

Forskjeller i utdanningsnivå blant innvandrere i Norge

av

Ragnhild Brevig Aarseth

Masteroppgave

Masteroppgaven er levert for å fullføre graden

Master i samfunnsøkonomi

Universitetet i Bergen, Institutt for økonomi

Desember 2008

UNIVERSITETET I BERGEN



Forord

Veilder: Arild Aakvik

Jeg vil først å fremst gi en stor takk til min veileder Arild Aakvik, for god hjelp med oppgaven underveis.

Jeg vil også takke mine alltid hjelpsomme og hyggelige medstudenter på institutt for økonomi, dere har i stor grad bidratt til å gjøre oppholdet mitt her til en kjempeflott tid!

En spesiell takk går også til Rikard, Stine og Rebekka for korrekturlesning på oppgaven min.

De data som er benyttet i denne publikasjonen er hentet fra Statistisk sentralbyrås database FD – trygd. Norsk samfunnsvitenskapelig datatjeneste har tilrettelagt datauttaket for analyseformål. Ingen av de overnevnte institusjoner er ansvarlig for de analyser eller tolkninger som er gjort her.

Ragnhild Brevig Aarseth, Bergen 1. desember 2008

Sammendrag

Forskjeller i utdanningsnivå blant innvandrere i Norge

av

Ragnhild Brevig Aarseth, Master i samfunnsøkonomi

Universitetet i Bergen, 2006

Veileder: Arild Aakvik

Gjennom denne empiriske analysen gjør jeg rede for forskjeller i utdanningsnivå blant innvandrerbefolkningen i Norge, og sammenligner disse forskjellene med utdanningsnivå hos de uten innvandrerbakgrunn. Det blir også undersøkt hvilke faktorer som kan være med på å påvirke disse forskjellene, med spesielt vekt på om antall barn og sivilstand kan ha noe å si for hvor mye utdanning man velger å ta. Analysen bygger på to paneldatasett, et over innvandrerbefolkningen og ett over de uten innvandrerbakgrunn, i perioden 1993-2003, og for å gjennomføre den økonometriske analysen brukes programmet STATA. Hovedmodellen som er brukt i analysen er tilfeldig effekt modellen for paneldata, men denne blir supplert med vanlig minste kvadrats metode og en ordnet probit modell. Resultatene i oppgaven viser at det er store forskjeller i utdanningsnivå mellom de forskjellige innvandrergруппene i Norge. Mens vestlige innvandrere og innvandrere fra Øst-Europa faktisk har et høyere utdanningsnivå enn nordmenn, skiller spesielt innvandrere fra Vest-Asia seg ut som en gruppe med svært lavt utdanningsnivå. For denne regionene ser det dessuten også ut til at det å være gift, samt det å ha barn under 16 år, kan ha en negativ sammenheng med utdanningsnivå.

Innholdsfortegnelse

Forord.....	i
Sammendrag.....	ii
1 INTRODUKSJON.....	- 1 -
1.1 PROBLEMSTILLING.....	- 1 -
1.2 MOTIVASJON FOR OPPGAVEN.....	- 1 -
1.3 METODE OG OPPBYGNING AV OPPGAVEN.....	- 3 -
2 INSTITUSJONELL BAKGRUNN	- 5 -
2.1 INNVANDRERBEFOLKNINGEN I NORGE	- 5 -
2.1.1 <i>Hvor kommer de fra?</i>	- 6 -
2.1.2 <i>Hvorfor kommer de til Norge?</i>	- 7 -
2.2 UTDANNINGSPOLITIKK FOR INNVANDRERE.....	- 9 -
2.2.1 <i>Introduksjonsloven</i>	- 10 -
2.2.2 <i>Strategi for bedre læring blant etniske minoriteter</i>	- 10 -
2.3 NORSK STANDARD FOR UTDANNINGSGRUPPERING (NUS2000)	- 11 -
2.3.1 <i>Utdanningsnivå</i>	- 12 -
2.4 KARTLEGGING AV INNVANDRERES UTDANNINGSNIVÅ.....	- 14 -
3 LITTERATURGJENNOMGANG OG TEORI	- 16 -
3.1 MODELL FOR UTDANNING.....	- 16 -
3.2 IMMIGRASJON	- 18 -
3.2.1 <i>Kohort effekter</i>	- 19 -
3.2.2 <i>Teoretisk betraktning på hvorfor emigrasjon forekommer</i>	- 22 -
3.3 TIDLIGERE FORSKNING INTERNASJONALT	- 26 -
3.4 TIDLIGERE FORSKNING I NORGE	- 28 -
4 DATABESKRIVELSE.....	- 32 -
4.1 DATAMATERIALET	- 32 -
4.2 VARIABELSPESIFIKASJON	- 35 -
4.2.1 <i>Den avhengige variabel</i>	- 35 -
4.2.2 <i>Forklaringsvariabler</i>	- 43 -
5 EMPIRISK MODELLERING	- 50 -
5.1 MINSTE KVADRATS METODE (MKM)	- 50 -
5.2 PANELDATA	- 52 -
5.2.1 <i>Regresjonsmodeller for paneldata</i>	- 52 -
5.2.2 <i>Valg av modell</i>	- 54 -
5.2.3 <i>Tilfeldig effekt modellen (RE)</i>	- 54 -
5.3 BINÆRT VALG MODELLER.....	- 55 -

5.3.1	<i>Modell med ordnet responsvariabel</i>	- 56 -
5.4	ØKONOMETRISKE PROBLEMER.....	- 58 -
5.4.1	<i>Endogenitet</i>	- 58 -
5.4.2	<i>Autokorrelasjon</i>	- 58 -
5.4.3	<i>Heteroskedastisitet</i>	- 59 -
5.4.4	<i>Multikollinearitet</i>	- 60 -
5.4.5	<i>Ubalanserte panel og selekterte utvalg</i>	- 61 -
6	EMPIRISK ANALYSE	- 62 -
6.1	RESULTATER FRA ESTIMERING MED MINSTE KVADRATS METODE, TILFELDIG EFFEKT OG ORDNET PROBIT - 63 -	
6.1.1	<i>Tolkning av resultater, landbakgrunn</i>	- 63 -
6.2	RESULTATER FRA ESTIMERING MED REGIONSMODELLENE.....	- 68 -
6.2.1	<i>Tolkning av resultater</i>	- 70 -
6.3	DISKUSJON AV RESULTATENE	- 74 -
7	KONKLUSJON	- 78 -
	Referanser.....	- 96 -
	APPENDIKS A	- 82 -
	APPENDIKS B	- 85 -
	APPENDIKS C	- 89 -
	APPENDIKS D	- 92 -
	APPENDIKS E	- 94 -

Tabeller

Tabell 1: NUS2000s inndeling av utdanningsnivå (Statistisk sentralbyrå, 2000).....	- 13 -
Tabell 2: Gjennomsnittsalderen i de forskjellige regionene	- 44 -
Tabell 3: Gjennomsnittlig botid i Norge for de forskjellige regionene.	- 44 -
Tabell 4: Familietype for de forskjellige regionene, oppgitt i prosent.....	- 45 -
Tabell 5: Sivilstand for de forskjellige regionene, i prosent.	- 46 -
Tabell 6: Antall barn under 16 år for de forskjellige regionene, i prosent.	- 46 -
Tabell 7: Andel av befolkningen i hvert fylke, basert på regioner.	- 47 -
Tabell 8: Andel av befolkningen i tettbygde/ spredtbygde strøk, basert på regioner.	- 48 -
Tabell 9: Andel av etterkommere i utvalget som kommer fra de forskjellige regionene.	- 49 -
Tabell 10: Andel av innvandrerbefolkningen med norsk statsborgerskap.	- 49 -
Tabell 11: Resultater fra estimering med MKM, Tilfeldig effekt og Oprobit, menn og kvinner hver for seg, med utdanningsnivå som avhengig variabel.....	- 64 -
Tabell 12: Resultater fra estimering med tilfeldig effekt regionsmodeller, utdanningsnivå som avhengig variabel- 69 -	
Tabell B. 1: Oppføring av antall personer i hvert fylke med uoppgitt utdanningsnivå.....	- 88 -
Tabell D. 1: Resultater fra regresjon med probitmodell, positiv utdanning som avhengig variabel (i forhold til ingen utdanning og uoppgitt utdanningsnivå). Resultatene er oppgitt i marginaleffekter.	- 92 -
Tabell E. 1: Resultater fra regresjon med probit modell, uoppgitt utdanningsnivå som avhengig variabel i forhold til hvert av utdanningsnivåene.....	- 95 -

Figurer

<i>Figur 1: Kohort effekter og innvandreres lønnsprofil basert på alder</i>	- 21 -
<i>Figur 2: Andel innvandrere fra hver region basert på landbakgrunn</i>	- 34 -
<i>Figur 3: Andel innvandrere i Norge basert på statsborgerskap</i>	- 35 -
<i>Figur 4: Forskjeller i utdanningsnivå blant innvandrere og de uten innvandrerbakgrunn i Norge</i>	- 36 -
<i>Figur 5: Forskjeller i utdanningsnivå blant vestlige innvandrere, ikke-vestlige innvandrere og de uten innvandrerbakgrunn i Norge</i>	- 37 -
<i>Figur 6: Befolkningen med ingen utdanning, barneskole, ungdomskole eller med uoppgitt utdanningsnivå etter landbakgrunn</i>	- 38 -
<i>Figur 7: Befolkningen med utdanning på universitets- og høyskolenivå eller forskerutdanning</i>	- 39 -
<i>Figur 8: Befolkningen uten utdanning eller med grunnskole som høyeste fullførte utdanning, etter landbakgrunn og kjønn</i>	- 39 -
<i>Figur 9: Befolkningen oppført med høyere utdanning, etter landbakgrunn og kjønn</i>	- 40 -
<i>Figur 10: Innvandrernes utdanningsnivå basert på botid</i>	- 41 -
<i>Figur 11: Innvandrerbefolkningens utdanningsnivå basert på aldersgrupper</i>	- 41 -
<i>Figur 12: Andel barn på hvert utdanningsnivå hos innvandrerbefolkningen</i>	- 42 -
<i>Figur 13: Andel gifte på hvert utdanningsnivå hos innvandrerbefolkningen, etter kjønn</i>	- 42 -
<i>Figur A. 1: Utdanningsnivå blant innvandrere basert på periode</i>	- 82 -
<i>Figur A. 2: Utdanningsnivå hos vestlige innvandrere basert på perioder</i>	- 82 -
<i>Figur A. 3: Utdanningsnivå hos ikke-vestlige innvandrere basert på periode</i>	- 83 -
<i>Figur A. 4: Andelen med uoppgitt utdanningsnivå basert på verdensregion og periode</i>	- 83 -
<i>Figur A. 5: Andelen med ingen utdanning eller kun grunnskoleutdanning basert på landbakgrunn og periode</i> -	84
-	
<i>Figur A.6: Andelen med høyere utdanning basert på verdensregioner og periode</i>	- 84 -
<i>Figur B.1: Forskjeller i utdanningsnivå i tettbygde strøk basert på årsintervaller</i>	- 85 -
<i>Figur B.2: Forskjeller i utdanningsnivå i spredtbygde strøk basert på årsintervaller</i>	- 85 -
<i>Figur B.3: Innvandrerbefolkningen med ingen utdanning, grunnskoleutdanning eller uoppgitt utdanningsnivå basert på fylker</i>	- 86 -
<i>Figur B. 4: Innvandrerbefolkningen med høyere utdanning basert på fylker</i>	- 87 -
<i>Figur C. 1: Gjennomsnittlig inntekt basert på utdanningsnivå</i>	- 89 -
<i>Figur C. 2: Gjennomsnittlig inntekt basert på utdanningsnivå, skille mellom vestlige- og ikke-vestlige innvandrere</i>	- 90 -
<i>Figur C. 3: Gjennomsnittlig inntekt basert på utdanningsnivå og kjønn</i>	- 90 -
<i>Figur C. 4: Gjennomsnittlig inntekt basert på utdanningsnivå og landbakgrunn</i>	- 91 -

1 Introduksjon

1.1 Problemstilling

Jeg skal i denne oppgaven gjøre rede for forskjeller i utdanningsnivå for ulike innvandrergupper i Norge, og sammenligne disse med utdanningsnivået til de uten innvandrerbakgrunn. Ved hjelp av en empirisk analyse vil jeg se på hvilke faktorer som påvirker disse forskjellene, med spesielt vekt på om sivilstand og antall barn man får, har innvirkning på hvor mye utdanning individene velger å ta.

1.2 Motivasjon for oppgaven

Det har de siste årene vært mye fokus på forskjeller i utdanningsnivå blant innvandrere i Norge. Innvandrerbefolkningen i Norge defineres ofte som de som er født med to utenlandske foreldre. Herunder kommer både de som selv har innvandret til Norge, som ofte kalles førstegenerasjonsinnvandrere, og de som er født i Norge av to utenlandske foreldre, altså såkalte andregenerasjonsinnvandrere eller etterkommere. I følge statistisk sentralbyrå er mange innvandrere mer utdannet enn resten av den norske befolkningen. Både innvandrergunn og hvilke land innvandrerne kommer fra er dokumentert som viktige årsaker til disse forskjellene, samtidig som etterkommere ofte har større deltakelse i utdanning enn foreldrene (Lie, 2003). Det er imidlertid lite av denne forskningen som har fokus på hvordan underliggende faktorer, som for eksempel hvordan det være gift og antall barn, har påvirkning på utdanningsnivået til de forskjellige innvandrerguppene.

Innvandringen til Norge har økt kraftig de siste tiårene (Statistisk sentralbyrå, 2004). I tillegg til denne økningen, har innvandringen til Norge også endret mønster. Før 1960-tallet besto innvandringen til Norge nesten utelukkende av personer med vestlig landbakgrunn, og i hovedsak også personer fra land med nær geografisk beliggenhet til Norge, som for eksempel Sverige, Danmark, Storbritannia og Tyskland. Fra 1960-tallet og utover har vi imidlertid hatt ulike bølger av innvandring til Norge. På begynnelsen av 1960-årene begynte innvandringen av arbeidsinnvandrere fra ikke-vestlige land, som Tyrkia og Pakistan, før Norge innførte innvandringsstopp i 1975. I årene som fulgte kom en bølge av familieinnvandring til disse, før flyktninger fra land som Vietnam og Chile begynte å immigrere til Norge på slutten av 70-

tallet (Djuve og Friberg, 2004). Både endringen i den norske innvandringspolitikken i 1975, og endringen i grupper innvandrere som har kommet til Norge, kan føre til at vi får såkalte kohort effekter¹ når det gjelder utdanning og kvalifikasjonsnivå. Dette vil si at innvandrere som kommer i forskjellige ”innvandringsbølger” kan ha vesentlige forskjeller i for eksempel utdanning, og dermed at tidspunktet innvandrere kom til Norge på også kan være med på å forklare forskjeller i utdanningsnivå.

Da mange studier, både i Norge og i utlandet, har konkludert med at antall år med utdanning har en svært positiv korrelasjon med inntekt senere i livet (Raaum, 1999), går mye av den tidligere forskningen på området ut på sammenhengen mellom utdanning og inntekt, og hvordan for eksempel landbakgrunn har innvirkning på forskjeller i inntekt for personer med lik utdanning. Borjas (1991) finner at forskjeller i hvilke verdensregioner innvandrerne kommer fra kan forklare store deler av forskjellene i utdanningsnivå mellom innvandrere i Canada og innvandrere i USA. Avkastningen av utdanning i landet man immigrerer til vil også være forskjellig for innvandrere med forskjellig landbakgrunn. I følge Friedberg (2000) som ser på arbeidsmarkedssituasjonen i Israel, har også innvandrere forskjellig avkastning på utdanning de har tatt i hjemlandet, avhengig av hvilken landbakgrunn de har.

Av tidligere forskning på emnet i Norge har Barth, Bratsberg og Raaum (2004) sammenlignet inntektsprofiler for innvandrere relativt til innfødte i Norge og i USA, og funnet at innvandrere i USA tjener mer, har en brattere lønnsvekst og har en raskere inntektsopphenting med de innfødte, enn innvandrere i Norge. Denne fordelene for innvandrere i USA er større, jo høyere utdannelse innvandrerne har. I følge Roys seleksjonsmodell (Borjas, 1987), vil innvandrere immigrere dersom de økonomiske incentivene til dette er store nok. Innvandrere med høy utdanning vil i så fall, i følge denne teorien, heller selektere seg til USA enn til Norge.

Noe som kan være problematisk når det gjelder å måle innvandrernes utdanningsnivå, er at det ikke er registrert opplysninger om utdanningsnivå for mange innvandrere. Dette gjelder spesielt for de som er nyankomne i Norge, og skyldes at det ikke utføres noen systematisk registrering av fullført utdanning ved innvandring. Alle som ikke har tatt utdanning etter

¹ Begrepet er hentet fra Borjas (1984), der han bruker det til å forklare forskjeller i inntekt/ kvalifikasjonsnivå for innvandring til USA. Gjennomgang av dette i kapittel 3.

ankomst til Norge blir da stående med ”uoppgitt” utdanningsnivå. I forbindelse med folketellingene i 1990 og 2000 ble det imidlertid utført en spørreskjemaundersøkelse for å forbedre statistikken (Dalheim, 2002). Disse har hjulpet til å gjøre andelen uten oppgitt utdanningsnivå betydelig mindre, men problemet med manglende utdanningsnivå for innvandrere eksisterer fortsatt i stor grad.

1.3 Metode og oppbygning av oppgaven

Datamaterialet som ligger til grunn for denne oppgaven er to paneldatasett som inneholder et utvalg på henholdsvis hele innvandrerbefolkningen i Norge og et tilfeldig utvalg av 20 prosent av befolkningen uten innvandrerbakgrunn. Begge er samlet inn over perioden 1993 til 2003². Vi følger altså individene over denne perioden, uten at det kommer noen nye individer inn i utvalget. Jeg har valgt å sette de to datasettene sammen, for å på best mulig måte kunne se på forskjellene i utdanningsnivå mellom de uten innvandrerbakgrunn og de med landbakgrunn fra de forskjellige regionene.

Siden innvandrergruppen i Norge både kommer fra mange forskjellige land, og har kommet til Norge av ulike grunner, kan ikke innvandrerbefolkningen i Norge sies å være noen homogen gruppe. Jeg vil på bakgrunn av dette også kjøre regresjoner for de ulike verdensregionene i utvalget hver for seg. Dette vil dermed gi meg anledning til å se på hvordan forklaringsvariabler som antall barn og sivilstand påvirker utdanningsnivået forskjellig innenfor de ulike gruppene. Hovedsakelig er resultatene funnet ved bruk av tilfeldig effektmodellen for paneldatasett, men standard minste kvadrats metode og en ordnet probitmodell benyttes for å supplere denne.

I kapittel 2 gjør jeg rede for innvandrerbefolkningen i Norge, og hvordan denne er bygd opp. Jeg vil i dette kapittelet også se på hvilken utdanningspolitikk vi har for innvandrere i Norge og gi en kort gjennomgang av den norske standarden for utdanningsgruppering, som jeg skal bruke i oppgaven. I kapittel 3 følger en litteraturoversikt med teorier om hva som gjør at man velger å ta høyere utdanning, samt generell teori om innvandring og utdanningsnivå.

² Datamaterialet i denne oppgaven er hentet fra databasen FD-trygd, Statistisk Sentralbyrå (SSB) og tilrettelagt av Norsk Samfunnsvitenskapelig Datatjeneste (NSD).

Jeg går i kapittel 4 gjennom datamaterialet som ligger til grunn for min analyse. Jeg ser her på variablene jeg skal bruke i analysen, og går samtidig gjennom deskriptiv statistikk for disse. Kapittel 5 er en gjennomgang av de empiriske modellene og estimatorene jeg bruker i oppgaven, og valg av modell. Deretter ser jeg kort på økonometriske problemer som kan oppstå i en slik analyse, som endogenitet, autokorrelasjon og selekterte utvalg. Den empiriske analysen blir gjennomgått i kapittel 6. Jeg begynner med en analyse som inkluderer alle regionene, samt de uten innvandrerbakgrunn, for å illustrere forskjeller mellom disse. Jeg går så videre til å se på de forskjellige regionene hver for seg, for så å gjennomgå en tolkning av resultater. Til slutt i kapittelet følger en kritisk vurdering av analysen, der jeg gjennomgår hvordan økonometriske problemer som simultanitet kan ha hatt en påvirkning på resultatene mine. Hovedresultatene vil bli oppsummert i kapittel 6.

Hovedresultatet viser at de med landbakgrunn i Nord-Amerika utmerker seg positivt ved å ha det desidert høyeste utdanningsnivået, både i forhold til de uten innvandrerbakgrunn og i forhold til de andre regionene. I store deler av innvandrerbefolkningen er utdanningsnivået generelt høyt i forhold til de uten innvandrerbakgrunn i Norge. Unntaket er imidlertid de med landbakgrunn fra de ikke-vestlige regionene Vest-Asia, Afrika og Øst-Asia. Gruppen med landbakgrunn fra Vest-Asia skiller seg ut som de med det laveste utdanningsnivået av alle regionene. Det er også store forskjeller når det gjelder kvinner og menn innenfor denne regionen, og vi ser at kvinner har et klart lavere utdanningsnivå enn menn. Det å ha barn under 16 år og det å være gift ser dessuten ut til å ha en negativ sammenheng med utdanningsnivået til kvinner fra Vest-Asia, selv om disse gir signifikante positive verdier, i absolutt forstand, for kvinner fra andre regioner.

