

Intergenerasjonell inntektsmobilitet: En undersøkelse av persistens i inntektsulikhet mellom innvandrere og nordmenn i Norge

Magnus Pettersen Nytnun

Masteroppgave

Masteroppgaven er levert for å fullføre graden

Master i samfunnsøkonomi

Universitetet i Bergen, Institutt for økonomi

[Juni 2019]



UNIVERSITETET I BERGEN

Forord

Jeg vil begynne med å takke Espen Bratberg for mange gode og konstruktive tilbakemeldinger. Samtidig vil jeg takke NSD og SSB for data, og de ansatte som drifter Microdata.

Jeg vil også takke gutta på datarommet. Dette semesteret har gått fort, og dette skyldes i høy grad den gode stemningen på datarommet. I tillegg har det også vært mulig å diskutere faglige spørsmål, dette har vært til stor hjelp. Jeg vil også takke Oscar M. Johannessen som tok seg tid til å korrekturlese oppgaven.

Til slutt vil jeg rette en takk til Karoline, og familien min for støtten de har bidratt med i dette, og tidligere semestre.

Sammendrag

Norge har i de siste årene hatt en økning i sosial ulikhet. En andel av denne økningen i sosial ulikhet skyldes innvandring fra lavinntektsland. Dette gjelder spesielt i Oslo. Ulike faktorer gjør at innvandrere fra lavinntektsland oftere har lavere lønn enn en gjennomsnittlig annengenerasjonsnordmann. Intergenerasjonell inntektsmobilitet er knyttet tett opp mot sosial ulikhet, da spesielt inntektsulikhet. Dette kommer av at intergenerasjonell inntektsmobilitet vil si noe om hvor vedvarende inntektsulikhet er over generasjoner. Denne oppgaven ønsker å undersøke hvor vedvarende inntektsulikheten mellom nordmenn og innvandrere i Norge er. Dette gjøres ved å sammenligne annengenerasjonsnordmenn med annengenerasjonsinnvandrere i Norge, og Oslo.

For å besvare problemstillingen brukes fire ulike målemetoder. Disse målemetodene omfatter intergenerasjonell inntektselastisitet, overgangsmatriser, overgangssannsynligheter og retningsrankmobilitet. Disse målemetodene har ulike egenskaper som komplementerer hverandre godt, i tillegg maler de et noe ulikt bilde. Den intergenerasjonelle inntektselastisiteten gir estimater som er lett å sammenligne med resultater fra tidligere forskning. De tre resterende målemetodene gir estimater som illustrerer individenes bevegelse i inntektsfordelingen relativt til far. De vil også si noe om i hvilken retning og hvor langt individer beveger seg i inntektsfordelingen. Dette er elementer som ikke har blitt kartlagt tidligere

Resultatene fra den intergenerasjonelle inntektsmobiliteten tilsier at annengenerasjonsinnvandrere er mye mer mobile enn annengenerasjonsnordmennene i utvalget. Resultatene fra overgangsmatrise, overgangssannsynligheter og retningsrankmobilitet tilsier at innvandrere er mindre tilbøyelig til å foreta korte byks i inntektsfordelingen. Dette betyr at en større andel annengenerasjonsinnvandrere forblir i bunnen av inntektsfordelingen, sammenlignet med annengenerasjonsnordmenn. Målemetodene viser også annengenerasjonsinnvandrere har en større tilbøyelighet til å foreta lengre byks i inntektsfordelingen, sammenlignet med annengenerasjonsnordmenn. Det er altså en todeling blant innvandrere. Dette endrer seg når en sammenligner annengenerasjonsnordmenn og -innvandrere i Oslo. I Oslo er annengenerasjonsnordmenn mer tilbøyelige til å foreta korte, og lengre byks i inntektsfordelingen.

Innholdsfortegnelse

Innholdsfortegnelse	iv
Tabelliste	vi
Figurliste.....	vii
Kapittel 1	1
Innledning.....	1
Kapittel 2	3
Målemetoder.....	3
2.1 Intergenerasjonell inntektselastisitet	3
2.1.1 Intergenerasjonell inntektselastisitet.....	3
2.1.2 Korrelasjonskoeffisienten.....	4
2.1.3 Kritikk	5
2.2 Målefeil	5
2.2.1 Klassisk målefeil	6
2.2.2 Livssyklusshjevhet	7
2.3 Overgangsmatriser.....	8
2.4 Overgangssannsynlighet.....	9
2.5 Retningsrankmobilitet	11
2.6 Sammenligning av overgangssannsynligheter og retningsrankmobilitet	13
Kapittel 3	14
Litteraturoversyn	14
3.1 Becker-Tomes og hovedlinjer	14
3.1.1 Becker-Tomes-modellen	14
3.1.2 Hovedlinjer i litteraturen	16
3.2 Norge sammenlignet med resten av verden.....	16
3.3 Utvikling i intergenerasjonell inntektsmobilitet i Norge og forskjeller i regioner	18
3.4 Innvandrere i Norge.....	20
3.5 Innvandring og mobilitet	21
3.6 Oppsummering	22
Kapittel 4.....	23
Data og utvalgsdesign	23
4.1 Data	23
4.2 Utvalgsdesign	24
4.2.1 Inklusjonskriterier.....	24
4.2.2 Begrunnelse av utvalg	24
4.3 Deskriptiv statistikk.....	28
4.3.1 Oppsummering av deskriptiv statistikk	35

Kapittel 5	36
Resultater.....	36
5.1 IGE spesifikasjoner og resultater.....	36
5.1.1 IGE spesifikasjoner	36
5.1.2 IGE resultater	38
5.2 Overgangsmatriser.....	41
5.3 Resultater fra overgangssannsynligheter	42
5.4 Resultater fra retningsrankmobilitet	47
5.5 Diskusjon.....	52
5.5.1 Forskjeller i geografiske inndelinger	54
Kapittel 6	54
Konklusjon	54
Litteraturliste	57
Appendiks A.....	60
Tester.....	60
F-test.....	63
Modell 1, uten bostedsvariabel:.....	63
Modell 2, med Oslo som bostedsvariabel:	63
Modell 3, med Stor-Oslo som bostedsvariabel:.....	63
Appendiks B	64
Diverse	64
Appendiks C	65
Microdata	65
Microdata	65
Konfidensialitet	65
Brukergrensesnitt.....	65
Variabler.....	66
Microdata skript	66

Tabelliste

Tabell 1: Opphavsland m.m.	29
Tabell 2: Diverse karakteristika.....	30
Tabell 3: Fordeling over inntektskvintiler.....	31
Tabell 4: Gjennomsnittslønn i ulike regioner.....	32
Tabell 5: Utdanning over region.....	34
Tabell 6 : Utdanning over inntektskvintil for innvandrere.....	35
Tabell 7: Utdanning over inntektskvintil for nordmenn.....	35
Tabell 8: Resultater fra IGE-estimer betinget på ulike regioner.....	38
Tabell 9: Resultater fra IGE-estimeringer med interaksjonsledd.....	39
Tabell 10: Resultater fra overgangsmatrise.....	41
Tabell 11: Sensitivitetstest for hvor mange år fars inntektsmål baserer seg på.....	60
Tabell 12: Sensitivitetstest for alder ved immigrasjon.....	61
Tabell 13: Sensitivitetstest for fars målealder.....	62
Tabell 14: Korrelasjonskoeffisienten.....	64
Tabell 15: Definisjon av Stor-Oslo.....	64

Figurliste

Figur 1: Innvandrerbefolkning fra 1970 til og med 2018.....	2
Figur 2: UTP, Norge, $\tau = 0$	43
Figur 3: UTP, Stor-Oslo, $\tau = 0$	43
Figur 4: UTP, Oslo, $\tau = 0$	43
Figur 5: UTP, Norge, $\tau = 0,3$	44
Figur 6: UTP, Stor-Oslo, $\tau = 0,3$	44
Figur 7: UTP, Oslo, $\tau = 0,3$	44
Figur 8: DTP, Norge, $\tau = 0$	45
Figur 9: DTP, Stor-Oslo, $\tau = 0$	45
Figur 10: DTP, Oslo, $\tau = 0$	45
Figur 11: URM, Norge, $\tau = 0$	48
Figur 12: URM, Stor-Oslo, $\tau = 0$	48
Figur 13: URM, Oslo, $\tau = 0$	48
Figur 14: URM, Norge, $\tau = 0,3$	49
Figur 15: URM, Stor-Oslo, $\tau = 0,3$	49
Figur 16: URM, Oslo, $\tau = 0,3$	49
Figur 17: DRM, Norge, $\tau = 0$	50
Figur 18: DRM, Stor-Oslo, $\tau = 0$	50
Figur 19: DRM, Oslo, $\tau = 0$	50

Kapittel 1

Innledning

Norge er et land med en sterk velferdsstat og høy grad av likhet mellom sine borgere, og et av grunnprinsippene er like muligheter uavhengig av individets bakgrunn. På tross av dette observerer vi at ulikhetene i Norge har økt de siste årene (SSB, 2018a, s.14). Dette har gitt opphav til debatt. Debatten omhandler i stor grad hva som driver den økte ulikheten i Norge. Mange peker på økt innvandring som en grunn. I 2017 publiserte SSB en artikkel i sitt tidsskrift *Økonomiske Analyser* som viser at økt innvandring er en vesentlig faktor i å forklare økningen i inntektsulikhetene fra 2004 til 2015 (SSB, 2017, s. 89). Innvandrerne som immigrerer til Norge har dårligere forutsetninger for å klare seg i det norske arbeidsmarkedet, de kan mangle relevant utdanning, de kan ha manglende språkkunnskaper, i tillegg kan kulturelle forskjeller gjøre at en lavere andel innvandrerkvinner sysselsettes sammenlignet med norske kvinner. Disse faktorene vil gi utslag i sosioøkonomiske forskjeller mellom innvandrere og nordmenn i Norge. Økning i inntektsulikhetene har hovedsakelig vært konsentrert i de store byene, dette er også der innvandringen har vært størst. Oslo-regionen, sammen med Stockholms-regionen, er de to regionene i Norden med størst ulikhet (SSB, 2014).

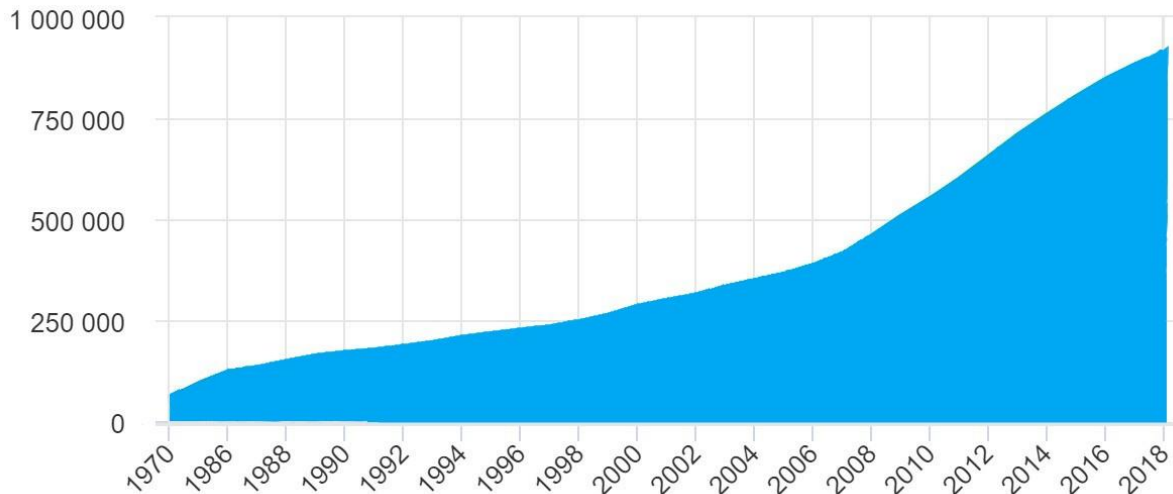
Intergenerasjonell inntektsmobilitet (IGM) er tett knyttet til ulikheter. IGM handler om i hvilken grad barn arver foreldrenes posisjon i inntektsfordelingen. Lav IGM tilsier at barn i høy grad arver foreldrenes posisjon i inntektsfordelingen. Dette betyr at barn født i hjem med lav inntekt oftere vil ha lav inntekt når de selv er blitt voksne. Høy IGM tilsier at denne sammenhengen er mindre viktig, altså vil barn født i lavinntektsfamilier ha en mulighet til å klatre i inntektsfordelingen. En observerer ofte at land med lav inntektsulikhet har høyere grad IGM (Bratberg, Nilsen & Vaage, 2005). Dette kan ha med hvilke politiske mål landet har, samtidig som høy mobilitet vil bety at økonomiske ulikheter er mindre vedvarende. Nettopp derfor er det viktig å undersøke IGM. Spesielt for innvandrere i Norge, når ulikheten mellom nordmenn og innvandrere er såpass påfallende. IGM hos innvandrere vil si noe om integreringsprosessen i det norske samfunnet, samtidig vil det si noe om hvor vedvarende den observerte inntektsforskjellen mellom innvandrere og nordmenn er.

Denne oppgaven har som formål å kartlegge forskjeller i IGM hos innvandrere og nordmenn i Norge, og undersøker også ulike regionale inndelinger. Grunnen til at ulike regionale inndelinger blir undersøkt er fordi innvandrerpopulasjonen er konsentrert i byer, hovedsakelig Oslo og Oslo-området. I tillegg til dette er også ulikhetene størst i disse områdene. På denne måten kan vi kartlegge hvor vedvarende ulikhetene er i Oslo-området. Samtidig kan vi peke på noen mulige mekanismer bak de observerte ulikhetene mellom innvandrere og nordmenn i Norge, Oslo og Oslo-området.

De største bølgene med innvandrere kom til Norge på 1970-tallet, og andelen innvandrere i Norge har siden den gang økt fra 1,5% til 17,3% i 2018 (SSB, 2018b). Dette betyr at innvandring i stor skala

fremdeles er et relativt nytt fenomen i Norge. Dette skaper noen utfordringer når en skal måle IGM mellom innvandrere og deres barn i Norge. Dette er hovedsakelig fordi gruppen barn av innvandrerforeldre i rett målealder er relativt liten. Dermed blir det også naturlig at en vesentlig del av oppgaven omhandler disse utfordringene, og hvordan en kan løse utfordringene på best mulig måte.

Figur 1: Innvandrerbefolkning fra 1970 til og med 2018.



I figuren ovenfor kan en se hvordan innvandrerbefolkningen har økt fra 1970 til og med 2018. Figur 1. Innvandrerbefolkning. Fra «Innvandrere og norskfødte med innvandrerforeldre», av SSB, 2018c. (<https://www.ssb.no/innvandring-og-innvandrere/faktaside/innvandring>).

I denne oppgaven sammenlignes barn av andregenerasjonsnordmenn og andregenerasjonsinnvandrere født 1977 – 1982. Andregenerasjonsinnvandrere omfatter i denne oppgaven også første generasjonsinnvandrere som innvandret før skolealder, slik at utvalget blir tilstrekkelig stort. Dataen oppgaven bygger på stammer fra et samarbeid mellom SSB og Norsk Senter for forskningsdata (NSD), og er av høy kvalitet.

For å kartlegge IGM benytter jeg meg av en rekke ulike målemetoder. Disse blir presentert i kapittel 2, og omfatter intergenerasjonell inntektselastisitet, overgangsmatriser, overgangssannsynligheter og retningsrankmobilitet. Ved hjelp av disse målemetodene blir det mulig å kartlegge de norske innvandrerbarnas bevegelse i inntektsfordelingen relativt til far, noe som ikke før har blitt gjort. For å undersøke ulike regionale inndelinger estimeres mål på IGM betinget på hele Norge, Oslo og en region definert som Stor-Oslo. Stor-Oslo består av omkringliggende kommuner rundt Oslo og Oslo, og er basert på Oslos pendleromland¹.

Oppgaven er bygd opp på følgende måte: I kapittel 2 presenteres målemetodene som blir benyttet. I kapittel 3 blir relevant litteratur presentert. I kapittel 4 gjøres det rede for data, utvalgsdesign og

¹ Pendleromlandet baserer seg på definisjonen til Storbymeldingen (Kommunal- og Moderniseringsdepartementet, 2003). Se appendiks for [definisjon av Stor-Oslo](#).

deskriptiv statistikk. I kapittel 5 presenteres og drøftes resultatene fra de ulike målemetodene. I kapittel 6 blir det presentert en konklusjon.

Kapittel 2

Målemetoder

Det er mange forskjellige metoder å måle IGM på. I denne oppgaven tar jeg utgangspunkt i hovedsakelig fire forskjellige metoder. Først presenteres intergenerasjonell inntektselastisitet som et mål på IGM. Intergenerasjonell inntektselastisitet viser hvor korrelert barnas inntekt er til foreldrenes inntekt, og i hvilken grad ulikheter vedvarer over tid. Målefeil, som er presentert i delkapittel 2,2, er et veldig viktig tema når en ønsker å undersøke IGM. Målefeil gjør seg enda mer gjeldende i denne oppgaven, hvor det er utfordringer knyttet til innvandrerutvalget. I delkapittel 2,3 presenteres overgangsmatriser. I delkapittel 2,4 presenteres overgangssannsynligheter. Deretter, i delkapittel 2,5 presenteres retningsrankmobilitet. Disse to målene ligner på hverandre, og de rapporteres ofte sammen. De gir estimater som illustrerer individenes bevegelse i inntektsfordelingen relativt til far. De vil også si noe om i hvilken retning og hvor langt individer beveger seg i inntektsfordelingen. Det vil også overgangsmatrisene gjøre. I delkapittel 2,6 sammenlignes overgangssannsynligheter og retningsrankmobilitet som mål på IGM.

2.1 Intergenerasjonell inntektselastisitet

I dette delkapitlet presenteres først IGE som mål på IGM. Deretter presenteres korrelasjonskoeffisienten, etterfulgt av kritikken rettet mot IGE.

2.1.1 Intergenerasjonell inntektselastisitet

Et enkelt og ofte brukt mål på IGM er intergenerasjonell inntektselastisitet (IGE). IGE viser hvor vedvarende forskjeller i inntekter er over generasjoner. Når en estimerer IGE er det følgende sammenheng en estimerer:

$$y_i^C = \alpha + \beta * y_i^P + \varepsilon_i \quad (1)$$

Her representerer y_i livstidsinntekt på log-form for individ i , og toppskrift C refererer til barn, toppskrift P til foreldre, og α er et konstantledd. β er koeffisienten til y_i^P , og er et estimat for IGE. Lavere β betyr at forskjeller i inntekt er mindre vedvarende over generasjoner, altså høyere grad av

IGM. Som mål på mobilitet kan en bruke $(1 - \beta)$, og i et samfunn med perfekt mobilitet er dette uttrykket lik én.

Som sagt er β et mål på hvor vedvarende inntektsulikhet er over generasjoner. Hvis $\beta < 1$ vil forskjeller i inntekt bli visket ut over tid. Hvis $\beta > 1$ vil forskjeller i inntekt øke over generasjoner.

Vi antar at $\beta = 0,15$ i et samfunn. Vi ønsker å undersøke hvor vedvarende inntektsulikheten er mellom to familier. Den ene familien tjener 100 % mer enn den andre. Ved hjelp av utregning kan vi vise at denne forskjellen kun vil utgjøre 11 %² i neste generasjon.

2.1.2 Korrelasjonskoeffisienten

Alternativt til IGE kan en rapportere korrelasjonskoeffisienten, som er gitt ved:

$$\rho = \beta * \frac{SD(y_i^P)}{SD(y_i^C)} \quad (2)$$

En kan merke seg at $\rho = \beta$ dersom $SD(y_i^P) = SD(y_i^C)$.

$SD(y_i^P)$ er standardavviket til foreldrenes inntektsfordeling, og $SD(y_i^C)$ er standardavviket til barnas inntektsfordeling. ρ blir kalt for korrelasjonskoeffisienten, og tolkes på lignende måte som IGE. Høyere korrelasjon betyr mindre mobilitet, og lavere korrelasjon betyr mer mobilitet. I motsetning til IGE er $\rho \in [0,1]$ ³. Hvis $\rho = 1$ betyr dette at barnets inntekt er helt avhengig av foreldrenes inntekt. $\rho = 0$ betyr at barnets inntekt er helt uavhengig av foreldrenes inntekt.

$SD(y_i^P)$ og $SD(y_i^C)$ blir ansett som et mål på inntektsulikhet i samfunnet, hvor høyere standardavvik betyr mer inntektsulikhet. Endringer i grad av inntektsulikhet i samfunnet påvirker IGE-estimatet. Hvis for eksempel barna opplever høyere grad av inntektsulikhet vil dette føre til en økning i IGE. Dette er en mekanisk sammenheng. Korrelasjonskoeffisienten på en annen side blir ikke påvirket av endringer i inntektsulikhet. Dette kan være en nyttig egenskap dersom inntektsfordelingen har endret seg drastisk fra en generasjon til en annen (Jäntti et al., 2006, s. 8).

² Y: Inntekt er på nivå. Ser bort fra ε . Dette gir: $Y_i^C = (Y_i^P)^\beta * \exp(\alpha)$. Forhold mellom familie A og B vil da

være $\frac{Y_A^C}{Y_B^C} = \left(\frac{Y_A^P}{Y_B^P}\right)^\beta = 2^{0,15} = 1,11$.

³ I teorien vil den være $\rho \in [-1,1]$, men en vil ikke observere en negativ korrelasjon mellom barn og foreldres inntekt (Black & Devereux, 2011).

2.1.3 Kritikk

IGE bygger på linearitet. Linearitet i denne sammenhengen vil bety at alle er like mobile uavhengig av hvor de er plassert i inntektsfordelingen. Dette er en sterk antakelse. For eksempel kan en ofte observere noe lavere intergenerasjonell inntektsmobilitet i bunnen av inntektsfordelingen, sammenlignet med toppen. Lavere mobilitet i bunnen av inntektsfordelingen kan reflektere nedarving av lavere preferanser for utdanning, og samtidig foreldrenes aspirasjoner for barna sine (Bratberg, Nilsen & Vaage, 2007, s. 114).

Et IGE-estimat vil ikke si noe om hvilken retning mobiliteten trekker. Dersom en finner ved hjelp av IGE at innvandrere er mer mobile enn nordmenn, kan det være nærliggende å tenke at dette er bra for innvandrere, som ofte kommer fra lavinntektsfamilier. I realiteten kan dette være svært negativt. For eksempel kan det tenkes at det høye IGE-estimatet reflekterer at en høyere andel innvandrere fra rikere familier faller nedover i inntektsfordelingen, sammenlignet med nordmenn. Samtidig kan det tenkes at en høyere andel forblir i lavere inntektsprosentiler. Dette tilsier en reduksjon i ulikhet hos innvandrere, men samtidig tilsier det også at innvandrerne har blitt relativt mer fattige. Ved kun å benytte IGE som mål på mobilitet vil en altså gå glipp av viktige aspekter ved IGM. Disse aspektene er meget viktige for å ha en fornuftig diskusjon rundt IGM, spesielt når en sammenligner to undergrupper innad i en populasjon.

Til slutt kan det nevnes at IGE har viktige begrensninger når en skal sammenligne to undergrupper i et utvalg. Dette er fordi IGE kun vil si noe om hvilken grad barnas inntekt i en spesifikk gruppe beveger seg til gruppens inntektsgjennomsnitt (Bhattacharya & Mazumder, 2011, s. 336). Dermed får en ikke et mål på barnas IGM relativt til samfunnet. En får heller et mål på i hvilken grad barn av fedre tilhørende en viss undergruppe konvergerer til inntektsgjennomsnittet innad i denne undergruppen. Dette kan være problematisk hvis inntektsgjennomsnittet i de to undergruppene er svært ulikt. Metodene presentert under gjør opp for dette ved å ta utgangspunkt i en inntektsfordeling som baserer seg på inntekten til alle fedrene fra utvalget uavhengig av karakteristika. På denne måten kan vi få mer utfyllende informasjon om forskjeller i mobilitet hos for eksempel annengenerasjonsinnvandrere og nordmenn i Norge.

2.2 Målefeil

Når en måler IGM i et samfunn ønsker en å kartlegge i hvilken grad barnets posisjon i inntektsfordelingen avhenger av foreldrenes inntektsposisjon. For å gjøre dette trenger en permanentinntekten eller et representativt mål på permanentinntekten til foreldre og barn. Det er sjeldent at en har permanentinntekten til et individ tilgjengelig, og det er heller ikke nødvendig. Det er viktigere å bruke et godt og representativt mål på permanentinntekt til foreldre og barn. Dersom en frastår fra dette kan det oppstå skjevheter i estimatet som følge av målefeil. For å unngå målefeil bør

en sørge for at 1) inntekten er målt over flere år, og 2) sørge for at alder til foreldre og barn på måletidspunktet er passende.

2.2.1 Klassisk målefeil

Hvis inntekten ikke er målt over flere år, slik som beskrevet i 1), kan en få problemer med at inntektsmålet blir påvirket av tilfeldige sjokk, såkalt klassisk målefeil. Dersom en baserer inntektsmålet sitt over kun et år, vil inntekten en måler, for foreldrene, være gitt ved:

$$y_{is}^P = y_i^P + w_{is}^P \quad (3)$$

Hvor y_{is}^P er målt inntekt for individ i , målt i tidspunkt s , der P indikerer at det gjelder foreldre. y_i^P er den faktiske permanentinntekten. w_{is}^P er et tilfeldig sjokk som påvirker inntektsmålet det året en foretar målingen. Altså vil den målte inntekten være ulik permanentinntekten avhengig av verdien til w_{is}^P .

Tilsvarende for barna har du:

$$y_{it}^C = y_i^C + w_{it}^C \quad (4)$$

Her blir tolkningen tilsvarende, y_{it}^C representerer inntektsmålet til individ i målt på tidspunkt s , der C indikerer at det er barnas inntekt som måles. y_i^C er den faktiske permanentinntekten til individet. w_{it}^C er igjen et tilfeldig sjokk på inntektsmålet.

Videre antar vi at variansen til w_{is}^P og w_{it}^C er gitt ved σ_{Pw}^2 og σ_{Cw}^2 , og at disse ikke er korrelert med hverandre eller y_i^P og y_i^C . Dette er sterke antakelser, men uten dem vil sammenhengen bli meget komplisert å utlede. Vi estimerer sammenhengen illustrert ved likning (1):

$$y_i^C = \alpha + \beta * y_i^P + \varepsilon_i \quad (1)$$

Tolkningen av sammenhengen i (1) er tilsvarende som ovenfor. Vi antar for enkelthetsens skyld at variansen til inntekt er lik for fedre og barn, dermed er variansen for y_i^P og y_i^C gitt ved σ_y^2 . Estimatet for β er gitt ved $\hat{\beta}$, og når vi bruker (3) og (4) som mål på inntekt vil sannsynlighetsgrensen for $\hat{\beta}$ være:

$$plim \hat{\beta} = \beta \frac{\sigma_y^2}{(\sigma_y^2 + \sigma_{Pw}^2)}$$

Det kan vises at:

$$\beta \frac{\sigma_y^2}{(\sigma_y^2 + \sigma_{Pw}^2)} < \beta$$

Som betyr at estimatoren vår er forventningsskjev mot null. Ofte vil en bruke et inntektsgjennomsnitt basert på flere år, anta at det baseres på T år. Da vil vi få:

$$\bar{y}_i^P = \sum_{s=1}^T \frac{y_i^P}{T}$$

Dette vil gi oss følgende:

$$plim \hat{\beta} = \beta \frac{\sigma_y^2}{(\sigma_y^2 + \frac{\sigma_{Pw}^2}{T})}$$

Her kan vi se at når T øker vil $plim \hat{\beta} \rightarrow \beta$, altså vil ikke IGE-estimatet lengre være forventningsskjev mot null når en måler inntekten over tilstrekkelig mange år.

