rindal
83	Surnadal	1571	Halsa

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunene.	Kommunenavn
84	Smøla	1573	Smøla
85	Aure	1576	Aure
86	Trondheim	1601	Trondheim
86	Trondheim	1624	Rissa
86	Trondheim	1648	Midtre Gauldal
86	Trondheim	1653	Melhus
86	Trondheim	1657	Skaun
86	Trondheim	1662	Klæbu
86	Trondheim	1663	Malvik
86	Trondheim	1664	Selbu
86	Trondheim	1714	Stjørdal
86	Trondheim	1718	Leksvik
87	Hemne	1612	Hemne
88	Hitra/Froya	1617	Hitra
88	Hitra/Froya	1620	Froya
89	Orland	1621	Orland
89	Orland	1627	Bjågn
90	Åfjord/Roan	1630	Åfjord
90	Åfjord/Roan	1632	Roan
91	Osen	1633	Osen
92	Oppdal/Rennebu	1634	Oppdal
92	Oppdal/Rennebu	1635	Rennebu
93	Orkdal	1613	Snållfjord
93	Orkdal	1622	Agdenes
93	Orkdal	1636	Meldal
93	Orkdal	1638	Orkdal
94	Roros	0441	Os
94	Roros	1640	Roros
94	Roros	1644	Holtålen
95	Tydal	1665	Tydal
96	Steinkjer	1702	Steinkjer
96	Steinkjer	1724	Verran
96	Steinkjer	1756	Inderøy
96	Steinkjer	1736	Snåsa
97	Namsos	1703	Namsos
97	Namsos	1725	Namdalseid
97	Namsos	1742	Grong
97	Namsos	1743	Høylandet
97	Namsos	1744	Overhalla
97	Namsos	1748	Fosnes
98	Meråker	1711	Meråker
99	Levanger/Verdal	1717	Frosta

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
99	Levanger/Verdal	1719	Levanger
99	Levanger/Verdal	1721	Verdal
100	Lierne	1738	Lierne
101	Røyrvik	1739	Røyrvik
102	Namsskogan	1740	Namsskogan
103	Flatanger	1749	Flatanger
104	Vikna/Nærøy	1750	Vikna
104	Vikna/Nærøy	1751	Nærøy
105	Leka	1755	Leka
106	Bodo	1804	Bodo
106	Bodo	1838	Gåldeskål
107	Narvik	1805	Narvik
107	Narvik	1854	Ballangen
107	Narvik	1919	Gratangen
108	Bindal	1811	Bindal
109	Bronnøy	1812	Sømna
109	Bronnøy	1813	Bronnøy
109	Bronnøy	1815	Vega
109	Bronnøy	1816	Vevelstad
110	Alstahaug	1818	Herøy
110	Alstahaug	1820	Alstahaug
110	Alstahaug	1822	Leirfjord
110	Alstahaug	1827	Donna
111	Vefsn	1824	Vefsn
111	Vefsn	1825	Grane
112	Hattfjelldal	1826	Hattfjelldal
113	Nesna	1828	Nesna
114	Rana	1832	Hemnes
114	Rana	1833	Rana
115	Lurøy	1834	Lurøy
116	Trøna	1835	Trøna
117	Rødøy	1836	Rødøy
118	Meløy	1837	Meløy
119	Beiarn	1839	Beiarn
120	Fauske	1840	Saltdal
120	Fauske	1841	Fauske
120	Fauske	1845	Sørfold
121	Steigen	1848	Steigen
122	Hamarøy Håbme	1849	Hamarøy Håbme
123	Tysfjord	1850	Tysfjord
124	Lødingen	1851	Lødingen
125	Røst	1856	Røst

(Gundersen & Juvkam 2013)

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
126	Værøy	1857	Værøy
127	Vestvågøy	1859	Flakstad
127	Vestvågøy	1860	Vestvågøy
128	Vågan	1865	Vågan
129	Sortland	1866	Hadsel
129	Sortland	1867	Bø
129	Sortland	1868	Øksnes
129	Sortland	1870	Sortland
130	Andøy	1871	Andøy
131	Moskenes	1874	Moskenes
132	Tromsø	1902	Tromsø
132	Tromsø	1936	Karlsøy
133	Harstad	1853	Evenes
133	Harstad	1852	Tjeldsund
133	Harstad	1903	Harstad
133	Harstad	1911	Kværfjord
133	Harstad	1913	Skånland
134	Ibestad	1917	Ibestad
135	Salangen	1920	Lavangen
135	Salangen	1923	Salangen
136	Målselv	1922	Bardu
136	Målselv	1924	Målselv
137	Torsken/Berg	1928	Torsken
137	Torsken/Berg	1929	Berg
138	Lenvik	1925	Sørreisa
138	Lenvik	1926	Dyrøy
138	Lenvik	1927	Tranøy
138	Lenvik	1931	Lenvik
139	Balsfjord/Storfjord	1933	Balsfjord
139	Balsfjord/Storfjord	1939	Storfjord
140	Lyngen	1938	Lyngen
141	Gáivuotna - Kålfjord	1940	Gáivuotna - Kålfjord
142	Skjervøy/Nordreisa	1941	Skjervøy
142	Skjervøy/Nordreisa	1942	Nordreisa
143	Kvænangen	1943	Kvænangen
144	Vardo	2002	Vardo
145	Vadsø	2003	Vadsø
145	Vadsø	2027	Unjárga - Nesseby
146	Hammerfest	2004	Hammerfest
146	Hammerfest	2017	Kvalsund
147	Guovdageaidnu-Kautokeino	2011	Guovdageaidnu-Kautokeino
148	Alta	2012	Alta

BA-regionnr.	BA-regionnavn	Kommunenr.	Kommunenavn
149	Loppa	2014	Loppa
150	Hasvik	2015	Hasvik
151	Måsøy	2018	Måsøy
152	Nordkapp	2019	Nordkapp
153	Porsanger Porsángu Porsanki	2020	Porsanger Porsángu Porsanki
154	Kárásjohka - Karasjok	2021	Kárásjohka - Karasjok
155	Lebesby	2022	Lebesby
156	Gamvik	2023	Gamvik
157	Berlevåg	2024	Berlevåg
158	Deatnu - Tana	2025	Deatnu - Tana
159	Båtsfjord	2028	Båtsfjord
160	Sør-Varanger	2030	Sør-Varanger

(Gundersen & Juvkam 2013)