2 Institusjonell bakgrunn

I dette kapitlet går jeg gjennom relevant bakgrunnsinformasjon. Jeg begynner med å se på sammensetningen i innvandrerbefolkningen i Norge, og går i kapittel 2.2 over til å se på utdanningspolitikk for innvandrere. I kapittel 2.3 skal jeg så beskrive den norske standarden for utdanningsgruppering, og hvilke inndelinger denne gjør for utdanningsnivå.

2.1 Innvandrerbefolkningen i Norge

Innvandrerbefolkningen i Norge defineres av SSB som personer med to utenlandskfødte foreldre, og består dermed av førstegenerasjons innvandrere og barn født i Norge av to utenlandskfødte foreldre. Førstegenerasjonsinnvandrere er, slik SSB definerer det, personer født i utlandet av to utenlandskfødte foreldre. Disse har altså på et tidspunkt innvandret til Norge. Barn født i Norge av to utenlandskfødte foreldre blir nå ofte kalt andregenerasjons innvandrere eller etterkommere (Daugstad, 2006). Jeg vil i denne oppgaven referere til de som etterkommere.

Mens noen innvandrere har kommet til Norge av arbeidsgrunner eller blitt familieforent med personer som har det, har andre kommet hit av fluktgrunner. Disse inngår i statistikken av innvandrerbefolkningen, og er førstegenerasjons innvandrere. Det viser seg imidlertid at mange i denne gruppen har betydelig dårligere levekår enn resten av befolkningen, især hvis de har bodd kort tid i landet. Det er også vanlig å skille mellom ikke-vestlige og vestlige innvandrere. Ikke-vestlige innvandrere har bakgrunn fra land i Asia med Tyrkia, Afrika, Sør- og Mellom-Amerika og Øst-Europa. Vestlige innvandrere kommer fra resten av verden, det vil si Vest-Europa, Nord-Amerika og Oseania.

Innvandrerbefolkningen i Norge hadde ved inngangen til 2006 bakgrunn fra 208 land og selvstyrte regioner (Østby, 2004). Det var på dette tidspunktet 387 000 personer som inngikk i innvandrerbefolkningen, hvorav 319 000 var førstegenerasjonsinnvandrere og 68 000 var etterkommere. Til sammen utgjorde denne innvandrerbefolkningen 8,3 prosent av hele befolkningen, mens førstegenerasjons innvandrere alene utgjorde 6,9 prosent. To av tre førstegenerasjonsinnvandrere kommer fra ikke-vestlige land (Daugstad, 2006).

Vi finner flest innvandrere i sentrale strøk og særlig i området rundt Østlandet. I Oslo er nesten 22 prosent av befolkningen innvandrere (Djuve og Friberg, 2004). Mange av innvandrerne er også norske statsborgere. I årene fra 1977 til 2003 fikk i alt 155.000 innvandrere innviglet norsk statsborgerskap. Av disse var 130.000 innvandrere med tidligere ikke-vestlig statsborgerskap (ibid).

2.1.1 Hvor kommer de fra?³

I 1993 var innvandrerbefolkningen totalt på 204 810 personer, og vesentlig flere enn i dag innvandret fra vestlige land. Vel halvparten av alle innvandrerne på denne tiden kom fra Europa. Her kom spesielt mange fra våre naboland Danmark og Sverige (over 20.000 fra Danmark og nesten 19.000 fra Sverige). Innvandringen fra Storbritannia og Tyskland var også betydelig. Innvandringen fra Vesteuropske land toppe statistikkene her. Fra Øst-Europa var innvandringen fra Jugoslavia den største. Det kom imidlertid også en del innvandrere til Norge fra Tyrkia og Polen på denne tiden.

Innvandringen til Norge fra Afrika var betydelig mindre enn den fra Europa. I 1993 var det kun 14.127 personer i Norge med fødeland i Afrika. De landene som skiller seg ut i Afrika med godt over 2.000 innvandrere til Norge hver, er Marokko, Somalia og Etiopia. Asia hadde imidlertid rimelig høy innvandring til Norge i årene før 1993. I 1993 var det totalt 59.534 med fødeland i Asia som bodde i Norge. Her skiller spesielt Pakistan, med 11.410 innvandrere, og Vietnam, med 9.957 innvandrere, seg ut. Også fra Iran og Sri Lanka kom det i årene før 1993 mange innvandrere til Norge.

I 1993 var det 17.915 personer i Norge som hadde fødeland i Nord-Amerika eller Mellom-Amerika totalt. Av disse kom flesteparten, omkring 15.000 fra USA, 1.887 fra Canada og kun 1.147 fra Mellom-Amerika. Fra Sør-Amerika kom det til sammen nesten ti tusen innvandrere til Norge før 1993, mens det fra Oseania kun kom 1.064.

³ Alle tallene under er hentet fra Statistisk sentralbyrå (2008b) Statistikkbanken.

<http://statbank.ssb.no/statistikkbanken/>.

I perioden fra 1993 til 2003 har innvandrerbefolkningen økt med 72 prosent. I denne perioden har alle innvandrergrupper vokst i antall, bortsett fra innvandrere fra Nord-Amerika og Oseania. Raskest var denne veksten for personer fra Øst-Europa, hvor antallet nesten er tredoblet siden 1993. I samme periode ble antall innvandrere med landbakgrunn fra Asia, Afrika, Sør- og Mellom-Amerika nesten fordoblet (Statistisk sentralbyrå, 2004). Ved utgangen av 2006 var de største gruppene i innvandrerbefolkningen personer med landbakgrunn fra Pakistan, Sverige, Irak og Danmark.

2.1.2 Hvorfor kommer de til Norge?

Innvandrere som kommer til Norge kan komme hit både som flyktninger fra land der det er krig eller uroligheter, som arbeidsinnvandrere, for å ta utdanning i Norge og som familieinnvandrere. Familieinnvandringen kan deles inn i de som kommer samtidig (medfølging) eller gjenforenes, og de som kommer til Norge for å etablere familie (Statistisk sentralbyrå, 2007).

Inn og utvandring har alltid eksistert i det norske samfunnet. Den såkalte nye innvandringen, av mennesker fra ikke-vestlige land gjorde seg imidlertid først gjeldene på slutten av 1960-tallet. Denne nye innvandringen kom da i tre bølger (Djuve og Friberg, 2004). Den første bølgen besto av arbeidsinnvandrere fra ikke-vestlige land, som ville prøve seg på det norske arbeidsmarkedet. I begynnelsen besto denne gruppen av tyrkere og marokkanere, men på begynnelsen av 70-tallet var pakistanere overrepresentert her. De fleste av disse fikk ufaglærte jobber i industri- og servicesektoren. I 1975 satte imidlertid Norge innvandringsstopp, og innvandringen av ufaglærte arbeidere ble dermed stanset (ibid).

Den andre bølgen kom i form av familieinnvandring. Denne gruppen besto av familiene til de arbeidsinnvandrerne som hadde etablert seg i Norge, og som nå søkte om familiegjenforening med sine nærmeste i hjemlandet. Selv om dette ikke endret sammensetningen av ulike nasjonaliteter i innvandrerbefolkningen, gjorde det naturlig nok at innvandrerfolkningen skiftet karakter. Istedenfor enslige menn besto denne gruppen nå av familier, med et helt annet behov når det gjaldt integrering i det norske samfunnet (ibid).

Fra slutten av 1970-tallet begynte flyktninger og asylsøkere å komme til Norge, noe som kan karakteriseres som den tredje innvandringsbølgen. De første var vietnamesiske båtflyktninger som kom i etterdønningene av Vietnam-krigen. Etter kuppet i Chile i 1973 kom også flyktninger som da var politisk forfulgte i hjemlandet (ibid). Stadig flere nasjonaliteter kom så med i gruppen av innvandrere i Norge. Utover 1980-tallet kom iranere på flukt etter revolusjonen i 1979, og fra krigen mot Irak, mens tamiler som hadde flyktet fra borgerkriger på Sri Lanka kom senere på 80-tallet. På 1990-tallet kom det blant annet mange fra borgerkrigene på Balkan og i Somalia. De senere årene har også utenlandske kvinner som gifter seg med norske menn utgjort en viktig del av innvandringen. Kvinner fra Thailand og Russland er særlig representert her (ibid).

I perioden 1990-2006 har det innvandret over 284 000 førstegenerasjonsinnvandrere med ikke-nordisk statsborgerskap til Norge. Av disse kom 122 000 (43 prosent) som følge av familieinnvandring. Nesten en tredjedel av innvandrerne kom som flyktninger (29 prosent), mens 16 prosent var arbeidsinnvandrere. 11 prosent kom for å ta utdanning (Statistisk sentralbyrå, 2007). Av de innvandrerne som kom i 1993, og tidligere, kom de aller fleste som flyktninger eller på grunn av familieinnvandring. I de senere årene har imidlertid arbeidsinnvandringen skutt i været, og i 2006 var de fleste innvandrerne som kom til Norge arbeidsinnvandrere (Statistisk sentralbyrå, 2007). Dette kan ha sammenheng med at alle med landbakgrunn innenfor EØS-området fritt kan søke arbeid og bosette seg i Norge. De fleste innvandrere som tidligere søkte arbeidstillatelse som sesongarbeidere i Norge, i tilknytning til jordbruket, kommer fra land som ble EU-medlemmer i 2004. Disse kan dermed ta del av det frie arbeidsmarkedet innenfor EU (Djuve og Friberg, 2004).

Ifølge Statistisk sentralbyrå (2008) har den største andelen av de som innvandret til Norge fra Europa i årene 1990 til 2007, innvandret på grunn av arbeid. Hovedårsaken til at innvandrere fra de asiatiske landene innvandret til Norge i samme tidsperiode var familieinnvandring, mens Afrika hadde en overvekt av flyktninger. Den største andelen fra de amerikanske landene kom også på grunn av familieinnvandring. Det er imidlertid store forskjeller mellom landene i disse regionene. Mens personer fra de østeuropeiske landene Polen og Litauen har innvandret til Norge på grunn av arbeidsårsaker eller for å ta utdanning i Norge, har

innvandrere fra andre østeuropeiske land, som Serbia og Montenegro og Bosnia-Hercegovina innvandret på grunn av flukt. Vi ser en lignende tendens for de asiatiske verdensregionene. Fra Vest-Asia kommer innvandrere fra land som Afghanistan, Iran og Irak på flukt, mens innvandrere fra for eksempel Pakistan og Tyrkia i hovedsak kommer på grunn av familiegjenforening. Fra Øst-Asia kommer imidlertid de fleste på grunn av familieinnvandring, men vi ser også skiller her, da for eksempel innvandrere fra Sri Lanka og Vietnam kommer for å forenes med en fra samme land, mens innvandrere fra Thailand og Filippinene kommer for å etablere familie i Norge. Mange innvandrere fra Filippinene kommer imidlertid også for å ta utdanning i Norge, og et flertall av innvandrerne fra Kina kommer av nettopp denne grunnen.

2.2 Utdanningspolitikk for innvandrere

Myndighetenes politikk ovenfor innvandrere er delt inn i to grupper; *innvandringspolitikk* og *integreringspolitikk*. Innvandringspolitikken består av regelverk og praksis innen utlendingsfeltet som bestemmer hvem og hvor mange som til enhver tid får opphold i Norge, og gir dermed myndighetene mulighet til å gi personer oppholdstillatelse på bakgrunn av utdanningsnivå. Eldrebølgen har gjort at behovet for kvalifiserte arbeidere er stort, og innvandring har i denne sammenhengen blitt trukket frem som en mulig løsning. I 2002 ble derfor regelverket for spesialistinnvandring myket opp. Dette vil si at det kan bli utstedt oppholdstillatelse til personer i Norge, kun på bakgrunn av deres utdanning og kvalifikasjoner. Det ble satt en kvote på 5000 spesialister årlig, som skulle velges ut på bakgrunn av kvalifikasjoner og arbeidskraftsbehov. Det har imidlertid vist seg vanskelig å tiltrekke seg denne ønskede arbeidskraften, da man ikke har klart å fylle den årlige kvoten (Djuve og Friberg, 2004).

Integreringspolitikken retter seg mot de som faktisk har fått opphold, og er bosatt i landet, og det er altså her *utdanningspolitikken* kommer inn. Det er en politisk målsetting at alle, uavhengig av opprinnelse og kjønn, skal ha like muligheter, rettigheter og plikter til å delta i samfunnet og å bruke sine ressurser, i utdanning, arbeidsliv og organisasjoner. Mange innvandrere trenger tilførsel av kompetanse og ferdigheter som setter dem i stand til å delta på lik linje med nordmenn (Djuve og Friberg, 2004). Norskopplæring er sentralt her, da det å ha

manglende kunnskaper i det norske språk vil være et av de største hindrene for deltakelse i arbeidslivet og andre sosiale arenaer.

2.2.1 Introduksjonsloven

Fra og med 1. september 2005 trådte en ny lov om introduksjonsordning og norskopplæring for nyankomne innvandrere i kraft. Formålet med loven var å styrke nyankomne innvandreres mulighet for deltakelse i yrkes- og samfunnslivet, og å styrke deres økonomiske selvstendighet. Loven går ut på at nyankomne innvandrere har en rett og/eller plikt til å delta på 300 timer opplæring i norsk og samfunnskunnskap. Opplæringen omfatter 250 timer norskopplæring og 50 timer samfunnskunnskap på et språk deltakerne forstår, og denne pliktige opplæringen skal gjennomføres i løpet av de første tre årene i Norge. Gjennomføring av denne opplæringen, eller dokumentasjon på tilstrekkelige kunnskaper i norsk eller samisk ble samtidig innført som et grunnlag for å få bosettingstillatelse (permanent oppholdstillatelse) i Norge. Kommunene har plikt til å gi opplæring opp til totalt 3 000 timer for de innvandrerne som trenger det, og denne plikten for kommunene gjelder i fem år etter innvandringen (Arbeids- og inkluderingsdepartementet, 2006).

De som har rett og plikt til denne ordningen er innvandrere mellom 16 og 55 år, som har fått opphold i Norge på bakgrunn av asyl, opphold på humanitært grunnlag, familiegjennforening eller kollektiv beskyttelse i en massefluktsituasjon. Dersom innvandrerne i de nevnte gruppene over er over 55 år, har de rett til denne opplæringen, men ikke plikt. Arbeidsinnvandrere og deres familiemedlemmer har plikt til deltakelse på ordningen, men ikke rett på gratis opplæring (Arbeids- og inkluderingsdepartementet, 2003).

2.2.2 Strategi for bedre læring blant etniske minoriteter

Regjeringen kom i 2004 med en ny strategi for bedre læring og større deltakelse av språklige minoriteter i barnehage, skole og utdanning. Strategiplanen har fem hovedmål som går på å bedre språkferdighetene, prestasjonene og andelen minoritetsspråklige i både førskole, grunnopplæring, videregående opplæring og høyere utdanning. Strategiplanen går også ut på å bedre norskerferdighetene til minoritetsspråklige voksne for å øke mulighetene for utdanning og aktiv deltakelse i samfunnsliv og arbeidsliv. Hovedmålet er her å sette i gang tiltak som skal

bidra til en likeverdig opplæring som gir alle like muligheter til læring, utdanning og arbeidsliv (Kunnskapsdepartementet, 2007).

Tiltakene som har blitt satt i gang går blant annet ut på å utvikle en mer fleksibel og målrettet opplæring for elever med mangelfull opplæring og kort botid i landet, økt rekruttering av studenter med minoritetsbakgrunn til høyere utdanning generelt og tiltak for å bedre arbeidet med godkjenning av høyere utdanning fra utlandet. Det vil også bli satt i gang tiltak for å styrke oppfølgingen av minoritetsspråklige studenter, for å bedre gjennomføringen av studiene (ibid).

2.3 Norsk standard for utdanningsgruppering (NUS2000)

Norsk standard for utdanningsgruppering er en inndeling av de forskjellige utdanningsnivåene og faginndelingene i Norge. Denne standarden er utarbeidet av Statistisk Sentralbyrå, og er først og fremst beregnet for bruk i offisiell norsk statistikk, og er dermed ment til å gi et godt bilde av det norske utdanningssystemet. Utdanningsstandarden er inndelt i seks siffer, der det første sifferet er en nivåinndeling, og forteller oss hvor lang utdanning personen har tatt. Andre til fjerde siffer er en faginndeling, som er lik på alle nivåer, og siffer nummer fem til seks presiserer hvilke fag hver enkeltutdanning omfatter, og ikke omfatter (Statistisk sentralbyrå, 2000).

Norsk standard for utdanningsgruppering ble første gang utarbeidet av SSB i 1970, og har senere blitt revidert i 1973 og 1989. Siden 1989 har det norske utdanningssystemet gjennomgått en rekke reformer, blant annet Reform 94 og Reform 97, som hadde stor innvirkning på utdanningsaktiviteter i Norge. Reform 97 førte til at den obligatoriske grunnskoleutdannelsen ble forlenget med ett år. Selv om grunnskolen i dag har en varighet på ti år, har de aller fleste som er født i Norge per i dag bare ni eller færre års obligatorisk grunnskoleutdanning. NUS2000 tar hensyn til Reform 97, og legger den derfor til grunn for klasstrinnsklassifiseringen. Av forenklingshensyn skiller de likevel ikke mellom dem som har ulik lengde på grunnutdanningen, noe som gjør at de som ikke har gjennomført ti års utdanning behandles som om de har gjort det. En person som har fullført grunnskolen før Reform 97, og som har treårig videregående skole, blir dermed plassert i klasstrinn 13, selv om personen i realiteten kun befinner seg på klasstrinn 12 (ibid).

SSB definerer *utdanning* som utdanningsaktiviteter som har en formell skolemessig ramme, innrettet mot en systematisk formidling av kunnskap og ferdigheter, og som har en viss lengde. NUS omfatter dessuten kun utdanningsaktiviteter som foregår innenfor utdanningssystemet. SSB definerer videre en *utdanningsaktivitet* som: et eller flere utdanningstilbud som er tilnærmet like med hensyn til faglig innhold og nivåmessig plassering (ibid).

2.3.1 Utdanningsnivå

De norske definisjonene av befolkningens utdanningsnivå har tidligere ikke vært helt i samsvar med de nåværende internasjonale retningslinjene. Det har derfor vært behov for endringer i nivådefinisjonene ved publisering av norske tall (Jørgensen, 2006), og utdanningsnivåene i NUS2000 er derfor satt opp på en slik måte at de kommer nærmere internasjonal standard. Variabelen *utdanningsnivå* blir gjerne sett i sammenheng med andre variabler, og brukes ofte internasjonalt som et mål på humankapital og som en indikator i levekårsundersøkelser. Statistikk om befolkningens utdanningsnivå er derfor en av de utdanningsstatistikkene som er mest etterspurt både nasjonalt og internasjonalt (ibid).

Norsk standard for utdanningsgruppering har i dag ni nivåer, pluss en verdi for uoppgitt. Denne nivåinndelingen er ment til å gi et best mulig bilde av strukturen i det norske utdanningssystemet, og er bygd opp slik som vist i tabell 1. Slik systemet er i dag, får alle som tar en grunnskoleutdanning bestått, uansett resultat. Før fram til slutten av 1980-tallet, var det imidlertid ikke slik, og de som ikke fikk bestått fram til da, og som heller ikke senere har blitt registrert med en fullført grunnskoleutdanning, er nå klassifisert i kategorien ”ingen utdanning eller førskoleutdanning”. Alle som har fullført en grunnskoleutdanning som sin høyeste utdanning, defineres på grunnskolenivå. De som har fullført barneskolen plasseres på ”Barneskoleutdanning” og de som har fullført ungdomskolen plasseres på ”Ungdomskoleutdanning”. På ungdomskoleutdanning kommer også i tillegg de som har fullført så lite videregående utdanning at de ikke har oppnådd en utdanning på videregående nivå etter de nye definisjonene.

Tidligere ble alle som har fullført en eller annen videregående utdanning, uansett lengde eller nivå plassert på videregående nivå. De som hadde en utdanning mellom videregående og universitets- og høyskolenivå inngikk også her. I den nye definisjonen av videregående nivå er nivået delt opp i ”Videregående, grunnutdanning”, ”videregående, avsluttende utdanning”

Tabell 1: NUS2000s inndeling av utdanningsnivå (Statistisk sentralbyrå, 2000)

Tredeling av nivå	Nivå	Nivånavn	Klassetrinn
	0	Ingen utdanning og førskoleutdanning	Under skoleplikt
Obligatorisk utdanning	1	Barneskoleutdanning	1.-7. klasstrinn
	2	Ungdomsskoleutdanning	8-10. klasstrinn
Mellomutdanning	3	Videregående, grunnutdanning	11.-12. klasstrinn
	4	Videregående, avsluttende utdanning	13. klasstrinn +
	5	Påbygging til videregående utdanning	14. klasstrinn +
Universitets- og høyskoleutdanning	6	Universitets- og høyskoleutdanning, lavere nivå	14. -17. klasstrinn
	7	Universitets- og høyskoleutdanning, høyere nivå	18.-19. klasstrinn
	8	Forskerutdanning	20. klasstrinn +
	9	Uppgitt	

og ”påbygging til videregående utdanning”. Under ”videregående, grunnutdanning” kommer grunnkurs og VK I i videregående opplæring, et års utdanning ved folkehøyskole og den gamle realskoleutdanningen inn. I tillegg gjelder også utdanning som ikke krever fullført treårig videregående opplæring før opptak, og utdanning fra utlandet som tilsvarer 11.-12. klasstrinn i videregående opplæring. Under ”videregående, avsluttende utdanning” gjelder VK II og VK III videregående utdanning og den gamle gymnasutdanningen, mens ”påbygging til videregående utdanning” gjelder de som har utdanning på 14. klasstrinn og høyere, altså de som har høyskoleutdanning som ikke er godkjent av Kirke-, utdannings- og forskningsdepartementet, forkurs som verken gir studiepoeng eller er en integrert del av en universitets- eller høyskoleutdanning og teknisk fagskole (Statistisk sentralbyrå, 2000). De som fra og med studieåret 1998/99 har tatt færre enn 120 studiepoeng i universitets- eller høyskolesystemet blir også definert her (Jørgensen, 2006).