Solon (1992) og Mazumder (2005) understreker hvor viktig det var å måle inntekten over flere år, og viste at målinger over en for kort periode kan føre til at IGE estimatet blir forventningsskjev mot null. Nilsen, Vaage, Aakvik og Jacobsen (2012) viser ved hjelp av norske data at antall år inntekten måles over påvirker verdiene til IGE i Norge, men at andre kilder til målefeil er mer utslagsgivende.

2.2.2 Livssyklusskjevhet

Det å ikke måle inntekten til far og sønn i riktig alder, som beskrevet i 2) kan føre til forventningsskjeve IGE estimat. Denne type målefeil blir ofte referert til som livssyklusskjevhet. Anta at inntekten blir målt i en gitt alder, a:

$$y_a^P = \omega_a y^P + w^P \quad (5)$$

$$y_a^C = \varphi_a y^C + w^C \quad (6)$$

ω_a og φ_a er koeffisienter som viser hvor godt y_a^P og y_a^C representerer permanentinntekten til individet. Variablene w^P og w^C representerer igjen tilfeldige sjokk. Dersom ω_a og φ_a er én, og w^P og w^C er null, vil y_a^P og y_a^C være perfekte mål på permanentinntekt. Det kan vises, dersom w^P og w^C er ukorrelert, at:

$$plim \hat{\beta} = \beta \frac{\omega_a \varphi_a \sigma_y^2}{(\omega_a^2 \sigma_y^2 + \sigma_{Pw}^2)}$$

Livssyklususkjevhet oppstår dersom φ_a eller ω_a endrer seg ved i hvilken alder en måler inntekten for foreldre og barn. Hvis disse er lik én vil inntektsmålene som sagt være helt representativt for permanentinntekten til et individ, dette er typisk ikke tilfelle⁴.

Dersom en har inntektsdata for individer gjennom livet deres er det mulig å estimere ω_a og φ_a . På denne måten kan en kartlegge når det er best å måle inntekt til foreldre og barn. Haider og Solon (2006) og Böhlmark og Lindquist (2006) har estimert ω_a og φ_a . De finner at det er optimalt å måle barnas inntekt når barna er eldre enn 30 år, og yngre enn 50 år. For foreldrene er optimal målealder et sted mellom 30 og 40 år. Nilsen et al. (2012) skriver også om livssyklususkjevhet, og finner tilsvarende funn som Haider og Solon (2006) og Böhlmark og Lindquist (2006) med norske data.

En kan merke seg at målefeil på fars inntekt kan skape forventningsskjevhet i IGE-estimatet, dette gjelder både klassisk målefeil og livssyklususkjevhet. For barna vil kun livssyklususkjevhet forårsake forventningsskjevhet i IGE-estimatet. Klassisk målefeil på barnets inntektsmål vil kun gi et mindre presist IGE-estimat. Dette kommer av at tilfeldige sjokk på barnets inntektsmål vil legge seg i feilledet i ligning (1).

2.3 Overgangsmatriser

I en overgangsmatrise viser hver celle sannsynligheten for at et barn havner i inntektskvantil j, gitt at foreldrene er i inntektskvantil k. Slike matriser spesifiseres ofte i kvartiler eller i kvintiler, hvor sistnevnte vil gi et mest detaljert bilde. I denne oppgaven fordeles foreldrene og barna inn i kvintiler, dette gir en (5 x 5) matrise, altså fem rader og fem kolonner. I en overgangsmatrise vil den vertikale dimensjonen representere foreldrene inntektskvantil, og den horisontale vil representere barnets inntektskvantil. Første kvantil vil representere de 20 % med lavest inntekt i inntektsfordelingen, og andre kvantil vil representere dem med inntekt som plasserer dem mellom de 21 – 40 % mest fattige i inntektsfordelingen, og så videre. Når en benytter seg av kvintiler vil en verdi på 0,20 i hver celle tilsi et samfunn med perfekt mobilitet. På motsatt side vil en matrise som gjenspeiler et samfunn med perfekt immobilitet vise at alle barn av fedre med for eksempel j = 1, selv havner i k = 1. Altså vil hver celle diagonalt i matrisen vise én.

Ved hjelp av overgangsmatriser kan en også beregne immobilitetsindeks. Dette gir et tall som er lett å sammenligne. I tillegg tar immobilitetsindeksen i betraktning at det er en større endring å for eksempel gå fra nederste kvantil til høyeste kvantil, sammenlignet med å gå fra nest øverste til øverste kvantil. Måten en estimerer denne verdien er ved å summere verdien til diagonalen, og verdiene til cellene rett over og/eller under diagonalen. I en situasjon med perfekt mobilitet vil denne indeksen gi et tall på 2,6, og i en situasjon med perfekt immobilitet vil indeksen gi et tall på fem.

⁴ Selv om disse er lik én vil en fremdeles ha problemer med klassisk målefeil dersom $\sigma_{0w}^2 \neq 0$.

Man kan også betinge overgangsmatrisene på at individene har spesielle karakteristika, for eksempel kan en betinge på at undergruppen er innvandrere eller nordmenn. Dermed får en to matriser som en kan sammenligne, hvor begge matrisene tar utgangspunkt i den samme inntektsfordelingen. En kan også merke seg at overgangsmatriser er intervallbaserte, med dette menes det at en studerer hvert kvintil hver for seg. Dette er i motsetning til retningsmobilitetsmålene som presenteres nedenfor⁵, dermed supplerer de ulike målemetodene hverandre godt. Det kan også nevnes at overgangsmatriser ikke bare vil gi et mål på oppover- og nedovermobilitet i et samfunn, de vil også vise til hvilket kvintil individet har beveget seg. I metodene presentert nedenfor vil en kun se at det har vært en bevegelse i inntektsfordelingen, og ikke hvor i inntektsfordelingen individet har havnet.

O'Neill, Sweetman og de gaer (2007) ser på konsekvenser av klassisk målefeil i overgangsmatriser. De undersøker hvordan klassisk målefeil i barnets inntekt, og i barnets og foreldrenes inntekt samtidig påvirker resultatene i overgangsmatriser. De finner at klassisk målefeil i barnets inntektsmål gjør at overgangsmatriser overestimerer mobiliteten i utvalget en studerer. Klassisk målefeil i både foreldres og barnets inntekt samtidig, skaper også forventingsskjeve resultater i overgangsmatriser. Retningen på skjevheten blir påvirket av i hvilken grad målefeilen forekommer samtidig i barnets inntekt, og foreldrenes inntekt. Dersom målefeilene sjeldent forekommer samtidig kan dette føre til skjevhet mot høyere mobilitet i matrisen. Hvis målefeilene ofte forekommer samtidig forårsaker dette en skjevhet mot mindre mobilitet i matrisen.

2.4 Overgangssannsynlighet

Oppoverovergangssannsynlighet (UTP⁶) viser sannsynligheten for at et barn tjener mer enn en gitt prosentil, ρ , gitt at foreldrene er posisjonert i en inntektsprosentil mindre eller lik ρ . Denne sannsynligheten kan formelt uttrykkes på følgende måte:

$$UTP_{\tau, \rho} = \Pr(Y_1 > \rho + \tau \mid Y_0 \leq \rho)$$

Her er Y_0 er barnas inntektsprosentil, Y_1 er foreldrenes inntektsprosentil, ρ er en gitt inntektsprosentil og τ er en vilkårlig grense. Hvis for eksempel $\rho = 20$ betyr dette at barnets foreldre må befinne seg i eller under den 20. inntektsprosentil. Og dersom en velger en $\tau = 0$ må barnet minst befinne seg i den 21. inntektsprosentilen for at det skal regnes som en bevegelse. Hvis en for eksempel setter $\tau = 0,1$ må barnet ha beveget seg minst til den 31. inntektsprosentilen for at det skal regnes som en bevegelse.

Slik UTP er formulert ovenfor er det et kumulativt mål, dette betyr at for eksempel med en $\rho = 0,1$ studerer vi alle barn med fedre som har en inntekt som plasserer dem i det laveste desilet. Dersom en justerer denne verdien til eksempelvis $\rho = 0,3$, vil vi nå studere alle barn med foreldre i de tre laveste

⁵ De kan også være intervallbaserte, men dette forårsaker mye støy når utvalget er lite. Dette forklares nærmere lengre nede i oppgaven.

⁶ UTP står for upward transition probability.

desil. En kan også studere intervall i inntektsfordelingen, dette kan gjøre det enklere å fastslå ved hvilken prosentil mobiliteten eventuelt endrer seg (Mazumder, 2014). Problemet er at intervallbaserte overgangssannsynligheter kan medføre mye støy, spesielt med færre observasjoner. Dette er grunnen til at det blir benyttet kumulativ UTP i denne oppgaven, da intervallbaserte blir preget av støy.

Nedoverovergangssannsynlighet (DTP⁷) viser sannsynligheten for at barnet havner i en inntektsprosentil lavere enn en gitt prosentil, ρ , gitt at foreldrene befinner seg i en inntektsprosentil større eller lik ρ . DTP kan uttrykkes på følgende måte:

$$DTP_{\tau,\rho} = \Pr(Y_1 < \rho + \tau \mid Y_0 \geq \rho)$$

Her har parameterne lik betydning som i UTP. Med verdier som $\rho = 0,9$ og $\tau = 0$ vil det regnes som en bevegelse nedover i inntektsfordelingen, dersom et barn faller til den 89. inntektsprosentilen eller lavere, når foreldrenes inntekt plasserer dem i den 90. inntektsprosentilen eller høyere. Med $\tau = 0,3$ må barnet plasseres i den 59. inntektsprosentilen eller lavere for at det skal regnes som en nedover bevegelse i inntektsfordelingen.

UTP og DTP kan, som med overgangsmatriser, betinges på ulike karakteristika, som bosted og hvorvidt personen er innvandrere eller ei. Når en betinger UTP eller DTP på ulike faktorer, kan de formuleres på følgende måter:

$$UTP_{\tau,\rho} = \Pr(Y_1 > \rho + \tau \mid Y_0 \leq \rho, x_j = X)$$

Og for DTP:

$$DTP_{\tau,\rho} = \Pr(Y_1 < \rho + \tau \mid Y_0 \geq \rho, x_j = X)$$

Hvor x_j representerer ulike karakteristika som er spesifikk med undergruppen en ønsker å undersøke.

I denne oppgaven foretas det estimeringer av overgangssannsynligheter for innvandrere og nordmenn separat. Disse blir betinget på ulike geografiske inndelinger. Disse geografiske inndelingene er definert som hele Norge, Stor-Oslo og Oslo. Selv om utvalget blir betinget på ulike karakteristika vil en i de to gruppene ta utgangspunkt i samme inntektsfordeling.

Overgangssannsynlighetene rapporteres med en grense, τ , hvor $\tau = 0$ og $\tau = 0,3$. Når $\tau = 0$ vil overgangssannsynlighetene måle kortere bevegelser i inntektsfordelingen. Når $\tau = 0,3$ vil en måle byks på minst tre desil i inntektsfordelingen. På denne måten kartlegges hvilken av undergruppene som er mest tilbøyelig å foreta små og store bevegelser i inntektsfordelingen. I valget av τ står en ovenfor en

⁷ DTP står for downward transition probability.

avveining, hvis τ er for lav vil en ikke i stor nok grad måle lengre byks i inntektsfordelingen. På en annen side vil en for høy τ ikke måle nok bevegelser, og resultatene kan bli preget av støy.

Overgangssannsynlighetene tar utgangspunkt i samme inntektsfordeling som overgangsmatrisen. Men i motsetning til overgangsmatrisen deles UTP og DTP opp i desil. Når en deler inntektsfordelingen opp i mindre enheter kan en se kortere bevegelser i inntektsfordelingen. På denne måten kan en kartlegge hvilke undergrupper som er mer tilbøyelig til å foreta små bevegelser i inntektsfordelingen. Det kan dog merkes at mindre måleenheter fører til mer støy i resultatene. Det blir derfor en avveining en må ta mellom å måle kortere bevegelser og støy. Ettersom innvandrerguppen er noe liten, er desil det mest nøyaktige målet som ikke blir for preget av støy.

Overgangssannsynlighetene rapporteres med 95% konfidensintervall i hvert målepunkt. Konfidensintervallet er konstruert ved hjelp av empiriske standardfeil⁸. Tilsvarende er også gjort for retningsrankmobilitet.

2.5 Retningsrankmobilitet

Retningsrankmobilitet (RM), utviklet av Bhattacharya og Mazumder (2011), er en ikke-parametrisk metode som måler retningsmobilitet. RM viser sannsynligheten for at et barn skal forbipassere foreldrenes posisjon i inntektsfordelingen, betinget på at foreldrenes inntekt er mindre eller lik et gitt prosentil. Oppoverretningsrankmobilitet (URM⁹) kan formuleres på følgende måte:

$$URM_{\tau,\rho} = \Pr(Y_1 - Y_0 > \tau \mid Y_0 \leq \rho)$$

Her representerer Y_i , $i = 0,1$, inntektsprosentilet for foreldre når $i = 0$, og barnas inntektsprosentil når $i = 1$. ρ bestemmer hvilken inntektsprosentil foreldrene til barna tilhører, dette betyr at hvis en setter $\rho \leq 0,01$ vil en måle oppovermobilitet for barn med foreldre i den aller laveste inntektsprosentilen. Med en $\rho \leq 0,5$ vil en måle oppovermobilitet for barn med foreldre i en av de 50. laveste inntektsprosentilene. På samme måte som i overgangssannsynligheter kan en måle mobilitet for mindre intervaller, dette kan gjøres ved å sette ρ til for eksempel $0,45 \leq \rho \leq 0,5$. Her vil en måle oppovermobiliteten til barn av foreldre fra og med 45. til og med 50. inntektsprosentil.

Variabelen τ indikerer igjen hvor mange prosentiler individet må bevege seg opp fra foreldrenes inntektsprosentil. τ kan altså bli ansett som en grense. Eksempelvis betyr dette at dersom τ er satt til 0 vil alle oppoverbevegelser bli målt. Ta for eksempel et barn som har foreldre i den første inntektsprosentilen, og som selv har havnet i den andre inntektsprosentilen. Dette barnet har med en τ

⁸ Dette er en tilnærming til konfidensintervallene benyttet i tidligere artikler. I tidligere artikler benyttes bootstrapping-metoden for å estimere konfidensintervall (Bhattacharya & Mazumder, 2011; Corak, Lindquist & Mazumder, 2014; Mazumder, 2014). Dette lar seg ikke gjøre i microdata.

⁹ URM står for upward rank mobility.

= 0 steget i rank, og vil være derfor være med å øke oppovermobiliteten. Med en $\tau = 0,1$ vil det samme barnet måtte stige helt til den 11. inntektsprosentilen for at barnets bevegelse i inntektsfordelingen skal regnes som en bevegelse. Altså vil RM ved $\tau = 0,05$, og $10 \leq \rho < 15$ måle sannsynligheten for at barn med foreldre med inntekt i 10. inntektsprosentil til og med 15. inntektsprosentil stiger med minst 5 inntektsprosentiler. Dette betyr at barnets inntekt må plassere barnet i minst den 19. inntektsprosentilen for at det skal telles som en oppoverbevegelse. Kortere bevegelser enn dette vil ikke regnes med som oppoverbevegelser. Når en studerer to undergrupper, kan en ved å justere τ opp vise hvilken gruppe som er mest tilbøyelige til å gjøre store byks eller kortere bevegelser i inntektsfordelingen.

Nedoverretningsrankmobilitet (DRM^{10}) formuleres på følgende måte:

$$DRM_{\tau,\rho} = \Pr(Y_0 - Y_1 > +\tau \mid Y_0 \geq \rho)$$

Altså måler en sannsynligheten for at barnets inntekt er lavere enn foreldrenes inntekt gitt at foreldrenes inntekt er større enn en gitt prosentil ρ . Dersom en setter $\rho = 0,9$, og $\tau = 0,05$ vil DRM vise hvor mange barn med foreldre i den 90. høyeste inntektsprosentilen som har falt ned til den 85. inntektsprosentilen og lavere.

Det er mulig å betinge URM og DRM på ulike karakteristika, slik at en måler mobilitet hos ulike undergrupper. La X_j representere ulike karakteristika, som for eksempel at individet er annengenerasjonsinnvandrere, eller at individet er bosatt i Oslo. Da kan en formulere URM og DRM på følgende måte:

$$URM_{\tau,\rho} = \Pr(Y_1 - Y_0 > +\tau \mid Y_0 \leq \rho, X_j = x)$$

$$DRM_{\tau,\rho} = \Pr(Y_0 - Y_1 > +\tau \mid Y_0 \geq \rho, X_j = x)$$

Dette er egenskaper som er svært gunstig i denne oppgaven, hvor nettopp sammenligning av ulike undergrupper er hovedfokuset.

Betingelsene og spesifikasjonene som gjøres i målingene RM er hovedsakelig de samme som i overgangssannsynlighetene. Altså er estimeringene av RM betinget på samme geografiske områder. RM vil også bli rapportert med tilsvarende grenser som ved overgangssannsynligheter, altså med $\tau = 0$ og $\tau = 0,3$. Dette gjør at de to metodene er mest mulig sammenlignbare. En kan merke seg at i RM blir inntektsfordelingen målt i vintil¹¹, dette diskuteres grundigere nedenfor.

¹⁰ DRM står for downward rank mobility.

¹¹ Dette er en fornorskning av det engelske ordet ventile som betyr tyvendedel.

2.6 Sammenligning av overgangssannsynligheter og retningsrankmobilitet

Grunnen til at jeg velger både å benytte meg av overgangssannsynligheter og RM i denne oppgaven er fordi disse komplementerer hverandre godt. Samtidig er det i tråd med tidligere artikler som Bhattacharya og Mazumder (2011), Corak et al. (2014) og Mazumder (2014). I tillegg har overgangssannsynligheter blitt kritisert for at en setter et vilkårlig avkuttingspunkt i inntektsfordelingen gjennom variabelen ρ . Dette gjør at det i teorien er uendelig mange ulike overgangssannsynligheter avhengig av grensen en setter. Dette er i motsetning til RM som sier at inntektsprosentilet til barnet må være under eller over farens inntektsprosentil i inntektsfordelingen for at det skal telles som en bevegelse. Videre kan det også merkes at Mazumder (2014, s. 4) skriver at RM ikke er et perfekt mål, og at det gjerne burde komplementeres med overgangssannsynligheter. En kan også merke seg at RM måler flere bevegelser enn overgangssannsynligheter. Dette kommer av at det ofte er lettere å overgå (eller falle under) fars inntektsprosentil enn grensen en benytter seg av i overgangssannsynligheter. Dette betyr at RM vil gi noe høyere sannsynlighet for bevegelse oppover eller nedover, avhengig av hva en måler.

I denne oppgaven blir RM spesifisert i vintil istedenfor desil. Dette blir gjort fordi RM måler flere bevegelser i inntektsfordelingen. Når RM måler flere bevegelser blir resultatene mindre utsatt for støy. Mindre støy betyr følgelig at det blir mulig å måle kortere bevegelser i inntektsfordelingen, dette gjøres ved å dele inntektsfordelingen i vintil. For å forklare hvordan vintilspesifisering måler kortere bevegelser, kan en se for deg at en sammenligner resultater fra to ulike RM-estimeringer. Begge har tilsvarende grense på null. Den ene derimot er spesifisert i vintil, og den andre i desil. I et slikt tilfelle må individene bevege seg lengre opp i inntektsfordelingen i estimeringen med desil for at det skal telles som en oppover bevegelse. Dette kommer av at individene målt ved hjelp av desil må bevege seg ti prosentiler for at det skal telles som en bevegelse. For at det skal telle som en bevegelse med vintil må individet kun bevege fem prosentiler. Dette betyr at individene må bevege seg fem ekstra prosentiler ved desil, sammenlignet med vintil. Det kan også merkes at grensen τ alltid representerer prosentiler. Dette betyr for eksempel at når $\tau = 0,3$ vil individene måtte bevege seg 30 inntektsprosentiler ekstra uavhengig om inntektsfordelingen er spesifisert i vintil eller desil.

En annen forskjell det er verdt å merke seg mellom overgangssannsynligheter og RM, er at RM blir påvirket annerledes enn overgangssannsynligheter når en bruker kumulative mål. Dette kommer av at i RM trenger en kun å slå faren din sin rank gitt at han plasserer i eller lavere enn et gitt prosentil. I overgangssannsynligheter må individet overstige et gitt prosentil betinget på at faren din befinner seg under eller i en gitt prosentil. Konsekvensen av dette når en bruker kumulative mål i RM, er at en inkluderer stadig flere individer som har oppnådd høyere rank enn faren sin når en studerer høyere inntektsprosentiler. Dette vil ikke gjelde for overgangssannsynligheter når grensen individet må overstige stadig blir høyere. Dette påvirker resultatene for innvandrerne mer enn nordmennene. Dette kommer av at de fleste av innvandrerfedrene befinner seg i bunnen av inntektsfordelingen. Dermed,

når en beveger deg oppover i inntektsfordelingen og det blir vanskeligere å overstige faren din sin inntektsprosentil, vil innvandrerne som har en liten andel fedre i de øvre prosentilene virke mer mobile sammenlignet med nordmenn.

Kapittel 3

Litteraturoversyn

I dette kapitlet presenteres litteratur som er relevant for problemstillingen oppgaven ønsker å besvare. Først, i delkapittel 3.1 blir det gjort rede for hovedtrekk i Becker-Tomes-modellen. Denne modellen forklarer overføringer av inntekt over generasjoner. For at leseren skal få en oversikt over den historiske utviklingen i IGM blir det også gjort rede for hovedlinjene i litteraturen. I delkapittel 3.2 presenteres litteratur som skal gjøre det klart hvordan IGM er i Norge, sammenlignet med resten av verden. Deretter, i delkapittel 3.3, presenteres litteratur som ser på utviklingen i IGM i Norge over tid. Disse artiklene fremhever også mekanismer som kan forklare hva som driver utviklingen i IGM over tid. Samtidig presenteres litteratur som har utforsket regionale forskjeller innad i Norge med IGM i øyemed. Det er også interessant å få en oversikt over hvordan de ulike sosioøkonomiske forholdene er hos innvandrere i Norge, og litteratur om dette temaet presenteres i delkapittel 3.4. I delkapittel 3.5 presenteres litteratur om IGM hos innvandrere i verden og Norge, og effekten innvandring har hatt på IGM hos nordmenn. Til slutt, i delkapittel 3.6 oppsummeres kapittel 3.

3.1 Becker-Tomes og hovedlinjer

I delkapittel 3.1.1 presenteres Becker-Tomes-modellen som skal forklare mekanismene bak IGM. Deretter, i delkapittel 3.1.2 blir det en gjennomgang av utviklingen til IGM, med spesielt fokus på målefeil.

3.1.1 Becker-Tomes-modellen

I Solon (1999) presenteres Becker-Tomes-modellen (1979) som kan brukes til å forklare mekanismene bak den intergenerasjonelle overføringen av inntekt mellom foreldre og barn. Modellen tar utgangspunkt i at hver familie består av en forsørger og et barn. Forsørgeren er nyttemaksimerende, og sprer inntekten sin utover enten konsumgoder eller investeringer i humankapitalen til barnet. Hvor mye som blir investert i hver avhenger av forsørgerens preferanser for konsum i nåtiden kontra preferanser for barnets fremtidige inntekt.

Barnets fremtidige inntekt er en funksjon av to ledd. Det ene leddet er forsørgerens inntekt. Dersom forsørgeren har høyere inntekt betyr dette ene og alene at barnet også vil ha høyere fremtidig inntekt. I hvilken grad dette leddet bestemmer barnets fremtidige inntekt avhenger av avkastningen på investeringene foreldrene gjør i barnets humankapital, og i hvor stor grad foreldrene har preferanser for å investere i barnets humankapital. Det andre leddet består av to komponenter. Den ene komponenten kan kalles natur og miljø. Natur og miljø skal representere overføringer av genetiske og ikke-genetiske egenskaper som gjør barnet mer eller mindre tilbøyelig til å ha høy inntekt. Eksempler kan være intelligens, psykisk og fysisk helse, barnets oppdragelse, oppmuntring til utdanning, og så videre. Disse egenskapene overføres over generasjoner i mer eller mindre grad. Den andre komponenten er flaks. Flaks blir ikke påvirket av forsørgerens inntekt eller bakgrunn, og kan anses som tilfeldig.

Komponentene som påvirker barnets fremtidige inntekt påvirker også grad av IGM. Høyere preferanser for å investere i barnets humankapital vil føre til lavere grad av IGM. Når preferansene for å investere i barnets humankapital øker, vil økningen i investeringer være relativt større hos mer ressurssterke foreldre. Dermed blir den intergenerasjonelle overføringen av inntekt sterkere. En annen faktor som påvirker IGM er hvor høy avkastningen på humankapital er. Høyere avkastning på investeringer i barnets humankapital betyr mindre IGM. Dersom avkastningen på humankapital er høy vil dette gi insentiver til å investere mer i barnets humankapital. Samtidig vil barn fra mer ressurssterke familier høste relativt mer fordeler av foreldrenes investeringer i deres humankapital. Dette betyr at ressursene til foreldrene blir mer avgjørende i å forklare barnets arbeidsmarkedsutfall. I hvilken grad ikke-genetiske og genetiske egenskaper blir overført mellom foreldre og barn påvirker også grad av IGM. Dersom disse egenskapene i stor grad blir overført over generasjoner vil dette gi utslag i lavere IGM. Eksempelvis kan en se på intelligens. Intelligens er en arvelig faktor som også er en god indikator på arbeidsmarkedsutfall. Foreldre med høy intelligens vil i gjennomsnitt tjene mer enn mindre intelligente foreldre. Vi ser for oss at intelligens i høy grad blir overført til barnet fra foreldrene. Dette vil gi utslag i at barn av mer intelligente foreldre vil få høyere inntekt enn barn av mindre intelligente foreldre. På denne måten blir høyere inntekt overført over generasjoner, som betyr lavere grad av IGM. Tilfeldigheter spiller også inn, dette gjennom flaks. Hvis flaks i høy grad bestemmer barnets fremtidige inntekt, vil grad av IGM følgelig bli høyere. Dette kommer av at andre faktorer enn foreldrenes inntekt blir viktig for å forklare barnets fremtidige inntekt (Becker & Tomes, 1979; Solon, 1999)

Solon (2004) presenteres en utvidet Becker-Tomes-modell. Denne utvidede modellen inkluderer statlige investeringer i barnets humankapital. Statlige investeringer påvirker IGM gjennom to kanaler. Den ene kanalen er at foreldrenes optimale investeringsnivå er fallende i statlige investeringer. Dette kommer av at investeringer i barnets humankapital har avtakende utbytte. Dermed vil private investeringer i barnets humankapital gi lavere avkastning sammenlignet med en situasjon uten statlige

investeringer. Dette betyr isolert sett mindre private investeringer. Samtidig vil barn fra familier med lavere inntekt ha høyere nytte av statlige investeringer, sammenlignet med barn fra rikere familier. Dette kommer av at statlige investeringer vil utgjøre relativt mer i en familie hvor inntekten er lav. Mer statlige investeringer i barns humankapital kan føre til at noen land har mer intergenerasjonell inntektsmobilitet enn andre land. Dette kommer av at statlige investeringer reduserer viktigheten av foreldrenes inntekt, spesielt hos ressursvake familier. Statlige investeringer i barns humankapital tar ofte form i offentlig utdanningstilbud. Dette kan forklare hvorfor land som for eksempel Norge har høy grad av IGM, da et godt offentlig utdanningstilbud vil utjevne muligheter for individer fra ulike sosioøkonomiske forhold (Markussen & Røed, 2017).