Tidligere omfattet Universitets- og høyskolenivå alle de som var registrert med en eller annen universitets- eller høyskoleutdanning. Nivået omfattet da alle som hadde tatt noen vekttall/studiepoeng på universitet eller høyskole, fra de som kun hadde tatt noen få vekttall/studiepoeng og helt til doktorgradsnivå. I den nye definisjonen av Universitets- og høyskolenivå går grensen for å bli definert her på 120 studiepoeng. Man må altså ha fullført to år med høyere utdanning for å bli definert på dette nivået. Dette gjelder imidlertid kun de som

har fullført sin utdanning fra og med studieåret 1998/99, da datasituasjonen ikke gjør det mulig å benytte denne definisjonen for de som har fullført sin høyeste utdanning tidligere (Jørgensen, 2006). Vi skiller altså her mellom ”Universitets- og høyskoleutdanning, lavere nivå”, der man må ha minst to år med høyere utdanning for å bli definert, ”Universitets- og høyskoleutdanning, høyere nivå”, der man må ha høyere utdanning med minst fire års varighet, som for eksempel hovedfags- og mastergradsutdanning, og ”Forskerutdanning”, der blant annet doktorgrader blir definert.

2.4 Kartlegging av innvandreres utdanningsnivå

Opplysninger om utdanning hentes fra registeret over befolkningens høyeste utdanning (BHU). Det største kvalitetsproblemet med BHU har vært at det manglet utdanningsopplysninger om de aller fleste som har innvandret til Norge, men ikke har vært i kontakt med det norske utdanningssystemet (Dalheim, 2001). Dette har gjort det vanskelig å foreta omfattende undersøkelser av forskjeller i utdanningsnivå mellom innvandrere fra forskjellige land og landgrupper. Et annet problem har vært at det er vanskelig å si noe om forskjellene i utdanningsnivå for de det er registrert utdanningsnivå for, og de det mangler opplysninger om. I forbindelse med folke- og boligtellingsen i 1999/2000 gjennomførte SSB en spørreskjemaundersøkelse blant innvandrere for å komplettere utdanningsstatistikken slik at denne kunne bli bedre for hele befolkningen i Norge. Det samme ble også gjort i forbindelse med folke- og boligtellingsen i 1990 (ibid).

For å utføre undersøkelsen ble det gått grundig til verks, med en målsetting om at alle skulle kunne forstå spørreskjemaet, uavhengig av om de hadde skolegang eller ei. Et kort spørreskjema og informasjonsbrev hvor mottakerne ble orientert om undersøkelsen ble deretter sendt ut både på arabisk, engelsk, tyrkisk, urdu, vietnamesisk, albansk og serbokroatisk. Undersøkelsen hadde hjemmel i lov om offisiell statistikk, som sier at alle som mottar skjemaet har plikt til å svare. Innvandrersorganisasjoner ble kontaktet, og et godt samarbeid med disse sørget for at mange sendte inn spørreskjema til SSB om den utdanningen de hadde tatt i utlandet (ibid).

Til slutt hadde hele 84 prosent av de som fikk spørreskjemaet returnert det til SSB, og det ble funnet at det var flere blant personer med innvandrerbakgrunn som hadde utdanning på

universitets- og høyskolenivå enn blant personer med norsk bakgrunn. Noen har aldri fullført noen utdanning i det hele tatt, men utdanningsnivået varierer mye etter landbakgrunn.

Frafallet ble her fulgt nøye opp, og det viste seg at det var vanskeligst å få inn svar fra de med kun grunnskoleutdanning (ibid).

3 Litteraturgjennomgang og teori

Jeg skal i dette kapittelet gå gjennom tidligere forskning og litteratur på området. Jeg starter med å gå gjennom en enkel modell for hvorfor noen velger å ta høyere utdanning, for så å gå videre på generelle teorier om immigrasjon. I kapittel 3.3 går jeg gjennom internasjonal forskning på området, for så til slutt å ta en gjennomgang av tidligere forskning i Norge.

3.1 Modell for utdanning⁴

Forskjeller i antall år med utdanning, eller utdanningsnivå, er av stor betydning i den forstand at utdanning er sterkt korrelert med deltakelse på arbeidsmarkedet, arbeidsledighetsrater og inntekt. Dersom man ser på det amerikanske arbeidsmarkedet, har personer som ikke har fullført high school vesentlig lavere andel deltakelse i arbeidsmarkedet enn personer med vitnemål fra college. De som ikke har fullført high school har også en arbeidsledighetsrate som er fire ganger så høy, og en inntekt som er ca. tre ganger lavere enn for de som har fullført college (Borjas, 1996). Studier av inntektseffekter på lønn i Norge, har dessuten også vist at et ekstra års utdanning i gjennomsnitt gir mellom 4.5 og 6 prosent høyere inntekt (Raaum, 1999). Utdanning kan utgjøre en enda større forskjell i for eksempel inntekt for kvinner og etniske minoriteter, da deres inntekt vil være vesentlig lavere enn den gjennomsnittlige innfødtes inntekt i utgangspunktet.

På bakgrunn av at utdanning er forbundet med lavere arbeidsledighet og høyere lønn, hvorfor velger ikke alle da å ta høyest mulig utdanning? Utdanningsmodellen (først analysert av Jacob Mincer i 1958) antar at personer velger å ta utdanning frem til det utdanningsnivået som maksimerer fremtidsverdien av personens inntekt gjennom hele livsløpet. Modellen forutsetter at man kun tar utdanning for å øke fremtidig inntekt, at man ikke jobber under studietiden og at kunnskapen man inntar i skolen ikke depresierer over tid, slik at realinntekten er konstant gjennom hele livet. Lønnsnivået blir dermed høyere jo mer utdanning man tar. En person vil velge å ta utdanning så lenge kostnadene ved utdanning er lavere enn den økte fremtidige inntekten man får av utdanningen. Kostnadene må her deles opp i alternativkostnad, altså det en alternativt kunne tjent ved å jobbe istedenfor å gå på skole

⁴ Modellen, og hele dette avsnittet, er hentet fra Borjas (1996), kapittel 7

i de årene utdanningen pågår, samt de direkte kostnadene ved utdanningen som for eksempel skolepenger, bøker, skrivemateriale etc.

Fremtidsverdien av inntektsstrømmen til en person med kun videregående utdanning⁵ vil være:

$$PV_{VS} = w_{VS} + \frac{w_{VS}}{(1+r)} + \frac{w_{VS}}{(1+r)^2} + \dots + \frac{w_{VS}}{(1+r)^{46}} \quad (1)$$

Her er r arbeidernes diskonteringsrate, og w_{VS} er lønnen man får etter kun fullført videregående utdanning. Det er 47 ledd i likningen over, noe som tilsier at fremtidsverdien er beregnet fram til man er 64 år (dersom man var 18 når man var ferdig med videregående). Fremtidsverdien av inntektsstrømmen dersom personen får for eksempel en bachelorgrad⁶, vil se slik ut:

$$PV_{BG} = \underbrace{-H - \frac{H}{(1+r)} - \frac{H}{(1+r)^2}}_{\text{Direkte kostnader}} + \underbrace{\frac{w_{BG}}{(1+r)^3} + \frac{w_{BG}}{(1+r)^4} + \dots + \frac{w_{BG}}{(1+r)^{46}}}_{\text{Inntektsstrøm etter bachelorgraden}} \quad (2)$$

De første tre leddene beskriver her fremtidsverdien av de direkte kostnadene ved å ta en bachelorgrad, mens de gjenværende 43 leddene gir oss fremtidsverdien av inntektsstrømmen etter fullført grad. Arbeidstakeren vil etter fullført videregående skole, velge å ta en bachelorgrad så lenge $PV_{BG} > PV_{VS}$. Altså så lenge den fremtidige inntektsstrømmen er høyere dersom man tar en bachelorgrad, enn dersom man ikke gjør det.

Vi ser av uttrykkene over at diskonteringsraten spiller en viktig rolle i avgjørelsen om man skal gå videre på skole eller ikke. Jo høyere denne er, jo mindre sannsynlig er det at personen vil velge å investere i utdanning. Det er enkelt å anta at denne diskonteringsraten er lik markedsrenten, da diskonteringen av fremtidig inntekt i nåverdikalkulasjoner delvis oppstår på grunn av at en krone mottatt i år, ikke er det samme som en krone mottatt neste år. En persons diskonteringsrate avhenger imidlertid også av hvordan han/hun prefererer inntekt *i dag* framfor inntekt *i fremtiden*. En person som er ”nåtid” orientert, altså har sterkere

⁵ I den opprinnelige modellen gjelder dette personer som har fullført high school, her litt fornorsket.

⁶ College diploma i den opprinnelige modellen.

preferanser for inntekt i dag, enn inntekt i fremtiden, vil ha en høy diskonteringsfaktor. Personen vil dermed ha mindre sannsynlighet for å investere i utdanning.

Dersom alle arbeidere har en konstant diskonteringsrate som er uavhengig av antall år med skolegang, burde arbeiderne avslutte sin utdanning i det punktet der den marginale avkastningsraten av utdanning er lik diskonteringsraten, r . Det vil si, i det punktet der ett år med ekstra utdanning vil gi en økning i inntekt som er like stor som personens diskonteringsrate. I dette punktet vil arbeideren ha maksimert sin fremtidsverdi av inntekt over livsløpet. Vi går her ut fra at den marginale avkastningen av utdanning avtar jo mer utdanning man tar.

Noe som er vanskeligere å observere, men også har innvirkning på både utdanningsnivå og inntekt, er forskjeller i personers *evner*. Jørgensen (1995). Det er ofte antatt at høyere evner skifter den marginale avkastningen av utdanning oppover, slik at de med mye evner får større utbytte av et års ekstra utdanning enn de med mindre evner. Inntektsforskjeller kommer dermed ikke bare av forskjeller i utdanningsnivå, men også av denne forskjellen i evner. Innvandrere vil for eksempel ha språkbarrierer og kulturbarrierer som gjør at de ofte må bruke flere år på utdanning i Norge, noe som igjen vil føre til at kostnadene ved å ta utdanning øker. De vil også kunne oppleve å ikke få brukt evnene sine fullt ut, og dermed ikke få samme avkastning på utdanningen sin som andre med samme evner. Effekten av dette vil være, alt annet likt, at innvandrere tar mindre utdanning enn de innfødte.

3.2 Immigrasjon

Det er ikke bare Norge som for tiden har en stor innvandrerstrøm, men også mange andre land. FN rapporterte i 2006 at innvandring har blitt et viktig kjennetegn ved det internasjonale liv. Personer som bodde utenfor sitt hjemland hadde i 2005 nådd 191 millioner, og 115 millioner av disse bor nå i utviklede land, mens 75 millioner har innvandret til utviklingsland. En tredjedel av disse innvandrerne har flyttet fra et utviklet land til et annet utviklet land, mens omtrent samme antall innvandrere har flyttet fra et utviklingsland til et utviklet land. Det er altså like vanlig med innvandring mellom rike land, som med innvandring fra et fattig til et rikt land (United Nations, 2006).

Det er imidlertid store forskjeller i innvandringspolitikken mellom land. Noen land, som for eksempel USA, gir oppholdstillatelse så å si kun til innvandrere som har familie i landet fra før. Andre land, slik som Australia og Canada, gir opphold til personer som har de sosioøkonomiske karakteristikene som er ønskelige i landet. Dette vil for eksempel være utdanning, type arbeidserfaring etc. Andre land igjen, slik som Tyskland, åpnet sine dører for ”midlertidige” arbeidsimmigranter på 60-tallet, for så å finne ut at disse midlertidige arbeiderne ble en permanent del av den tyske befolkningen (Borjas, 1994). Etter innvandringsstoppen i 1975 tok Norge, som nevnt, for det meste inn innvandrere som kom på grunn av familiegjenforening. De siste årene har imidlertid arbeidsinnvandring fra de østeuropeiske EU-landene skutt i været, da de åpne grensene mellom EU-land tillater alle som ønsker det å søke jobb i Norge (Djuve og Friberg, 2004).

Innvandringspolitikken kan altså ha en god del å si for den økonomiske aktiviteten, både på kort og lang sikt. Jeg skal derfor se litt på hvordan forskjeller mellom lands innvandringspolitikk, altså hvilken type innvandrerbefolkning de tar inn i landet, får innvirkning på utdanningsnivået og inntekten til disse innvandrergruppene. Jeg begynner imidlertid med å se på hvordan det kan være forskjeller på innvandrernes utdanningsnivå og kvalifikasjoner, avhengig av hvor innvandrerne kommer fra, og når de kommer til landet. Jeg skal så se litt på hvilke personer i et land som faktisk velger å immigrere til et annet land ved bruk av den såkalte Roy modellen. I utdanningsmodellen over så vi at høyere inntekt senere i livet er det største incentivet til å ta utdanning. Siden sammenhengen mellom disse to er så sterk, vil jeg derfor sidestille utdanning/kvalifikasjoner og inntekt i de neste avsnittene.

3.2.1 Kohort effekter

Flere studier i den moderne økonomiske litteraturen om innvandrere fokuserer på å bestemme trender i kvalifikasjonsnivå og inntekt hos innvandrere (Borjas, 1994). Med kvalifikasjoner menes både utdanningsnivå og andre relevante kvalifikasjoner, som språkkunnskap og erfaring. Disse studiene ser på hvordan innvandrerne klarer seg i arbeidsmarkedet i vertslandet, som er det landet de innvandrer til, da innvandrerne har en påvirkningskraft på kvalifikasjonsnivået og produktiviteten i økonomien. Det tidlige arbeidet til Chriswick (1978) analyserte for eksempel hvordan innvandrere tilpasset seg på det amerikanske arbeidsmarkedet ved hjelp av en tverrsnitts regresjonsmodell (Borjas, 1994). Han fant at på innvandringstidspunktet, tjente innvandrerne ca. 17 prosent mindre enn de innfødte i landet.

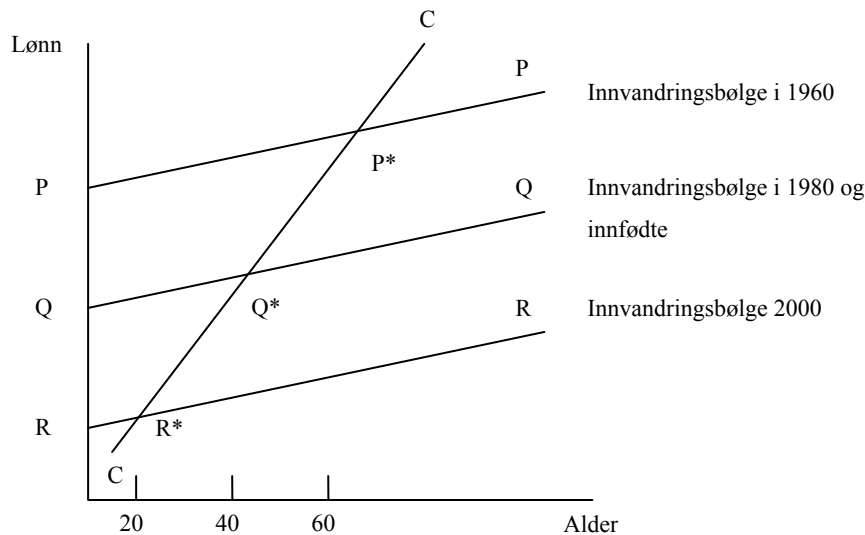
Innvandrerne opplevde imidlertid en raskere lønnsvekst, og innvandrernes lønninger ville derfor gå forbi lønningene til de innfødte innen femten år etter innvandringstidspunkt. Etter 30 års botid i USA ville den gjennomsnittlige innvandrer tjene 11 prosent mer enn den gjennomsnittlige amerikanske arbeideren. Dette ble forklart med at innvandrere på innvandringstidspunktet manglet en del spesifikke kvalifikasjoner som er nødvendig på det amerikanske arbeidsmarkedet, som for eksempel engelske språkkunnskaper. De hadde derfor en lavere lønn enn resten av befolkningen tidlig etter innvandringstidspunktet. Når innvandrerne etter hvert inntok disse kvalifikasjonene, vokste humankapitalnivået til innvandrerne relativt til de innfødtes, slik at de opplevde en sterkere lønnsvekst. Siden humankapitalhypotesen ikke i seg selv genererte et punkt der innvandrernes lønn går forbi de innfødtes, med samme utdanningsnivå, ble dette forklart med at innvandrerne var ”mer egnet og mer motivert” til å jobbe enn de innfødte (Borjas, 1994).

Det grunnleggende problemet med denne tverrsnittsanalysen er imidlertid at det blir dratt konklusjoner om hvordan inntekten til innvandrere utvikler seg over tid, kun ved å studere innvandrerbefolkningen over *ett* år. Borjas (1985) argumenterte derfor for at det er mulig at innvandrere som kommer til USA i dag er iboende forskjellig fra innvandrere som innvandret til USA for tjue år siden. Dette gjør det umulig å bruke erfaringene fra arbeidsmarkedet for de som kom til landet for tjue år siden, til å analysere fremtidige inntekter for de som kommer i dag. Implikasjonene med dette blir illustrert i figur 1, der vi antar at det var tre forskjellige innvandringsbølger som kom til landet i henholdsvis år 1960, 1980 og 2000.

Vi antar at innvandrerne som kommer til landet er 20 år ved innvandringstidspunkt, og at den innvandrergruppen som kom på 1960-tallet er den gruppen i befolkningen med det høyeste produktivitetnivået i økonomien, inkludert innfødte arbeidere. Dersom vi kunne observert inntekten deres hvert år etter at de kom til USA, ville lønnslinjen deres vært som linje PP i figuren. Vi antar også at den siste innvandrergruppen er den minst produktive gruppen i befolkningen, og dermed har lønnslinje RR i figuren. Til slutt antar vi at innvandrerbølgen som kom i 1980 har de samme kvalifikasjonene som innfødte, og dermed har lønnslinje QQ. Disse forskjellene i kvalifikasjoner blant de forskjellige innvandringsbølgene, blir kalt kohort effekter (Borjas, 1996). Ved en tverrsnittsanalyse, der vi kun vil kunne observere ett punkt på hver av inntektsprofilene til disse innvandrerne, vil vi observere regresjonslinje CC. Vi ser at denne linjen er brattere enn inntektsprofilen til de innfødte, noe som får det til å se ut som om

lønnen til innvandrere konvergerer den til de innfødte i en alder av 40 år, selv om dette ikke er tilfelle.

Figur 1: Kohort effekter og innvandreres lønnsprofil basert på alder⁷



Dette viser altså at en tverrsnittsanalyse kan gi feilaktige estimater angående hvordan innvandrere tilpasser seg arbeidsmarkedet, dersom det er forskjeller i kvalifikasjoner mellom innvandrere som kommer i en periode og innvandrere som kommer i en annen. Disse kohort effektene kan for eksempel komme på bakgrunn av endringer i innvandringspolitikk. I 1965 ble det vedtatt en ny lov i USA, Immigrasjons- og nasjonalitetsloven av 1965, som endret innvandringspolitikken i USA drastisk. Fra å først å fremst gi innvandrere opphold på bakgrunn av nasjonalitet, var det nå familiegjenforening med innvandrere som allerede bodde i landet, som ble den viktigste grunnen til å gi opphold. Dersom dette skiftet i politikk genererte en ny innvandringsstrøm med mindre kvalifikasjoner, ville vi altså ikke oppleve en inntektskonvergens mellom innvandrere og innfødte, men vi ville isteden oppleve forskjeller i evner eller utdanningsnivå blant de forskjellige innvandringsstrømmene (Borjas, 1994).

⁷ Siden utdanning påvirker innvandreres lønnsprofil/kvalifikasjoner, avhenger kohort effektene også av utdanning. Borjas (1994) har vist at innvandreres utdanningsnivå har sunket fra eldre til nyere innvandringsbølger, så vell som lønnen. På grunn av at avkastningen av utdanning på lønn sank på 80-tallet, er imidlertid ikke forskjellene i utdanning like store som forskjellene i lønn.

Endring i landbakgrunn for innvandrerstømmene vil generere kohort effekter dersom kvalifikasjonsnivået varierer mellom land, eller dersom det ikke er like lett å ta i bruk de kvalifikasjonene man har tilegnet seg i hjemlandet for innvandrere med ulik landbakgrunn. Før immigrasjons- og nasjonalitetsloven i 1965 kom to tredjedeler av innvandrerne til USA fra Europa eller Canada, mens kun 6 prosent hadde landbakgrunn i Asia. Etter 1965 så man et stort skifte i innvandrerstømmen, da nå kun 13 prosent av innvandrerne på 80-tallet hadde landbakgrunn i Europa eller Canada, mens 37 prosent av innvandrerne nå kom fra asiatiske land. Det virker sannsynlig at kvalifikasjoner som er påkrevd i andre utviklede økonomier er lettere å ta i bruk på det amerikanske arbeidsmarkedet, enn kvalifikasjoner som er påkrevd i utviklingsland. Det er faktisk en sterk positiv korrelasjon mellom innvandreres inntekter i USA og nivået av økonomisk utvikling i landet de kommer fra, målt ved landets BNP (Borjas, 1994). Dette kan implisere forskjeller i lønnsnivå mellom forskjellige innvandrerstømmer.