3.1.2 Hovedlinjer i litteraturen

Solon (1999) gir en oversikt over den historiske utviklingen til IGM, og i tillegg presenteres resultater fra en rekke tidligere studier. Solon (1999) skriver at nyere forskning på IGM har basert seg på data som inneholder informasjon om flere individer, og samtidig strekker seg over lengre perioder. På grunn av dette blir det mulig å belyse problemer tilknyttet det å ta utgangspunkt i et for heterogent utvalg, måle barnets inntekt i for ung alder og det å basere inntektsmålet over for få år. Måten han gjør dette på er ved å sammenligne resultater fra en rekke artikler om IGM. Samtidig kartlegger Solon i hvilken alder forfatterne måler barnets inntekt, og over hvor mange år farens inntektsgjennomsnitt er basert. Han skriver både det å basere fars inntekt over for få år, og det å måle barnets inntekt i for ung alder gjør IGE estimatet forventningsskjev mot null.

Black og Devereux (2011) tar for seg utviklingen som har skjedd innenfor IGM etter Solon (1999). I tillegg til å fokusere på hvordan en kan få mest mulige presise estimat av IGE, fokuserer også Black og Devereux på de kausale mekanismer bak IGM. Siden 1999 har det blitt publisert en rekke artikler om målefeil. Dette gir utslag i at mekanismene bak målefeil er mer nøyaktig kartlagt i Black og Devereux (2011) sammenlignet med Solon (1999). Det refereres til ulike artikler om klassisk målefeil, og viktigheten av å basere inntektsmålet over flere år er nøye understreket. Samtidig vektlegges også viktigheten av å måle inntekten i en passende alder for både barn og foreldre, dersom en frastår fra dette kan estimatene bli påvirket av livssyklus-skjevhet. De vier også mye av kapitlet sitt til forskningen som har forsøkt å kartlegge hvor mye av den intergenerasjonelle inntektskorrelasjonen som skyldes arv, og hvor mye som skyldes miljø.

3.2 Norge sammenlignet med resten av verden

Jäntti et al. (2006) sammenligner IGM mellom USA, Storbritannia, Norge, Danmark, Sverige og Finland. De ønsker at datasettene skal være bygd opp på samme måte, og ha tilsvarende definisjoner. Dette er ønskelig slik at resultatene skal være mest mulig sammenlignbare over landegrensene. De

benytter seg av overgangsmatriser, korrelasjoner og IGE som mål på IGM. Resultatene deres tilsier at USA er minst mobil, etterfulgt av Storbritannia, og deretter de nordiske landene. Videre skriver de at Norden, USA og Storbritannia har omtrentlig lik mobilitet i midten av inntektsfordelingen, men at Storbritannia og USA har mye lavere nedovermobilitet fra toppen av inntektsfordelingen. I tillegg til dette har også USA, i motsetning til resten av landene, svært lav oppovermobilitet fra bunnen av inntektsfordelingen. Det kan påpekes at Jäntti et al. (2006) benytter seg av et inntektsmål som baserer seg kun på et år, men at sensitivitetstestene deres tilsier at klassisk målefeil ikke skal forårsake store problemer.

Bratberg et al. (2017) sammenligner estimat for IGM fra Tyskland, Norge, Sverige og USA. De fokuserer på to metoder, retningsrankmobilitet og inntektsandelsmobilitet¹². Disse målemetodene skal gjøre opp for at inntektsmobilitet kan være ulik i forskjellige deler av inntektsfordelingen. I tillegg gjør inntektsandelsmobiliteten opp for at avstanden mellom inntektsranger kan være ulike i forskjellige land. Når de bruker retningsrankmobilitet finner de at Norge og Sverige har høyere grad av oppovermobilitet i bunnen av inntektsfordelingen, og sterkere nedovermobilitet i toppen av inntektsfordelingen sammenlignet med USA og Tyskland. De påpeker dog at resultatene for Tyskland er mindre pålitelig ettersom datamengden fra Tyskland er begrenset. Resultatene endrer seg når de benytter seg av inntektsandelsmobilitet, og de finner at mobiliteten i de fire landene er omtrentlig lik. De konkluderer med at det kan være problematisk å trekke for bastante slutninger når en sammenligner intergenerasjonell inntektsmobilitet mellom ulike land. Dette baseres på at resultatene avhenger av målemetode og hvilken del av inntektsfordelingen en studerer. På tross av dette konkluderer de med at nordiske land som Norge, fremstår som mer mobile i bunnen av inntektsfordelingen sammenlignet med for eksempel USA.

Connoly, Corak og Haeck (2019) sammenligner IGM i USA og Canada. De fremhever noen interessante faktorer som er viktig ved sammenligning av mobilitet over landegrensener. Noen av disse faktorene er folkets tanker rundt statlig omfordeling og statlige inngrep, og samtidig hvilke egenskaper som vektlegges i arbeidslivet og andre sosiale dimensjoner. Connoly et al. argumenterer for at den «amerikanske drømmen» og den «canadiske drømmen» er ganske like, og at borgere i Canada og USA har noenlunde like tanker om hvor mye staten skal delta i omfordelingsprosesser¹³. Med dette i mente er det rimelig å anta at canadierne og amerikanerne ønsker omtrentlig lik grad av IGM, derfor blir også sammenligning mellom de to landene mer relevant. På tross av omtrentlig like verdier og syn på faktorer som påvirker IGM, tilsier resultater fra ulike estimeringer at Canada er mer mobilt enn USA. Dette leder videre til neste interessante faktor som er viktig når en sammenligner ulike land, nemlig at en burde ta høyde for regionale forskjeller innad i landene. Connoly et al. finner at mye av forskjellen i

¹² På engelsk kalt Income Share Mobility.

¹³ Canadierne er mer åpen for statlige inngrep enn hva amerikanerne er, men Connoly et al. (2019) argumenterer for at canadierne ligner mer på amerikanere enn hva europeere gjør.

IGM mellom USA og Canada stammer fra sørstatene i USA hvor det er vanskelig for innbyggerne å stige oppover i inntektsfordelingen, spesielt for afroamerikanere.

3.3 Utvikling i intergenerasjonell inntektsmobilitet i Norge og forskjeller i regioner

Bratberg et al. (2005) undersøker sammenhengen mellom IGM og inntektsulikhet i Norge. De skriver at en rekke artikler peker på en mulig sammenheng mellom lite inntektsulikhet, og høy grad av IGM. Dette baseres hovedsakelig på at skandinaviske land, altså land med liten grad av inntektsulikhet, har høy grad av IGM. Dette sammenlignet med land som USA og England hvor inntektsulikheten er større. Videre skriver de at sammenligninger over landegrensener kan være preget av støy. De argumenterer derfor for at det kan være mer givende å se på utviklingen i inntektsulikhet og IGM innad i et land, for studere sammenhengen mellom inntektsulikhet og IGM. For kohortene de studerer, altså 1950- til og med 1965-kohorten¹⁴, har inntektsulikheten holdt seg relativt stabil. Bratberg et al. (2005) skriver at dersom sammenhengen mellom inntektsulikhet og IGM er gjeldende, vil også grad av IGM ha holdt seg relativt stabil over denne perioden. De finner at IGM har vært stabil for 1950- til og med 1965-kohorten. Altså underbygger resultatene deres sammenhengen mellom inntektsulikhet og IGM.

Pekkarinen, Salvanes og Sarvimäki (2017) undersøker hvordan den sosiale mobiliteten i Norge har utviklet seg for individer født fra 1930- til og med midten av 1970-tallet. De knytter endringer i sosial mobilitet over tid til utvikling av velferdsstaten, og legger spesiell vekt på tilbudet og avkastningen til utdanning. Det kan merkes at analysen deres må anses som deskriptiv. De finner at den sosiale mobiliteten økte fra 1930- til 1940-tallet. Pekkarinen et al. (2017) skriver at en mulig grunn for den økte mobiliteten kan ha vært mer tilgjengelig utdanning uavhengig av fars økonomiske bakgrunn, sammen med lavere avkastning på utdanning. Mer tilgjengelig utdanning for befolkningen øker den intergenerasjonelle inntektsmobiliteten i et land ettersom barnets utdanning vil være mindre avhengig av foreldrenes sosioøkonomiske forhold. I tillegg vil mindre avkastning av utdanning føre til at sammenhengen mellom barnets utdanning og foreldrenes sosioøkonomiske forhold gjør seg mindre gjeldene i inntektssammenheng. Videre skriver Pekkarinen et al. (2017) at resultatene for barn født i slutten av 1940- til og med begynnelsen av 1970-tallet er mer tvetydige, men at inntektskorrelasjon mellom far og sønn forholdt seg relativt stabil. De peker på at avkastningen av utdanning økte, sammen med at utdanningsgraden hos individer fra lav- og middelklassen økte. Disse to faktorene kan ifølge Pekkarinen et al. (2017) ha nullt hverandre ut, ettersom de skal trekke den sosiale mobiliteten i ulike retninger.

¹⁴ De ser på 1950-, 1955, 1960 og 1965-kohorten.

Markussen og Røed (2017) undersøker velferdsstatens tilstand, og fokuserer på likhet. De finner at individer med foreldre i de laveste inntektsprosentilene gjør det dårligere på en rekke sosiale dimensjoner sammenlignet med individer fra øvrige inntektsprosentiler. Disse sosiale dimensjonene omfatter posisjon i inntektsfordelingen, intergenerasjonell inntektsmobilitet, tilbøyelighet for å bli ansatt, utdanning, helse og etablering av familie blant flere. De argumenterer for at dette ikke skyldes at individer fra de lavere inntektsprosentilene har svakere kognitive egenskaper, men peker heller på støtte fra familie. Da hovedsakelig støtte og oppmuntring til å ta en utdanning på høyere nivå. Markussen og Røed (2017) skriver at familier i øvrige inntektsprosentiler er mer tilbøyelige til å oppmuntre barna sine til å ta en utdanning. Dette sammen med at utdanning har blitt viktigere i samfunnet, fører til at individer i de laveste inntektsprosentilene gjør det dårligere på ulike sosiale dimensjoner.

Markussen og Røed (2018) ser på hvordan det å vokse opp i nabolag med ulik sosioøkonomisk status påvirker barns utdannings- og arbeidsmarkedsutfall tidlig og sent i livet. Dette gjør de ved å sammenligne helsøsken som har blitt utsatt for nabolag med ulik sosioøkonomisk status. Måten de har blitt utsatt for nabolag med ulik sosioøkonomisk status er gjennom flytting eller ved at nabolaget endrer seg. De finner at nabolagssegregering har økt. Dette betyr at individer fra ulike sosioøkonomiske klasser i mindre grad bor samlet nå, sammenlignet med tidligere. Dette betyr igjen at andelen nabolag som kan anses som middelklassenabolag har sunket, og andelen som vokser opp i nabolag som klassifiseres som mindre velstående og mer velstående har økt. De finner også en konkav sammenheng mellom resultater fra utdanning og arbeidsmarkedsutfall, og hvor velstående nabolaget er. Dette betyr altså at barn fra nabolag som er mindre velstående og mer velstående typisk gjør det dårligere med tanke på utdanning og arbeidsmarkedsutfall, sammenlignet med barn fra middelklassenabolag. Videre skriver de at det er positive eksternaliteter ved å flytte fra et mindre velstående til et middelklasse nabolag. Når en fra et mindre velstående nabolag flytter til et middelklassenabolag vil personen bli eksponert for gode forbilder og ressurssterke mennesker. Dette påvirker utdanning og arbeidsmarkedsutfall positivt. I motsetning vil det å flytte fra et mindre til et mer velstående nabolag gjøre at en blir relativt sett mer fattig. Dette kan påvirke sosial dimensjoner negativt, som igjen påvirker utdannings- og arbeidsmarkedsutfall.

I 2019 publiserte Samfunnsøkonomisk Analyse en rapport der de så på forskjeller i IGM mellom regioner og innad i byer i Norge (Norberg-Schulz & Østhus, 2019). De finner at byer på Vestlandet er spesielt mobile i Norge, og argumenterer for at dette skyldes oljenæringen.

3.4 Innvandrere i Norge

Bratsberg, Raaum og Røed publiserte i 2010 en artikkel om arbeidsinnvandrerne¹⁵ som kom til Norge på begynnelsen av 1970-tallet. Her undersøker de hvordan arbeidsinnvandrerne har tilpasset seg i det norske arbeidsmarkedet. De finner at arbeidsdeltakelsen til arbeidsinnvandrerne overgår arbeidsdeltakelsen til den norske sammenligningsgruppen i de ti første årene etter ankomst. Etter ti år faller arbeidsdeltakelsen hos arbeidsinnvandrerne kraftig, og i 2000 var kun 50% av denne gruppen sysselsatt. Ektefellene¹⁶ til arbeidsinnvandrerne gjør det ifølge Bratsberg al. (2010) enda dårligere enn arbeidsinnvandrerne. Bratsberg et al. (2010) viser at sysselsetningen hos ektefellene aldri overstiger 40%. Tilsvarende tall for ektefellene til den norske sammenligningsgruppen er 80%. Bratsberg et al. (2010) finner også at i år 2000 er 74% av dem som forlot arbeidsmarkedet uføretrygdet, og at over 90% av dem mottok en form for sosial stønad. Bratsberg et al. (2010) finner også at når disse arbeidsinnvandrerne trer ut av arbeidsmarkedet vil det være vanskeligere for dem å tre inn igjen sammenlignet med den norske sammenligningsgruppen. Videre finner de også at arbeidsinnvandrerne må ha hatt jobben lengre enn nordmennene i sammenligningsgruppen for å oppnå samme jobbsikkerhet. Det ser også ut til at arbeidsinnvandrerne er mer utsatt for konjunktursvingninger. Altså er det mer sannsynlig at en arbeidsinnvandrer mister jobber i en nedgangskonjunktur. Dette kombinert med at det fremstår som vanskeligere for arbeidsinnvandrerne å entre arbeidsmarkedet igjen, fører til at økonomiske nedganger har sterkere og lengre effekt på arbeidsinnvandrerne.

Bratsberg, Raaum og Røed (2012) studerer utfallet i utdanning til barn av innvandrere fra land med lavt inntektsnivå. De fokuserer på forskjeller i utfall hos barn født i Norge og innvandrere som kom til Norge som barn. Bratsberg et al. (2012) finner at barn av innvandrere nærmer seg barn av andregenerasjonsnordmenn når det kommer til utdanning. De skriver at dette gjelder for både andregenerasjonsinnvandrere og barn født i utlandet, men at det gjør seg mest gjeldene for førstnevnte. I artikkelen gjør de det også klart at foreldrenes ressurser har mye å si for utfallet i barnas utdanning, men at det ikke er dette som er driveren til konvergeringen i utdanningsutfallet for barna. Bratsberg et al. (2012) finner at gapet i inntekt mellom innvandrerforeldre og norske foreldre øker, og derfor heller skulle tilsi større forskjeller i utdanningsutfall mellom innvandrerbarn og norske barn.

Bratsberg, Raaum og Røed (2014) har studert utfallet til innvandrerne i Norge etter immigrasjon. De fokuserer på innvandrere som har kommet til Norge med de største innvandringsbølgene. Disse bølgene omfatter arbeidsinnvandrerne som kom på 1970-tallet til og med 2004. De gjør en rekke interessante funn, og skriver at opphavslandet til arbeidsinnvandrerne er meget viktig i å bestemme hvordan det går med arbeidsinnvandrerne i arbeidsmarkedet i Norge. Innvandrere som kommer fra land som er på nivå med Norge økonomisk sett gjør det typisk like godt som nordmenn i arbeidsmarkedet. Dersom innvandrerne har immigrert fra et mindre utviklet land er det større sjanse

¹⁵ Disse kommer hovedsakelig fra Pakistan, Tyrkia, India og Marokko (Bratsberg et al., 2010).

¹⁶ Ektefelle er ofte ikke arbeidsinnvandrer siden disse ankom Norge på et senere tidspunkt gjennom familiejenforening.

for at innvandreren vil havne opp som arbeidsledig eller som mottaker av sosialstønning. De skriver også at innvandrere fra mindre utviklede land har en tendens til å bosette seg permanent i Norge, i motsetning til innvandrere fra rikere land. For innvandrere som har kommet til Norge av andre grunner enn arbeid, som for eksempel som flyktninger eller på grunn av familiegjening, er bilde noe annerledes. Videre finner de at disse innvandrerne, i begynnelsen av sitt opphold, presterte på lik linje som nordmenn i arbeidsmarkedet, men at etter 10 til 15 år begynte prestasjonene deres i arbeidsmarkedet å svekke seg. Bratsberg et al. (2014) skriver at disse innvandrerne etter en viss periode hadde lavere deltakelse i arbeidslivet, og høyere andel på uføretrygd sammenlignet med nordmenn. Bratsberg et al. (2014) ser også på barna til arbeidsinnvandrerne som kom på 1970-tallet. De finner at disse i gjennomsnitt hadde lavere utdanning, lavere sysselsetting og lavere gjennomsnittslønn enn nordmenn i samme alder. De finner også en tendens til at en høyere andel av innvandrerbarna mottar uføretrygd, og at denne andelen øker. Til tross for dette konkluderer Bratsberg et al. (2014) med at barn av arbeidsinnvandrerne viser en assimileringstendens. Altså konvergerer arbeidsmarkeds- og utdanningsutfallet til innvandrerbarna og de norske barna. Samtidig viser det seg at annengenerasjonsinnvandrere gjør det bedre enn førstegenerasjonsinnvandrere på samme alder.

3.5 Innvandring og mobilitet

Aydemir, Corak og Chen (2009) studerer IGM mellom første- og andregenerasjonsinnvandrere i Canada. De ønsket også å kartlegge mulige mekanismer som skal hjelpe med å forklare den intergenerasjonelle overføringen av inntekt hos innvandrere i Canada. Aydemir et al. (2009) påpeker også at IGM hos innvandrere er spesielt interessant. Dette kommer av at andelen førstegenerasjonsinnvandrere i Canada med lav inntekt er høy og økende. Det å vokse opp i en husholdning med lav inntekt er en ulempe, og grad av IGM hos innvandrere vil si noe om hvor vedvarende denne ulempen er over generasjoner. Funnene deres tilsier at innvandrerne i Canada er omtrent like mobile som andregenerasjonscanadiere. De finner også at mye av den intergenerasjonelle inntektskorrelasjonen hos innvandrere drives av at de med best arbeidsmarkedsutfall av innvandrere fra lav-inntektsfamilier tjener mindre enn de med best arbeidsmarkedsutfall av innvandrere fra høy-inntektsfamilier.

Hermansen (2016) måler IGM hos innvandrere i Norge, for henholdsvis 1973 – 1982 kohortene. Hermansen (2016) benytter seg av rankmobilitet for å kartlegge IGM. Han finner at gapet i utdanning og inntekt som var til stede mellom innvandrere og nordmenn i foreldregenerasjonen, er sterkt redusert mellom deres barn. Resultatene til Hermansen (2016) tilsier også at barn av innvandrere oppnår tilsvarende, og eller høyere grad av utdanning og inntekt når en kontrollerer for forskjeller i sosioøkonomiske forhold og nabolagseffekter. Til tross for dette skriver Hermansen at barn av innvandrere med opphav til Pakistan, Afrika og Midtøsten henger noe etter. Altså er utdannings- og

inntektsnivået noe lavere i disse gruppene, sammenlignet med innvandrere fra andre regioner. Han skriver videre at det kan tenkes at barn av innvandrere fra Pakistan, Afrika og Midtøsten ofte er fra familier med lav inntekt og lite utdanning, og at disse vil trenge en ekstra generasjon for å få utdanning og inntekt på nivå med andregenerasjonsnordmenn.

Hoen, Markussen og Røed (2018) undersøker om økt innvandring fra Øst-Europa og lavinntektsland kan forklare hvorfor vi ser økte forskjeller mellom individer fra lavere sosiale klasser og øvrige sosiale klasser. Økonomisk teori tilsier at individer med arbeidsferdigheter som komplementeres av innvandring vil tjene på arbeidsinnvandring. Samtidig vil individer der arbeidsferdighetene substitueres av innvandrere tape på innvandring. Hoen et al. (2019) skriver at brorparten av innvandring har kommet fra lavinntektsland. Disse innvandrerne innehar ferdigheter som typisk substituerer ferdighetene til individer i de lavere sosiale klassene i samfunnet, og komplementerer de fra høyere klasser. Dette har ifølge Hoen et al. (2019) ført til at den sosiale ulikheten i Norge har økt. Dette kommer av at individer i de lavere sosiale klassene opplever sterkere konkurranse i arbeidslivet. Samtidig vil de fra øvrige klasser, som komplimenteres i arbeidssammenheng, oppleve økt produktivitet og bedret arbeidsmarkedsutfall.

3.6 Oppsummering

Ut ifra litteraturen presentert ovenfor er det tydelig at Norge er et land med høy grad av IGM. Mye av dette kan skyldes en sterk velferdsstat, sammen med preferanser for likhet. Som forklart ovenfor vil avkastning av, sammen med tilgjengeligheten til utdanning også være viktig når en ønsker å forklare utviklingen i IGM over tid innad i Norge.

Litteraturen om innvandrere i Norge tilsier at førstegenerasjonsinnvandrere har dårligere arbeidsmarkedsutfall sammenlignet med nordmenn i Norge. De har også en høyere andel som mottar stønader fra det offentlige, og barna deres har for eksempel høyere frafallsrate fra utdanning. I tillegg er innvandrerforeldrene mer utsatt for konjunkturer i arbeidsmarkedet. Dette kan bety at inntektsmålet til innvandrerne er mer utsatt for klassisk målefeil enn hva nordmennene sitt inntektsmål er. Til tross for dette er det en assimileringstendens hos andregenerasjonsinnvandrere. Dette betyr at andregenerasjonsinnvandrere har blitt likere andregenerasjonsnordmenn når det kommer til arbeidsmarkedsutfall og utdanning. Dette støttes av Hermansen (2016) som skriver at gapet mellom førstegenerasjonsinnvandrere og andregenerasjonsnordmenn er sterkt redusert hos barna deres. Ved hjelp av målemetodene benyttet i denne oppgaven vil vi få et mer nøyaktig bilde av hvordan denne intergenerasjonelle assimileringprosessen foregår.

Kapittel 4

Data og utvalgsdesign

I dette kapittelet presenteres først datagrunnlaget til oppgaven, og deretter utvalgsdesignet. Til slutt presenteres relevant deskriptiv statistikk for første- og andregenerasjonsinnvandrere og andregenerasjonsnordmenn i Norge.

4.1 Data

Dataen jeg benytter meg av i denne oppgaven stammer fra SSB, og er gjort tilgjengelig for meg gjennom en nettbasert database- og statistikkverktøy kalt Microdata. Microdata driftes og videreutvikles av Norsk Senter for Forskningsdata (NSD), og er utviklet gjennom et infrastrukturprosjekt kalt RAIRD¹⁷. Dataen tilgjengelig på Microdata er av høy kvalitet ettersom det er data fra ulike registre. Disse registrene omfatter folkeregisteret, nasjonal utdanningsdatabase, selvangivelsesregisteret, arbeidsmarkedsdata og forløpsdatabasen FD-trygd (Microdata, 2019a). Disse registrene inneholder en rekke variabler om ulike forhold. Disse forholdene omfatter, men er ikke begrenset til, utdanningsgrad, hvorvidt personen mottar sosialhjelp fra staten, kjønn, inntekt, innvandrer kategorier, om personen har dødd et gitt år eller emigrert. Det er også mulig å koble foreldre og barn gjennom nøkkelvariabler tilgjengelig fra NSD, på denne måten slipper vi å estimere IGM ved hjelp av konstruerte foreldre-barn forhold som hadde gjort analysen mindre presis.

Infrastrukturprosjektet RAIRD har som hensikt å gjøre mikrodata mer tilgjengelig i forskningssammenhenger, og samtidig opprettholde krav om konfidensialitet. For å opprettholde konfidensialitet har NSD gjennomført ulike tiltak, disse omfatter 1: det er ikke mulig å benytte seg av utvalg spesifisert på en slik måte at de består av mindre enn 1000 individer. 2: dataen er winsorert til 2%. Dette betyr at de 1% laveste verdiene settes til 1-prosentilen, og de de 1% høyeste settes til 99-prosentilen. Dermed anonymiseres ekstremverdier på begge sider av fordelingen. 3: dataen er støybelagt med x , som er stokastisk og har $E(x) = 0$ (forventning lik null). Dette innebærer at det minste observerbare tallet i tabuleringer kan være 5, foruten 0. I tillegg støylegges verdier med maksimalt 5, og minimalt -5. 4: alle plotter er glattet ut med en glatteteknikk. Denne glattetekningen kalles hexbin, og resultatet er plotter som er anonyme og mindre presise enn et standard plot (Microdata, 2019b). Det er også andre begrensninger ved bruk av Microdata, disse omfatter blant annet at brukeren ikke kan se på eller laste ned dataen. Det er heller ikke mulig å bruke egne statistikkpakker og antall tilgjengelige statistiske analyser er begrenset. I tillegg til dette er brukergrensesnittet nyutviklet, og det er heller ikke tilsvarende til andre statistikkverktøy. Altså vil nye

¹⁷ RAIRD står for Remote Access Infrastructure for Register Data.

brukere av microdata være nødt til å lære seg et nytt grensesnitt, uavhengig av hvilket verktøy de har benyttet tidligere¹⁸.

4.2 Utvalgsdesign

I denne delen av oppgaven presenteres utvalgsdesignet. Først, i delkapittel 4.2.1, presenteres inklusjonskriteriene. Deretter, i delkapittel 4.2.2, begrunnes utvalgskriteriene, og mulige konsekvenser av kriteriene diskuteres.

4.2.1 Inklusjonskriterier

Utvalget består av individer født i 1977 til og med 1982, og deres fedre tilhørende 1938- til og med 1955-kohorten. Barnas inntekt måles fra 2007 – 2015, når de er 30 – 33 år gamle. Dette betyr at inntekten måles over en fireårs periode for hver kohort. For fedrene i utvalget måles inntekten fra og med 1993 til og med 1998. For fedrene betyr dette at inntekten måles over seks år, og i varierende alder. For de yngste fedrene i utvalget måles inntekten i en alder av 38 – 43 år, og for de eldste i en alder av 55 – 60 år.

Opgavens formål er å sammenligne ulike mål på IGM hos innvandrere og nordmenn i Norge. Ved hjelp av SSB sine innvandringskategorier definerer jeg innvandrerbarn som individer født i Norge med to innvandrerforeldre, eller individer som har immigrert til Norge senest i en alder av 6 år med far eller begge foreldre. Det kan også nevnes at norskfødte med en utenlandsfødt forelder og utenlandsfødte med to norskfødte foreldre blir ekskludert av analysen. Analysen omfatter i alt 212 003 ulike far-sønn og far-datter koblinger. 4 477 av disse er definert som innvandrer koblinger ut ifra kriteriene nevnt ovenfor. Ut av disse 4 477 innvandrerfar-datter og innvandrerfar-sønn koblingene er 1 647 av barna førstegenerasjonsinnvandrere som innvandret før en alder av 7 år.

4.2.2 Begrunnelse av utvalg

Definisjonen av andregenerasjonsinnvandrere utvides som sagt med innvandrere som innvandret før en alder av syv år. Dette blir gjort for å utvide et ellers begrenset utvalg av andregenerasjonsinnvandrere. Tidligere forskning har vist at alder ved immigrasjon er viktig for integrering i samfunnet, og indikerer at barn som immigrerer før skolealder gjør det bedre enn de som immigrerer senere (Böhlmark, 2008; Bratsberg et al., 2012, s. 244; Bratsberg et al., 2014, s. F671). Derfor velger jeg å ekskludere innvandrere som immigrerte eldre enn 6 år¹⁹, slik utvides utvalget uten å inkludere undergrupper som er altfor ulik norskfødte med to innvandrerforeldre. Det er dog viktig å påpeke at denne gruppen er vist

¹⁸ I [appendiks](#) blir det presentert et skript som viser hvordan brukergrensesnittet i microdata er.

¹⁹ I 1997 slo en lovendring til som endret skolealder fra 7 til 6 år. Dette betyr at barna inkludert i utvalget begynte på grunnskolen i en alder av 7 år (Nilsen, 2014).

til å gjøre det relativt dårligere i sosiale dimensjoner, som yrkesliv og utdanning sammenlignet med andregenerasjonsinnvandrere (Bratsberg et al., 2014, s. F674). På tross av dette viser sensitivitetstester at inklusjonen av førstegenerasjonsinnvandrere ikke påvirker IGE-estimer i særlig grad²⁰. Som nevnt ekskluderes individer med en utenlandsfødt forelder sammen med utenlandsfødte med to norskfødte foreldre. Disse blir ekskludert for å unngå definisjonsproblemer. Hermansen (2016) som også undersøker grad av IGM hos innvandrerbarn foretar tilsvarende valg. Hermansen (2016, s. 678) velger altså å inkludere førstegenerasjonsinnvandrere som innvandret før skolealder, og ekskludere norskfødte med en utenlandsfødt forelder og utenlandsfødte med to norskfødte foreldre.

Inntekten til fedrene er målt i alderen, for de yngste, 38 – 43 år, og for de eldste, 55 – 60 år. Dette kan påvirke analysen gjennom livssyklusshjevhet på fars alder. Hvis fedrene i de to undergruppene har omtrentlig likt aldersgjennomsnitt, samt fordeling rundt aldersgjennomsnittet, kan det argumenteres for at forventningsskjevheten påvirker de to undergruppene symmetrisk. Dette bygger på at livssyklusshjevhet påvirker inntektsmålet til innvandrerfedre og norske fedre likt. Dette trenger nødvendigvis ikke å gjelde. Det kan godt tenkes at optimal målealder for innvandrerfedrene og de norske fedrene er ulike. Dette kan gi utslag i asymmetrisk påvirkning av livssyklusshjevhet på fedrenes inntektsmål i de to gruppene. Dette vil igjen kunne skape problemer når en ønsker å sammenligne grad av IGM hos innvandrere og nordmenn i Norge. Det er viktig med et tilstrekkelig stort utvalg innvandrere, og begrensinger på fars alder fører til at mange observasjoner forsvinner. Dette taler for å være liberal med tanke på fars målealder. Ut ifra diskusjonen i dette avsnittet blir det klart at en ikke skal utelukke at estimatene er påvirket av livssyklusshjevhet i fars inntektsmål. Hvorvidt denne påvirker de to undergruppene på samme måte er også uvisst²¹.

Barnas inntekt måles når barna er mellom 30 – 33 år gamle. Dette er ofte anerkjent som en grei alder å måle inntekten til barnet i (Black & Devereux, 2011, s. 1493). Selv om en optimalt sett gjerne skulle ha målt barnas inntekt i slutten av 30-årene for å minimere sannsynligheten for livssyklusshjevhet (Nilsen et al., 2012, s. 10). Dersom en skulle gjort dette i denne oppgaven hadde utvalget blitt svært redusert. Dette kommer av at innvandring til Norge i stor skala er et relativt nytt fenomen. Dette betyr at barna til førstegenerasjonsinnvandrere fremdeles er meget unge. Hvorvidt livssyklusshjevhet påvirker de to gruppene på samme måte er også en dimensjon i denne diskusjonen. Dersom effektene av livssyklusshjevhet påvirker gruppene symmetrisk vil problemene tilknyttet målealder være redusert. På en annen side kan det godt tenkes at den ene gruppen kommer inn i arbeidsmarkedet i en senere alder, dette kan føre til skjevheten i den ene gruppen er sterkere enn i den andre.

Barnas inntekt måles over fire år, dette betyr at det kan være problemer med såkalt klassisk målefeil som kan skape mindre presise estimer, se delkapittel 2.2.1. I litteraturen ser en ofte at inntekten måles over fem år (Connolly et al., 2019, s. 12; Chetty, Hendren, Kline & Saez, 2014, s. 1564).

²⁰ Se appendiks for sensitivitetstest for inklusjon [Sensitivitetstest for førstegenerasjonsinnvandrere og alder ved innvandring](#).

²¹ Se appendiks for [sensitivitetstest for fars målealder](#).

Grunnen til at inntektsmålet strekkes over kun fire år i denne oppgaven er for å maksimere antall innvandrere i utvalget. Samtidig ønsker en å unngå mindre presise estimater ved å basere inntektsmålet over for få år. Dersom en ønsker å måle inntekten over flere år kan en velge å måle inntekten til barna når de enten er yngre eller eldre. For eksempel dersom inntekten måles for de inkluderte kohortene i en alder av 29 – 33 år ville en ha et inntektsgjennomsnitt på fem år. Dette kan skape problemer med livssyklususkjevhet, og en blir nødt til å ta en avveining. Ettersom livssyklususkjevhet forårsaker forventingskjevhet, og klassisk målefeil kun forårsaker mindre presise estimater, blir det naturlig å måle inntekten for barna i en noe kortere periode. En annen måte en kunne basert inntektsgjennomsnittet på flere år er ved å måle inntekten i en alder av 30 – 34. Da ville en blitt nødt til å inkludere eldre, og ekskludere yngre kohorter. Dette ville redusert utvalget av innvandrere betraktelig, siden annengenerasjonsinnvandrere i Norge fremdeles er relativt unge. Det blir igjen en avveining en må ta, og et rimelig stort utvalg veier tyngre enn eventuelle problemer med klassisk målefeil i barnas inntekt.

Fedrenes inntektsmål baseres på seks år. Dette gjøres for å minimere risiko tilknyttet klassisk målefeil i fedrenes inntektsmål. Klassisk målefeil oppstår som en følge av tilfeldige sjokk i inntektsmålet, og kan for eksempel være en konsekvens av at individet har vært arbeidsledig i kort periode. Fra litteraturen presentert i delkapittel 3.4 er det tydelig at innvandrerfedre er utsatt for konjunkturer i arbeidslivet. Dette kan bety at inntektsmålene til innvandrerfedrene er mer utsatt for tilfeldige sjokk, og dermed også klassisk målefeil. Dette taler for å basere inntektsmålet over flere år, slik at klassisk målefeil ikke degraderer analysen. Det er mulig å argumentere for å basere inntektsmålet til fedrene over flere år. Dette blir ikke gjort av hensyn til fedrenes målealder. Gjennom sensitivitetstester²² fremstår IGE-estimatene som noenlunde stabile når inntektsmålet baseres over seks år, dette underbygger valget om å måle inntekt over seks år videre.

Siden de største bølgene av annengenerasjonsinnvandrerne i Norge, i all hovedsak, fant sted i nyere tid, blir jeg nødt til å gjøre visse kompromisser for å få tilstrekkelig stort utvalg. Disse kompromissene, som diskutert ovenfor, kan resultere i at analysen blir mindre robust. Tiltak som kunne motvirket dette er å redusere aldersspennet til fedrene, samtidig som en ekskluderer innvandrere som immigrerte yngre enn 7 år. Slik vil en i høyest mulig grad analysere IGM mellom innvandrere i Norge, og annengenerasjonsinnvandrere. Optimalt sett skulle en i tillegg ønske å basere inntektsmålet til barna og fedrene på flere år for å unngå problemer med klassisk målefeil. Innvandrerutvalget kunne også gjerne vært større, slik at en i større grad kan dele dette utvalget inn i flere undergrupper basert på kjønn, opphavsland, fedrenes alder ved immigrasjon for å nevne noen av mange ulike dimensjoner.

I oppgaven bruker jeg fars inntekt som inntektsmål. Grunnen til at jeg bruker fars inntekt er fordi nøkkelvariabelen som kobler husholdninger sammen ikke strekker seg lengre tilbake enn 2004. Dette

²² Se [sensitivitetstest for antall måleår på fars inntektsmål](#).

betyr at dersom en skulle brukt husholdningsinntekt ville en vært nødt til å bruke mor og fars fødselsnummer som en kobling til barnet. Dette kan være problematisk, ettersom det ikke nødvendigvis er sann at disse vil bo sammen, som igjen betyr at deres samlede inntekt ikke representerer husholdningsinntekt. Det kan også merkes at fars inntekt som proxy for husholdningsinntekt ofte er standarden i artikler som studerer IGM. Dermed blir det også lettere å sammenligne resultater fra denne oppgaven med resultater fra andre relevante artikler.

Som et mål på inntekt brukes variabelen samlet inntekt inflasjonsjustert til 2015-nivå. Denne er definert som yrkesinntekt, kapitalinntekt, skattepliktige og skattefrie overføringer i løpet av kalenderåret (SSB, 2018d). Dette betyr at for eksempel inntekter fra aksjer, forrentninger av bankinnskudd og stønader som uføretrygd også inkluderes som et mål på inntekten til individet. Mange norske artikler om IGM har brukt pensjonsgivende inntekt som inntektsvariabel (Markussen og Røed, 2017, s. 10; Bratberg et al., 2005, s. 424). Det som kan være en ulempe med pensjonsgivende inntekt som inntektsvariabel er at den ikke inkluderer uføretrygd (Folketrygdloven, 1997, § 3-15), og som en kan se fra Tabell 5 er 14 % av innvandrede uføretrygdet. Dette er 11 prosentpoeng flere enn de norske fedrene. Dette betyr at en betraktelig andel av innvandrede lever av inntekter som ikke er pensjonsgivende. Dersom en ønsker et inntektsmål som representerer de uføretrygdedes evne til å investere i barnet sitt, blir en nødt til å bruke et mål på inntekt som inkluderer uføretrygd. Det er viktig å påpeke at barnets inntekter fra for eksempel aksjer og forrentninger ikke nødvendigvis gjenspeiler foreldrenes investeringer i barnets evner, men heller kan gjenspeile direkte arv fra foreldrene til barn. Det kan tenkes at det er forskjeller mellom innvandrere og innfødte med direkte arv fra foreldre til barn i øyemed. På en annen side er det ikke nødvendigvis sann at penger fra kapitalavkastning er arv. Det kan godt være at inntekten fra kapitalavkastning er tjent på egenhånd. Altså vil det ikke være hensiktsmessig å ekskludere kapitalavkastning fra inntektsmålet. Inntektsmål som tilsvarende samlet inntekt er brukt i tidligere artikler, se for eksempel Chetty et al. (s. 1566, 2014).

Databasen inneholder informasjon om statusen til et gitt individ, altså om personen er bosatt i Norge, død eller emigrert. Dette gjør at en kan omkode manglende inntektsobservasjoner i datasettet til null, dette er et tiltak for å gjøre databruken mest mulig effisient. Dette er i tråd med tidligere artikler (Chetty et al., s. 1565, 2014; Connolly et al., s. 11, 2019). Før manglende observasjoner omkodes, sørges det for at individene er i live, og er bosatt i Norge i relevante år. Dersom en ikke gjør dette vil individer med manglende data i et av årene bli ekskludert fra analysen. Det skal nevnes at relativt få individer får deler av inntekten sin i observasjonsperioden omkodet.

I denne oppgaven er disse inntektsfordelingene konstruert ut i fra inntekten til individene som oppfyller kriteriene listet ovenfor. Disse inntektsfordelingene er hentet fra et prosenthistogram på microdata som er winsorisert i begge ender av fordelingen.

4.3 Deskriptiv statistikk

Tabell 1 viser opphavsland, andel første generasjonsinnvandrere av barna, gjennomsnittsalder for fedrene i 1994, alder ved og årstall for immigrasjon for innvandrernes fedrene i utvalget. Gruppen definert som andregenerasjonsinnvandrere er meget heterogen, og har opphav fra mange forskjellige regioner. Regionen som er sterkest representert er Sentral-, Sør- og Vest-Asia. Innad i denne regionen, og alle regioner sett sammen, er Pakistan det største opphavlandet med 1632 individer. Pakistan skiller seg ut, og 35,9% av innvandrerne i utvalget har Pakistan som opphavsland. Etter Pakistan kommer Vietnam med 545 individer, etterfulgt av India med 290, deretter Tyrkia med 274 og Marokko med 224 individer. Resten av landene overstiger ikke 200 individer. En kan også legge merke til hvordan sammensetningen av opphavsland endrer seg når en ser på de som regnes som første generasjonsinnvandrere. Spesielt vietnameserne består av en høy andel første generasjonsinnvandrere, og hele 77% av vietnameserne er første generasjonsinnvandrere. Vi kan også se fra Tabell 1 at foreldrene kom til Norge på ulike tidspunkt. Av de fem største landene er vietnameserne de som i gjennomsnitt har oppholdt seg kortest i Norge når inntekten måles, mens marokkanerne er de som har oppholdt seg lengst. Vi ser at vietnameserne i gjennomsnitt kom i 1980, og vi ser fra Tabell 1 at pakistannerne i gjennomsnitt kom i begynnelsen av 1972. Disse årstallene stemmer godt overens med norsk innvandringshistorie (Bratsberg et al., 2014, s. F646). Grunnen til å emigrere fra disse landene er ulike, noen land, som for eksempel Vietnam var rammet av krig, mens innvandrere fra land som Pakistan og Tyrkia hovedsakelig kom som arbeidsinnvandrere. I tillegg er kulturen svært annerledes for de ulike landene, dette kan bety at integreringsprosessen har ulik varighet og resultat. Det er derfor viktig å presisere at resultatene presentert i denne oppgaven vil gjelde en svært heterogen gruppe individer.

Tabell 1: Opphavsland m.m.

Land/Region	Antall individ	Andel av total innvandrerpopulasjon	Andel førstegenerasjonsinnvandrere	Gj.snittsalder ved immigrasjon	Gj.snittlig årstall for immigrasjon	Fars Gj.snittsalder i 1994
Pakistan	1632	35,9 %	26 %	24,9 (4,3)	1972	46,7
Vietnam	545	11,8 %	77 %	30,3 (3,9)	1980	44,7
India	290	6,4 %	14 %	25,9 (4,3)	1974	45,6
Tyrkia	274	6,2 %	33 %	24,7 (4)	1973	45,3
Marokko	224	5,1 %	23 %	23,8 (3,9)	1970	47,6
Norden	262	6,1 %	46 %	28 (5,3)	1976	45,7
Vest-Europa	187	4,4 %	50 %	27,8 (5,6)	1976	45,7
Øst-Europa	378	8,3 %	44 %	25,9 (6)	1975	44,7
Nord-, Sør- og Mellom-Amerika	170	4 %	69 %	30,5 (4,7)	1980	44,7
Sentral-, Sør-, og Vest-Asia	126	2,9 %	40 %	32,4 (4,1)	1982	45,1
Sørøst- og Øst-Asia	275	6 %	17 %	28,3 (5,1)	1977	44,9
Afrika	114	2,9 %	43 %	27,7 (4,9)	1975	46,8
Totalt	4477	100 %	1642	27,3	-	-

Notat: Sentral-, Sør- og Sørøst-Asia består Afghanistan, Irak, Iran, Israel, Jordan, Kuwait, Libanon og Syria. Sørøst- og Øst-Asia består av Thailand, Filippinene, Kambodsja, Kina, Japan, Taiwan, Hongkong, Singapore og Sri-Lanka. Vest-Europa u/ Norden består av Tyskland, Frankrike, Spania, Portugal, Belgia, Luxembourg, Nederland, Østerrike, Sveits, Liechtenstein, Monaco, Storbritannia, Irland, Italia og Hellas. Øst-Europa består av Bulgaria, Kroatia, Latvia, Polen, Romania, Russland, Ungarn, Bosnia, Makedonia, Tsjekkia, Serbia og Kosovo. Totalen er annerledes enn tidligere oppgitt på grunn av anonymisering gjennom microdata. Standardfeil rapporteres i parenteser.

Tabell 2: Diverse karakteristika

	Innvandrere		Nordmenn	
	Fedre	Barn	Fedre	Barn
Gj.snittsalder målt i 1994	45,9 (4,32)	-	44,7 (4,05)	-
Kjønns sammensetning	-	Menn: 51,65% Kvinner: 48,35%	-	Menn: 51,3% Kvinner: 48,7%
Andel bosatt i Oslo	-	49 %	-	14 %
Andel bosatt i Stor-Oslo	-	77 %	-	30 %

Notat: Standardfeil rapporteres i parenteser. Se appendiks for [definisjon av Stor-Oslo](#).

I Tabell 2 presenteres gjennomsnittsalderen for innvandrerfedre og norske fedre, kjønns sammensetning hos barna og andelen bosatt i Oslo og Stor-Oslo. Hvor Stor-Oslo er en samlebetegnelse for pendlerområdet til Oslo. Gjennomsnittsalderen til fedrene er 44,9 år for innvandrerne i utvalget, og 43,7 år for andregenerasjonsnordmennene i utvalget. Aldersgjennomsnittet er altså ganske likt, noe som er ønskelig i denne oppgaven. Gjennomsnittsalder presenteres ikke for barna, siden denne vil være lik i de to gruppene. Kjønns sammensetning i de to gruppene er meget lik. Dette er ønskelig ettersom oppgaven ikke foretar analyser på kjønnene hver for seg. Da vet vi at kjønnsbalanse i de to gruppene ikke skal påvirke resultatene. Med dette sagt kan det godt tenkes at det er forskjeller i kjønnene i de to undergruppene. Disse får en ikke kartlagt ettersom det kun gjennomføres kjønnsblandede analyser. Det er også slående å se forskjeller i bosetningsmønster hos innvandrerne og de innfødte. Andelen bosatt i Stor-Oslo er for annengenerasjonsinnvandrerne 75%, hos de innfødte er det samme tallet kun 29%. Bortsett fra dette er innvandrerne spredt jevnt utover Norge, og det kan nevnes at Stavanger og Sandnes, og Bergen bosetter omtrent bosetter 4% innvandrerne hver.

Tabell 3: Fordeling over inntektskvintiler

	Barn		Fedre	
	Innvandrere	Nordmenn	Innvandrere	Nordmenn
Laveste kvintil	29 %	19 %	54 %	20 %
Nest laveste kvintil	19 %	21 %	19 %	20 %
Midterste kvintil	15 %	19 %	12 %	20 %
Nest høyeste kvintil	17 %	21 %	9 %	20 %
Høyeste kvintil	20 %	20 %	7 %	20 %
N	4481	207251	4481	207251

Notat: Barnas inntekt måles fra 2007 – 2015, når de er 30 – 33 år gamle. Fedrenes inntekt måles fra og med 1993 til og med 1998. For fedrene betyr dette at inntekten måles over seks år, og i varierende alder. For de yngste fedrene i utvalget måles inntekten i en alder av 38 – 43 år, og for de eldste i en alder av 55 – 60 år. N viser antall observasjoner.

I Tabell 3 ovenfor kan en se hvordan utvalget er fordelt i inntektskvintiler. Tabellen viser en tendens til at innvandrerne befinner seg i bunnen av inntektsfordelingen, mens nordmennene er jevnere fordelt utover inntektsfordelingen²³. En kan observere at 29% av innvandrerbarna befinner seg i det laveste kvintilet, sammenlignet med 19% av nordmennene. Forskjellen mellom fedrene er mer dramatisk, og en kan se at 54% av innvandrerfedrene befinner seg i det laveste kvintilet. Hos de norske fedrene er dette tallet 20%. Sammenligner en de to nederste kvintilene hos fedrene ser en at 73% av innvandrerfedrene befinner seg i de to laveste inntektskvintilene, sammenlignet med 40% av de norske fedrene.

²³ Her skulle det vært like mange individer i hvert kvintil, men siden microdata ikke lar deg laste inn inntektsfordelingen må denne, som sagt, konstrueres manuelt. Dermed blir antall individer i hvert kvintil noe ujevnt.

Tabell 4: Gjennomsnittslønn i ulike regioner

	Barn		Fedre	
	Innvandrere	Nordmenn	Innvandrere	Nordmenn
Norge	457380 (280472)	485950 (270175)	320141 (159806)	451823 (208449)
Stor-Oslo	458718 (274985)	514777 (281424)	310641 (145817)	502492 (236667)
Oslo	444020 (250929)	537465 (299977)	310712 (148657)	530729 (249941)

Notat: Inntekt er målt i samme periode som i Tabell 3. Standardfeil rapporteres i parenteser.

I Tabell 4 presenteres gjennomsnittet av samlet inntekt til barn av første generasjonsinnvandrere og barn av andre generasjons nordmenn, sammen med gjennomsnittet av samlet inntekt til fedrene deres betinget på ulike regioner. Den samlede inntekten til barn av første generasjonsinnvandrere og fedrene deres er lavere enn nordmennene sin lønn uavhengig av region. Vi kan også se at barn av første generasjonsinnvandrere har høyest lønns gjennomsnitt dersom en betrakter hele Norge i ett, og minst dersom en kun betrakter Oslo. Denne sammenhengen er motsatt av det vi ser hos barna av andre generasjons nordmenn. For denne gruppen er gjennomsnittsinntekten høyest i Oslo, og lavest når en betrakter hele Norge. Vi ser tilsvarende sammenhenger for fedrene i utvalget også. Videre kan vi også legge merke til at forskjellene i inntekt mellom innvandrerfedrene og de norske fedrene er betraktelig større.

I Tabell 5, som vist på side 34, rapporteres utdanning, andel uføretrygdete og andel arbeidsledige sammen med andel utenfor arbeidsstyrken²⁴, for innvandrere og nordmenn betinget på bosted. Utdanningsmessig er ikke forskjellene betinget på bosted oppsiktsvekkende for innvandrerbarna eller innvandrerfedrene i utvalget.

For nordmennene i utvalget er forskjellene i utdanning større når en betinger på bosted. En kan se at for de norske barna er andelen som har videregåendeutdanning som sin høyeste utdanning mye lavere i Oslo og Stor-Oslo, sammenlignet med hele Norge. I gjengjeld er andelen med mastergrad som høyeste fullførte utdanning mye høyere i Stor-Oslo og Oslo.

Når en observerer forskjeller i utdanning for de norske fedrene er det slående hvor mye utdanning varierer med bosted. Andelen uten utdanning er noe større i Oslo sammenlignet med hele Norge. Videre kan en merke seg at andelen som har videregåendeutdanning som høyeste utdanning i Oslo er høyere enn i hele Norge. Det er også andelen som har bachelor og/eller master som høyeste utdanning, som altså betyr at fedrene i Oslo er høyere utdannet enn snittet i Norge.

Fra Tabell 5 kan en også se at de norske fedrene har et mye høyere utdanningsnivå sammenlignet med innvandrerfedrene i utvalget. 23 % av innvandrerfedrene har ingen utdanning i Norge, det samme tallet for norske fedre er 8 %. En kan også merke seg at større andel av innvandrerfedrene er uføretrygdet sammenlignet med norske fedre. Hele 14 % av innvandrerfedrene sammenlignet med kun 3 % av de norske er uføretrygdte, lignende funn blir gjort av Bratsberg et al. (2014, s. F657).

Forskjellene i utdanning jevnes noe ut hos barna i utvalget når vi ser på hele Norge. På tross av dette virker det som om det generelle utdanningsnivået fremdeles er noe høyere for norske barn sammenlignet med innvandrerbarn. Dette gjelder for videregåendeutdanning og opp, ekskludert andel med mastergrad, hvor innvandrerne overgår nordmennene i utvalget.

I Oslo er bildet noe annerledes. En ser at andelen av norske barn med bachelorgrad og/eller mastergrad er mye høyere for nordmennene, sammenlignet med innvandrerne i utvalget. Her er det også verdt å merke seg forskjellen i andelen av innvandrere født i Oslo, sammenlignet med andelen av norske barn født i Oslo. Denne forskjellen i fødested tilsier at langt flere av de norske barna har flyttet til Oslo, som igjen betyr at langt flere av innvandrerbarna har blitt født i Oslo. Når en studerer denne forskjellen må en også ha i mente at 37% av innvandrerbarna er førstegenerasjonsinnvandrere, og dersom disse ekskluderes er andelen født i Oslo av de som er bosatt i Oslo 77,5% for innvandrerne.

²⁴ SSB skriver at personer utenfor arbeidsstyrken er personer som ikke er i jobb, og som samtidig ikke kvalifiserer som arbeidsledige. Grunnen til at de ikke kvalifiserer som arbeidsledige er fordi de ikke fort nok kan entre en ny jobb, og ikke aktivt søker etter jobb. Personer utenfor arbeidsstyrken vil omfatte studenter, uføre, hjemmearbeidene, personer uten arbeid (SSB, 2019).