Borjas (1994) finner at det er signifikante forskjeller i inntekt for innvandrerstømmer som har kommet i de senere år, sammenlignet med eldre innvandringsstrømmer. Innvandrere som har kommet i de senere år har lavere inntekt enn innvandrere som kom tidligere, også når man tar hensyn til at innvandrerne som har kommet nylig har et lavere gjennomsnittlig utdanningsnivå enn tidligere immigranter. Resultatet kan imidlertid være litt misvisende, da det er mulig at de innvandrerne som tjener dårlig i vertslandet, flytter tilbake til hjemlandet sitt. Selv om det ikke hadde vært noen returnigrasjon, har forskning vist at utvalgssammensetningen av en bestemt innvandringskohort endres over tid, fordi arbeidsstyrken i de senere årene inkluderer et stort antall personer som immigrerte som barn (Borjas, 1994).

3.2.2 Teoretisk betraktning på hvorfor emigrasjon forekommer

Borjas (1994) finner at det var store forskjeller i utdanningsdeltakelse og inntekt mellom de forskjellige landgruppene i USA i 1990. Disse forskjellene kan imidlertid ikke forklares ved at noen grupper har bodd lenger i landet enn andre grupper. Det er store forskjeller både i utdanningsnivå og inntekt selv for grupper som har bodd i landet i over ti år. Forskjellene i landbakgrunn kan faktisk forklare over 90 prosent av nedgangen i utdanning og inntekt mellom innvandringsbølgene i 1960 og 1980.

Det vil imidlertid også være forskjeller i kvalifikasjoner mellom innvandrere med samme landbakgrunn, da ikke alle fra samme land har for eksempel samme utdanningsnivå og

arbeidserfaring. Hvilken del av befolkningen vil finne det mest lønnsomt å immigrere til USA, de med høy utdanning, eller de med lave kvalifikasjoner? På bakgrunn av A. Roys (1951) selvseleksjonsmodell, som beskriver hvordan arbeidere sorterer seg selv mellom arbeidsmuligheter, argumenterer Borjas (1987) for at innvandrerstømmen ikke er et tilfeldig utvalg av personer fra befolkningen i landet den kommer fra.

Vi antar at befolkningen i land 0 (hjemlandet) vurderer om de skal migrere til land 1 (vertslandet). Dersom hele befolkningen valgte å bli i hjemlandet, eller hele befolkningen valgte å immigrere til vertslandet, ville inntektsdistribusjonene være gitt ved henholdsvis ligning (3) og (4):

$$\log w_0 = \mu_0 + \varepsilon_0 \quad (3)$$

$$\log w_1 = \mu_1 + \varepsilon_1 \quad (4)$$

der μ_0 og μ_1 er henholdsvis gjennomsnittsinntekten i hjemlandet og gjennomsnittsinntekten i vertslandet, og ε_0 og ε_1 måler avviket fra gjennomsnittslønnen i hjemlandet og vertslandet, og er antatt å være normalfordelt med gjennomsnitt lik null og varians lik σ^2 . Korrelasjonskoeffisienten mellom de tilfeldige variablene ε_1 og ε_0 er lik ρ . Vi antar her at gjennomsnittspersonen i land 0 har like kvalifikasjoner/lik utdanning som gjennomsnittspersonen i land 1, slik at μ_1 også gir gjennomsnittsinntekten til de innfødte i vertslandet, selv om dette vanligvis ikke er tilfelle. Denne forutsetningen hjelper oss å isolere innvirkningen seleksjonsprosessen har på sammensetningen av kvalifikasjoner i innvandrerstømmen.

Migrasjonsavgjørelsen kommer da an på en sammenligning mellom inntektsmuligheter i hjemlandet og i vertslandet, minus migrasjonskostnader (C). Indeks funksjonen kan da defineres som:

$$I = \left(\frac{w_1}{w_0 + C} \right) \approx (\mu_1 - \mu_0 - \pi) + (\varepsilon_1 + \varepsilon_0) \quad (5)$$

Her er $\pi = C/w_0$ et "tidsekvivalent" mål på migrasjonskostnader. En arbeider vil altså innvandre til vertslandet så lenge $I > 0$, og bli i hjemlandet ellers. Migrasjonskostnaden, C, er forskjellig for forskjellige arbeidere, og er både transportkostnad, samt kostnaden en får

dersom man må gå arbeidsledig en stund i det nye landet mens man ser seg om etter jobb. Vi antar at den tidsekvivalente kostnaden, π , er konstant i populasjonen, slik at migrasjonskostnaden er proporsjonal med inntekt. Sannsynligheten for at en person velger å immigrere til vertslandet er da:

$$P = Pr\{v > (\mu_0 + \pi - \mu_1)\} = 1 - \Phi(z) \quad (6)$$

der $v = \varepsilon_1 - \varepsilon_0$, $z = (\mu_0 + \pi - \mu_1) / \sigma_v$, og Φ er funksjonen av standard normalfordelingen. Vi kan da vise at:

$$\frac{\partial P}{\partial \mu_0} < 0, \frac{\partial P}{\partial \mu_1} > 0 \text{ og } \frac{\partial P}{\partial \pi} < 0 \quad (7)$$

Emigrasjonsraten er altså her negativt korrelert med gjennomsnittsinntekten i hjemlandet, og med migrasjonskostnaden, mens den er positivt korrelert med gjennomsnittsinntekten i vertslandet. De potensielle innvandrerne er inntektsmaksimerende individer, og vil derfor velge det alternativet som gir den høyeste inntekten. Vi antar de betingede gjennomsnittene $E(\log w_0 | \mu_0, I > 0)$ og $E(\log w_1 | \mu_1, I > 0)$, som er henholdsvis den gjennomsnittlige inntekten i hjemlandet og i vertslandet for de som innvandrer. På grunn av normalfordelingen har vi da at:

$$E(\log w_0 | I > 0) = \mu_0 + \frac{\sigma_0 \sigma_1}{\sigma_v} \left(\rho - \frac{\sigma_0}{\sigma_1} \right) \lambda \quad (8)$$

$$E(\log w_1 | I > 0) = \mu_1 + \frac{\sigma_0 \sigma_1}{\sigma_v} \left(\frac{\sigma_0}{\sigma_1} - \rho \right) \lambda \quad (9)$$

der $\lambda = \varphi(z) / (1 - \Phi(z))$, og φ er standard normalfordelingen. Variabelen λ er inverst relatert til emigrasjonsraten, og vil være positiv så lenge noen finner det mest lønnsomt å bli i hjemlandet ($P < 1$). Vi lar så $Q_0 = E(v_0 | I > 0)$ og $Q_1 = E(v_1 | I > 0)$. Ligning (8) og (9) indikerer at det er tre mulige typer seleksjon som karakteriserer innvandrerstrømmene:

$$Q_0 > 0 \text{ og } Q_1 > 0 \text{ dersom } \rho > \frac{\sigma_0}{\sigma_1} \text{ og } \frac{\sigma_1}{\sigma_0} > 1 \quad (10)$$

$$Q_0 < 0 \text{ og } Q_1 < 0 \text{ dersom } \rho > \frac{\sigma_1}{\sigma_0} \text{ og } \frac{\sigma_0}{\sigma_1} > 1 \quad (11)$$

$$Q_0 < 0 \text{ og } Q_1 > 0 \text{ dersom } \rho < \min\left(\frac{\sigma_1}{\sigma_0}, \frac{\sigma_0}{\sigma_1}\right) \quad (12)$$

Positiv seleksjon (10) forekommer når innvandrere har over gjennomsnittlig inntekt i både hjemlandet og vertslandet. Negativ seleksjon (11) forekommer derimot når innvandrere har under gjennomsnittlig inntekt i begge landene. Begge disse typene seleksjon krever at kvalifikasjoner er positivt korrelert med inntekt/utdanning mellom landene. Variansene σ_0 og σ_1 måler "prisen" på kvalifikasjoner: jo større avkastning man får av utdanning, jo større er inntektsforskjellene mellom individene (Borjas, 1999). De med høy utdanning vil altså øke inntekten i forhold til de med lav utdanning, jo høyere avkastningen av utdanning er. Vi har da positiv seleksjon blant innvandrere når hjemlandet, relativ til vertslandet, er et land med lite ulikheter i inntekt blant befolkningen. Gode eksempler på slike land kan tenkes å være Norge, og de andre skandinaviske velferdsstatene. Mens vi har negativ seleksjon når hjemlandet, relativt til vertslandet, har stor ulikhet i befolkningens inntekt, etter ulikhet i utdanning. USA kan eksempelvis være et godt eksempel på et slikt type land. Jamfør (Barth et al., 2004). Det kan derfor argumenteres for at de med høy utdannelse i Norge ville hatt økonomiske insentiver til å immigrere til USA, mens de med lav utdannelse i USA ville hatt insentiver til å immigrere til Norge.

Ligning (12) viser imidlertid et tilfelle der innvandrere som har under gjennomsnittlig inntekt i hjemlandet, får over gjennomsnittlig inntekt i vertslandet. Borjas (1987) kaller denne seleksjonen for en "flyktningseleksjon", og vi vil ha denne typen seleksjon så lenge korrelasjonskoeffisienten, ρ , er liten eller negativ. Ifølge Borjas (1987) kan denne korrelasjonen for eksempel tenkes å være negativ dersom et nytt kommunistregime tar over styringen i hjemlandet. Dette politiske systemet vil omfordele inntekt ved å konfiskere midlene til relativt suksessfulle personer. Innvandrere fra slike systemer vil da kunne være i den nedre delen av inntektsdistribusjonen i hjemlandet, men på den øvre delen av inntektsdistribusjonen i vertslandet. Ligningene over viser at det verken er forskjellene i gjennomsnittsinntekt mellom landene, eller migrasjonskostnaden som bestemmer hvilken type seleksjon som bestemmer innvandringsstrømmen. Gjennomsnittsinntekter og migrasjonskostnader har betydning for størrelsen på innvandringsstrømmen, men de avgjør ikke om innvandreren er trukket hovedsakelig fra den øvre eller den nedre delen av kvalifikasjonsdistribusjonen, altså om de har høy eller lav utdanning.

3.3 Tidligere forskning internasjonalt

Friedberg (2000) har gjort en undersøkelse på i hvor stor grad utdanning som er tatt utenlands kan brukes i Israel. Resultatene er her svært like de man kan finne i litteraturen for USA, noe som impliserer at observasjonene fra Israel kan bli brukt til å kaste lys over dette spørsmålet mer generelt (Friedberg, 2000). Hun finner at fra hvilket land en persons utdanning og arbeidserfaring er fra har stor betydning for deres verdi i arbeidsmarkedet. Utdanning er derfor ikke perfekt overførbart mellom land. Med få unntak, gir utdanning som man har tilegnet seg i utlandet lavere avkastning på arbeidsmarkedet i vertslandet, enn utdanning som er tilegnet hjemme. Avkastningen av utdanning er høyere for innvandrere fra Europa, Amerika og Oseania, enn for innvandrere fra Afrika og Asia. Disse forskjellene kan reflektere forskjeller i kvalitet på utdanning i de forskjellige verdensdelene, men kan også reflektere hvor lett utdannelsen tatt i utlandet kan tilpasses til kravene på arbeidsmarkedet i vertslandet (ibid).

Friedberg finner også at den direkte avkastningen av et års utdanning eller et års erfaring, også er høyere for de innfødte i landet enn for innvandrere. En naturlig forklaring på dette kan være at språkkvalifikasjonene til de innfødte er vesentlig bedre. Hvor ”transportabel” utdanningen er over grensene kommer også an på hvilket nivå utdannelsen er på. Grunnskole er noenlunde likt evaluert, nesten uansett hvor man har tatt denne utdannelsen. Verdien av videregående skole på arbeidsmarkedet er sterkt avhengig av hvor denne utdanningen ble tatt, og videregående tatt hjemme gir her den høyeste lønnen. Avkastningen av høyere utdanning som er tatt i utlandet, kommer sterkt an på i hvilket land denne utdannelsen ble tilegnet. Dersom innvandrere tar videre utdanning i Israel etter ankomst, ser dette ut til å gi en signifikant økning i avkastningsraten på utdannelsen tatt i utlandet for asiatiske og afrikanske innvandrere, med en svakere effekt på innvandrere fra Amerika og Europa (Friedberg, 2000).

Både Canada og USA opplevde store endringer i innvandringspolitikken på 1960-tallet. Før dette brukte begge landene landbakgrunnskriterier for å dele ut oppholdstillatelser blant de mange søkende. De fra nordvest Europa ble da preferert (Borjas, 1991). Som nevnt tidligere, gjorde skiftet i den amerikanske innvandringspolitikken i 1965 at USA begynte å gi oppholdstillatelse hovedsakelig til de som hadde familietilknytning til USA. Fra 1962/1967 endret også Canada sin innvandringspolitikk, og i 1967 begynte de med et såkalt poengsystem for hvilke innvandrere de tok inn i landet. Poengene ble da delt ut etter observerbare

kvalifikasjoner, der innvandrerne kunne bli tildelt opp til hundre poeng. Innvandrere fikk da poeng for blant annet utdanning, hvor etterspurt deres yrke var i på arbeidsmarkedet og alder. Dersom søkeren hadde femti av hundre mulige poeng, fikk søkeren oppholdstillatelse. I 1976 ble imidlertid denne ordningen utvidet, slik at de som hadde familie som allerede hadde fått opphold i landet, fikk ekstra poeng for dette (Borjas, 1991).

Selv om trenden fra europeisk immigrasjon mot asiatisk immigrasjon var sterk i begge landene, var det fortsatt store forskjeller i landbakgrunn for innvandrerne som kom til USA og Canada på 70-tallet. Canada hadde en større andel av innvandrere fra Europa, mens USA hadde en mye større andel innvandrere fra Sør- og Mellom-Amerika. Borjas (1991) fant at innvandrere i Canada, i gjennomsnitt, hadde høyere kvalifikasjoner enn innvandrere i USA. Ved å sammenligne utdanningsnivå i de to landene, fant han at innvandrere i Canada har omtrent et år mer med gjennomsnittlig skolegang enn innvandrerne i USA. Når det gjaldt inntekt hadde også innvandrerne i Canada høyere lønn relativt til befolkningen ellers, enn innvandrere i USA (Borjas, 1991). Den empiriske analysen impliserer imidlertid en enkel forklaring på disse forskjellene når det gjelder kvalifikasjoner. Det gjennomsnittlige utdanningsnivået for spesifikke landgrupper er omtrent det samme i USA og Canada. Dette betyr at Canadas poengsystem ikke tiltrekker seg personer med høyere utdanningsnivå fra et gitt land. Innvandrerbefolkningen i Canada består imidlertid i større grad av landgrupper som pleier å gjøre det bra på arbeidsmarkedet både i USA og Canada. Det er derfor nettopp denne sammensetningen som forklarer mesteparten av den observerte forskjellen i utdanning og lønn for innvandrere i USA og Canada.

Kalena Cortes (2004) mener at innvandrere ikke kan bli behandlet som en homogen gruppe, og deler derfor innvandrergruppen i USA opp i to deler: flyktninger og økonomiske innvandrere. Flyktninger er her de som flykter fra forfølgelse i landet de kommer fra, og dermed ikke har mulighet til å returnere til hjemlandet. Økonomiske innvandrere er, på den andre siden, de som kommer til USA på søken etter en bedre jobb og økonomisk trygghet. Disse kommer ofte til landet for å jobbe noen år, og spare opp kapital, for så å reise tilbake til hjemlandet igjen. Flyktninger vil dermed mest sannsynlig ha mindre sosial kontakt med hjemlandet sitt, da de ikke bare kan sette seg på et fly og dra på besøk. Det at flyktningene har en lengre tidshorisont i landet og mindre sosiale bånd med hjemlandet, gjør at de har større insentiver til å investere i humankapital i vertslandet, enn andre innvandrere har. Cortes (2004) identifiserer flyktninggrupper og de økonomiske innvandrerne som kom til USA mellom 1975

og 1980, og finner at flyktninger i gjennomsnitt har lavere årlige inntekter ved innvandringstidspunktet, men at inntekten øker mer over tid for flyktninger enn for økonomiske innvandrere. Innvandrere tenderer også til å investere mer i landspesifikk humankapital enn de økonomiske innvandrerne, som for eksempel tilegning av det engelske språket. Da begge innvandrergruppene hadde like gode språkkunnskaper på innvandringstidspunktet, var bedringen i språkkunnskaper betydelig høyere for flyktninger enn for de økonomiske innvandrerne etter ti år (Cortes, 2004).

3.4 Tidligere forskning i Norge

Det har blitt dokumentert store forskjeller i utdanningsnivået til innvandrerne i Norge. OECD publiserte i 2006 en rapport som viste at 40 prosent av utenlandske statsborgere i Norge har høyere utdanning.⁸ Geir Nygård og Gunnlaug Daugstad poengterte imidlertid i SSBs magasin at OECD da ikke tok med de innvandrerne som har fått norsk statsborgerskap. Tallene hadde i så fall bli langt lavere. Grunnen til dette er at utdanningsnivået vanligvis er høyt blant vestlige innvandrere, men at disse sjelden tar norsk statsborgerskap. Det motsatte er tilfelle for ikke-vestlige innvandrere, da deres utdanningsnivå er lavere, samtidig som de langt oftere blir norske statsborgere (Nygård og Daugstad, 2006). I følge statistisk sentralbyrå er likevel mange ikke-vestlige innvandrere mer utdannet enn resten av den norske befolkningen. Andelen høyt utdannede er for eksempel langt større blant filippinere, polakker, russere og indere, som alle er en del av ikke-vestlige innvandrergrupper, enn gjennomsnittet i Norge. Samtidig kommer innvandrere fra andre ikke-vestlige områder dårlig ut på utdanningsstatistikken. Dette gjelder for eksempel innvandrere fra Somalia, Pakistan og Thailand, som kommer dårligst ut når det gjelder høyere utdanning. Det er også mange ikke-vestlige kvinner som har lav utdanning, eller ikke utdanning i det hele tatt (Statistisk sentralbyrå, 2001).

Dælen (2001) undersøker rekruttering til høyere utdanning blant innvandrere ved hjelp av registerdata. Hun finner at ungdommer med ikke-vestlig innvandrerbakgrunn sjeldnere begynner i høyere utdanning enn øvrige ungdommer. Ungdommer med vestlig innvandrerbakgrunn har derimot tilnærmet lik tendens til rekruttering i høyere utdanning som norskfødte av norske foreldre. Jenter begynner oftere i høyere utdanning enn gutter, men for

⁸ publikasjonen Trends in International Migration (OECD)

ungdommer med ikke-vestlig innvandrerbakgrunn er denne tendensen omvendt (Dæhlen, 2001). Olsen (2006) undersøker om ikke-vestlig innvandrerungdom faller utenfor jobb og utdanning. Når det gjelder ikke-vestlige innvandrere finner han imidlertid at førstegenerasjonsinnvandrere fra Øst-Europa er de som gjennomgående har den høyeste andelen i jobb eller utdanning av de i aldersgruppen 16 til 29 år. Innvandrere fra Sør- og Mellom-Amerika kommer også høyt opp på statistikken, og er bare fem prosentpoeng under østeuropeerne i andelen aktive. Førstegenerasjonsinnvandrere under 25 år fra Asia kommer her dårligst ut når det gjelder andelen aktive. Blant etterkommerne har imidlertid disse forskjellene jevnet seg betraktelig ut (Olsen, 2006). Han finner også at en økning i botid høyner aktivitetsnivået, og at det fremkommer et markant økning i aktivitetsnivå etter fire års botid.

Lie (2003) finner at *innvandringsgrunn* har mye å si i forhold til forskjellene i utdanningsnivå. De som kommer til Norge fra vestlige land, kommer ofte nettopp i forbindelse med jobber som krever en viss utdanning, selv om en del innvandrere fra disse regionene også kommer for å få midlertidig arbeid som ikke krever spesielle kvalifikasjoner. Uansett innvandringsgrunn vil disse vestlige innvandrerne ofte være langt mer ressurssterke enn innvandrere fra ikke-vestlige land. Mange av innvandrerne fra ikke-vestlige land har kommet som arbeidsimmigranter før innvandringstoppen i 1975, og har i følge Lie (2003) som regel lav utdanning. Omtrent halvparten av de ikke-vestlige innvandrerne har kommet som flyktninger. Flyktninger som kommer til Norge for å søke politisk asyl har ofte høy utdannelse, mens mange av flyktningene fra den tredje verden ofte har lav utdanning eller dårlig dokumentasjon på utdanningen sin (Lie, 2003, Østby, 2004). De som kommer for familiegjenforening, ligner ofte på dem de gjenforenes med når det gjelder utdanningsnivå. I tillegg er det en god del innvandrere som kommer til Norge for å ta høyere utdanning. Disse kommer ofte fra land i Øst-Europa eller den tredje verden, og for land med relativt mange slike studenter i forhold til antall personer fra landet som er bosatt i Norge, vil denne gruppen kunne trekke opp utdanningsnivået betraktelig (ibid).