Tabell 5: Utdanning over region

	Innvandrere		Nordmenn	
	Fedre	Barn	Fedre	Barn
Norge				
Ingen utdanning	23 %	6 %	8 %	0 %
Grunnskole	48 %	29 %	46 %	22 %
Videregående	17 %	30 %	26 %	35 %
Bachelor	7 %	18 %	12 %	29 %
Master	5 %	17 %	8 %	14 %
Andel uføre	14 %	2,4 %	3 %	2,3 %
Andel arbeidsledig og utenfor arbeidsstyrken	-	23 %	-	14 %
Stor-Oslo				
Andel født i bosted	-	56 %	-	57 %
Ingen utdanning	23 %	4 %	11 %	0 %
Grunnskole	50 %	30 %	37 %	22 %
Videregående	16 %	31 %	23 %	28 %
Bachelor	6 %	18 %	15 %	28 %
Master	5 %	17 %	14 %	22 %
Andel uføre	18 %	2,6 %	2,5 %	0 %
Andel arbeidsledig og utenfor arbeidsstyrken	-	23 %	-	12 %
Oslo				
Andel født i bosted	-	52 %	-	37 %
Ingen utdanning	24 %	8 %	13 %	1 %
Grunnskole	50 %	30 %	29 %	19 %
Videregående	16 %	29 %	49 %	19 %
Bachelor	6 %	17 %	17 %	30 %
Master	4 %	16 %	19 %	31 %
Andel uføre	16 %	2,4 %	2 %	0 %
Andel arbeidsledig og utenfor arbeidsstyrken	-	23 %	-	11 %

Tabell 6 : Utdanning over inntektskvintil for innvandrere

Inntektskvintil	Ingen Utdanning	Grunnskole	VGS	Bachelorgrad	Mastergrad
1	10 %	42 %	30 %	10 %	8 %
2	5 %	32 %	37 %	18 %	7 %
3	4 %	24 %	36 %	25 %	10 %
4	3 %	20 %	28 %	27 %	22 %
5	2 %	17 %	22 %	20 %	40 %

Tabell 7: Utdanning over inntektskvintil for nordmenn

Inntektskvintil	Grunnskole	VGS	Bachelorgrad	Mastergrad
1	40 %	34 %	19 %	7 %
2	22 %	41 %	31 %	6 %
3	17 %	34 %	39 %	10 %
4	15 %	36 %	32 %	18 %
5	14 %	32 %	25 %	29 %

Tabell 6 og 7 ovenfor viser hvordan utdanning fordeler seg over inntektskvintilene til innvandrere og nordmenn i utvalget. En kan merke seg at i motsetning til innvandrerne, har alle nordmennene i utvalget minimum fullført grunnskoleutdanning. En kan også legge merke til at en veldig høy andel av innvandrerne i øverste inntektskvintil har mastergrad. Andelen med mastergrad i øverste inntektskvintil hos innvandrerne er 11 prosentpoeng større enn hos nordmennene i utvalget. Det samme gjelder i fjerde inntektskvintil, hvor fire prosentpoeng flere av innvandrerne har mastergrad.

4.3.1 Oppsummering av deskriptiv statistikk

Som vi har sett i dette kapittel er innvandrergruppen en svært heterogen gruppe. De stammer fra mange ulike land, og ankom Norge på ulike tidspunkt. En kan også merke seg at en betraktelig andel av innvandrerne er fra Pakistan og Vietnam. Hvor en høy andel vietnamesere er førstegenerasjonsinnvandrere som innvandret før en alder av syv år. Det kommer også frem at en stor andel av innvandrerne er bosatt i Oslo, og Stor-Oslo. Videre kan vi legge merke til at en svært høy andel av innvandrerfedrene har inntekt som plasserer dem i det laveste inntektskvintilet. Innvandrerfedrene har også betraktelig lavere inntektsgjennomsnitt enn de norske fedrene. Innvandrerbarna er mye jevnere fordelt over inntektsfordelingen. Samtidig kan vi merke oss at gjennomsnittsinntekten til innvandrerbarna ligner mye mer på nordmennene sitt gjennomsnitt, sammenlignet med innvandrerfedrene i utvalget. På tross av dette er det fremdeles inntektsulikheter mellom innvandrere og nordmenn, og den er størst når vi kun betrakter Oslo. Vi kan også merke oss at innvandrerbarna har mye høyere utdanning sammenlignet med fedrene sine. Det er også slik at en høy andel av nordmennene i Oslo har høyere utdanning. Det er også her forskjellene i utdanning mellom innvandrere og nordmenn gjør seg mest gjeldende. Til slutt kan en også merke seg at en høy andel av innvandrerne som har inntekt som plasserer dem i de to øverste inntektskvintilene også har høy grad av utdanning.

Kapittel 5

Resultater

I dette kapittelet presenteres IGE spesifikasjoner og resultatene fra de ulike målemetodene. Først, i delkapittel 5.1.1 presenteres empiriske spesifikasjoner av IGE-modellene. Deretter presenteres resultatene fra IGE i delkapittel 5.1.2, samtidig diskuteres det om eventuelle forskjeller er statistisk signifikante. I delkapittel 5.2 presenteres og sammenlignes resultater fra overgangsmatrisene for de to undergruppene. Deretter presenteres estimerte overgangssannsynligheter i delkapittel 5.3, og retningsrankmobilitet i delkapittel 5.4. Til slutt diskuteres resultatene i delkapittel 5.5.

5.1 IGE spesifikasjoner og resultater

I dette delkapittelet presenteres først de ulike spesifikasjonene for IGE-modellene i delkapittel 5.1.1. Disse IGE spesifikasjonene omfatter modeller betinget på ulike regionale inndelinger og hvorvidt individet er innvandrer eller ei. I tillegg presenteres det modeller med interaksjoner, disse skal undersøke om eventuelle forskjeller i IGE-estimat med ulike betingelser er statistisk signifikante. I delkapittel 5.1.2 presenteres resultatene fra disse modellene.

5.1.1 IGE spesifikasjoner

For å studere forskjeller i IGE hos innvandrere og nordmenn estimeres det en enkel IGE betinget på individene enten er innvandrere eller nordmenn, det blir også betinget på bosted:

$$y_{i,s,b}^C = \alpha + \beta y_{i,s,b}^P + \varepsilon_i \quad (7)$$

y er et mål på permanentinntekt, hvor liten y indikerer at målet er på log-form. C viser til barns inntekt, og P til fedres inntekt, s indikerer innvandrerstatus. Ved $s = 0$ er individet nordmann, og ved $s = 1$ er individet innvandrer. b er bosted, og denne vil variere mellom hele Norge, Stor-Oslo og Oslo. Det betinges på bosted for å gjøre opp for at en stor andel av innvandrerne er bosatt i Oslo-området, og for å måle hvor vedvarende inntektsforskjeller er i Oslo-området.

Hvis en observerer forskjeller i β , hos for eksempel innvandrere og nordmenn bosatt i Oslo, i modellen presentert ovenfor, kan en ikke med sikkerhet fastslå at disse forskjellene er statistisk signifikante. For å undersøke dette benyttes det to modeller med interaksjonsledd:

$$y_i^C = \alpha + \delta * inn_i + \beta_1 * y_i^P + \beta_2 * inn_i * y_i^P + \varepsilon_i \quad (8)$$

Modell (8) tilsvarende Modell (7) presentert ovenfor. Forskjellen er inklusjonen av to nye ledd. Den binære variabelen inn_i , som er én dersom individet er innvandrer, inkluderes for å ta høyde for

eventuelle forskjeller i inntekt mellom innvandrere og nordmenn. Interaksjonen mellom y_i^P og inn_i , undersøker om eventuelle forskjeller mellom innvandrere og nordmenn med IGE i øyemed er statistisk signifikante.

Modellen presentert i ligning (8) kan enkelt utvides med en bostedsvariabel. Ved hjelp av bostedsvariabelen undersøkes det om eventuelle forskjeller i IGE betinget på bosted er statistisk signifikante. Under presenteres en utvidet interaksjonsmodell med bostedsvariabel:

$$y_i^C = \alpha + \partial_1 * inn_i + \partial_2 * bosted_i + \beta_1 * y_i^P + \beta_2 * inn_i * y_i^P + \beta_3 * bosted_i * y_i^P + \partial_3 * inn_i * bosted_i + \beta_4 * inn_i * y_i^P * bosted_i + \varepsilon_i \quad (9)$$

Modellen ovenfor utvides med en binær variabel for bosted. Denne representer enten Oslo eller Stor-Oslo, hvor $bosted_i = 1$ indikerer at individet er bosatt i en av de to. Denne binære variabelen må interagere med alle de andre observerte variablene i modellen, som betyr at antall ledd i ligningen øker eksponentielt med inklusjon av kun én ny variabel.

Tolkningen av Modell (9) vil være tilsvarende som Modell (8), bortsett fra de to nye interaksjonsleddene, $(bosted_i * y_i^P)$ og $(inn_i * y_i^P * bosted_i)$. Dersom $(bosted_i * y_i^P)$ er statistisk signifikant betyr dette at eventuelle forskjeller i IGE observert mellom geografiske områder i Modell (7) også er statistisk signifikante. Dette interaksjonsleddet skal, alt annet likt, undersøke signifikans ved forskjeller i ulike geografiske inndelinger for både innvandrere og nordmenn i utvalget. Vær oppmerksom på at resultatene kan bli påvirket av det skjeve forholdet mellom innvandrere og nordmenn i utvalget. Interaksjonsleddet $(inn_i * y_i^P * bosted_i)$ undersøker om det er noen særegne effekter for innvandrere bosatt i området angitt av bostedsvariabelen utover effekten fanget av $(bosted_i * y_i^P)$.

Dummyer som inn_i , $bosted_i$ og interaksjonen dem imellom inkluderes, på tilsvarende måte som i Modell (7). Altså inkluderes de for å ta høyde for at innvandrere bosatt i og utenfor Oslo eller Stor-Oslo kan ha varierende inntekt, og at nordmenn bosatt i og utenfor Oslo også kan ha forskjeller i inntekt.

5.1.2 IGE resultater

Under rapporteres ulike IGE-estimer representert med β . IGE rapporteres for innvandrere og nordmenn bosatt i henholdsvis hele Norge, Stor-Oslo og Oslo hver for seg. IGE estimatet for alle innvandrere i Norge er på 0,13, dette tallet er svært lavt, og tilsier en meget høy mobilitet for innvandrere i Norge. En kan også observere at innvandrere, gitt at forskjellene er statistisk signifikante, er mest mobile i området definert som Stor-Oslo. Som en kan se fra tabellen under er forskjellene relativt små. Dette kan tilsi at region ikke påvirker innvandrernes IGE-estimat i særlig grad, noe som er logisk ettersom så mange av innvandrerne i utvalget er bosatt i Oslo-området.

Tabell 8: Resultater fra IGE-estimer betinget på ulike regioner

Område	Innvandrere		Nordmenn	
	β	N	β	N
Norge	0,13*** (0,02)	4477	0,18*** (0,003)	207256
Stor-Oslo	0,12*** (0,02)	3463	0,17*** (0,005)	61814
Oslo	0,13*** (0,01)	2191	0,16*** (0,007)	29263
Norge uten Oslo	0,12*** (0,028)	2286	0,18*** (0,0027)	177993

Notater: Standardfeil er rapportert i parenteser. * tilsvarer signifikansnivå på 0,1. ** tilsvarer signifikansnivå på 0,05. *** tilsvarer signifikansnivå på 0,01.

For nordmennene i utvalget blir den estimerte IGE 0,18, dette er i tråd med tidligere artikler som omhandler intergenerasjonell inntektsmobilitet i Norge (Bratberg et al, 2005, s. 427; Black & Devereux, 2011, s. 1496). En IGE på 0,18 er meget lavt, altså tilsier IGE at Norge er et land med høy grad av IGM. Vi merker oss også at innvandrere, ifølge IGE, har høyere grad av IGM enn nordmenn i utvalget. Videre kan vi merke oss at nordmennene i utvalget fremstår som mest mobile når en kun studerer Oslo, og minst mobile når vi studerer hele Norge i ett. Region fremstår også som mer utslagsgivende for nordmenn sin estimerte IGE, sammenlignet med innvandrere. En kan merke seg at forskjellene er ganske små også for nordmennene i utvalget.

For å undersøke om forskjellene i resultatene presentert i tabellen ovenfor er statistisk signifikante rapporteres modeller med interaksjonsledd. I Modell (1) undersøkes det om forskjellen mellom nordmenn og innvandrere er signifikant. Dette gjøres ved å se på signifikansnivået til interaksjonen mellom β og Inn. Vi ser at denne interaksjonen har et signifikansnivå på 99%, dette indikerer at forskjellen mellom nordmenn og innvandrere i Tabell 8 er signifikant. Ved hjelp av F-tester blir det også fastslått at de inkluderte innvandrersvariablene har samlet signifikans i de tre rapporterte modellene. Det samlede signifikansnivået er på 0,01

Tabell 9: Resultater fra IGE-estimeringer med interaksjonsledd

	Modell 1	Modell 2	Modell 3
y_i^P	0,18*** (0,0025)	0,18*** (0,003)	0,18*** (0,003)
Inn	0,64*** (0,20)	0,76*** (0,27)	0,36 (0,38)
y_i^P * inn	-0,06*** (0,016)	-0,06*** (0,022)	-0,04 (0,03)
Oslo		0,33*** (0,089)	
SO			0,21*** (0,07)
y_i^P * Oslo		-0,02*** (0,01)	
y_i^P * SO			-0,012** (0,005)
Inn * Oslo		-0,63 (0,40)	
SO * Inn			0,05 (0,45)
y_i^P * Inn * Oslo		0,041 (0,032)	
y_i^P * Inn * SO			-0,005 (0,034)

Notater: Standardfeil er rapportert i parenteser. N = 211733 i alle de tre modellene. * tilsvarer signifikansnivå på 0,1. ** tilsvarer signifikansnivå på 0,05. *** tilsvarer signifikansnivå på 0,01. Resultater fra F-tester tilsier at innvandrersvariablene har samlet signifikans i de tre modellene. F-verdiene gir et signifikansnivå på 0,01, [se appendiks](#).

I Modell (2) undersøkes signifikansen ved å være innvandrere, bosatt i Oslo og interaksjonen dem imellom. Vi observerer fremdeles at innvandrere har høyere grad av IGM, og at denne igjen har et signifikansnivå på 0,01. Med interaksjonen mellom β og Oslo undersøkes det om effekten av å være bosatt i Oslo er signifikant. Vi ser at det å bo i Oslo har en signifikant negativ effekt på IGE. Signifikansnivået til denne interaksjonen er på 0,05. Dette betyr at forskjellene i IGE, observert i Tabell 8, for nordmenn bosatt i og utenfor Oslo er signifikant. På tross av signifikant interaksjon skal en være forsiktig med å konkludere med at denne effekten også er gjeldende for innvandrerne. Grunnen til dette er at $\beta * \text{Oslo}$ interaksjonen trolig kan bli drevet av en stor mengde nordmenn i forhold til innvandrere i utvalget. Interaksjonen mellom β , Inn og Oslo undersøker effekten av å være innvandrere bosatt i Oslo, vi ser at denne effekten er ikke-signifikant. Dette betyr at det ikke er noen signifikant særegen effekt av å være innvandrere i Oslo på IGE-estimatet.

I Modell (3) undersøkes effekten av å være innvandrere, bosatt i Stor-Oslo og interaksjonen dem imellom. Her ser vi at effekten av å være innvandrere på IGE ikke lengre er signifikant. Dette kan være grunnet det store overlappet mellom de to variablene SO og Inn. Som tidligere nevnt bor hele 75% av innvandrerne i området definert som Stor-Oslo. Dette betyr altså at det er 75% overlapp mellom Inn og SO. En konsekvens av dette kan være at effekten av å være innvandrere blir fanget opp i bostedsvariabelen SO, som igjen fører til at $(\beta * \text{Inn})$ interaksjonen blir ikke-signifikant. Igjen kan det være nyttig å påpeke at innvandrervariablene i Modell (3) har et samlet signifikansnivå på 0,01. Effekten av å være bosatt i Stor-Oslo på IGE, ligner på effekten av å være bosatt i Oslo. Denne er også signifikant ved 0,05. Igjen skal en være forsiktig med å trekke slutningen at denne effekten også er gjeldende for innvandrerne bosatt i Stor-Oslo. Til slutt tolkes interaksjonen mellom β , Inn og SO, hvor vi undersøker særegne effekter av å være innvandrere i Stor-Oslo. Denne interaksjonen tilsier at en særegen effekt av å være innvandrere i Stor-Oslo, utover effekten av å være bosatt i Stor-Oslo fanget av interaksjonen ($bosted_i * y_i^P$), ikke er signifikant.

For å oppsummere resultatene ovenfor ser en at IGE-estimatene er betydelig lavere for innvandrerne i utvalget, sammenlignet med nordmenn i utvalget. Når vi deler inn etter bosted ser vi at resultatene er mer tvetydige, bosted ser ut til å påvirke nordmenn mer enn innvandrere. Dette er ikke urimelig, ettersom en stor andel av innvandrerne er bosatt i Oslo-området. I Modell (3) ser vi også at effekten av å være innvandrere ikke lengre er signifikant.

5.2 Overgangsmatriser

Nedenfor presenteres resultatene fra overgangsmatriser delt opp i kvintiler. Her er utvalget betinget på hvorvidt individene er nordmenn eller innvandrere. Resultatene for nordmennene rapporteres i den venstre matrisen, mens resultatene for innvandrerne rapporteres til høyre. På den vertikale akse er fedrenes inntektskvintil representert. På den horisontale akse er barnas inntektskvintil representert. Det presenteres også et tilhørende mål på immobilitet, kalt immobilitetsindeks.

Tabell 10: Resultater fra overgangsmatrise

	Nordmenn					Innvandrere				
	1	2	3	4	5	1	2	3	4	5
1	0,27	0,25	0,19	0,17	0,12	0,32	0,19	0,15	0,16	0,18
2	0,20	0,24	0,21	0,20	0,15	0,26	0,20	0,16	0,20	0,17
3	0,19	0,21	0,21	0,21	0,18	0,29	0,18	0,11	0,20	0,23
4	0,17	0,19	0,20	0,23	0,22	0,24	0,18	0,15	0,18	0,24
5	0,16	0,15	0,17	0,22	0,30	0,22	0,17	0,12	0,15	0,34

Immobilitetsindeks: 2,96

Immobilitetsindeks: 2,67

Fra matrisen kan vi se at nordmenn med fedre blant de 20% mest fattige i fedrenes inntektsfordeling selv har en tendens til å havne i en av de laveste prosentilene i sin egen inntektsfordeling. Dette er å forvente, og en observerer ofte at det blir fortetning i bunnen og toppen av inntektsfordelingen når en analyserer barn av fedre i det laveste og høyeste inntektskvintilet. Dette er fordi disse barna kun kan bevege seg i en retning i inntektsfordelingen. Videre kan det merkes at nordmennene har litt større sannsynlighet for å havne et sted mellom de 41-60 % rikeste, sammenlignet med sannsynligheten for å havne blant de 61-80 % rikeste. Til slutt kan en merke seg at de har minst sannsynlighet for å havne i det øverste kvintilet, og kun 12% av barna med fedre i det laveste inntektskvintilet presterer å gjøre spranget til topps i sin egen inntektsfordelingen.

Resultatene for innvandrerbarna med fedre i første inntektskvintil er noe annerledes sammenlignet med tilsvarende nordmenn. Vi ser at fortetningen i det aller laveste kvintilet er tydeligere hos innvandrerne enn hos nordmennene. Samtidig ser vi også at flere innvandrerbarn hopper til topps i inntektsfordelingen.

Dersom en sammenligner resultatene for innvandrere og nordmenn med fedre i det andre kvintilet ser en lignende tendens. Altså at innvandrerbarna har en sterkere tendens til å havne i det nederste eller øverste inntektskvintilet. Nordmennene har en tydeligere tendens til å bevege seg inn mot midten av inntektsfordelingen.

Denne tendensen gjentar seg for øvrige kvintiler. En kan dog merke seg at det er få innvandrerfedre i de øverste inntektskvintilene. Dette betyr at resultatene for innvandrerbarn av fedre i de øvrige inntektskvintilene ikke er like robuste.

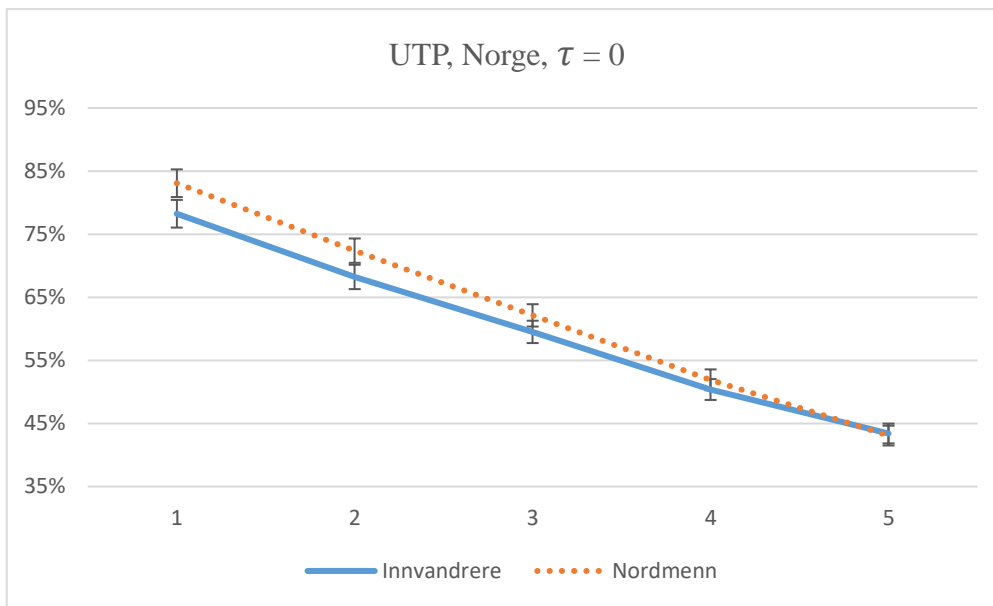
Det grove bildet matrisen tegner ser ut til å være at nordmennene i større grad konvergerer til midten av inntektsfordelingen uavhengig av fars tilhørende kvintil. Overgangsmatrisen maler et annerledes bilde for innvandrerne. En kan se at det er større sannsynlighet for at innvandrerbarna, sammenlignet med nordmenn, havner i et ekstremt kvintil. Altså er sannsynligheten høyere for at en innvandrer havner i det første eller femte kvintilet, uavhengig av fars plassering i inntektsfordelingen. En mulig grunn til at innvandrerbarn foretar lengre bevegelser opp i inntektsfordelingen kan være at innvandrerfars inntektskvintil i mindre grad reflekterer fars evner. Innvandrerfedrene har ofte kommet fra land der tilgangen til utdanning er lavere enn i Norge, dermed har de ikke fått samme grad av utdanning som innfødte nordmenn. I tillegg kan de oppleve en språkbarriere og, ofte, en vesentlig kulturforskjell. Barna til innvandrerfedrene derimot vil ha hatt større tilgang til utdanning, vil ofte kunne snakke norsk og samtidig ha større mulighet for å forstå den norske kulturen. Med dette tatt i betraktning vil de innvandrerbarna som har arvet gode vaner, og evner fra foreldrene sine ha større mulighet til å plassere i en inntektskvintil som reflekterer deres evner. Dermed ser vi en høyere andel av innvandrerbarna som hopper til det øverste inntektskvintilet sammenlignet med nordmenn.

Bratsberg et al. (2014) har studert mange ulike aspekter med innvandring i Norge. De skriver at innvandrerbarna overgår nordmenn i begge sider av utdanningsskalaen. Med dette menes det at det er en høyere prosentandel av annengenerasjonsinnvandrere som fullfører doktorgrader, og en høyere andel som ikke fullfører videregåendeutdanning sammenlignet med nordmenn. Gitt at utdanning kan trekkes som en parallell for inntekt, vil Bratsberg et al. (2014, s. F673) resultat samsvare med funnene i overgangsmatrisene ovenfor. Altså er det en tendens til at innvandrerbarna enten gjør det veldig bra eller mindre bra i både utdanningssammenheng og i arbeidsmarkedsutfall.

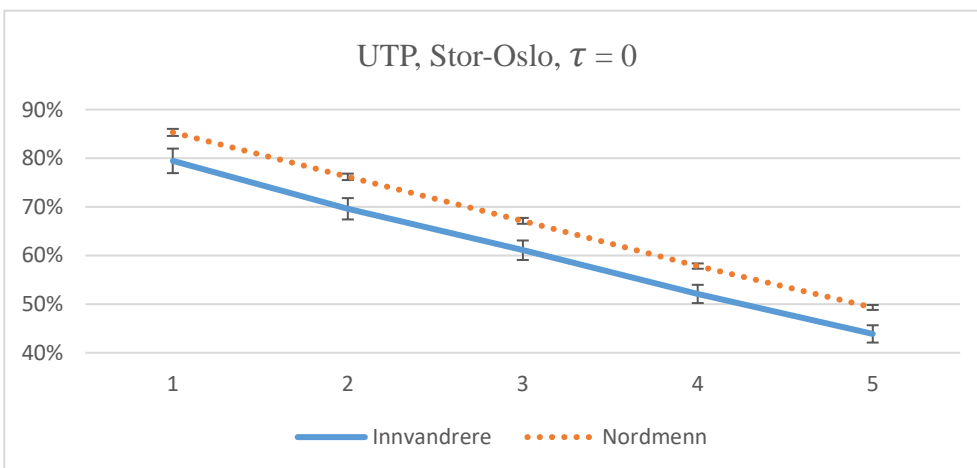
5.3 Resultater fra overgangssannsynligheter

I denne seksjonen presenteres resultatene fra de estimerte overgangssannsynlighetene med tilhørende grense (τ) for henholdsvis hele Norge, Oslo og Stor-Oslo. Først tolkes resultatene fra oppoverovergangssannsynlighetene (UTP) med en $\tau = 0$ som vist på neste side. Deretter tolkes resultatene fra UTP med $\tau = 0,3$, som betyr at individet må ha foretatt et sprang på minst fire desil for at det skal tolkes som en oppoverbevegelse. Ved hjelp av en høyere τ ser en hvem som er mest tilbøyelig til å foreta større byks i inntektsfordelingen. På den vertikale akse vises tilbøyelighet for å gjøre bevegelser opp eller ned i inntektsfordelingen. På den horisontale akse vises hvilket inntektskvantil en studerer.