I følge Lie (2003) er det blant etterkommerne i Norge kun 4 prosent som er over 30 år, mens 60 prosent er over ti år. Det er derfor for tidlig å få et fullstendig bilde av utdanningsnivået i forhold til førstegenerasjonsinnvandrere ellers. Hun finner imidlertid at etterkommerne har større deltakelse i både videregående skole og høyere utdanning enn førstegenerasjons-

innvandrere. Ved en alder av 25 år er det likevel de uten innvandrerbakgrunn som i størst grad har fullført en universitets- eller høyskolegrad på lavt nivå. Førstegenerasjonsinnvandrere og etterkommere ser altså ut til å fullføre utdannelsen sin på et senere tidspunkt enn nordmenn uten innvandrerbakgrunn.

Barth, Bratsberg og Raaum (2004) har sammenlignet inntektsprofiler for innvandrere relativt til innfødte i Norge og i USA, ved bruk av sammenlignbare mikrodata fra begge landene. Funnene deres er entydige; Innvandrere i USA tjener mer, har en brattere lønnsvekst og har en raskere inntektsopphenting med de innfødte, enn innvandrere i Norge. Denne fordelene for innvandrere i USA er større, jo høyere utdanning innvandrerne har. En del av disse forskjellene kan forklares ved sammensetning av innvandrerens utdanning og landbakgrunn, men selv når det blir kontrollert for dette, ser vi betydelige forskjeller mellom de to arbeidsmarkedene. Det ser altså ikke ut som om en sammenpresset lønnsstruktur og omfattende reguleringer i arbeidslivet fører til noen merkbar beskyttelse av innvandrere som en gruppe, men tvert imot ser det ut til at disse forholdene kan være et hinder for inntektsopptjening. Dette fordi innvandrere med høyt utdanningsnivå vil kunne tjene mer i et land som USA, der lønnsspredningen er større, og dermed ønske seg til USA fremfor Norge (Barth et al., 2004).

Bratsberg, Raaum og Røed (2007) har gjort en analyse av hvordan tidlige arbeidsinnvandrere har klart seg på arbeidsmarkedet i Norge. I årene rett før innvandringsstoppen trådte i kraft foregikk det, som nevnt tidligere, en betydelig arbeidsinnvandring til Norge. Utvalget i analysen består av mannlige innvandrere fra Pakistan, Tyrkia, India og Marokko som kom til Norge mellom 1971 og 1975, og som fortsatt var bosatt i Norge i år 2000, og funnene her er at disse arbeidsinnvandrerne har svært korte yrkeskarrierer i forhold til like gamle personer født i Norge, med samme utdanningsnivå. De første årene etter ankomst hadde de en sysselsettingstilbøyelighet på 96 prosent, som var høyere enn den norske sammenligningsgruppen. Om lag 10 år etter ankomst til Norge begynte sysselsettingen hos denne gruppen å falle markant. Dette både i absolutt forstand, og relativt til sammenligningsgruppen. I år 2000, da personene i analyseutvalget var mellom 45-64 år gamle, var sysselsettingsraten blant innvandrere falt til 50 prosent, mot 87 prosent i sammenligningsgruppen. I tillegg mottok 73 prosent av innvandrerne en trygdeytelse og/eller økonomisk sosialhjelp i år 2000, mot kun 37 prosent av medlemmene i sammenligningsgruppen. Mange av ektefellene til arbeidsinnvandrerne kom på grunnlag av familiegjenforening på slutten av 70-tallet. Hos disse har

yrkesdeltakelsen vært gjennomgående lav. I år 2000 lå sysselsettingsraten for disse på omtrent 30 prosent, i forhold til 80 prosent for ektefellene til den norske sammenligningsgruppen (Bratsberg et al., 2007).

Mulige årsaker til denne lave sysselsettingen og høye trygdeavhengigheten blant arbeidsinnvandrere er ifølge Bratsberg, Raaum og Røed (2007), konjunkturforhold og økonomiske incentiver. Etterspørselen etter innvandrernes arbeidskraft har vært ustabil og konjunkturfølsom, og det er for mange innvandrere lite økonomisk lønnsomt å holde seg i arbeid. Den siste årsaken skyldes både det at innvandrere har lavere lønn enn nordmenn, og at de faktisk ofte kan oppnå *høyere* ytelser fra trygdesystemet dersom de først kvalifiserer for slike ytelser. Som vi har sett tidligere blir ofte inntektsgapet mellom innvandrere og nordmenn mindre jo lengre botid innvandrerne har i landet. I dette tilfellet er det med andre ord omvent (Bratsberg et al., 2006).

4 Databeskrivelse

I dette kapitlet skal jeg gå gjennom datamaterialet jeg bruker i oppgaven min. Kapittel 4.1 er en gjennomgang av datasettet mitt generelt, mens kapittel 4.2 og 4.3 tar for seg den avhengige variabelen jeg skal bruke i oppgaven, og de uavhengige variablene. Jeg viser den deskriptive statistikken til variablene underveis.

4.1 Datamaterialet

Datamaterialet i denne oppgaven er hentet fra databasen FD-trygd, Statistisk Sentralbyrå (SSB) og tilrettelagt av Norsk Samfunnsvitenskapelig Datatjeneste (NSD). Dataene er fordelt på to ulike paneldatasett. Utvalget i datasettene består av bosatte i Norge per 1.1.1993, som er mellom 16 og 66 år. Det vil si at de er født i årene 1926-1976, og kan også omtales som den arbeidsføre delen av befolkningen. Det første datasettet består av alle med innvandrerbakgrunn på dette tidspunktet, og det andre datasettet er et tilfeldig utvalg fra omtrent 20 prosent av befolkningen uten en slik bakgrunn. Til sammen ble populasjonen i utvalget på 673.598 individer, hvorav 142.614 har innvandrerbakgrunn, og 530.984 er etnisk norske.

Jeg har her funnet det hensiktsmessig å stramme inn aldersgruppen i utvalget mitt fra 22-54 år, da jeg mener dette er den gruppen som representerer utdanningsnivået i dagens samfunn best. Begrunnelsen for dette valget er at mange av de under 22 år har nok ikke gjort seg ferdig med den utdanningen de kommer til å ta, mens for de over 54 år vil trolig utdanningsnivået være vesentlig lavere enn det er i dag, da høyere utdanning ikke var like vanlig før som nå. Innvandringen til Norge fra Oseania og Mellom-Amerika før 1993 var svært lav (se avsnitt 2.1.1). Dette kommer også frem i datasettet mitt, da det her er svært få observasjoner fra disse regionene. Jeg har derfor valgt å ta også de med landbakgrunn fra disse regionene ut av utvalget mitt. Det samme gjelder for de med ”uoppgitt” landbakgrunn.

Dataene er tatt ut over en elleveårsperiode, fra og med 1993 til og med 2003. Ingen av de to datasettene er balanserte, det vil si at jeg følger individene over hele denne perioden, uten at det kommer til nye individer. I følge Mads Ivar Kirkeberg økte innvandrerbefolkningen med 72 prosent fra 1993 til 2003 (Statistisk sentralbyrå, 2004). Datasettet mitt kommer altså ikke

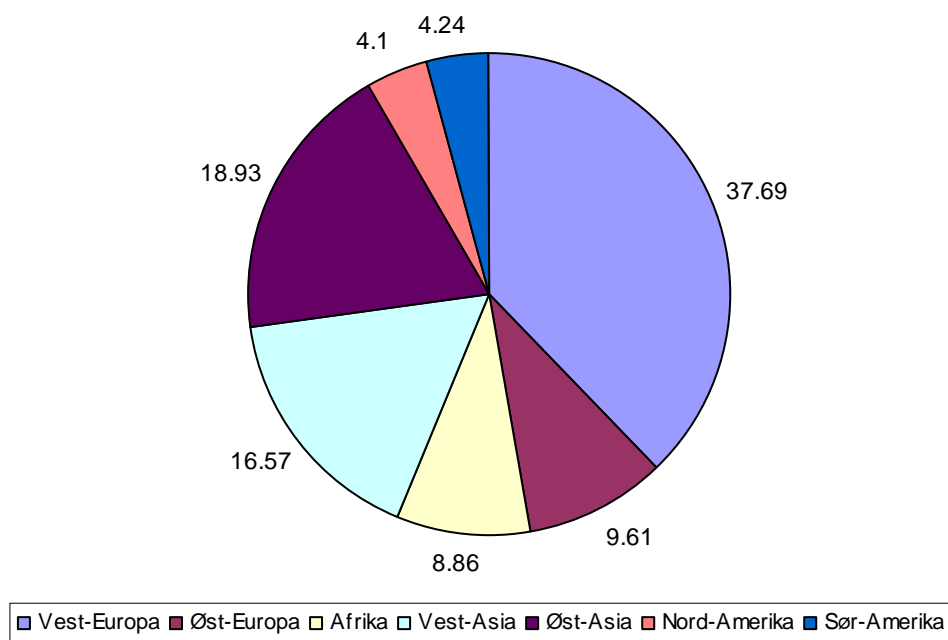
til å fange opp slike endringer, men vil derimot gi meg et bilde av innvandrerbefolkningen i Norge slik den var i 1993. Datasettet for innvandrere opererer med fem ulike innvandrer kategorier: Førstegenerasjonsinnvandrere (uten norsk bakgrunn), de som er født i Norge av to utenlandske foreldre (etterkommerne), de som er født i utlandet med én norsk forelder, de som er norskfødt med én utenlandsk forelder og de som er utenlandskfødt med to norske foreldre (herunder også adopterte). SSB definerer de to første av kategoriene som innvandrere, og det er derfor kun disse kategoriene det vil være ønskelig for meg å ha med i oppgaven. Jeg står dermed igjen med et utvalg på 118.018 personer med innvandrerbakgrunn.

Datasettet inneholder informasjon både om personens landbakgrunn, fødeland og statsborgerskap, i tillegg til detaljert bakgrunnsinformasjon om individene og deres levekår. I henhold til avtale med Statistisk Sentralbyrå, kan NSD kun levere ut anonymiserte datasett til forskere og studenter for å ivareta personvernet. Dette datasettet som brukes i masteroppgaven min er derfor anonymisert, med helt tilfeldig lagrede løpenummer. På grunn av dette har jeg heller ikke tilgang til hvilket land innvandrerne i datasettet mitt kommer fra, men har oppgitt verdensregionene *Vest-Europa, Øst-Europa, Afrika, Vest-Asia med Tyrkia, Øst-Asia, Nord-Amerika og Sør-Amerika*⁹, i tillegg til Norge.

Variabelen fødeland viser til hvilken verdensdel individet er født i, mens landbakgrunn går på "eget, eventuelt mors, eventuelt fars utenlandske fødeland" (Akselsen et al., 2007). Forskjellene på fødeland og landbakgrunn i mitt datasett vil for det meste gå på om individene er førstegenerasjons innvandrere eller om de er etterkommere, da etterkommerne vil ha fødeland i Norge, mens de som er førstegenerasjonsinnvandrere vil ha fødeland i det landet de har sin landbakgrunn. Det er imidlertid noen små forskjeller her, som tyder på at foreldrene til noen av innvandrerne i datasettet kan ha innvandret til et annet land, før de kom til Norge. Denne forskjellen er imidlertid svært liten, og jeg velger her å bruke kategorien "landbakgrunn".

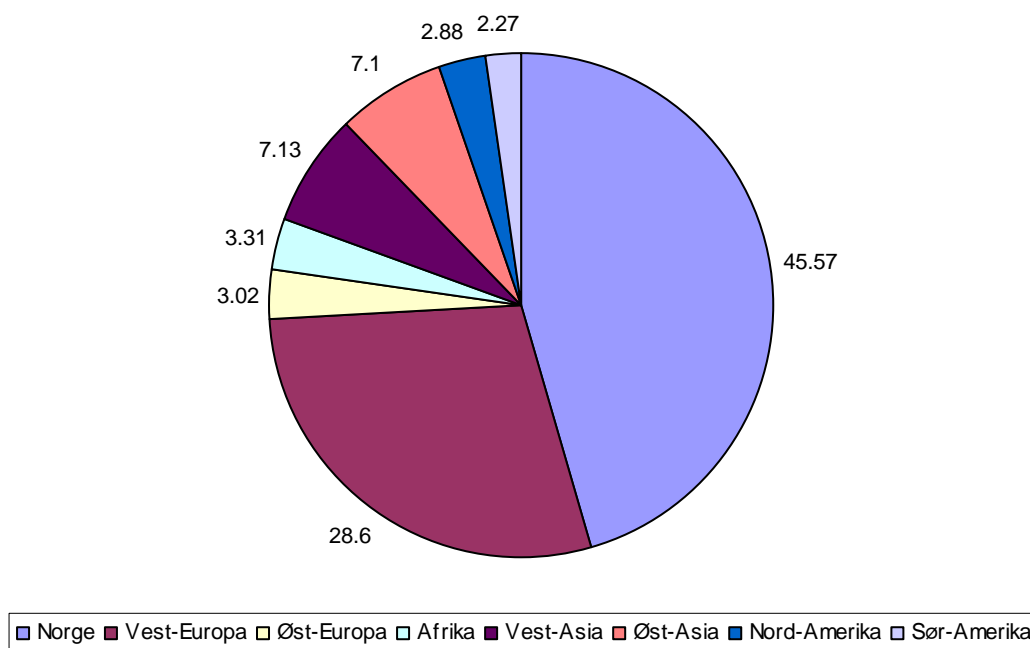
Figuren viser at omtrent 38 prosent av utvalget mitt har landbakgrunn fra Vest-Europa. Dette gir mening i og med at Vest-Europa geografisk sett ligger nært Norge. Vi ser imidlertid at det

⁹ Mellom-Amerika (med Mexico) og Karibia, Oseania og uoppgitt landbakgrunn er også oppgitt i datasettet, men på grunn av svært få observasjoner av disse har jeg valgt å ta disse bort.

Figur 2: Andel innvandrere fra hver region basert på landbakgrunn

i 1993 var mindre enn ti prosent av innvandrerne som hadde landbakgrunn i Øst-Europa, selv om Øst-Europa også har en nær geografisk beliggenhet til Norge. Innvandringen av personer med landbakgrunn fra Asia var derimot høy, mens innvandrere fra Nord-Amerika og Sør-Amerika kun utgjorde en andel på litt over åtte prosent til sammen.

Datasettet tar også med kategorien ”statsborgerskap”, som ikke nødvendigvis trenger å være lik de andre kategoriene. Dette fordi det ikke er uvanlig at innvandrere tar norsk statsborgerskap etter å ha bodd en viss tid i Norge. Figur 3 viser at nesten femti prosent av innvandrerne har norsk statsborgerskap, og at nesten 30 prosent av innvandrerne har Vesteuropeisk statsborgerskap. Variabelen statsborgerskap sier med andre lite om hvor personene kommer fra, men kan kanskje si noe om hvor godt integrert i det norske samfunnet innvandrerne er, eller hvor lenge de har vært i Norge.

Figur 3: Andel innvandrere i Norge basert på statsborgerskap

4.2 Variabelspesifikasjon

4.2.1 Den avhengige variabel

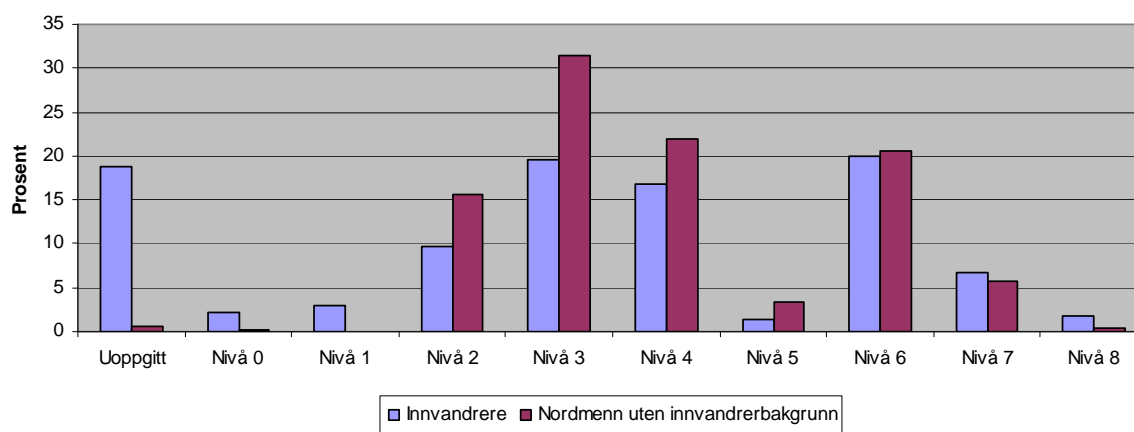
Den avhengige variabelen i denne oppgaven er *utdanning*, og er oppgitt i utdanningsnivåene fra NUS2000. Siden utdanningen ofte endres over tid i datasettet, på grunn av at noen av individene i utvalget tar mer utdanning i løpet av paneldataperioden, velger jeg i den deskriptive statistikken å bruke det maksimale utdanningsnivået individene i utvalget har mellom 1993 og 2003 som den avhengige variabelen.

Som nevnt i kapittel 2 inneholder NUS2000 ni ulike utdanningsnivåer. Nivå 0 er for de som har ingen utdanning eller kun førskoleutdanning. Nivå 1 og 2 er for henholdsvis barneskole og ungdomskole. Nivå 3, 4 og 5 er forskjellige nivåer på videregående skole, mens nivå 6 og 7, går på høyere utdanning, henholdsvis lavere og høyere grad. Nivå 8, som er det høyeste nivået er for de som tar enda høyere utdanning enn dette, som for eksempel doktorgradsutdanning.

Andelen med uoppgitt utdanning varierer i stor grad med landbakgrunn. Da det nesten er en komplett registrering av utdanningen til nordmenn, inneholder registreringen av de med landbakgrunn i andre verdensregioner mange med uoppgitt utdanning. Vi ser av figur 4 at ca.

16 prosent av alle innvandrerne i Norge er registrert med uoppgitt utdanningsnivå. Som vi så i avsnitt 2.4, er det ingen systematisk registrering av utdanningsnivået til innvandrere som kommer til Norge, noe som kan ha sammenheng med dette høye nivået av uoppgitt utdanning. Siden andelen uoppgitt ikke er definert etter utdanning, og mange innvandrere som har kommet de senere årene har høyere utdanning (Nygård og Daugstad, 2006), vil tallene som blir presentert her derfor være minimumstall. Gjennom kartleggingen av innvandrernes utdanningsnivå, fant imidlertid Dalheim (2001) ut at det var vanskeligst å få postale svar fra de med utdanning på grunnskolenivå. Jeg vil derfor i den deskriptive statistikken gå ut fra at det er en stor andel av de med uoppgitt utdanningsnivå som har en lav utdannelse, altså enten ingen utdanning, eller kun utdanning på grunnskolenivå.

Figur 4: Forskjeller i utdanningsnivå blant innvandrere og de uten innvandrerbakgrunn i Norge

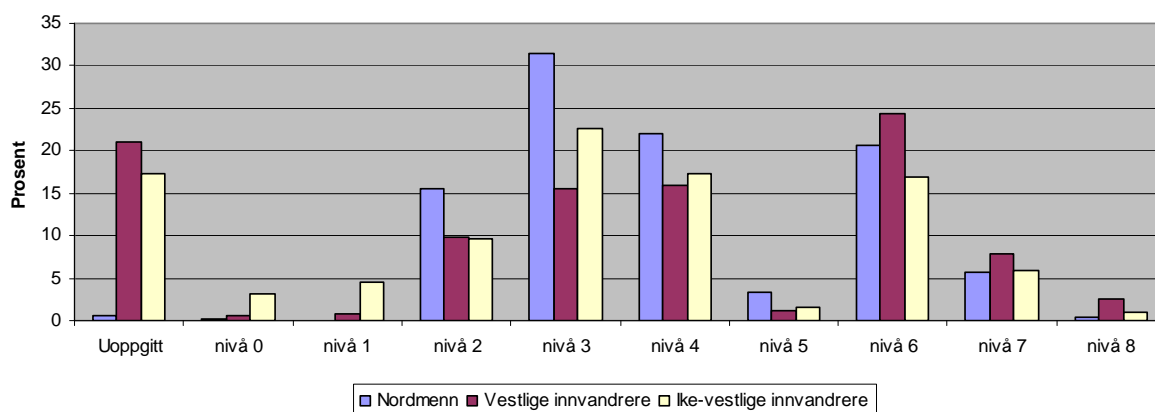


Når vi sammenligner innvandrere og nordmenn uten innvandrerbakgrunn under ett, ser vi at de med innvandrerbakgrunn har de største andelen på de laveste og de høyeste nivåene, mens de uten innvandrerbakgrunn topper statistikken på ungdomskolenivå og nivåene for videregående skole, prosentvis. Som vi ser av figur 4 har så å si alle nordmenn minst tatt ungdomskoleutdanning. Dette er et naturlig resultat, siden vi har obligatorisk grunnskoleopplæring i Norge til og med ungdomskolen. Alle som tar en grunnskoleutdanning i dag får dessuten bestått, uansett resultat. Kun 0,2 prosent nordmenn er registrert på ingen utdanning eller førskoleutdanning. Dette kan komme av at man før slutten av 1980-tallet ikke nødvendigvis fikk bestått grunnskoleopplæringen. De som ikke fikk bestått fram til da, og som heller ikke senere har bestått grunnskoleutdanningen, er derfor klassifisert i nivå 0 – ingen utdanning eller førskoleutdanning (se avsnitt 2.3.1). Det er likevel verdt å merke seg at ca. 19 prosent av alle nordmenn kun har fullført ungdomskolen.

Over 30 prosent av nordmennene er derimot klassifisert i nivå 3 – videregående grunnutdanning. Det vil si at de kun har tatt ett eller to år på videregående skole, men ikke fullført. En høy gruppe av innvandrerne er også plassert her, ca. 20 prosent. Vi ser imidlertid at innvandrerne har høyere statistikk enn nordmenn på alle nivåene av høyere utdanning.

Blant annet Østby (2004) skriver at det er store forskjeller i utdanningsnivået til vestlige og ikke-vestlige innvandrere. Forskjellene i utdanningsnivå for disse gruppene er dermed illustrert i figur 5. Vestlige land omfatter her landene i Vest-Europa og Nord-Amerika, mens de ikke-vestlige landene omfatter landene i Øst-Europa, Asia (inkludert Tyrkia), Afrika og Sør-Amerika.

Figur 5: Forskjeller i utdanningsnivå blant vestlige innvandrere, ikke-vestlige innvandrere og de uten innvandrerbakgrunn i Norge

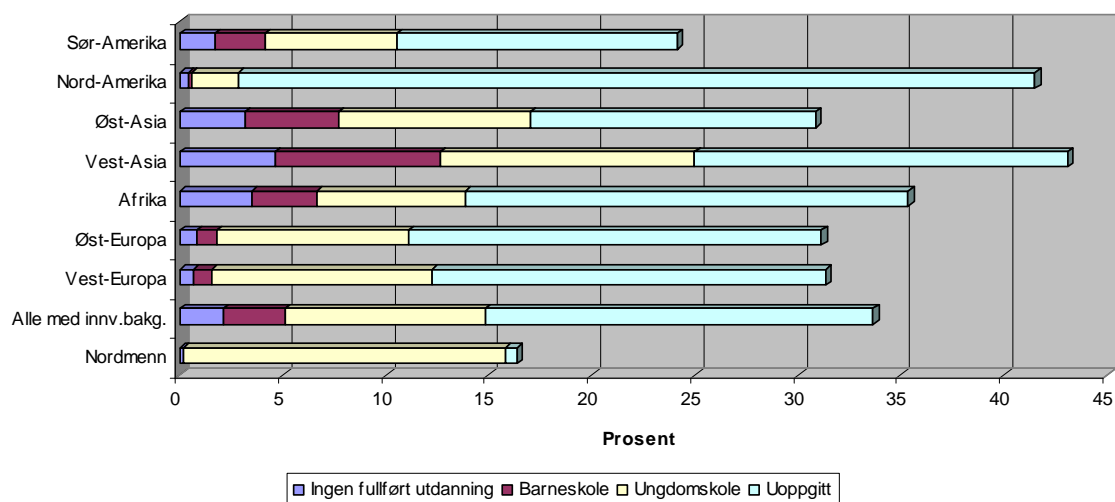


Av figuren ser vi at det er en markant forskjell på vestlige og ikke-vestlige innvandrere når det gjelder høyere utdanning. Mens vestlige innvandrere troner høyt over nordmenn når det gjelder høyere utdanning, har flere nordmenn utdanning på universitets- og høyskole, lavere grad, enn ikke-vestlige innvandrere. De ikke-vestlige innvandrerne har imidlertid litt flere på universitets- og høyskoleutdanning, høyt nivå, og forskerutdanning enn nordmenn.

Variablene landbakgrunn, kjønn, aldersfordeling og botid i Norge vil være av spesiell interesse her, da alle disse faktorene kan ha innvirkning på utdanningen man velger å ta. Jeg vil derfor se nærmere på disse, samt sette opp statistikker for antall barn og sivilstand, for å kontrollere for om disse kan ha en innvirkning på utdanningsnivået. Jeg begynner med en deskriptiv statistikk over landbakgrunn. Jeg vil her fokusere på hvor mange som har et lavt og

hvor mange har et høyt utdanningsnivå, og sammenligne landene på bakgrunn av dette. Med lavt utdanningsnivå legger jeg utdanning til og med grunnskolen til grunn. Jeg velger også å ta de med uoppgitt utdanningsnivå med her, da mange av disse trolig vil ha utdanning på dette nivået, som redegjort for tidligere i oppgaven. Med høyt utdanningsnivå mener jeg de med utdanning på universitets- og høyskolenivå eller over. Vi ser av figur 6 at de fra Vest-Asia har spesielt mange med lavt utdanningsnivå, her over 40 prosent.

Figur 6: Befolkningen med ingen utdanning, barneskole, ungdomskole eller med uoppgitt utdanningsnivå etter landbakgrunn

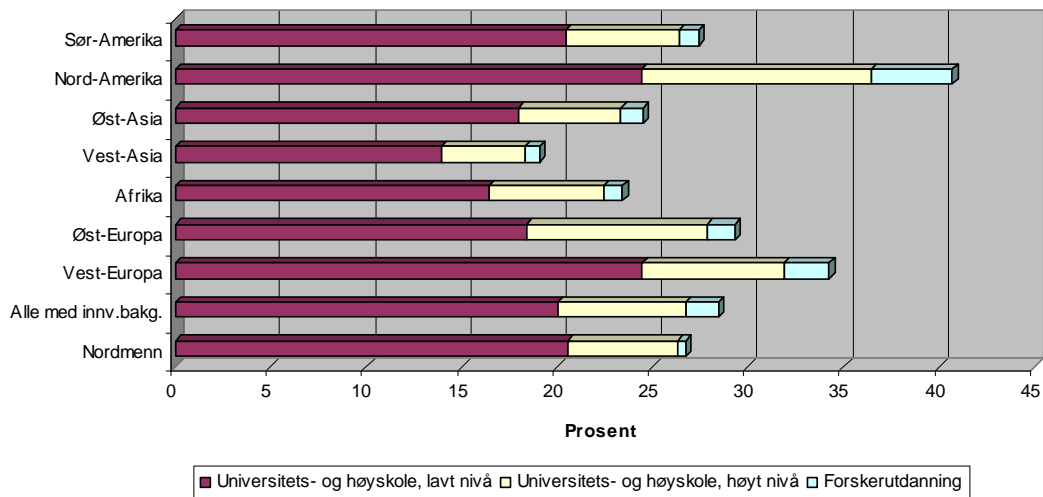


Nord-Amerika kommer også høyt opp her, men dette kan forklares ved at de har veldig høye tall på uoppgitt utdanning. Over 32 prosent av innvandrerne med landbakgrunn i Nord-Amerika har altså ikke registrert utdanningsnivå i Norge. Dette kan komme av for eksempel kort botid i Norge, og trenger dermed nødvendigvis ikke å bety at disse har et lavt utdanningsnivå. Videre kommer også Afrika og Øst-Asia høyt opp på denne statistikken, mens Sør-Amerika har overraskende få i denne kategorien. Nordmenn har her den laveste andelen av personer med lav utdanning, men dette må sees i sammenheng med at så å si alle nordmenn er oppført med registrert utdanning.

Av figur 7 ser vi at Nord-Amerika er på et helt eget nivå når det gjelder høyere utdanning, da 45 prosent av innvandrerne med landbakgrunn i Nord-Amerika har en eller annen form for høyere utdanning. På grunn av den høye andelen med uoppgitt utdanningsnivå for de med landbakgrunn i Nord-Amerika, må vi i tillegg huske at dette kun er minimumstall, da andelen med høyere utdannelse faktisk kan være høyere. Vest-Europa ligger på plass nummer to med

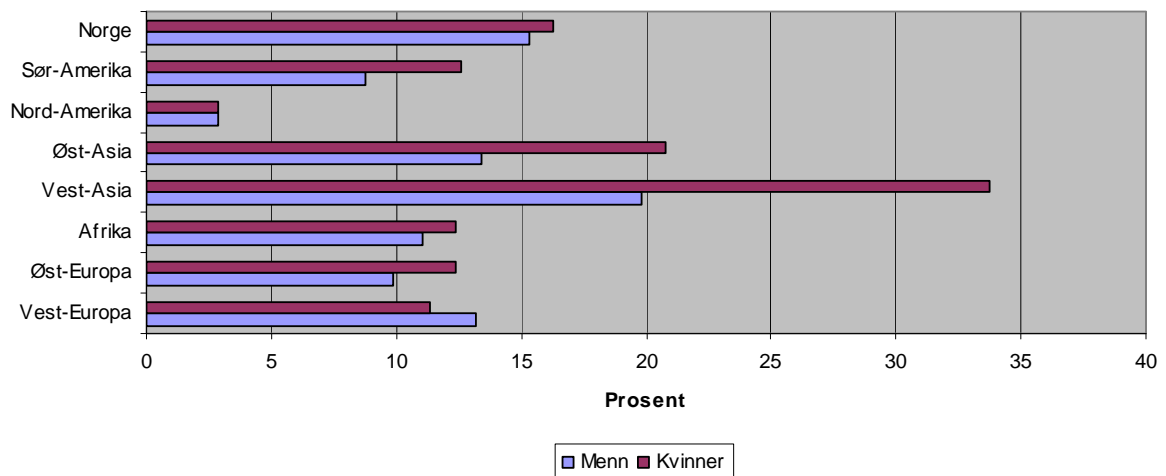
over 35 prosent med høyere utdanning. De med landbakgrunn i Vest-Asia er de som prosentvis tar minst høyere utdanning.

Figur 7: Befolkningen med utdanning på universitets- og høyskolenivå eller forskerutdanning



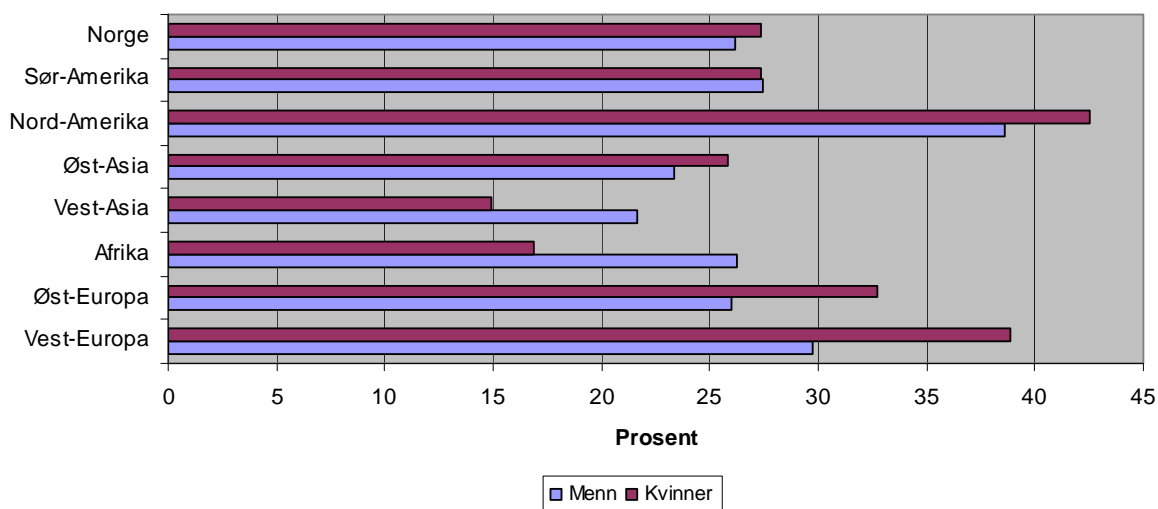
Jeg ønsker videre å se hvordan kjønns sammensetningen er når det gjelder utdanning fra de forskjellige verdensregionene. Av figur 8 kan vi se at for de ikke-vestlige regionene er det størst andel kvinner som enten ikke har utdanning eller med grunnskole som høyeste fullførte utdanning. Denne forskjellen er spesielt stor i de asiatiske landene og i Afrika. I de vestlige regionene er det derimot menn som utgjør den største andelen med utdanning på lavt nivå, selv om forskjellen er betydelig mindre her enn for de ikke-vestlige regionene.

Figur 8: Befolkningen uten utdanning eller med grunnskole som høyeste fullførte utdanning, etter landbakgrunn og kjønn

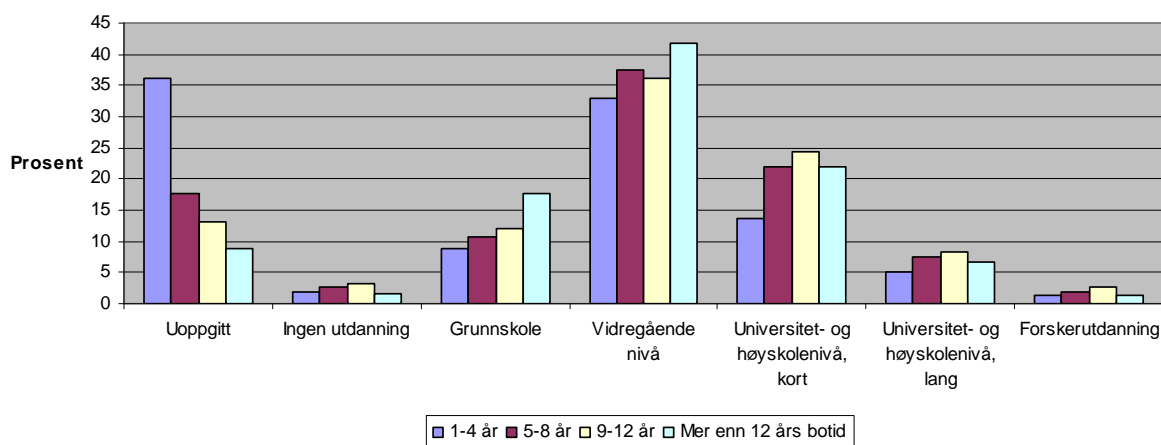


Når det gjelder høyere utdanning, kan vi lese av figur 9 at det i de fleste regionene er kvinner som i størst grad tar høyere utdanning. Dette ser vi spesielt i de vestlige regionene, men også i de ikke-vestlige regionene, Øst-Asia og Øst-Europa, er dette tilfelle. For innvandrere fra Afrika og Vest-Asia ser vi imidlertid at menn har en vesentlig større andel med høyere utdanning enn kvinner.

Figur 9: Befolkningen oppført med høyere utdanning, etter landbakgrunn og kjønn

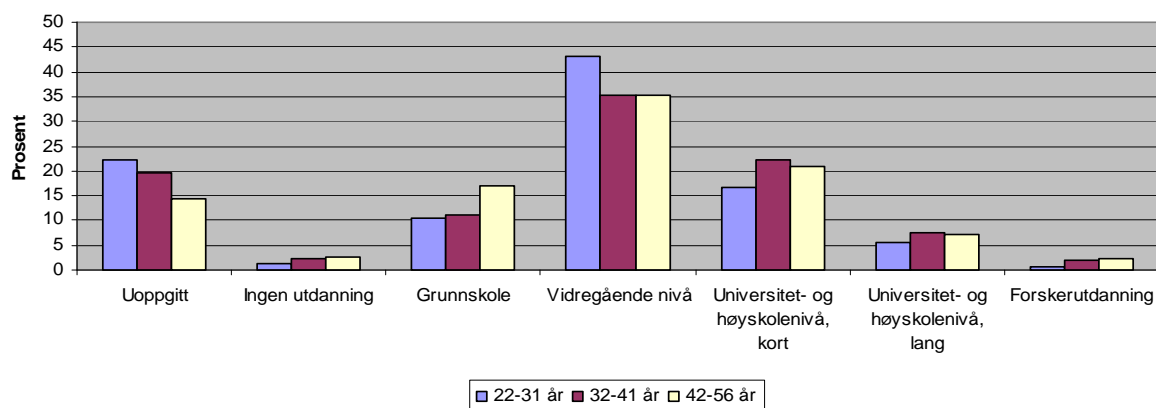


For å kontrollere for om botid har noen innvirkning på utdanningsnivået har jeg delt utvalget opp i grupper etter botiden i Norge. De tre første gruppene inneholder de med henholdsvis 1-4 års botid, 5-8 års botid og 9-12 års botid, mens den siste gruppen inneholder alle innvandrere med lenger botid i Norge enn 12 år. Botiden er registrert i 1993, slik at de som kom til Norge i 1992 er registrert med et års botid etc. Vi ser tydelig av figur 10 at andelen med uoppgitt utdanningsnivå synker jo lenger personen har bodd i Norge. Dette kan både komme av at innvanderne deltar i spesialundersøkelser om fullført utdanning etter at de kommer til Norge, og at de kommer i kontakt med utdanningssystemet i Norge. Høyere utdanning stiger med botid, bortsett fra for de som har mer enn 12 års botid i Norge, der nivået går litt ned igjen. Dette kan bety at mange faktisk tar utdannelse i Norge, men kan også komme av såkalte kohort effekter (se figur 1), altså at innvandrere som kommer på forskjellige tidspunkter, typisk har forskjellige kvalifikasjonsnivå. Den økte andelen utdanning for økt botid henger imidlertid utvilsomt sammen med den sterke nedgangen i uoppgitt utdanningsnivå, og siden andelen på alle utdanningsnivåene øker med botiden, kan kanskje mesteparten av denne økningen forklares ut fra det. Den siste spesialundersøkelsen om utdanningsnivå før 1993,

Figur 10: Innvandrernes utdanningsnivå basert på botid

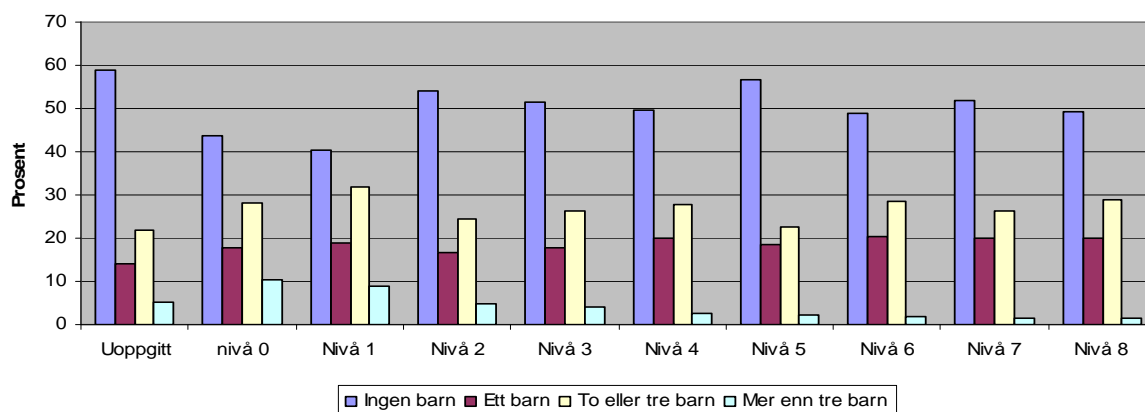
skjedde i sammenheng med folke- og bolig tellingen i 1990, noe som kan forklare den sterke nedgangen i antallet med uoppgitt utdanningsnivå for de som har bodd over 4 år i Norge.

Alderen er som botiden registrert i 1993, slik at de som var 22 i 1993 her er registrert som 22 år gjennom hele perioden. Vi ser av figur 11 at utdanningsnivået stiger med alderen, mens antallet med uregistrert utdanning synker med alderen. Dette er den samme trenden som vi ser i figuren for botid, og kan bety at det er en sammenheng mellom botid og alder, altså at de yngste gjerne har bodd kortest tid i Norge etc. Når det gjelder høyere utdanning ser vi at de i gruppen 32-41 år er de med den høyeste andelen med høyere utdanning. Innvandrerne i aldersgruppen 22-31 år er her den gruppen med minst andel høyere utdanning. Dette kan komme av at de er såpass unge, slik at ikke alle i denne gruppen er ferdigutdannet ennå. Vi ser også at den yngste gruppen er de som i minst grad er representert i gruppen for 'ingen utdanning' og på grunnskolenivå.

Figur 11: Innvandrerbefolkningens utdanningsnivå basert på aldersgrupper

I figur 12 vises det hvordan andelen barn hos innvandrerbefolkningen er fordelt på hvert utdanningsnivå. Jeg har her brukt fordelingen av personer med ingen barn, ett barn, to eller tre barn og mer enn tre barn, innenfor hver kategori, slik at søylene summerer seg til hundre for hvert utdanningsnivå.

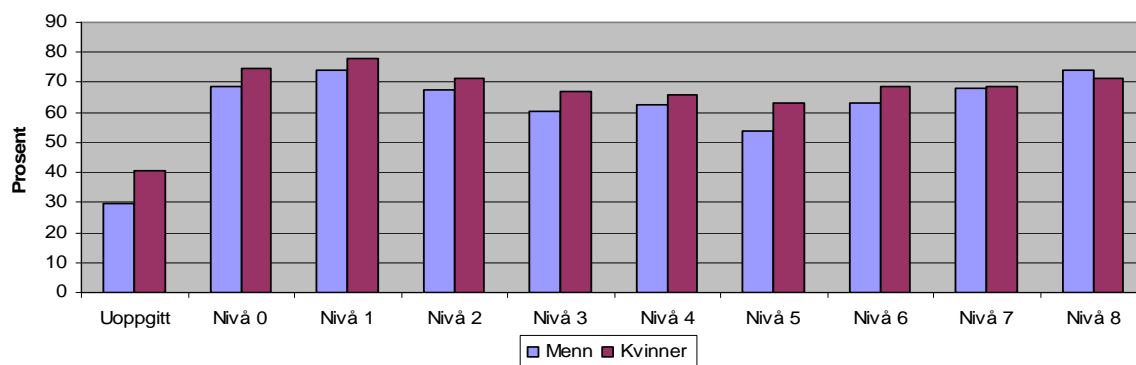
Figur 12: Andel barn på hvert utdanningsnivå hos innvandrerbefolkningen



For de som har fra null til tre barn kan vi egentlig ikke se noe mønster i variasjonen av antall barn over utdanningsnivå. Vi ser imidlertid at andelen med mer enn tre barn synker med utdanningsnivåene. Mens andelen personer som har mer enn tre barn er på over ti prosent for de uten utdanning, er den kun på litt over én prosent for de som har utdanning på doktorgradsnivå. Dette indikerer at antall barn har en negativ innvirkning på utdanningsnivået, men dette kun dersom antall barn overstiger et visst antall.