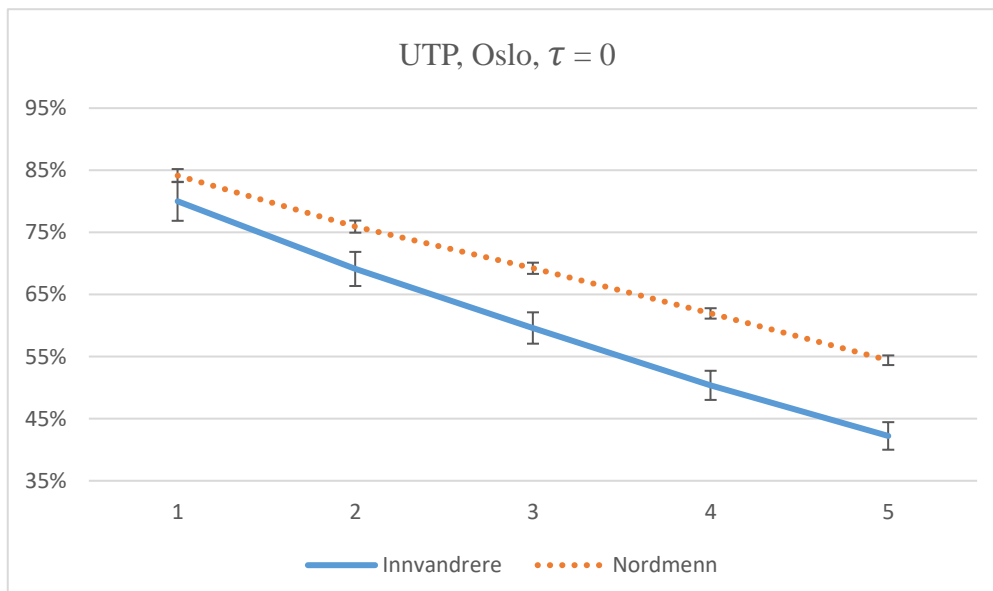
Figur 2: UTP Norge, $\tau = 0$



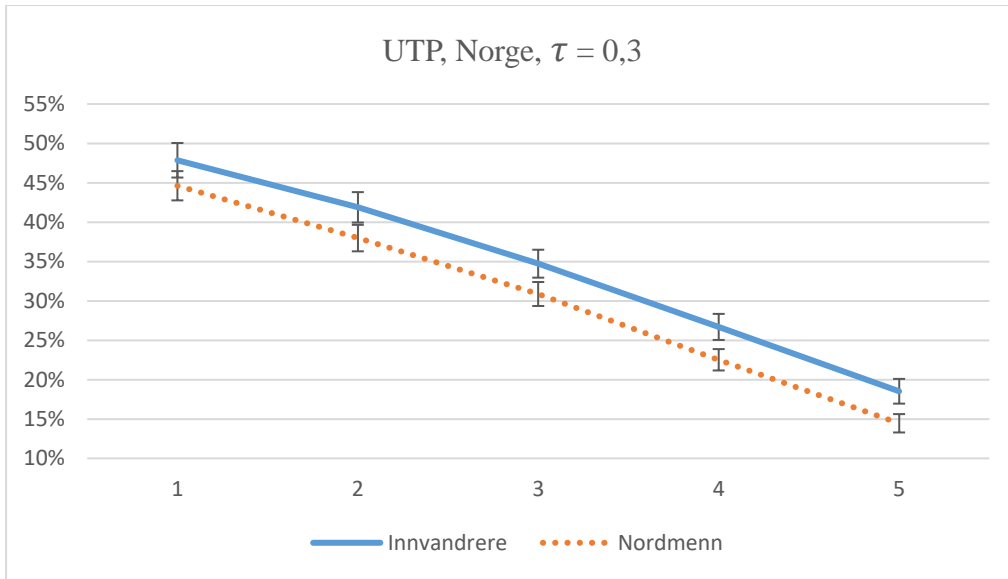
Figur 3: UTP, Stor-Oslo, $\tau = 0$



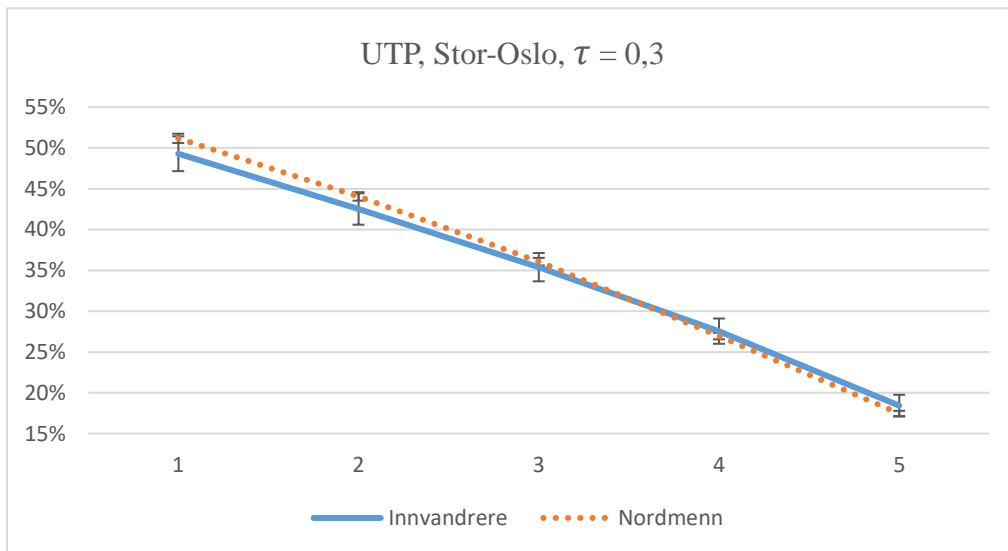
Figur 4: UTP, Oslo, $\tau = 0$



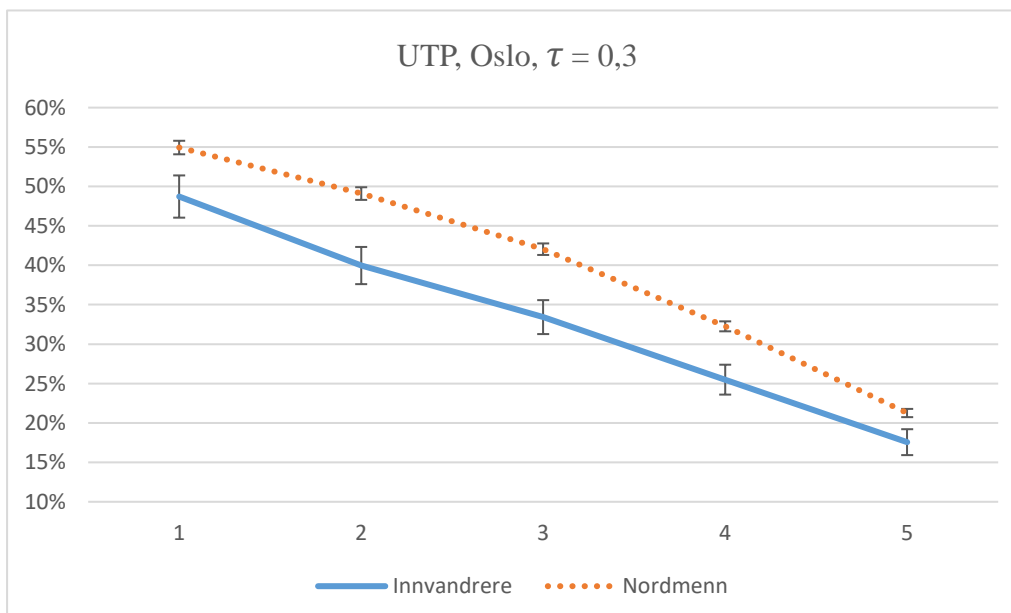
Figur 5: UTP, Norge, $\tau = 0,3$



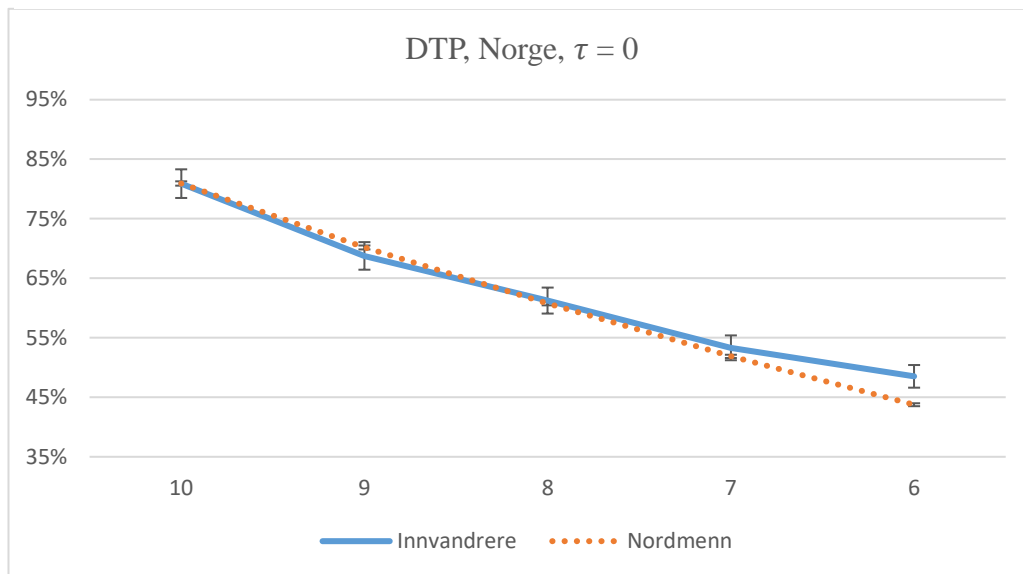
Figur 6: UTP, Stor-Oslo, $\tau = 0,3$



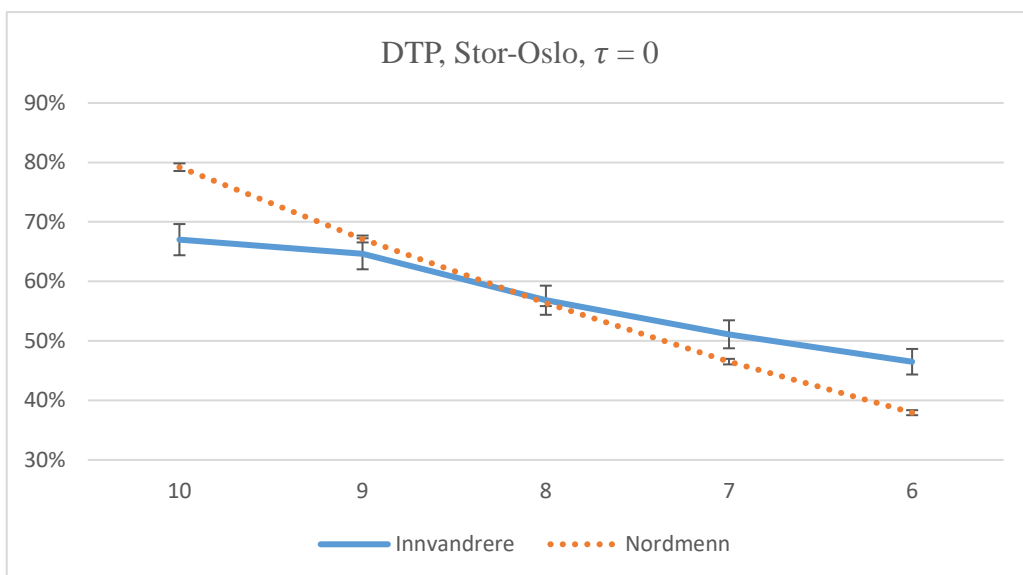
Figur 7: UTP, Oslo, $\tau = 0,3$



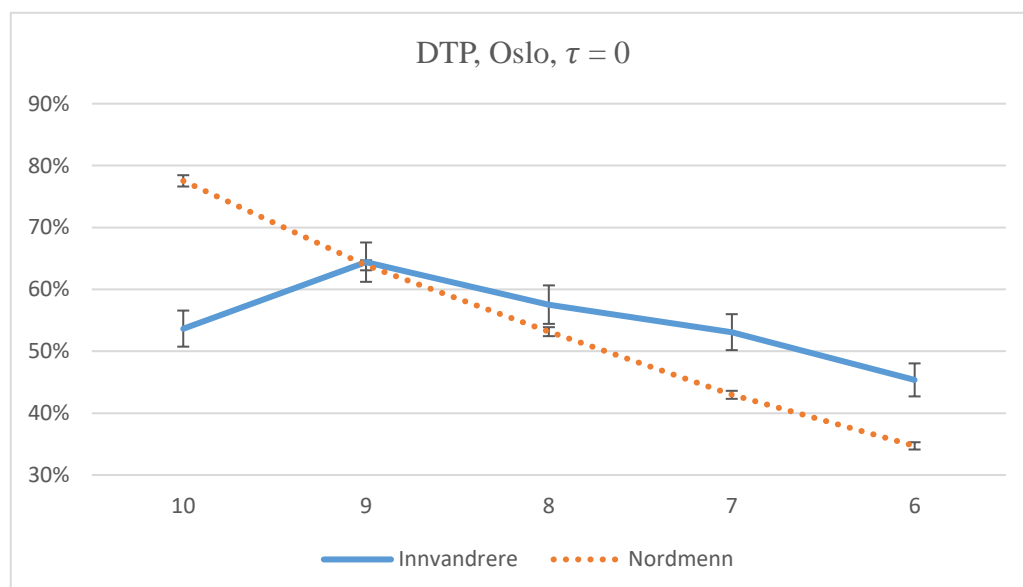
Figur 8: DTP, Norge, $\tau = 0$



Figur 9: DTP, Stor-Oslo, $\tau = 0$



Figur 10: DTP, Oslo, $\tau = 0$



Resultatene når hele Norge blir tatt i betraktning viser at nordmenn med fedre i de fire laveste desil har marginalt høyere sannsynlighet for å bevege seg oppover sammenlignet med innvandrerne i samme desil. En kan merke seg at forskjellene kun er statistisk signifikant i det første desilet. Når det siste desilet blir tatt i betraktning ser vi at forskjellen i oppovermobilitet forsvinner helt, og tilbøyeligheten til å bevege seg opp er helt lik for de to gruppene.

Tilsvarende resultater for Stor-Oslo tilsier at nordmenn er mer oppovermobile enn innvandrerne. Vi ser at dette gjelder i alle inntektsdesil, hvorpå forskjellene også er statistisk signifikant. Vi kan også merke oss at både innvandrerne og nordmennene er noe mer oppovermobile i Stor-Oslo, sammenlignet med hele Norge.

Resultatene fra Oslo maler et ganske likt bilde som resultatene fra Stor-Oslo. Her kan vi se at sannsynligheten for å stige opp i første desil er noe lavere for nordmennene, sammenlignet med resultatene for nordmenn i Stor-Oslo. I det første desilet har nordmenn 4 prosentpoeng større tilbøyelighet til å stige oppover i inntektsfordelingen, denne forskjellen er ikke statistisk signifikant. Vi kan også se at forskjellen i oppovermobilitet øker og blir statistisk signifikant utover i diagrammet, og i det siste inkluderte desilet har denne forskjellen økt til godt over ti prosentpoeng.

Dersom en studerer diagrammet nøye ser en at forskjellene i oppovermobilitet mellom innvandrerne og nordmenn i utvalget er noe høyere i Oslo, sammenlignet med Stor-Oslo. Dette kan være fordi Stor-Oslo er et større geografisk område. Dermed vil resultatene ligne mer på tilsvarende resultater fra hele Norge.

Når grensen, τ , justeres opp blir bildet for hele Norge helt annerledes. Vi ser at innvandrerne fremstår som mer oppovermobile sammenlignet med nordmenn. Kurven til innvandrerne ligger over nordmennene sin kurve i alle desil. Dette kan tolkes som at innvandrere har høyere sannsynlighet for å foreta større byks oppover i inntektsfordelingen sammenlignet med nordmenn i utvalget. Det kan merkes at forskjellene kun er statistisk signifikante i fjerde, femte og sjette desil.

Resultatene fra Stor-Oslo, med en $\tau = 0,3$, tilsier at innvandrerne og nordmennene er omtrent like oppovermobile i alle inntektsdesil.

Når en betinger på at individene skal være bosatt i Oslo er igjen nordmennene mer oppovermobile enn innvandrerne i utvalget, disse forskjellene er statistisk signifikant. Nordmennene i utvalget har omtrent fem til ti prosentpoeng høyere sannsynlighet til å foreta en oppoverbevegelse i Oslo, avhengig av hvilket desil en observerer. Igjen kan en merke seg at resultatene fra Stor-Oslo legger seg mellom resultatene Oslo og Norge. Intuisjonen bak dette er trolig lik som ovenfor.

Med resultatene fra UTP $\tau = 0$ og $\tau = 0,3$ tatt i betraktning ser det ut til at innvandrere i hele Norge er litt mindre tilbøyelig til å foreta kortere bevegelser i inntektsfordelingen. Derimot kan det virke som

om innvandrerne, når hele Norge blir tatt i betraktning, er mer tilbøyelig til å foreta større byks i inntektsfordelingen.

Når utvalget betinges på at individene er bosatt i Stor-Oslo, fremstår nordmenn som mer mobile enn innvandrere når $\tau = 0$. Denne forskjellen forsvinner når grensen økes til 0,3, og innvandrerne og nordmennene fremstår som like oppovermobile.

Resultatene endres ytterligere når vi betinger utvalget på at de må være bosatt i Oslo. Vi ser at for både $\tau = 0$ og $\tau = 0,3$ at nordmenn er mer mobile oppover enn innvandrere. Mulige bakenforliggende mekanismer til dette vil bli diskutert i slutten av dette kapittelet.

Her diskuteres resultatene fra de estimerte nedoverovergangssannsynligheter (DTP). Når en studerer disse resultatene kan en ha i mente at det er veldig få innvandrerfedre i toppen av inntektsfordelingen.

Resultatene fra DTP i hele Norge tilsier at innvandrere og nordmenn er omtrent like tilbøyelig til å falle ned i inntektsfordelingen. Dette endrer seg noe utover i diagrammet. I det sjette inntektsdesilet har nordmenn mindre tilbøyelighet til å falle nedover i inntektsfordelingen, sammenlignet med innvandrerne i utvalget. Denne forskjellen er statistisk signifikant.

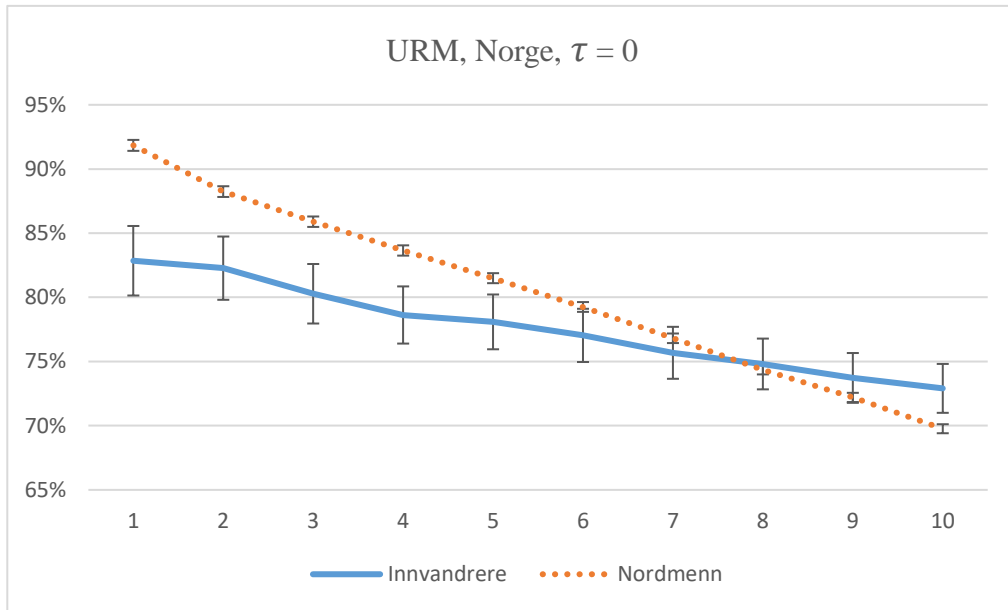
Resultatene fra Stor-Oslo tilsier at nordmenn i det øverste inntektsdesilet er statistisk signifikant mer tilbøyelig til å falle ned i inntektsfordelingen sammenlignet med innvandrerne. Dette endrer seg, og innvandrerne og nordmennene er omtrent like tilbøyelig til å falle ned i inntektsfordelingen fra det niende og åttende inntektsdesilet. I det syvende inntektsdesilet er innvandrerne statistisk signifikant mer tilbøyelig til å falle ned i inntektsfordelingen, dette gjelder i det sjette inntektsdesilet også.

Resultatene fra Oslo ligner på resultatene fra Stor-Oslo, men forskjellene i mobilitet er mer markante. I de øverste desilet er nordmenn statistisk signifikant mer nedovermobile enn innvandrerne. Dette snur i det niende desilet, og innvandrerne og nordmennene er omtrent like nedovermobile. Fra det åttende inntektsdesilet blir innvandrerne statistisk signifikant mer nedovermobile enn nordmennene.

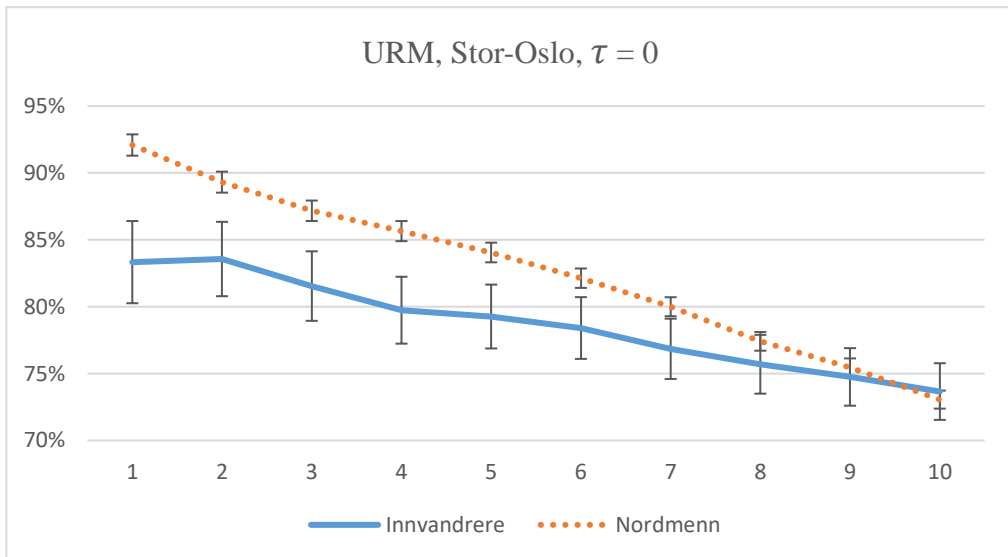
5.4 Resultater fra retningsrankmobilitet

Under presenteres retningsrankmobilitetsresultatene fra hele Norge, Stor-Oslo og Oslo. Først presenteres oppovermobilitetsresultatene med $\tau = 0$ og $\tau = 0,3$. Her kan en ha i mente at inntektsfordelingen er i vintil, og ikke i desil. Når en inntektsfordelingen er i vintil og $\tau = 0,3$, vil en bevegelse på minst 35 prosentiler telle som en oppoverbevegelse. Det blir også presentert nedovermobilitetsresultater, her benyttes desilfordeling på grunn av få innvandrerfedre i toppen av inntektsfordelingen.