Vi ser av figur 13 at det er en overvekt av gifte personer i forhold til ugifte på alle utdanningsnivåene, både for menn og kvinner. Dette gjelder imidlertid ikke for de med uoppgitt utdanningsnivå.

Figur 13: Andel gifte på hvert utdanningsnivå hos innvandrerbefolkningen, etter kjønn



Vi ser at andelen gifte er høyest for de med ingen utdanning og kun grunnskole, samt for de med utdanning på doktorgradsnivå. De i utvalget mitt som i størst grad gifter seg, er altså de med svært lav eller svært høy utdannelse, men siden andelen gifte er såpass høy for alle utdanningsnivåene er det imidlertid vanskelig å se noe bestemt mønster her. Andelen gifte er desidert lavest for de med uoppgitt utdanningsnivå, men vi ser også en nedgang i andelen gifte for de på videregående nivå, sett i forhold til de andre utdanningsnivåene. Andelen gifte er også høyere for kvinner enn for menn på alle nivåene, bortsett fra nivå 7 og nivå 8. Dette kan indikere at gifte kvinner i mindre grad tar utdanning på de høyeste nivåene enn gifte menn.

4.2.2 Forklaringsvariabler

Etter å ha presentert den avhengige variabelen, og variabler som det kan tenkes at denne avhenger spesielt av, vil jeg gå over til å presenterer de forskjellige forklaringsvariablene hver for seg. Jeg vil her oppgi enkel statistikk, som gjennomsnitt, for de forskjellige variabelene.

Kjønn

Dersom vi ser på hele innvandrerbefolkningen under ett ser vi at den består av 54 prosent menn og 46 prosent kvinner. Det er altså omtrent åtte prosent flere menn enn kvinner i utvalget for innvandrere. Det er også flere menn i utvalget for nordmenn, men denne forskjellen er klart mindre påfallende, med 51 prosent menn mot 49 prosent kvinner.

Dersom vi ser på andelen menn og kvinner i utvalget på regionsnivå ser vi at det i all hovedsak er to regioner som trekker opp mannsandelen i utvalget for innvandrere. Både Afrika og Vest Asia (med Tyrkia) har stor overvekt av menn i utvalget, med henholdsvis prosentandeler på 70 og 63. Vi ser at Nord-Amerika er den eneste av regionene der det er hovedvekt av kvinner, med en andel kvinner på 54 prosent.

Alder

På grunn av anonymitetshensyn får jeg ikke tilgang på direkte alder til individene i utvalget mitt, kun aldersgrupper med femårsintervaller. Den yngste aldersgruppen jeg opererer med her var fra 22-26 år i 1993, mens den eldste var fra 52-56 år i 1993. Alle i disse

aldersgruppene vil mest sannsynlig ha fullført sin utdanning innen 2003, dersom vi ser bort fra eventuell etterutdanning. Når jeg her regner ut gjennomsnittsalder er den nok av denne årsaken ikke helt nøyaktig, men gir oss likevel et relativt godt bilde på hvordan aldersgrupperingene er for de forskjellige regionene.

Når det gjelder innvandrerbefolkningen sett under ett, har de en gjennomsnittsalder på litt over 37 år. Ser av tabell 2 at de vestlige regionene trekker opp her, da snittalderen fra Vest-Europa er på hele 40,4 år, mens snittalderen fra Nord-Amerika er på 37,8 år. Den regionen med lavest snittalder er Afrika, med en snittalder på 33,6 år.

Tabell 2: Gjennomsnittsalderen i de forskjellige regionene

Regioner	Vest-Europa	Øst-Europa	Afrika	Vest-Asia	Øst-Asia	Nord-Amerika	Sør-Amerika	Nordmenn
Snittalder (1993)	40,1 år	36,7 år	33,6 år	34,7 år	34,4 år	37,8 år	36,1 år	38 år

Botid

Når det gjelder botid er det også forskjell blant de vestlige og de ikke-vestlige regionene. Vi ser at begge de vestlige regionene har vesentlig lengre botid enn de andre. Tabell 3 viser at dette gjelder spesielt for Vest-Europa, med en gjennomsnittlig botid i Norge på 14,2 år, mens den gjennomsnittlige botiden for innvandrere med landbakgrunn i Nord-Amerika er på 10,7 år. Den korteste botiden finner vi hos de med landbakgrunn i Afrika, da disse kun har en botid på 7,8 år. Det ser ut som det kan være en korrelasjon mellom alder og botid her, da de med den høyeste alderen har den lengste botiden, og omvendt.

Tabell 3: Gjennomsnittlig botid i Norge for de forskjellige regionene.

Regioner	Vest-Europa	Øst-Europa	Afrika	Vest-Asia	Øst-Asia	Nord-Amerika	Sør-Amerika
Botid (1993)	14,2 år	9,4 år	7,8 år	9,5 år	8 år	10,7 år	8,2 år

Familietype

Ser av tabell 4 at nordmenn er de som i størst grad er gift uten å ha barn, da dette gjelder hele 18 prosent av den norske befolkningen uten innvandrerbakgrunn.

Tabell 4: Familietype for de forskjellige regionene, oppgitt i prosent.

<i>Region/Familietype prosent</i>	<i>Ektepar uten barn</i>	<i>Ektepar med barn</i>	<i>Samboere med barn</i>	<i>Andre med flere</i>
Nordmenn	18	43,2	6,1	32,6
Innvandrerbefolkningen	12,2	50,4	3,7	33,7
Vest-Europa	15,8	44,6	5,4	34,6
Øst-Europa	13,4	51,5	3	34,2
Afrika	8,3	40,2	3,1	48,4
Vest-Asia	8,2	61,3	1,8	28,7
Øst-Asia	10,6	57,1	2,6	29,7
Nord-Amerika	14,9	50,2	2,2	32,7
Sør-Amerika	7,7	41,5	6,2	44,5

6,1 prosent av nordmennene er også samboerpar med barn. Dette er også flere enn i alle de andre verdensregionene, med unntak av de med landbakgrunn fra Sør-Amerika, der tallet er omtrent like stort. Vi ser at de med innvandrerbakgrunn fra de vestlige regionene i større grad er i ekteskap uten barn enn de med innvandrerbakgrunn fra de ikke-vestlige regionene. De med landbakgrunn i Vest-Asia er de som har den største andelen ekteskap med barn.

Sivilstand

For sivilstand får vi i datasettet oppgitt om de i utvalget har status som ugift, gift, enke/enkemann, skilt eller separert. I tillegg får vi også oppgitt om man har status som registrert partner, separert partner, skilt partner eller gjenlevende partner. Det er imidlertid svært få observasjoner i utvalget mitt over de fire siste kategoriene, noe som kan komme av at ordningen med registrerte partnerskap først trådte i kraft fra 1. august 1993 (Barne- og likestillingsdepartementet, 2006). Jeg velger derfor å fjerne disse observasjonene fra statistikken, og kun fokusere på de fem første kategoriene av sivilstand.

Tabell 5 viser at de uten innvandrerbakgrunn er de som i størst grad er ugift, med hele 31,6 prosent. Dette tallet skiller seg markant ut fra de andre regionene, da Sør-Amerika, med nest størst andel her, kun har en andel av befolkningen på 20,3 prosent som er ugift. Når det gjelder sivilstand er det ellers ikke noe klart mønster for de forskjellige regionene, og vi ser ikke særlig forskjell på de vestlige og de ikke-vestlige regionene. Det er noen flere enker/enkemenn av de uten innvandrerbakgrunn enn de fra de andre regionene. Vi ser at både skilsmisse og separasjonsprosenten er størst blant de fra Sør-Amerika, og at de uten innvandrergrunn faktisk har en lav prosent her.

Tabell 5: Sivilstand for de forskjellige regionene, i prosent.

<i>Region/Sivilstand prosent</i>	<i>Ugift</i>	<i>Gift</i>	<i>Enke/Enkemann</i>	<i>Skilt</i>	<i>Separert</i>
Nordmenn	31,6	54,1	3,4	8,9	2
Innvandrerbefolkningen	17,3	65,6	1,5	11,1	4,5
Vest-Europa	19,3	62,1	1,8	13,6	3,3
Øst-Europa	12,3	67,6	1,6	13,1	5,4
Afrika	19,9	55,6	1	14,9	8,6
Vest-Asia	14,2	73,7	1,3	6,3	4,5
Øst-Asia	17,3	70,6	1,5	7	3,6
Nord-Amerika	16,7	67,1	1,3	11,5	3,4
Sør-Amerika	20,3	54	1,2	15	9,4

Antall barn under 16 år

Tabell 6 viser hvor stor andel av befolkningen som har barn under 16 år. Vi ser at nordmenn uten innvandrerbakgrunn er den største andelen av de som ikke har noen barn under 16 år. Nordmenn er også de som i minst grad har flere enn tre barn over 16 år. De vestlige regionene, Vest-Europa og Nord-Amerika, skiller seg ikke vesentlig ut fra de uten innvandrerbakgrunn her, men vi ser store forskjeller når det gjelder de ikke-vestlige regionene. Mens kun 2,51 prosent av de med landbakgrunn i Sør-Amerika har fire barn under 16 år eller mer, gjelder dette 9,77 prosent av afrikanerne, samt hele 12,38 prosent av vestasiatene.

Tabell 6: Antall barn under 16 år for de forskjellige regionene, i prosent.

<i>Region/ Antall barn</i>	<i>Ingen barn</i>	<i>Ett barn</i>	<i>To barn</i>	<i>Tre barn</i>	<i>Fire barn eller flere</i>
Nordmenn	58,51	17,37	16,89	6,1	1,13
Innvandrerbefolkningen	38,15	21,75	24,06	10,94	5,11
Vest-Europa	51,01	18,42	21,1	7,68	1,78
Øst-Europa	39,68	25	22,71	9,06	3,55
Afrika	31,9	20,88	23,47	13,99	9,77
Vest-Asia	25,79	20,97	24,73	16,13	12,38
Øst-Asia	30,16	26,65	28,22	11,33	3,63
Nord-Amerika	40,43	19,85	26,42	10,43	2,87
Sør-Amerika	35,51	23,79	27,13	11,07	2,51

Fylke

Vi ser av tabell 7 at de uten innvandrerbakgrunn i utvalget er rimelig godt fordelt over fylkene i landet. Det er selvfølgelig flere som er bosatt i fylker med stort innbyggertall, slik som Oslo, Akershus og Hordaland, enn i fylker med lavt innbyggertall, slik som Finnmark, noe som gjør at utvalget fordeler seg ganske likt befolkningen i Norge. Blant innvandrerbefolkningen ser vi imidlertid at et svært høyt antall er bosatt i Oslo, da hele 35,5 prosent av innvandrerbefolkningen totalt holder til her, mot 10,1 prosent av de uten innvandrerbakgrunn.

Hele 48,1 prosent av innvandrerne er bosatt i Stor-Oslo, som tilsvarer Oslo og Akershus. I tillegg er over syv prosent bosatt i både Rogaland og Hordaland, der vi kan anta at de fleste bor i de store byene Stavanger og Bergen. Det er imidlertid få med innvandrerbakgrunn (kun 3,3 prosent) som er bosatt i Sør-Trøndelag til tross for at Norges tredje største by, Trondheim, ligger her. Likevel tyder tallene på at de fleste med innvandrerbakgrunn bosetter seg i større byer/ tettbygde strøk.

Tabell 7: Andel av befolkningen i hvert fylke, basert på regioner.

Fylke/Region prosent	Nordmenn	Innvandrerbefolkningen	Vestlige	Ikke-vestlige
Østfold	5,6	5,1	5,7	4,7
Akershus	10,2	12,7	16,5	10,2
Oslo	10,1	35,4	22,1	44,1
Hedmark	4,4	2,3	3,1	1,7
Oppland	4,4	1,9	2,4	1,5
Buskerud	5,3	5,6	5,7	5,5
Vestfold	4,8	3,5	4,8	2,7
Telemark	3,8	2,5	3	2,2
Aust-Agder	2,2	1,5	2,2	1
Vest-Agder	3,3	3,2	3,4	3
Rogaland	8	7,5	9,2	6,4
Hordaland	9,7	7,1	6,9	7,1
Sogn og Fjordane	2,5	1,1	1,1	1
Møre og Romsdal	5,6	2,1	2,2	2
Sør-Trøndelag	6,2	3,3	3,3	3,3
Nord-Trøndelag	3	0,9	1,2	0,6
Nordland	5,7	1,8	2,6	1,3
Troms	3,6	1,6	2,5	1,1
Finnmark	1,8	1,2	2	0,7

Vi kan imidlertid se en relativt stor forskjell når det gjelder bosettingen til innvandrere med vestlig landbakgrunn og innvandrere med ikke-vestlig landbakgrunn. Selv om også de vestlige innvandrerne i større grad bor i Oslo enn de uten innvandrerbakgrunn, er de i mye mindre grad sentralisert her enn de ikke-vestlige innvandrerne. Da 44,3 prosent av de ikke-vestlige innvandrerne holder til i hovedstaden, er det kun 22 prosent av de vestlige innvandrerne som bor her. Vi ser imidlertid at flere av de vestlige innvandrerne holder til i Akershus enn av de ikke-vestlige innvandrerne, og i Stor-Oslo totalt har vi da 54,4 prosent av innvandrerne med ikke-vestlig landbakgrunn og 38,4 prosent av innvandrerne med vestlig landbakgrunn.

Tettbygde strøk

Tettbygde strøk blir av FD-trygd definert som en plass der det bor minst 200 personer, og avstanden mellom husene ikke overstiger 50 meter. Tabell 8 bekrefter det vi fant i tabell 7, om at innvandrerne i større grad holder til i tettbygde strøk enn nordmenn. Mens kun 74,4 prosent av alle uten innvandrerbakgrunn holder til i tettbygde strøk, holder heler 91,8 prosent av innvandrerbefolkningen til i disse områdene. Tallet blir enda høyere når vi ser på de ikke-vestlige regionene hver for seg, da alle disse i svært høy grad holder til i tettbygde strøk. Størst andel har Vest-Asia, da hele 98,3 prosent av innvandrerne med landbakgrunn i Vest-Asia holder til i tettbygde strøk.

Tabell 8: Andel av befolkningen i tettbygde/ spredtbygde strøk, basert på regioner.

Region/type bosted i prosent	Tettbygd strøk	Spredtbygd strøk
Nordmenn	74,4	25
Innvandrerbefolkningen	91,7	7,9
Vest-Europa	85,5	14
Øst-Europa	92,7	6,9
Afrika	97	2,7
Vest-Asia	98,3	1,5
Øst-Asia	94,8	4,8
Nord-Amerika	86,5	13,1
Sør-Amerika	94,1	5,5

Etterkommere

Etterkommere defineres av SSB som personer født i Norge av to foreldre som har fødeland i et annet land enn Norge. Disse har med andre ord bodd i Norge hele sitt liv, men på grunn av at foreldrene ikke er norske, kan de møte på vanskeligheter som barn med norske foreldre ikke har. Dette kan for eksempel være at foreldrene kanskje ikke er så godt kjent med det norske språket, og derfor ikke kan hjelpe dem med leksene. Det gir derfor mening å skille ut etterkommerne som en egen gruppe, da det kan sies at de ofte kan ha en ulempe utdanningsmessig i forhold til de uten innvandrerbakgrunn, samtidig som de vil ha en fordel i forhold til førstegenerasjonsinnvandrerne. På grunn av få observasjoner av etterkommere fant jeg det lite hensiktsmessig å kjøre en egen regresjon på disse. Siden jeg synes dette er en viktig variabel, velger jeg likevel å ta dummyvariabelen *etterkommere* med i datasettet som en av forklaringsvariablene.

Ser av tabell 9 at av de personene i utvalget som karakteriseres som etterkommere, har hele 66,7 prosent landbakgrunn fra Vest-Europa. Videre har 24,3 prosent landbakgrunn fra Øst-Europa. Omtrent 91 prosent av etterkommerne har altså landbakgrunn fra et land i Europa. Siden 5,2 prosent av etterkommerne også har landbakgrunn fra Nord-Amerika, forteller dette oss at det er svært få etterkommere med landbakgrunn fra Asia, Sør-Amerika og Afrika. Dette kommer nok av at innvandringen til Norge fra Europa går lenger tilbake i tid enn fra de andre verdensregionene, og at det derfor omtrent bare er disse som hadde rukket å få voksne barn i 1993.

Tabell 9: Andel av etterkommere i utvalget som kommer fra de forskjellige regionene.

<i>Verdens-region</i>	<i>Vest-Europa</i>	<i>Øst-Europa</i>	<i>Afrika</i>	<i>Vest-Asia</i>	<i>Øst-Asia</i>	<i>Nord-Amerika</i>	<i>Sør-Amerika</i>
Andel	66,7	24,3	0,48	1,35	1,45	5,2	0,53

Norsk statsborgerskap

Ser av tabell 10 at det er svært få av de som kommer fra de vestlige verdensregionene som har tatt norsk statsborgerskap, bare 20,3 prosent av de med landbakgrunn fra Vest-Europa og kun 9,7 prosent av de med landbakgrunn fra Nord-Amerika. For de ikke-vestlige landene er denne andelen betydelig større, da over femti prosent av de med landbakgrunn fra Øst-Europa og de asiatiske landene, samt nesten femti prosent av de med afrikansk landbakgrunn, har tatt norsk statsborgerskap.

Tabell 10: Andel av innvandrerbefolkningen med norsk statsborgerskap.

<i>Andel/verdensregion</i>	<i>Vest-Europa</i>	<i>Øst-Europa</i>	<i>Afrika</i>	<i>Vest-Asia</i>	<i>Øst-Asia</i>	<i>Nord-Amerika</i>	<i>Sør-Amerika</i>
Utenlandsk statsborger	79,7	43,6	50,4	46,7	44,5	90,3	58,9
Norsk statsborger	20,3	56,4	49,6	53,3	55,5	9,7	41,1

5 Empirisk modellering

I dette kapittelet skal jeg gjennomgå de empiriske modellene jeg skal bruke i oppgaven. I kapittel 5.1 går jeg gjennom estimering med minste kvadrats metode som et utgangspunkt til oppgaven. Jeg har så, i kapittel 5.2, en kort presentasjon av ulike paneldata modeller, for så å velge den paneldata modellen som vil være best å bruke i min oppgave. I kapittel 5.3 er en ordnet probit modell forklart, og jeg avslutter kapittelet med en gjennomgang av økonometriske problemer. Teorien i dette kapittelet er i hovedsak hentet fra Veerbeek (2004) og Wooldridge (2006), der ikke annet er referert til.

5.1 Minste kvadrats metode (MKM)

En av grunnsteinene i økonometrisk analyse er den såkalte lineære regresjonsmodellen, og den vanlige minste kvadrats metode (MKM) estimeringsmetoden. MKM estimerer koeffisienter ved å minimere summen av de kvadrerte avvikene. Vi går her ut fra populasjonsmodellen:

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \\ y_i &= x_i' \beta + \varepsilon_i \end{aligned} \quad (13)$$

Der y_i og x_i er observerbare variabler og ε_i er et uobserverbart feilledd. Elementene i β er ukjente populasjonsparametre. Sammenhengen i (13) skal gjelde for alle mulige observasjoner, selv om vi kun har tilgang på et visst antall observasjoner, N .

MKM-estimatoren er gitt ved:

$$\hat{\beta} = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})(x_i - \bar{x})'} \quad (14)$$

For å estimere sammenhengen mellom den avhengige variabelen og forklaringsvariablene må fem vilkår, eller antakelser, være oppfylt. MKM bygger altså på de klassiske Gaus-Markov vilkårene (Verbeek, 2004):

- (A1) $E(\varepsilon_i) = 0, \quad i = 1, \dots, N$
- (A2) $\{\varepsilon_1, \dots, \varepsilon_N\}$ og $\{x_1, \dots, x_N\}$ er uavhengige
- (A3) $Var(\varepsilon_i) = \sigma^2, \quad i = 1, \dots, N$

$$(A4) \quad \text{cov}(\varepsilon_i, \varepsilon_j) = 0, \quad i, j = 1, \dots, N \quad i \neq j$$

$$(A5) \quad \varepsilon \sim N(0, \sigma^2 I_N)$$

Vilkår (A1) forteller oss her at den forventede verdien til feilleddet er lik null. Dette betyr at dersom vi ser på *gjennomsnittet* av alle observasjoner, skal regresjonslinjen bli korrekt spesifisert. (A2) impliserer at x_i og ε_i er uavhengige av hverandre for alle individ i . Dette innebærer at vi har forklaringsvariabler som er eksogene, og at vi dermed har fravær av endogenitet. Vilkår (A3) sier at variansen til feilleddet skal være lik for alle verdier av x , noe som gjør at vi får et *homoskedastisk* feilledd, altså et feilledd som ikke varierer over tid. Den nest siste antakelsen, (A4), forteller oss at ulike feilledd ikke skal være korrelert med hverandre. Dette ekskluderer enhver form for autokorrelasjon.

Dersom disse vilkårene er oppfylt vil MKM- estimatoren være BLUE (Best linear biased estimator), og dermed være effisient. Det at en estimator er effisient betyr at det er den estimatoren som gir minst varians blant forventningsrette estimatorene, og dermed også de mest korrekte koeffisientestimaterne. MKM-estimatoren er altså den beste estimatoren å bruke så lenge disse Gauss-Markov forutsetningene er oppfylt. Den siste forutsetningen er om fordelingen til feilleddet, og forteller oss at residualene skal være normalfordelt. Dette er en viktig forutsetning når vi skal avgjøre om koeffisientestimaterne er statistisk signifikante i små utvalg.