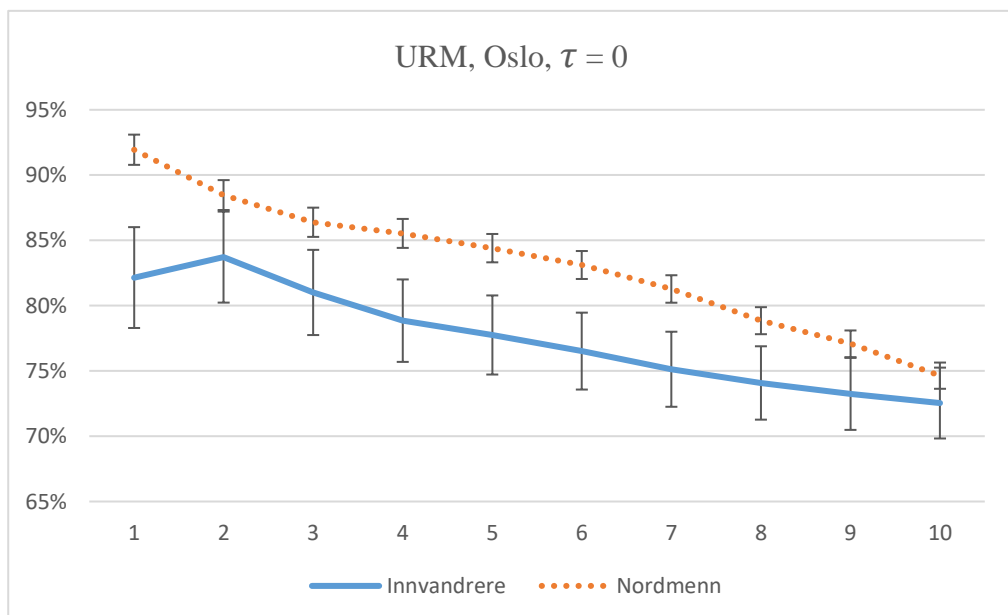
Figur 11: URM, Norge, $\tau = 0$



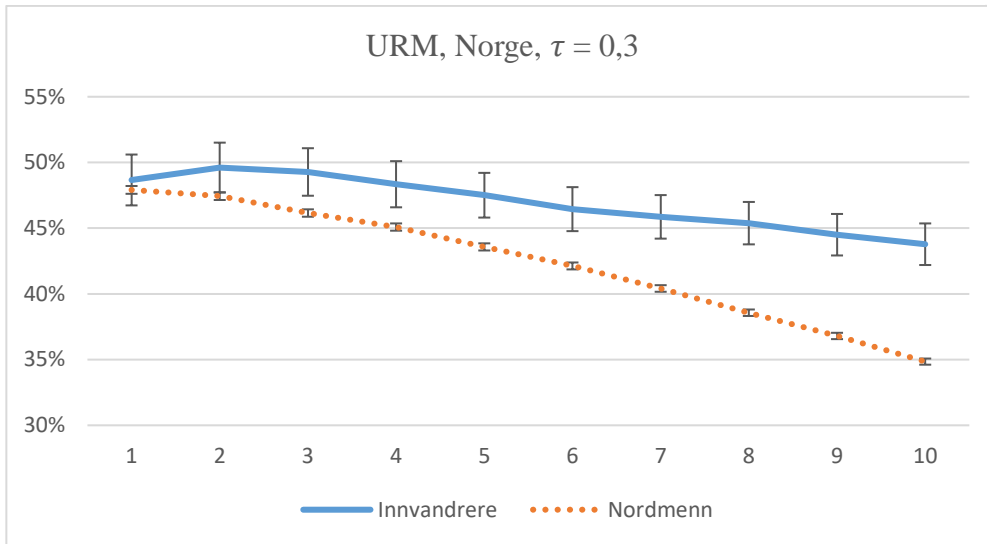
Figur 12: URM, Stor-Oslo, $\tau = 0$



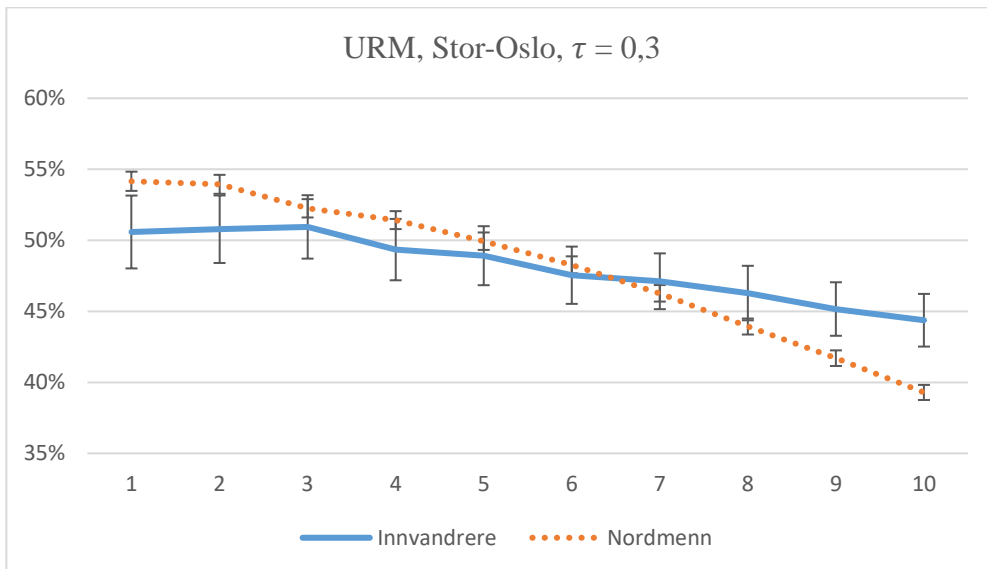
Figur 13: URM, Oslo, $\tau = 0$



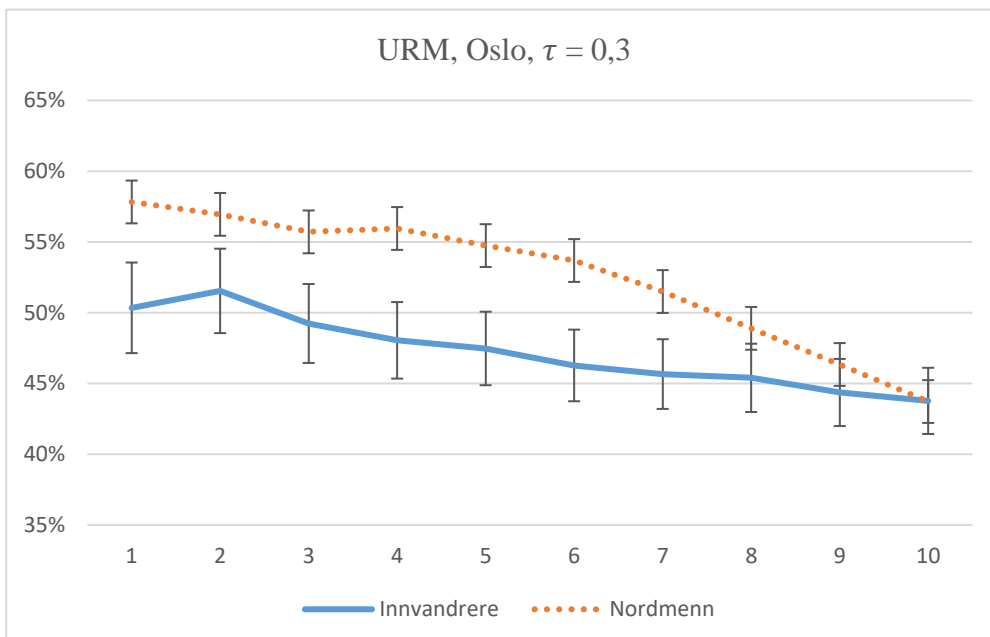
Figur 14: URM, Norge, $\tau = 0,3$



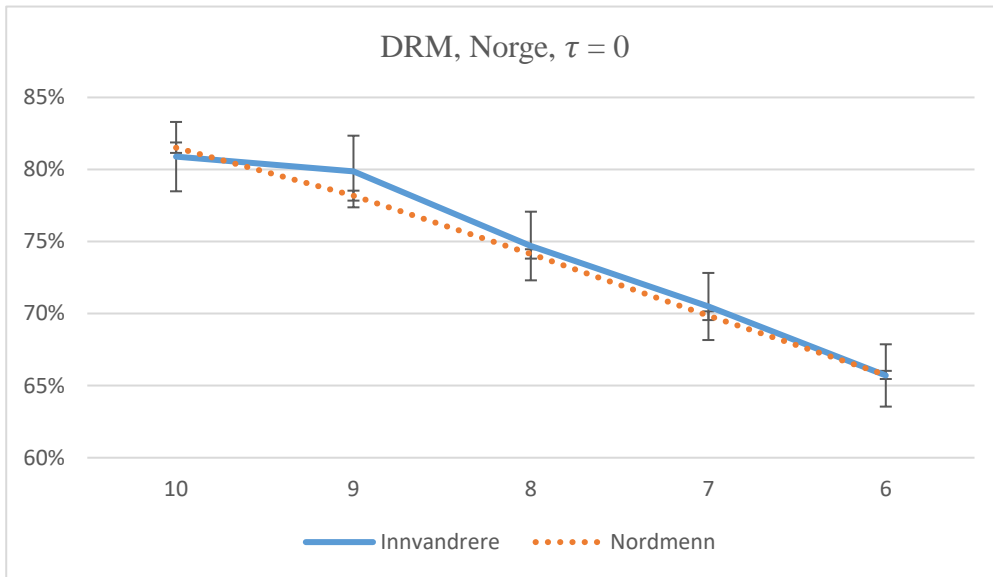
Figur 15: URM, Stor-Oslo, $\tau = 0,3$



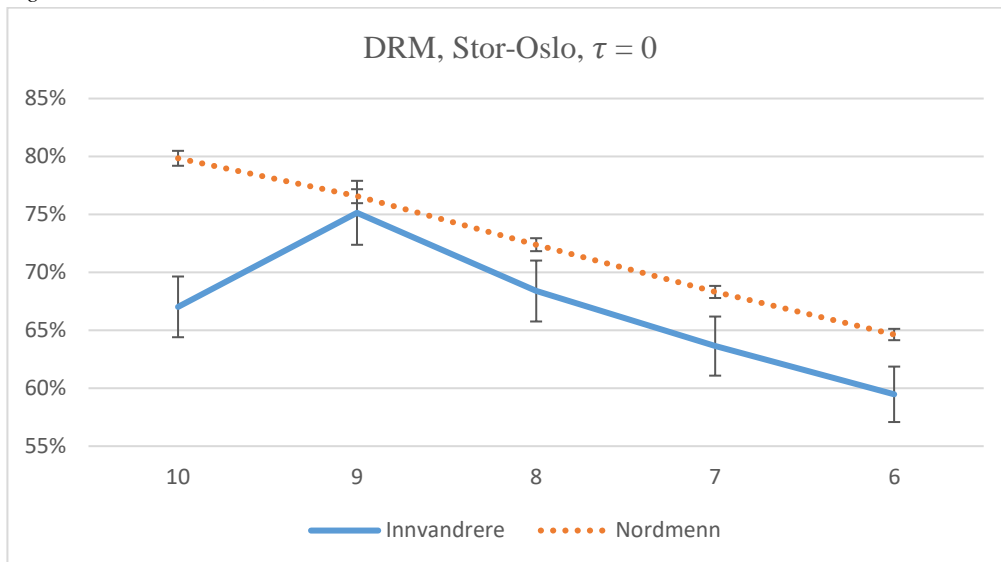
Figur 16: URM, Oslo, $\tau = 0,3$



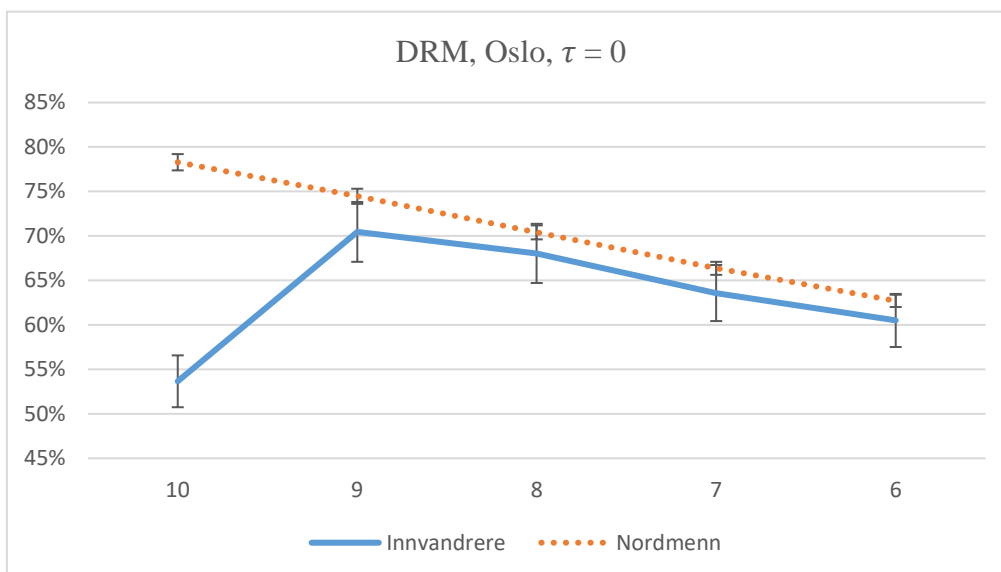
Figur 17: DRM, Norge, $\tau = 0$



Figur 18: DRM, Stor-Oslo, $\tau = 0$



Figur 19: DRM, Oslo, $\tau = 0$



Som vi kan se fra diagrammene ovenfor danner retningsrankmobilitetsmålene et lignende bilde som overgangssannsynlighetene. Når en studerer hele Norge med $\tau = 0$, er nordmennene i utvalget mer mobile oppover enn innvandrerne i de syv laveste vintilene. Forskjellen er kun statistisk signifikant fra første til og med femte vintil. Utover i diagrammet endrer dette seg, og vi kan se at oppovermobiliteten blir lik mellom de to gruppene i det åttende vintilet. I det niende og tiende vintilet er innvandrerne mer oppovermobile enn nordmennene. Denne forskjellen er kun statistisk signifikant i det siste inkluderte vintilet.

Resultatene fra URM i Stor-Oslo viser en økning i forskjellene mellom innvandrere og nordmennene. Vi ser at nordmennene er mer oppovermobile i de åtte første vintilene. Det kan merkes at forskjellene kun er statistisk signifikante i de seks første vintilene. I det niende vintilet er tilbøyeligheten for å bevege seg opp omtrent lik for de to gruppene. Forskjellene i oppovermobilitet øker enda mer når utvalget blir betinget på Oslo, og vi ser at oppovermobiliteten er betydelig høyere for nordmenn i de fleste vintilene. Vi kan se at nordmennene er statistisk signifikant mer tilbøyelig til å stige opp i første vintil, og fra tredje til og med åttende vintil.

Når τ justeres opp fra 0 til 0,3 ser vi en lignende tendens som ved overgangssannsynligheter. Når hele Norge studeres er innvandrerne mer oppovermobile enn nordmennene. Dette gjelder i alle vintil bortsett fra det første. Vi kan og merke oss at forskjellene er statistisk signifikant fra tredje vintil og utover. Dersom en studerer diagrammet nøye kan en legge merke til at forskjellene i oppovermobilitet øker utover i diagrammet. Dette kan ha med egenskapene til RM å gjøre. Individuer som har beveget seg forbi farens inntektsprosentil og den tilhørende grensen vil bli inkludert utover i diagrammet når RM er kumulativ. De norske fedrene er jevnere fordelt over inntektsfordeling enn innvandrerfedrene. Dette betyr at nordmennene som overgikk farens inntektsprosentil og grensen vil utgjøre en stadig mindre andel utover i diagrammet for nordmennene. De fleste innvandrerfedrene befinner seg i bunnen av inntektsfordelingen. Dermed vil ikke effekten av å bevege seg ut i diagrammet redusere andelen innvandrerne som har foretatt et byks i samme grad som hos nordmennene. Samtidig vil det være relativt vanskeligere å foreta byks lengre opp i inntektsfordelingen. Dette kan forklare hvorfor vi ser større forskjeller mellom nordmenn og innvandrere utover i diagrammet.

Når vi studerer Stor-Oslo med $\tau = 0,3$ er nordmennene og innvandrerne noe mer oppovermobile, men økningen i mobilitet er størst for nordmennene. Dette gir utslag i at nordmennene, i noen vintil, er mer tilbøyelig til å bevege seg oppover i inntektsfordelingen sammenlignet med innvandrerne. Dette gjelder fra og med første til og med femte vintil. Men forskjellen er kun signifikant i første vintil. I det sjette vintilet snur dette, og vi ser at nordmennene og innvandrerne er omtrent like tilbøyelig til å foreta byks. I det syvende vintilet er innvandrerne mer tilbøyelig til å foreta byks oppover i inntektsfordelingen. Denne forskjellen øker utover i diagrammet, og er signifikant i de to siste observerte vintilene.

Til slutt tolkes resultatene fra Oslo sett alene. Her kan en se at nordmennene er statistisk signifikant mer oppovermobile enn innvandrerne for alle vintil, bortsett fra de tre siste.

Ved estimering av retningsrankmobilitet dannes et bilde som ligner på resultatene fra overgangssannsynlighetene. Nemlig at innvandrerne virker mindre tilbøyelig eller like tilbøyelig som nordmennene til å foreta kortere bevegelser i inntektsfordelingen. Når grensen justeres høyere er innvandrerne, når hele Norge betraktes, mer tilbøyelige til å stige i inntektsfordelingen. Dette gjelder ikke i Stor-Oslo og Oslo, hvor nordmennene er mest tilbøyelige til å foreta byks.

DRM er fordelt i desil, altså en grovere inntektsfordeling. Dette blir gjort siden vintilfordeling skaper mye støy når det er så få innvandrerfedre i toppen av inntektsfordelingen.

Resultatene fra DRM i Norge tilsier at nordmennene og innvandrerne har omtrent lik sannsynlighet for å falle under inntektsprosentilet til faren sin. Dette gjelder for de fleste desil, bortsett fra det andre hvor innvandrerne har noe høyere tilbøyelighet for å falle nedover. Denne forskjellen er ikke statistisk signifikant.

Resultatene fra Stor-Oslo reflekterer at det er få innvandrerfedre i toppen av inntektsfordelingen, og ser altså ut til å bli preget av støy. Dette er naturlig å tenke siden nedovermobiliteten for innvandrerne er markant økende fra tiende til niende desil. Etter niende desil er støyen mindre tydelig, og vi ser at nedovermobiliteten faller gradvis utover i diagrammet. Vi kan også se ut i fra resultatene at innvandrere er mindre nedovermobile enn nordmenn i alle inntektsdesil, dette er statistisk signifikant i alle vintil utenom det niende.

Resultatene fra Oslo er noe forskjellig fra resultatene til Stor-Oslo. Vi kan se at innvandrerne er litt mindre nedovermobile enn nordmennene i alle desil. En kan merke seg at denne forskjellen kun er signifikant i første desil. Disse resultatene ser også ut til å være preget av støy, som en kan se på knekken i resultatene til innvandrerne.

5.5 Diskusjon

Som vi har sett fra de rangbaserte resultatene ovenfor er innvandrere mer tilbøyelig til å foreta lengre bevegelser oppover i inntektsfordelingen enn nordmenn. En mulig grunn til dette kan være at innvandrerfedrenes posisjon i inntektsfordelingen er mindre representativ for deres evner, sammenlignet med de norske fedrene i utvalget. Dette betyr at det finnes mange innvandrerfedre i bunnen av inntektsfordelingen som innehar evner som kunne plassert dem høyere opp i inntektsfordeling om de hadde vært født i Norge. Når disse fedrene får barn, vil barna arve fedrenes evner. Samtidig vil disse barna møte færre barrierer i samfunnet sammenlignet med fedrene sine. Dette kan gi utslag i høyere tilbøyelighet til å foreta lengre byks i inntektsfordelingen hos innvandrerbarna. Utdanning er en viktig mekanisme i IGM, og kanskje ekstra viktig for innvandrerne. Vi kan for

eksempel observere at en betraktelig større andel av innvandrerne i øverste inntektskvintil har mastergrad eller høyere, sammenlignet med nordmenn i samme kvintil. I Norge er utdanning svært tilgjengelig, og ofte mer tilgjengelig enn i opphavslandet til innvandrerfedrene. Dette gir utslag i at utdanningsnivået til innvandrerbarna er relativt høyere enn hos fedrenes deres. De med høyere utdanning, alt annet likt, vil være kvalifisert for flere høyt betalende jobber. Når disse innvandrerbarna kommer fra lavinntektsfamilier vil resultatet være et lengre byks oppover i inntektsfordelingen. Altså virker utdanning som en katalysator for lengre byks oppover i inntektsfordelingen for innvandrerbarn fra lavinntektsfamilier.

Fra resultatene er det også tydelig at innvandrerbarn er mindre tilbøyelig til å foreta kortere bevegelser oppover i inntektsfordelingen. Dette betyr at en større andel av innvandrerbarna forblir i lavere inntektsprosentiler sammenlignet med nordmenn. Fra overgangsmatrisene ser vi også at det er en tydelig tendens til at innvandrerbarna beveger seg til eller forblir i det laveste inntektskvintilet sammenlignet med nordmenn i tilsvarende kvintil. Dette tyder til at utfallet for innvandrerne er sprikende. Noen innvandrerbarn foretar lengre byks oppover i inntektsfordelingen, samtidig vil en større andel innvandrere bli værende igjen i bunnen av inntektsfordelingen.

Det kan være mange ulike faktorer som gjør at en høyere andel innvandrere forblir i bunnen av inntektsfordelingen. Tidligere artikler tyder på at innvandrerforeldrenes manglende språkkunnskaper og innsyn i samfunnet de har innvandret til kan være til hinder for barna deres (Bleakley & Chin, 2008; Dustmann & Glitz, 2011, s. 405). Dette kan gjøre at innvandrerbarn ikke får utnyttet mulighetene som ligger til rette for dem i Norge på grunn av foreldrenes manglende integrering. Norske artikler har også vist at de med synlig minoritetsbakgrunn kan oppleve diskriminering i arbeidslivet. Dette kan virke som en barriere mellom visse innvandrere og muligheten for å klatre i arbeidsmarkedet (Midtbøen, 2015).

IGE-resultatene for innvandrerne tilsier at innvandrerne er mer mobile enn nordmennene i utvalget. Disse resultatene stemmer overens med Hermansen (2016), som benytter seg av en lineær rankregresjon for å måle IGM. På en annen side tilsier resultatene fra RM, og de estimerte overgangssannsynlighetene med $\tau = 0$ at nordmennene er mer mobile oppover. Resultatene er altså noe motstridene. Som vi kan se fra den deskriptive statistikken presentert i kapittel 4.3, har det vært en kraftig intergenerasjonell assimilering hos innvandrerne i Norge. Innvandrerbarna er mye jevnere fordelt over inntektsfordelingen enn sine fedre, som er sentrert i bunnen²⁵. Denne assimileringssprosessen gir utslag i et lavt IGE-estimat, som altså tilsier høy grad av IGM. IGE bygger på linearitet, og kan bli ansett som et gjennomsnittlig mål på IGM. På grunn av dette greier ikke IGE å fange opp at en høyere andel innvandrerbarn forblir i bunnen av inntektsfordelingen, sammenlignet med norske barn. Dette kan gjøre at resultatene fra de ulike målemetodene fremstår som sprikende.

²⁵ Dette er i tråd med funnene til Bratberg et al. (2014), som finner lignende tendenser.

Samtidig kan det tenkes at innvandrerfars inntekt ikke er den mest avgjørende faktoren for å forklare innvandrerbarnas inntekt. Altså er det ikke nødvendigvis sånn at barna til innvandrerfedrene med høyest lønn er de som selv ender opp med høyest lønn. Dette vil også gi utslag i et lavt IGE-estimat. Dette underbygges av lav grad av korrelasjon mellom innvandrerfars og innvandrerbarnas inntekt, se [Tabell 13](#).

5.5.1 Forskjeller i geografiske inndelinger

Som vi kan se fra resultatene ovenfor er nordmenn bosatt i Oslo mer oppovermobile enn innvandrere i Oslo. Tilsvarende resultater finner vi ikke i samme grad når vi studerer hele Norge. Arbeidsmarkedet i Oslo kan trolig forklare disse regionale forskjellene i mobilitet, hovedsakelig mellom Oslo og hele Norge. En betraktelig høyere andel innvandrere bosatt i Oslo er også født der, sammenlignet med nordmenn bosatt i Oslo²⁶. En kan også merke seg at utdanningsnivået hos nordmennene er betraktelig høyere i Oslo, sammenlignet med resten av Norge. Disse faktorene kan tilsi at arbeidsmarkedet i Oslo tiltrekker seg individer fra hele Norge med høy utdanning, disse vil trolig være svært oppovermobile. I tillegg vil de med høy utdanning ofte ha høy inntekt. Dette kan være en av grunnene til at vi observerer såpass store forskjeller i inntekt, og oppovermobilitet blant innvandrere og nordmenn i Oslo.

Kapittel 6

Konklusjon

Oppgavens formål er å undersøke forskjeller i IGM hos innvandrere og nordmenn i Norge. Dette er viktig å undersøke siden andelen innvandrere i Norge har økt betraktelig de siste tiårene. Førstegenerasjonsinnvandrere har også lavere inntekt enn innfødte nordmenn. En sammenligning av IGM mellom innvandrere og nordmenn i Norge vil si noe om hvor persistent inntektsulikheten er mellom disse gruppene. Samtidig vil en kunne si noe om effektiviteten til integreringsprosessen i Norge. Oppgaven undersøker også forskjeller i ulike regionale inndelinger, mellom Norge, Stor-Oslo og Oslo. Dette blir gjort siden en høy andel av innvandrerne er bosatt i Oslo-området, dette er også området hvor inntektsulikheten er størst.

Oppgaven bygger på offentlig registerdata levert av et samarbeid mellom NSD og SSB. Dataen er meget rik, og samtidig av høy kvalitet. Oppgaven studerer barn født 1977 – 1982, og deres fedre. Hvor innvandrerutvalget består av 4 477 forskjellige far-barn koblinger, og nordmennene består av 212 003

²⁶ Dette gjelder spesielt en ser bort ifra førstegenerasjonsinnvandrere som immigrerte til Norge før en alder av syv år.

ulike far-barn koblinger. Ettersom innvandring i stor skala er et relativt nytt fenomen i Norge kan det oppstå problemer med målefeil. Disse problemene blir unngått på best mulig måte, slik at datagrunnlaget er mest mulig robust.

Målemetodene som blir benyttet i oppgaven er IGE, overgangsmatriser, overgangssannsynligheter og retningsrankmobilitet. Disse har ulike egenskaper, og komplementerer hverandre godt. Overgangsmatriser, overgangssannsynligheter og retningsrankmobilitet har ikke blitt brukt til å studere forskjeller i IGM hos innvandrere og nordmenn i Norge før. Dette gjør at oppgaven har noen nye og interessante funn.

I oppgaven blir det tydelig at det er forskjeller mellom innvandrerne og nordmennene i utvalget. Nordmennene har høyere inntekt, dette gjelder spesielt i Oslo. Hvis vi ser bort ifra andelen med fullført mastergrad hos barna, har også nordmennene høyere utdanningsnivå enn innvandrerne. Forskjellen er mest vesentlig hos fedrene i utvalget. Og vi kan se at det har vært en intergenerasjonell assimileringssprosess hos innvandrerne. Med dette menes det at innvandrerbarna har inntekt og utdanningsnivå som er mer sammenlignbar med nordmennene, enn hva innvandrerfedrene har.

Denne assimileringssprosessen blir reflektert i resultatene fra IGE. IGE er et gjennomsnittlig mål på mobilitet, ettersom den bygger på at alle er like mobile uavhengig av plassering i inntektsfordelingen. Dermed kan en tolke resultatene fra IGE-estimeringene som innvandrerne i gjennomsnitt er mer mobile enn nordmennene i utvalget. Altså er funnene i oppgaven i tråd med tidligere artikler med sammenlignbare målemetoder (Hermansen, 2016).

IGE danner ikke et fullstendig bilde av assimileringssprosessen. Ved hjelp av overgangssannsynligheter og retningsrankmobilitet blir det kartlagt at innvandrerne i utvalget er mindre tilbøyelig til å foreta korte bevegelser i inntektsfordelingen. En følge av dette er at en høyere andel innvandrere forblir i bunnen av inntektsfordelingen, sammenlignet med nordmenn. Dette kan være en gruppe som sliter med assimileringssprosessen. I tillegg kan dette reflektere at innvandrerbarn møter flere barrierer i utdanningssammenheng, og i arbeidsmarkedet enn hva nordmenn gjør.

Ved hjelp av overgangsmatriser, overgangssannsynligheter og retningsrankmobilitet blir det også klart at innvandrere er mer tilbøyelig til å foreta lengre byks oppover i inntektsfordelingen. Altså er det høyere sannsynlighet for at en innvandrer fra en lavinntektsfamilie ender opp med høy inntekt selv, sammenlignet med en tilsvarende nordmann. Disse resultatene tyder på at det er en todeling mellom innvandrerne. Den ene siden er de som sliter med assimileringssprosessen, og som en følge av dette ikke klatrer i inntektsfordelingen. På den andre siden er innvandrerne som klarer å benytte seg av mulighetene som ligger til rette for dem i den norske velferdsstaten. Disse foretar lengre byks oppover i inntektsfordelingen.

Oppgaven viser tydelige forskjeller mellom nordmenn og innvandrere i Oslo. Gjennomsnittsinntekten og utdanningsnivået er for nordmenn høyest når en kun studerer Oslo. Det motsatte gjelder for innvandrere i Oslo. Samtidig er nordmenn i Oslo mer tilbøyelige til å foreta kortere og lengre byks i inntektsfordelingen sammenlignet med innvandrere. Dette reflekterer at Oslo har et arbeidsmarked tilpasset for individer med høy utdanning. Disse individene vil ofte ha høyere grad av IGM, og samtidig ha høyere gjennomsnittsinntekt enn resten av befolkningen. Dette forklarer trolig hvorfor vi ser så store inntektsulikheter og grad av IGM mellom nordmenn og innvandrere i Oslo.

Litteraturliste

- Aakvik, A., Salvanes, K. G. & Vaage, K. (2010). Measuring Heterogeneity in the Returns to Education in Norway Using Educational Reforms. *European Economic Review*, 54, 483 – 500. <https://doi.org/10.1016/j.euroecorev.2009.09.001>
- Aydemir, A., Corak, M. & Chen, W. (2009). Intergenerational Earnings Mobility among the Children of Canadian Immigrants. *The Review of Economics and Statistics*, 91, 377-397. <https://doi.org/10.1162/rest.91.2.377>
- Baker, M. & Solon, G. (2003). Earnings dynamics and inequality among Canadian men, 1976-1992. *Journal of Labor Economics* 21, 289–321. <https://doi.org/10.1086/345559>
- Becker, G. S., & Tomes, N. (1979). An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, 87, 1153-189. <http://www.jstor.org/stable/1833328>
- Bhattacharya, D. & Mazumder, B. (2011). A nonparametric analysis of black–white differences in intergenerational income mobility in the United States. *Quantitative Economics*, 2, 335-379. [10.3982/QE69](https://doi.org/10.3982/QE69)
- Black, S. & Devereux, P. (2011). Recent Developments in Intergenerational Mobility. I: Ashenfelter, O. & Card, D. (Red.), *Handbook of labor economics* (s. 1487 – 1541). USA, Chicago: North Holland.
- Bleakley, H. & Chin, A. (2008). What holds back the second generation? The intergenerational transmission of language human capital among immigrants. *Journal of Human Resources*, 43, 267–298. [10.3368/jhr.43.2.267](https://doi.org/10.3368/jhr.43.2.267)
- Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D. D. & Vaage, K. (2017). A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US. *Scandinavian Journal of Economics*, 119, 72-101. [10.1111/sjoe.12197](https://doi.org/10.1111/sjoe.12197)
- Bratberg, E., Nilsen Ø. A. & Vaage, K. (2005). Intergenerational Earnings Mobility in Norway: Level and Trends. *The Scandinavian Journal of Economics*, 107, 419 – 435. <https://www.jstor.org/stable/3440966>
- Bratberg, E., Nilsen, Ø. A. & Vaage, K. (2007). Trends in Intergenerational Mobility across Offspring's Earnings Distribution in Norway. *Industrial Relations: A Journal of Economy and Society*, 46, 112-129. [10.1111/j.1468-232X.2007.00459.x](https://doi.org/10.1111/j.1468-232X.2007.00459.x)
- Bratsberg, B., Raaum, O. & Røed, K. (2010). When Minority Labor Migrants Meet the Welfare State. *Journal of Labor Economics*, 28, 633-676. <https://doi.org/10.1086/650546s>
- Bratsberg, B., Raaum, O. & Røed, K. (2012) Educating children of immigrants: Closing the gap in Norwegian schools. *Nordic Economic Policy Review*, 3(1), 211-251.
- Bratsberg, B., Raaum, O. & Røed, K. (2014), Immigrants, Labour Market Performance and Social Insurance. *Economic Journal*, 124, F644-F683. <https://doi.org/10.1111/econj.12182>
- Böhlmark, A. (2008). Age at immigration and school performance: A siblings analysis using Swedish register data. *Labour Economics*, 15, 1366-1387. <https://doi.org/10.1016/j.labeco.2007.12.004>.
- Böhlmark, A & Lindquist, M. J. (2006). Life-cycle variations in the association between current and lifetime income: replication and extension for Sweden. *Journal of Labor Economics*, 24, 879–896. [10.1086/506489](https://doi.org/10.1086/506489)
- Chetty, R., Hendren, N., Kline, P. & Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129, 1553–1623. <https://doi.org/10.1093/qje/qju022>
- Connolly, M., Corak, M. & Haeck, C. (2019). *Intergenerational Mobility between and within Canada and the United States*. Upublisert manuskript. National Bureau of Economic Research.
- Corak, M., Lindquist, M. J. & Mazumder, B. (2014). A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Economics*, 30, 185-200. <https://doi.org/10.1016/j.labeco.2014.03.013>
- Dustmann, C. & Glitz, A. (2011). Migration and education. I E. A. Hanushek, S. J. Machin & L. Woessmann (Red.), *Handbook of the Economics of Education* (s. 327-439). Amsterdam: Elsevier.

- Folketrygdloven. (1997). Lov om folketrygd (LOV-1997-02-28-19). Hentet fra: https://lovdata.no/dokument/NL/lov/1997-02-28-19/KAPITTEL_3-1-2#%C2%A73-15
- Haider, S. J. & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review*, 96, 1308–1320. <https://www.jstor.org/stable/30034342>
- Hermansen, A. S. (2016). Moving Up or Falling Behind? Intergenerational Socioeconomic Transmission among Children of Immigrants in Norway. *European Sociological Review*, 32, 675 - 689. <https://doi.org/10.1093/esr/jcw024>
- Hoën, M., Markussen, S. & Røed, K. (2018). *Immigration and Social Mobility*. Upublisert manuskript. IZA Institute of Labor Economics.
- Jäntti, M., Bratsberg, B., Røed, K., Raaum, O., Naylor, R. A., Osterbacka, E., Bjorklund, A. & Eriksson, T. (2006). *American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States*. Upublisert manuskript. IZA Institute of Labor Economics.
- Kao, G. & Tienda, M. (1995). Optimisim and achievement: the educational performance of immigrant youth. *Social Science Quarterly*, 76, 1 - 19. <https://www.jstor.org/stable/44072586>
- Kommunal- og Moderniseringsdepartementet (2003). *Storbymeldingen* (Meld. St. 31 (2002-2003)). Hentet fra: <https://www.regjeringen.no/no/dokumenter/stmeld-nr-31-2002-2003-/id402979/sec10>
- Markussen, S. & Røed, K. (2017). *Egalitarianism under Pressure*. Upublisert manuskript. IZA Institute of Labor Economics.
- Markussen, S. & Røed, K. (2018). *The Golden Middle Class Neighborhood: Trends in Residential Segregation and Consequences for Offspring Outcomes*. Upublisert manuskript. IZA Institute of Labor Economics.
- Mazumder, B. (2005). Fortunate sons: new estimates of intergenerational mobility in the US using social security earnings data. *Review of Economics and Statistics*, 87, 235–255. <https://doi.org/10.1162/0034653053970249>
- Mazumder, B. (2014). Black-White Differences in Intergenerational Economic Mobility in the U.S.. *Economic Perspectives*, 38(1).
- Mazumder, B. (2015). *Estimating the intergenerational elasticity and rank association in the US: Overcoming the current limitations of tax data*. Upublisert manuskript. Federal Reserve Bank of Chicago, Chicago.
- Midtbøen, A. H. (2015). The context of employment discrimination: interpreting the findings of a field experiment. *The British Journal of Sociology*, 66, 193 – 214. [10.1111/1468-4446.12098](https://doi.org/10.1111/1468-4446.12098)
- Mitnik, P. A., Bryant, V., Weber, M. & Grusky, D. B. (2015). *New Estimates of Intergenerational Mobility Using Administrative Data*. Upublisert manuskript. Stanford Center on Poverty and Inequality.
- Nilsen, F. H. (2014). Grunnskolen, SNL. Hentet fra: <https://snl.no/grunnskolen>
- Nilsen, Ø. A., Vaage, K., Aakvik, A. & Jacobsen, K. (2012). Intergenerational Earnings Mobility Revisited: Estimates Based on Lifetime Earnings. *The Scandinavian Journal of Economics*, 114, 1 - 23. [10.1111/j.1467-9442.2011.01672.x](https://doi.org/10.1111/j.1467-9442.2011.01672.x)
- Norberg-Schulz, M. & Østhus, S. (2019). *Geografiske forskjeller i inntektsmobilitet i Norge* (Rapport 31-2018). Hentet fra: <http://www.samfunnsokonomisk-analyse.no/nye-prosjekter/2019/4/3/geografiske-forskjeller-i-inntektsmobilitet-i-norge>
- Microdata. (2019a). Ofte stilte spørsmål. Hentet fra: <https://microdata.no/faq/>
- Microdata. (2019b). Brukermanual for microdata.no. Hentet fra: <https://microdata.no/brukermanual.pdf>
- O'Neill, D., Sweetman, O. & de Gaer, D. V. (2007). The effects of measurement error and omitted variables when using transition matrices to measure intergenerational mobility. *Journal of Economic Inequality* 5(2), 159 - 178.

- Pekkarinen, T., Salvanes, K. G. & Sarvimäki, M. (2017). The Evolution of Social Mobility: Norway during the Twentieth Century. *Scandinavian Journal of Economics*, 119, 5 - 33. [10.1111/sjoe.12205](https://doi.org/10.1111/sjoe.12205)
- Solon, G. (1992). Intergenerational income mobility in the United States. *American Economic Review* 82, 393–408. <https://www.jstor.org/stable/2117312>
- Solon, G. (1999). Intergenerational Mobility in the Labor Market. I K. J. Arrow & M. D. Intriligator (Red.), *Handbook of Labor Economics* (s. 1761 – 1796). Amsterdam: Elsevier.
- Solon, G. (2004). A model of intergenerational mobility variation over time and place. I M. Corak (Red.), *Generational income mobility in North America and Europe* (s. 38–47). Cambridge: Cambridge University Press.
- SSB. (2014). Store regionale forskjeller i husholdningsinntekt innenfor OECD. Hentet fra: <https://www.ssb.no/inntekt-og-forbruk/artikler-og-publikasjoner/store-regionale-forskjeller-i-husholdningsinntekt-innenfor-oecd>
- SSB. (2017). Økonomisk utsyn over året 2016. *Økonomiske analyser* 36(1), 86 – 99. Hentet fra: https://www.ssb.no/nasjonalregnskap-og-konjunkturer/oa/_attachment/299128?_ts=15c1a23d558#page=89
- SSB. (2018a). *Dette er Norge 2018*. Hentet fra: https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/359877?_ts=166ca623630
- SSB. (2018b). Immigrants and Norwegian-born to immigrant parents. Hentet fra: <https://www.ssb.no/en/innvbef>
- SSB. (2018c). Innvandrerbefolkning fra 1970 til og med 2018. Hentet fra: <https://www.ssb.no/innvandring-og-innvandrere/faktaside/innvandring>
- SSB. (2018d). Samlet inntekt. Hentet fra: <https://www.ssb.no/a/metadata/conceptvariable/vardok/570/nb>
- SSB. (2019). Fakta om Arbeid. Hentet fra: <https://www.ssb.no/arbeid-og-lonn/faktaside>
- Zimmerman, D., J. (1992). Regression toward mediocrity in economic stature. *American Economic Review* 82, 409–429. <https://www.jstor.org/stable/2117313>

Appendiks A

Tester

Tabell 11: Sensitivitetstest for hvor mange år fars inntektsmål baserer seg på

Aldersspenn	38-55	39-56	40-57	41-58	42-59	43-60	44-61	45-62
Antall år	1 år	2 år	3 år	4 år	5 år	6 år	7 år	8 år
Måleår	1993	1993-1994	1993-1995	1993-1996	1993-1997	1993-1998	1993-1999	1993-2000
IGE for Innvandrere	0,06	0,11	0,11	0,11	0,12	0,13	0,13	0,14
	(0,01)	(0,02)	(0,02)	(0,02)	(0,02)	(0,021)	(0,021)	(0,021)
N	4712	4655	4591	4538	4504	4477	4441	4404
IGE for Nordmenn	0,14	0,16	0,17	0,18	0,18	0,18	0,19	0,19
	(0,002)	(0,002)	(0,002)	(0,002)	(0,002)	(0,002)	(0,003)	(0,003)
N	21742	210164	209506	208759	208085	207256	206466	205712

Notat: Standardfeil rapporteres i parenteser. N er antall observasjoner.

I Tabell 11 kan en se hvordan IGE-estimatet endrer seg når en endrer antall år fars inntektsmål blir basert på. Andre dimensjoner blir også påvirket, som for eksempel hvilke år en observerer inntekten, aldersspennet til fedrene og antall observasjoner. Vi ser at IGE-estimatet til innvandrerfedrene og de norske fedrene er spesielt lavt når en kun baserer inntektsmålet over et år. Dette er mye tydeligere hos innvandrerne i utvalget, enn hos nordmennene. Når vi baserer inntektsmålet over flere år ser vi at IGE-estimatet øker gradvis. Dette er også mest tydelig for innvandrerne i utvalget.

Tabell 12: Sensitivitetstest for alder ved immigrasjon

Alder ved immigrasjon	IGE	N	Alder ved immigrasjon	IGE	N
Kun andre-generasjons-innvandrere	0,13*** (0,02)	2830	4 år	0,13*** (0,02)	3888
0 år	0,14*** (0,02)	3017	5 år	0,13*** (0,02)	4167
1 år	0,13*** (0,02)	3234	6 år	0,13*** (0,02)	4477
2 år	0,13*** (0,02)	3448	7 år	0,12*** (0,02)	4846
3 år	0,13*** (0,02)	3652	8 år	0,13*** (0,02)	5214

Notat: Standardfeil rapporteres i parenteser. N er antall observasjoner. * tilsvarer signifikansnivå på 0,1. ** tilsvarer signifikansnivå på 0,05. *** tilsvarer signifikansnivå på 0,01.

I Tabell 12 vises hvordan IGE-estimatet til innvandrerne blir påvirket av alder ved immigrasjon. En kan også se hvordan antall observasjoner blir påvirket ved inklusjon førstegenerasjonsinnvandrere som innvandret i ulike aldre. Vi kan se at estimatet forholder seg relativt stabilt for alle aldre ved immigrasjon.

Tabell 13: Sensitivitetstest for fars målealder

Kohorter		Innvandrere	Nordmenn
1938-1955	IGE	0,13*** (0,02)	0,18*** (0,003)
	Gjennomsnittsalder i 1994	46	45
	N	4477	207256
1941 - 1955	IGE	0,13*** (0,03)	0,18*** (0,003)
	Gjennomsnittsalder i 1994	44	43
	N	4199	200368
1944-1955	IGE	0,13*** (0,03)	0,19*** (0,003)
	Gjennomsnittsalder i 1994	43	43
	N	3712	186571
1947 - 1955	IGE	0,11*** (0,03)	0,19*** (0,003)
	Gjennomsnittsalder i 1994	42	42
	N	2985	156832
1950-1955	IGE	0,08*** (0,03)	0,19*** (0,003)
	Gjennomsnittsalder i 1994	41	41
	N	1919	109741

Notat: Standardfeil rapporteres i parenteser. N er antall observasjoner. * tilsvarende signifikansnivå på 0,1. ** tilsvarende signifikansnivå på 0,05. *** tilsvarende signifikansnivå på 0,01.

I Tabell 13 rapporteres sensitivitetstester for fars alder. Når 1938 – 1955-kohortene er inkludert er den estimerte IGE 0,13 hos innvandrerne, og 0,18 hos nordmennene. De eldste fedrene er da 55 – 60 år gamle i måleperioden. For 1941 – 1955-kohortene er de eldste fedrene 52 – 57 år gamle. Vi kan se at antall observasjoner synker når de eldste kohortene blir ekskludert. Dette skaper mye støy på IGE-estimatet til innvandrerne, og det blir vanskelig å se en klar effekt av livssyklus-skjevhet på innvandrerfedrenes inntektsmål som en følge av høy målealder.

F-test

Her undersøkes det om innvandrervariablene presentert i Tabell 9 har samlet signifikans.

Modell 1, uten bostedsvariabel:

$$y_i^C = \alpha + \partial * inn_i + \beta_1 * y_i^P + \beta_2 * inn_i * y_i^P + \varepsilon_i$$

$$H_0: \partial = \beta_2 = 0$$

H_a : H_0 stemmer ikke

$$F_1 = 37$$

Variablene har samlet signifikansnivå på 0,01. Forkaster H_0 .

Modell 2, med Oslo som bostedsvariabel:

$$y_i^C = \alpha + \partial_1 * inn_i + \partial_2 * bosted_i + \beta_1 * y_i^P + \beta_2 * inn_i * y_i^P + \beta_3 * bosted_i * y_i^P + \partial_3 * inn_i * bosted_i + \beta_4 * inn_i * y_i^P * bosted_i + \varepsilon_i \quad (9)$$

$$H_0: \partial_1 = \partial_3 = \beta_2 = \beta_4 = 0$$

H_a : H_0 stemmer ikke

$$F_2 = 48$$

Variablene har samlet signifikansnivå på 0,01. Forkaster H_0 .

Modell 3, med Stor-Oslo som bostedsvariabel:

$$y_i^C = \alpha + \partial_1 * inn_i + \partial_2 * bosted_i + \beta_1 * y_i^P + \beta_2 * inn_i * y_i^P + \beta_3 * bosted_i * y_i^P + \partial_3 * inn_i * bosted_i + \beta_4 * inn_i * y_i^P * bosted_i + \varepsilon_i \quad (9)$$

$$H_0: \partial_1 = \partial_3 = \beta_2 = \beta_4 = 0$$

H_a : H_0 stemmer ikke

$$F_3 = 36$$

Variablene har samlet signifikansnivå på 0,01. Forkaster H_0 .

Appendiks B

Diverse

Tabell 14: Korrelasjonskoeffisienten

	Innvandrere	Nordmenn
ρ	0,09	0,16
$SD(y_i^p)$	0,47	0,43
$SD(y_i^c)$	0,65	0,5

I Tabell 13 vises korrelasjonen mellom far og barns inntekt hos innvandrere og nordmenn i utvalget, sammen med standardavviket til inntekt for far og barn. Vi kan se at korrelasjonen er svært lav mellom inntekten til innvandrerfar og innvandrerbarn. Vi kan også se at standardavviket til inntekten til fedrene er lavere enn barna, og innvandrerbarna har det høyeste standardavviket i inntekt.

Tabell 15: Definisjon av Stor-Oslo

Kommuner:

Asker	Nes
Askim	Nesodden
Aurskog-Høland	Nittedal
Bærum	Oppegård
Drammen	Oslo
Eidsberg	Rygge
Eidsvoll	Rælingen
Enebakk	Rømskog
Fet	Røyken
Frogn	Råde
Gjerdrum	Sande
Hobøl	Skedsmo
Hof	Ski
Holmestrand	Skiptvet
Hurdal	Spydeberg
Hurum	Svelvik
Lier	Sørums
Lunner	Trøgstad
Lørenskog	Ullensaker
Marker	Vestby
Moss	Våler
Nannestad	Øvre Eiker
Nedre Eiker	Ås

I Tabell 14 vises hvilke kommuner som inngår i definisjonen av Stor-Oslo.

Appendiks C

Microdata

Microdata

Microdata er et nettbasert statistikkverktøy driftet og utviklet av SSB og NSD. Microdata er designet for å gi studenter og forskere rask tilgang til registerdata fra SSB.

Konfidensialitet

Dataen en får tilgang til inneholder i utgangspunktet sensitiv informasjon på individnivå. Får unngå at dataen blir misbrukt er dataen sterkt anonymisert ved hjelp av fire ulike tiltak. Disse omfatter:

- 1: det er ikke mulig å benytte seg av utvalg spesifisert på en slik måte at de består av mindre enn 1000 individer.
- 2: dataen er winsoriert til 2%. Dette betyr at de 1% laveste verdiene settes til 1-prosentilen, og de de 1% høyeste settes til 99-prosentilen. Dermed anonymiseres ekstremverdier på begge sider av fordelingen.
- 3: dataen er støybelagt med x , som er stokastisk og har $E(x) = 0$ (forventning lik null). Dette innebærer at det minste observerbare tallet i tabuleringer kan være 5, foruten 0. I tillegg støylegges verdier med maksimalt 5, og minimalt -5.
- 4: alle plotter er glattet ut med en glatteteknikk. Denne glatteteknikken kalles hexbin, og resultatet er plotter som er anonyme og mindre presise enn et standard plot (Microdata, 2018).

I tillegg er det heller ikke mulig å se direkte på dataen som i andre statistikkverktøy. Men som et substitutt til dette kan en lage tabuleringer og histogram, sammen med oppsummerende statistikk.

Brukergrensesnitt

Får å benytte seg av microdata må en logge seg inn på nettsiden deres. Dette gjøres ved for eksempel bankID eller MinID. Etter dette vil en ha tilgang til databasen, og statistikkverktøyet microdata. Ettersom det er et databaseverktøy byggeren opp egne arbeidsdatasett. Dette betyr at dersom en for eksempel laster inn inntektsvariabelen samlet inntekt målt i 01.01.2015, vil datasettet ditt bestå av alle individer som er registret til denne variabelen. Deretter kan en for eksempel laste inn en fødselsdatovariabel, og ut i fra denne konstruere en aldersvariabel. Ut ifra denne kan en snevre inn datasettet ytterligere ved å for eksempel forkaste alle individer yngre enn 30, og eldre enn 40. Hvis en ønsker å koble disse individene opp til for eksempel fedrene deres blir en nødt til å lage et nytt datasett. Deretter laste inn relevante variabler i riktig tidsrom, dette kan være utfordrende ettersom en

ikke kan se direkte på dataen. Når en har hentet inn relevante variabler kan en overføre disse variablene til datasettet som omfatter sønner ved hjelp av en familiekoblingsnøkkel.

En kan ikke benytte seg av egne statistikkpakker på microdata, dette kommer av kravene til konfidensialitet. Det er heller ikke mulig å koble egne variabler opp til variablene på microdata. Dette medfører at mulige funksjoner tilbudt på microdata er begrenset. Utfallet av dette er ofte mye nitidig manuelt arbeid. Dette gjør at mange arbeidsoppgaver tar lang tid å gjennomføre, samtidig som grad av nøyaktighet reduseres.

Ettersom microdata er nettbasert vil det være utsatt for sammenbrudd, dette forekommer med jevne mellomrom. Hvis dette forekommer i arbeidstiden løses eventuelle problemer raskt. Problemet oppstår når eventuelle sammenbrudd forekommer på fridager eller etter kontortid. Da blir en nødt til å vente til kontorene igjen er bemannet. En annen konsekvens av microdata er nettbasert er at variabler må lastes inn hver gang skriptet kjøres. Det er heller ikke mulig å kjøre en del av skriptet om gangen, derfor må alle variabler lastes inn på nytt dersom en foretar en liten endring i skriptet.

Variabler

Microdata har informasjon om en rekke forskjellige variabler fra ulike registre. Disse registrene omfatter folkeregisteret, nasjonal utdanningsdatabase, selvangivelsesregisteret, arbeidsmarkedsdata og FD-trygd. Alle variablene står listet på microdata sine hjemmesider. I disse listene står det hvor lenge variablene strekker, hvilken type det er og i hvilken måleenhet de er oppført i.

Det kan også merkes at det ikke er mulig å laste tidsvariabler inn i samme datasett som statiske variabler. Det en kan gjøre er å laste inn den relevante tidsvariabelen for hvert år, for deretter å aggregere den på egenhånd.

Microdata skript

Her presenteres et enkelt skript fra microdata som viser hvordan en gjennomfører en IGE-estimering betinget på at individene er innvandrere eller nordmenn²⁷.

```
create-dataset barn
```

```
import BEFOLKNING_FOEDSELS_AAR_MND as fodmnd
```

```
generate kohort = int( fodmnd / 100)
```

```
drop if kohort > 1982
```

²⁷ Vær oppmerksom på at det kan være problematisk å kopiere direkte fra Word og inn i microdata.

```

drop if kohort < 1977

import BEFOLKNING_REGSTAT 2006-01-01 as stat06
keep if stat06 == '1'
import BEFOLKNING_REGSTAT 2007-01-01 as stat07
keep if stat07 == '1'
import BEFOLKNING_REGSTAT 2008-01-01 as stat1
keep if stat1 == '1'
import BEFOLKNING_REGSTAT 2009-01-01 as stat2
keep if stat2 == '1'
import BEFOLKNING_REGSTAT 2010-01-01 as stat3
keep if stat3 == '1'
import BEFOLKNING_REGSTAT 2011-01-01 as stat4
keep if stat4 == '1'
import BEFOLKNING_REGSTAT 2012-01-01 as stat5
keep if stat5 == '1'
import BEFOLKNING_REGSTAT 2013-01-01 as stat6
drop if stat6 != '1' & kohort > 1978
import BEFOLKNING_REGSTAT 2014-01-01 as stat7
drop if stat7 != '1' & kohort > 1979
import BEFOLKNING_REGSTAT 2015-01-01 as stat8
drop if stat8 != '1' & kohort > 1980

import BEFOLKNING_INVKAT as innv
import BEFOLKNING_INNALDER as innalder
drop if innalder > 6
generate inn = 0
replace inn = 1 if innv == 'C'
replace inn = 1 if innv == 'B'
//drop if innv == 'B'
drop if innv == 'F'
drop if innv == 'G'

```

```
drop if innv == 'E'
```

```
generate innt30 = 0
```

```
generate innt31 = 0
```

```
generate innt32 = 0
```

```
generate innt33 = 0
```

```
import INNTEKT_WSAMINNT 2015-01-01 as i15
```

```
replace i15 = 0 if sysmiss(i15)
```

```
import INNTEKT_WSAMINNT 2014-01-01 as i14
```

```
replace i14 = 0 if sysmiss(i14)
```

```
generate i14ij = i14*1.02145
```

```
import INNTEKT_WSAMINNT 2013-01-01 as i13
```

```
replace i13 = 0 if sysmiss(i13)
```

```
generate i13ij = i13*1.04275
```

```
import INNTEKT_WSAMINNT 2012-01-01 as i12
```

```
replace i12 = 0 if sysmiss(i12)
```

```
generate i12ij = i12*1.06496
```

```
import INNTEKT_WSAMINNT 2011-01-01 as i11
```

```
replace i11 = 0 if sysmiss(i11)
```

```
generate i11ij = i11*1.07181
```

```
import INNTEKT_WSAMINNT 2010-01-01 as i10
```

```
replace i10 = 0 if sysmiss(i10)
```

```
generate i10ij = i10*1.08578
```

```
import INNTEKT_WSAMINNT 2009-01-01 as i09
```

```
replace i09 = 0 if sysmiss(i09)
```

```
generate i09ij = i09*1.11235
```

```
import INNTEKT_WSAMINNT 2008-01-01 as i08
```

```
replace i08 = 0 if sysmiss(i08)
```

```
generate i08ij = i08*1.13636
```

```
import INNTEKT_WSAMINNT 2007-01-01 as i07
```

replace i07 = 0 if sysmiss(i07)

generate i07ij = i07*1.179

replace innt30 = i12ij if kohort == 1982

replace innt30 = i11ij if kohort == 1981

replace innt30 = i10ij if kohort == 1980

replace innt30 = i09ij if kohort == 1979

replace innt30 = i08ij if kohort == 1978

replace innt30 = i07ij if kohort == 1977

replace innt31 = i13ij if kohort == 1982

replace innt31 = i12ij if kohort == 1981

replace innt31 = i11ij if kohort == 1980

replace innt31 = i10ij if kohort == 1979

replace innt31 = i09ij if kohort == 1978

replace innt31 = i08ij if kohort == 1977

replace innt32 = i14ij if kohort == 1982

replace innt32 = i13ij if kohort == 1981

replace innt32 = i12ij if kohort == 1980

replace innt32 = i11ij if kohort == 1979

replace innt32 = i10ij if kohort == 1978

replace innt32 = i09ij if kohort == 1977

replace innt33 = i15 if kohort == 1982

replace innt33 = i14ij if kohort == 1981

replace innt33 = i13ij if kohort == 1980

replace innt33 = i12ij if kohort == 1979

replace innt33 = i11ij if kohort == 1978

replace innt33 = i10ij if kohort == 1977

generate ai = (innt30 + innt31 + innt32 + innt33) / 4

```

import BEFOLKNING_FAR_FNR as fnr

create-dataset fedre

import BEFOLKNING_FOEDSELS_AAR_MND as fodmndf
generate kohortf = int( fodmndf / 100)
drop if kohortf > 1955
drop if kohortf < 1938

import BEFOLKNING_REGSTAT 1998-01-01 as statf8
keep if statf8 == '1'
import BEFOLKNING_REGSTAT 1997-01-01 as statf7
keep if statf7 == '1'
import BEFOLKNING_REGSTAT 1996-01-01 as stat1
keep if stat1 == '1'
import BEFOLKNING_REGSTAT 1995-01-01 as statf
keep if statf == '1'
import BEFOLKNING_REGSTAT 1994-01-01 as stat2
keep if stat2 == '1'
import BEFOLKNING_REGSTAT 1993-01-01 as stat3
keep if stat3 == '1'

import INNTEKT_WSAMINNT 2000-01-01 as if00
replace if00 = 0 if sysmiss(if00)
import INNTEKT_WSAMINNT 1999-01-01 as if99
replace if99 = 0 if sysmiss(if99)
import INNTEKT_WSAMINNT 1998-01-01 as if98
replace if98 = 0 if sysmiss(if98)
import INNTEKT_WSAMINNT 1997-01-01 as if97
replace if97 = 0 if sysmiss(if97)
import INNTEKT_WSAMINNT 1996-01-01 as if96

```

```

replace if96 = 0 if sysmiss(if96)
import INNTEKT_WSAMINNT 1995-01-01 as if95
replace if95 = 0 if sysmiss(if95)
import INNTEKT_WSAMINNT 1994-01-01 as if94
replace if94 = 0 if sysmiss(if94)
import INNTEKT_WSAMINNT 1993-01-01 as if93
replace if93 = 0 if sysmiss(if93)
generate aif = ((if97 * 1.43) + (if95 * 1.49 ) + (if94 * 1.52 ) + (if93 * 1.54 ) + (if96 * 1.466)+ (if98 *
1.40))/6

merge aif into barn on fnr

use barn

drop if sysmiss ( ai )
drop if sysmiss ( aif )

generate lnai = ln (ai)
generate lnaiif = ln (aif)

regress lnai lnaiif if inn == 1
regress lnai lnaiif if inn == 0

```