Dersom noen av disse vilkårene ikke er oppfylt vil MKM-estimatoren gi forventningsskjeve og inkonsistente koeffisientestimater. Det at estimater er forventningsskjeve vil si at $E(\beta_1) \neq \beta_1$, med andre ord at forventningen til den estimerte koeffisienten er ulik den sanne verdien. Denne er viktig dersom utvalget vårt består av et lite antall enheter, og den blir da beregnet ved hjelp av gjennomsnittsverdien. Dersom vi har større utvalg, er det viktigste at estimatoren vår er konsistent, noe som vil si at sannsynligheten for at avstanden mellom den estimerte koeffisienten vår og den sanne verdien er stor, skal være lik null. Når utvalget blir større tilpasser estimatene seg altså den sanne fordelingen. Jeg kommer tilbake med eksempler på dette i avsnitt 5.4.

5.2 Paneldata

Et paneldatasett følger et gitt utvalg av økonomiske aktører over et visst tidsrom. Vi har altså flere observasjoner av hver aktør, noe som gjør paneldata til en kombinasjon av tverrsnitt og tidsserier. Selv om paneldata typisk blir samlet inn på et mikroøkonomisk nivå, har det blitt mer og mer praksis å sette sammen disse paneldata observasjonene for et gitt antall land, og analysere dem simultant. Mulighetene til å analysere gjentatte observasjoner for samme individ/enhet gjør det mulig for økonomer å spesifisere og estimere mer kompliserte og realistiske modeller enn hva som er mulig med vanlige tverrsnitt eller tidsserier (Verbeek, 2004). Andre positive sider ved å bruke paneldata er først og fremst at et økt antall observasjoner vil føre til mer presise estimater. For det andre gir et paneldatasett anledning til å identifisere for eksempel endringer på individnivå, og å se spuriøse sammenhenger over tid. Paneldata gjør det også mulig å kontrollere for individuell heterogenitet, som kan komme av utelatte individuelle variabler som er konstante over tid (evner, motivasjon osv). Dersom en overser denne heterogeniteten vil en få forventningsskjevne og inkonsistente koeffisientestimater.

Paneldata skiller seg notasjonsmessig noe fra vanlig minste kvadrats metode. Dette fordi vi her inkluderer individspesifikk heterogenitet i modellene. Denne inkluderes som et ekstra feilledd i disse modellene, α_i :

$$y_i = \beta_0 + x_i \beta + \alpha_i + \varepsilon_i \quad (15)$$

5.2.1 Regresjonsmodeller for paneldata

Det er fire forskjellige regresjonsmodeller som er vanlig å bruke i paneldata analyser. Her skal jeg kort presentere de fire modellene; Vanlig minste kvadrats metode på sammenslått panel, regresjon basert på gruppegjennomsnitt, fast effekt modellen og tilfeldig effekt modellen. Deretter vil jeg vurdere hvilken modell som vil være best å bruke i min analyse.

Dersom det ikke eksisterer individspesifikke komponenter, er *minste kvadrats metode (MKM) på sammenslått panel* den mest effisiente måten å estimere panelparametre på. Dette fordi en mister presisjon ved å estimere de individspesifikke parametrene dersom dette ikke er nødvendig. MKM estimatoren benytter både within- og between-dimensjonene i datasettet, altså både variasjonen over tid og variasjonen mellom individer. Den gjør imidlertid ikke

dette på en effektiv måte, og den tar ikke hensyn til panelstrukturen i data. De tre andre paneldata modellene som presenteres videre er alle mer avanserte versjoner av minste kvadrats metode.

Dersom vi skal utføre en regresjon basert på gruppegjennomsnitt vil vi kun utnytte variasjonen *mellom* individene, og ikke variasjonen *hos* individene over tid. En *between effekt estimator (BE)* vil da være den rette å bruke. Vi fokuserer altså kun her på individdimensjonen, og ser helt bort fra tidsdimensjonen. Vi finner BE-estimatoren ved å kjøre en vanlig MKM-regresjon på individuelle gjennomsnitt av y på individuelle gjennomsnitt av x . BE-estimatoren vil være konsistent dersom $E\{x_i\alpha_i\}=0$ og $E\{x_i\varepsilon_i\}=0$. Det vil si at de uavhengige variablene må være eksogene og ukorrelet med den individspesifikke variabelen, α_i .

Fast effekt modellen (FE) kontrollerer for uobservert heterogenitet, og fjerner derfor forventningsskjevheten som oppstår dersom vi estimerer paneldata uten å ta hensyn til den uobserverte heterogeniteten. Modellen gjør dette ved å behandle den individspesifikke heterogeniteten som en konstant. Denne estimatoren utnytter within-variasjonen i data, altså forskjellen over tid hos individene. Ved bruk av FE utfører vi en innen-gruppe transformasjon på variablene i populasjonsmodellen, slik at de individspesifikke komponentene ($\alpha_1, \alpha_2, \dots, \alpha_N$) blir fjernet fra modellen så lenge de er konstante over tid. Dette medfører at FE-estimatoren ikke kan brukes til estimering dersom nøkkelvariabelen i modellen er tidskonstant, da denne blir fjernet fra modellen. FE-estimatoren er forventningsrett og konsistent så lenge vi har uavhengighet mellom regressor og feilledd, altså $E\{x_{it}\varepsilon_{it}\}=0$. FE-estimatoren legger imidlertid ingen restriksjoner på forholdet mellom x_{it} og α_i , siden α_i fjernes via FE-transformasjonen.

Tilfeldig effekt modellen (RE) benytter variasjonen både *mellom* og *innad* individene på den mest effisiente måten. Tilfeldig effekt estimatoren kan derfor sees på som et vektet gjennomsnitt av BE- og FE-estimatorene. Estimatoren er forventningsrett og konsistent dersom $E\{x_{it}\varepsilon_{it}\}=0$ og $E\{x_{it}\alpha_{it}\}=0$, altså kun dersom uavhengighet mellom regressor, og feilledd og regressor og individspesifikk komponent, og er dermed mindre robust enn FE-estimatoren.

5.2.2 Valg av modell

Alle de omtalte modellene i foregående avsnitt har sine fordeler og ulemper. Det avgjørende for valget av modell er sannsynligheten for om det finnes uobservert individuell heterogenitet som en bør ta hensyn til, og om det er uavhengighet mellom den individuelle heterogeniteten og forklaringsvariablene.

Siden fast effekt estimatoren tillater korrelasjon mellom den individspesifikke komponenten, α_i , og forklaringsvariablene, x_{it} , kan den være bedre å bruke enn tilfeldig effekt estimatoren i sammenhenger der vi har mistanke om at det finnes uobservert heterogenitet. Tilfeldig effekt (eller MKM på sammenslått panel) vil imidlertid være bedre enn fast effekt estimatoren i tilfeller der nøkkelvariabelen er konstant over tid. I oppgaven min er blant annet variabelen *landbakgrunn* en sentral variabel som er konstant over tid, så fast effekt estimatoren vil dermed ikke være en bra estimator å bruke her. Tilfeldig effekt modellen er i det store og det hele foretrukket framfor minste kvadrats metode på sammenslått panel, da tilfeldig effekt estimatoren generelt er mer effisient (Wooldrige, 2006). Jeg vil derfor bruke tilfeldig effekt modellen videre, og presenterer denne litt nærmere i avsnittet under.

5.2.3 Tilfeldig effekt modellen (RE)

Uobserverte individspesifikke forhold blir i tilfeldig effekt modellen betraktet som stokastiske komponenter, og inkludert som en del av det stokastiske feilleddet i modellen:

$$y_{it} = \mu + \beta x_{it} + \overbrace{\alpha_i + \varepsilon_{it}}^{w_{it}}, \quad (16)$$

hvor μ er et konstantledd som ikke varierer over tid og $w_{it} = (\alpha_i + \varepsilon_{it})$ er det stokastiske feilleddet. MKM som estimator på denne modellen vil gi autokorrelert feilledd, w_{it} , siden den individspesifikke komponenten gir en kovarians ulik null. Autokorrelasjon representerer brudd på de klassiske forutsetningene for at MKM skal være BLUE. MKM vil altså ikke være effisient. For å få en effisient estimator bruker vi generaliserte minste kvadrats metode (GMKM), og transformerer den opprinnelige regresjonslikningen gjennom å vekte komponentene. Det kan i dette tilfellet vises at den nødvendige transformasjonen er:

$$(y_{it} - \mathcal{G}y_i) = \mu(I - \mathcal{G}) + \beta(x_{it} - \mathcal{G}x_i) + (w_{it} - \mathcal{G}w_i), \text{ der}$$

$$\rho = 1 - \sqrt{\psi} = 1 - \frac{\sigma_\varepsilon}{\sqrt{\sigma_\varepsilon^2 + \tau\sigma_\alpha^2}}$$

MKM på transformerte data vil så gi oss RE-estimatoren:

$$\hat{\beta}_{RE} = \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) + \psi T \sum_{i=1}^N (\bar{x}_i - \bar{\bar{x}})(\bar{y}_i - \bar{\bar{y}})}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 + \psi \sum_{i=1}^N (\bar{x}_i - \bar{\bar{x}})^2} \quad (17)$$

Vi ser at RE-estimatoren utnytter både variasjonen innad grupper (første ledd) og variasjonen mellom grupper (andre ledd). Estimatoren er forventningsrett og konsistent dersom $E\{x_{it}\varepsilon_{it}\}=0$ og $E\{x_{it}\alpha_{it}\}=0$, altså forklaringsvariablene er uavhengig av feilleddet og den individspesifikke komponenten. Den siste forutsetningen er her viktig for at modellen skal være konsistent. Dersom antakelsen er oppfylt vil tilfeldig effekt modellen være den mest effisiente av paneldatamodellene.

5.3 Binært valg modeller

En vanlig antakelse i økonometrisk analyse er at modellen vi skal estimere er lineær i parametrene, og at den avhengige variabelen, y_i , er en kontinuerlig variabel som i prinsippet kan ta enhver verdi i et gitt intervall. Eksempler på slike avhengige variabler er inntekt, arbeidsledighetsrate, aggregert forbruk osv. (Bratberg, 2004). Det er imidlertid også ofte tilfelle at vi studerer variabler som er diskrete av natur, altså kun kan ta verdi en eller null. Her kan vi for eksempel trekke inn deltakelse i arbeidsmarkedet, konkurser og også deltakelse ved utdanningsinstitusjoner. Alle disse kan ved spørsmål besvares med 'ja' eller 'nei' svar, også kalt binære valg.

Siden disse variablene kun kan ta verdien 1 eller 0, vil det å estimere disse med en lineær sannsynlighetsmodell ikke gi noen garanti for at den estimerte sannsynligheten vil ligge innenfor intervallet $[0,1]$. Dette gjør den lineære sannsynlighetsmodellen til en lite egnet modell, og det er her om å gjøre å finne en modell som løser dette problemet. Logit- og Probit-modellene er de mest vanlige å bruke her, da begge sørger for en S-formet kumulativ fordelingsfunksjon, som sørger for at de estimerte sannsynlighetene holder seg i $[0,1]$ -intervallet. Logit modellen er gitt ved den logistiske fordelingen, mens probit modellen

er gitt ved den kumulative standard normalfordelingen. Jeg skal her fokusere på probit-modellen. Modellen kan skrives som:

$$P(y_i = 1) = \Phi(x_i'\beta) \quad (18)$$

hvor Φ står for den kumulative distribusjonsfunksjonen til standard normalfordelingen;

$$\Phi(x_i'\beta) = \int_{-\infty}^{x_i'\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt \quad (19)$$

Det er mulig å utlede en binært valg modell fra de underliggende antagelser om nyttemaksimering. Dette gir oss en latent variabel presentasjon av modellen, som er vanlig å bruke selv når slike forutsetninger om nyttemaksimering ikke er tilstede. Anta for eksempel at nytten av å ta utdanning er gitt ved y_i^* , med fordeling:

$$y_i^* = \beta_0 + x_i'\beta + \varepsilon_i \quad (20)$$

der ε_i er et tilfeldig feilledd. Personer vil kun velge å ta utdanning dersom nytten av å ta utdanning $y_i^* > 0$. Siden nytten av utdanning, y_i^* , ikke kan observeres, blir den kalt en latent variabel. Det vi imidlertid kan observere, er om personen faktisk velger å ta utdanning eller ikke:

$$y_i = \begin{cases} 1 & \text{hvis } y_i^* > 0 \\ 0 & \text{hvis } y_i^* \leq 0 \end{cases}$$

I denne konteksten blir $x_i'\beta$ kalt index funksjonen. Vi har at:

$$P(y_i^* > 0) = P(x_i'\beta + \varepsilon_i > 0) = P(\varepsilon_i > -x_i'\beta)$$

Hvis distribusjonen av ε_i er symmetrisk er dette ekvivalent med:

$$P(y_i^* > 0) = P(\varepsilon_i < x_i'\beta) = F(x_i'\beta) \quad (21)$$

I dette tilfellet står $F(\bullet)$ for Probit-modellen.

5.3.1 Modell med ordnet responsvariabel

Det finnes imidlertid flere eksempler på variabler, som i tillegg til å være diskrete, også har en naturlig rangering av alternativer. Vi kan i stor grad si at *utdanningsnivå* er en slik ordnet variabel. Dette fordi nivåene kommer i en ordnet rekkefølge; Barneskolenivå,

ungdomskolenivå, videregående osv. Det har altså stor betydning for resultatene i regresjonen hvordan denne rekkefølgen er satt sammen. Samtidig er det *kun* rangeringen som teller her, ikke størrelsen på nivåene. Vi kan altså ikke si at nivå 2 er dobbelt så stort som nivå 1 etc.

Ordnete responsmodeller baserer seg på at vi har en underliggende latent variabel, og modellen er som i (20). Y_i^* er fortsatt uobserverbar, men vi kan observere om den ligger innenfor et nærmere angitt intervall. Vi lar her $\mu_j, j=0,1,\dots,J$ representere kutt punkt for J mulige intervaller for y_i^* . Verdien for den observerbare variabelen y_i kan dermed defineres som

$$y_i = j \text{ dersom } \mu_{j-1} < y_i^* \leq \mu_j \quad (22)$$

Dersom vi for eksempel har fire intervaller, får vi:

$$\begin{aligned} y_i &= 0 \text{ dersom } -\infty < y_i^* \leq \mu_0 \\ y_i &= 1 \text{ dersom } \mu_0 < y_i^* \leq \mu_1 \\ y_i &= 2 \text{ dersom } \mu_1 < y_i^* \leq \mu_2 \\ y_i &= 3 \text{ dersom } \mu_2 < y_i^* \leq \infty \end{aligned}$$

Vi ser altså som nevnt tidligere, at det *kun* er rangeringen av nivåene som teller her, ikke størrelsen. For å kunne estimere denne modellen er vi nødt til å normalisere den noe. Det er ikke mulig å observere både konstantleddet β_0 og kutt punktet μ_0 , slik at en av disse må bli normalisert til null. Stata bruker normaliseringen $\beta_0 = 0$ og estimerer μ_0 , mens andre programmer gjør det på motsatt måte.

Den ordnede probit modellen (OP) er utledet fra (20) og (22) og ut fra antakelsen om at feilleddet, ε_i , har en standard normal distribusjon, der den kumulative distribusjonsfunksjonen er gitt som

$$\phi(\varepsilon) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2}\right), \quad \Phi(\varepsilon) = \int_{-\infty}^{\varepsilon} \phi(t) dt \quad (23)$$

Modellen blir så estimert med maximum likelihood metoden (ML). ML-estimatene er de verdiene av forventningen, μ , og variansen, σ^2 , som beskriver sannsynlighetsfordelingen som med størst sannsynlighet er opphav til de observerte verdiene av x (Bratberg, 2004).

5.4 Økonometriske problemer

5.4.1 Endogenitet

Dersom feilledet i en regresjonsmodell er korrelert med en eller flere av forklaringsvariablene i modellen, har vi et endogenitetsproblem. Forutsetning (A2) om at alle forklaringsvariabler og alle feilledd skal være uavhengig for alle individ, er dermed brutt, og forklaringsvariablene er ikke lenger eksogene.

Det kan være flere årsaker til at vi får endogene forklaringsvariabler. Her er det verdt å nevne utelatte variabler, målefeil og simultane likninger. MKM vil i dette tilfellet gi forventningsskjevne og inkonsistente koeffisientestimater. Simultane likninger oppstår når en eller flere av forklaringsvariablene bestemmes samtidig som den avhengige variabelen (jamfør Wooldridge, 2006, kap. 16), og kan i dette tilfellet være ekstra aktuelt. Dette fordi verdien av forklaringsvariabler, som antall barn, slik vi får de oppgitt i datasettet, ikke nødvendigvis vil være bestemt før man tar utdanning. Dersom personer først tar utdanning, for så å få barn senere i livet, kan vi med andre ord ikke si at antall barn har en innvirkning på utdanningsnivå. Det å prøve å estimere effekten av barn på utdanningsnivå vil da helt klart gi inkonsistente koeffisientestimater.

En vanlig måte å korrigere for endogenitet, er å bruke en instrumentvariabel som er korrelert med den endogene variabelen, men som ikke korrelert med feilledet. Her kan en også bruke såkalt to-steps metode. I paneldatanalyser kan bruk av FE-estimatoren fjerne den enkle formen for endogenitet som tidskonstante individspesifikke variabler kan presentere, da disse fjernes ved hjelp av FE-trasformasjonen.

5.4.2 Autokorrelasjon

Autokorrelasjon, eller seriekorrelasjon, er et problem som lett kan oppstå i tidsserier eller paneldatamodeller, og vil si at vi har en korrelasjon mellom feilleddene i forskjellige tidsperioder. Dette er et brudd på Gauss-Markov antakelse (A4). Årsaker til autokorrelasjon kan være at vi har en feilspesifisert modell, hvor både utelatte variabler, utelatt dynamikk og feilspesifiserte funksjonsformer inngår. Dersom vi for eksempel utelater en relevant positiv forklaringsvariabel fra modellen over flere perioder, vil denne komme inn som en del av feilledet, og feilleddene i disse periodene vil dermed være korrelert. Feilleddene vil i tillegg

ha en positiv verdi, noe som også vil føre til brudd på forutsetning A(1), om at den forventede verdien til feilleddet skal være lik null.

Andre årsaker til autokorrelasjon kan være vedvarende virkninger av sjokk i økonomien, som for eksempel kriger, naturkatastrofer osv., eller treghet i atferd/vanedanning. Disse problemene vil imidlertid være mindre omfattende enn problemet med feilspesifikasjon i modellen, da estimatoren vil være forventningsrett så lenge forklaringsvariablene er strengt eksogene. Hovedproblemet med autokorrelasjon vil være at vi ikke lenger får oppgitt korrekte standardfeil, som igjen kan føre til feil i signifikanstestene.

I økonomiske analyser er det mest vanlig med positiv autokorrelasjon, det vil si at positive verdier blir etterfulgt av positive verdier, og negative verdier etterfulgt av negative. Etterfølgende verdier på feilleddet har altså en tendens til å ha samme fortegn. For å teste for 1. ordens autokorrelasjon i feilleddet kan man bruke en såkalt t-test for autokorrelasjon. Man kan da "lagge" residualen i regresjonen med en periode, for så å kjøre en regresjon med residualen som avhengig variabel og den "laggede" residualen som uavhengig variabel. Dersom den "laggede" variabelen da er signifikant forskjellig fra null, har vi et problem med autokorrelasjon i modellen. En annen vanlig test for autokorrelasjon er Durbin-Watson testen. Man kan også benytte seg av robuste standard feil for å korrigere for ukorrekte standardfeil som følge av autokorrelasjon. Robuste standardfeil er en type standardfeil for estimatorer som vil være asymptotisk gyldige uansett, dette uavhengig av om feilleddet i modellen er autokorrelert, eller ikke. Signifikanstestene vil dermed være pålitelige.

5.4.3 Heteroskedastisitet

Heteroskedastisitet i feilleddet er brudd på Gauss Markov vilkår A3, om at variansen til feilleddet, betinget på forklaringsvariablene, skal være konstant. Heteroskedastisitet vil dermed si at variansen til feilleddet varierer for forskjellige observasjoner, og kan for eksempel komme av at det er stor variasjonsbredde i forklaringsvariablene. MKM med heteroskedastiske feilledd vil fremdeles gi forventningsrette og konsistente koeffisient-estimer, men standardfeilene til estimatorene vil være forventningsskjeve. Dette innebærer at MKM ikke lenger vil være effisient.

Problemet med heteroskedastisitet kan løses ved å bruke robuste standardfeil. Dette kan imidlertid kun gjøres i store utvalg. Man kan også teste for heteroskedastisitet ved bruk av Breusch-Pagan eller White-test.

5.4.4 Multikollinearitet

Det er generelt ikke noe galt i å inkludere variabler i modellen som er korrelert. I mitt utvalg vil jeg for eksempel inkludere både en variabel for om man er gift, og en variabel for om man har barn under 16 år, selv om det kan forventes at de som er gift, i gjennomsnitt, i større grad har barn under 16 år enn de som ikke er gift. Det samme gjelder for eksempel for variablene norsk statsborger og botid i Norge, selv om vi kan forvente en viss korrelasjon her. Dersom denne korrelasjonen blir for stor, kan vi imidlertid få upresise og ustabile estimater, med høye standardavvik. Dette kommer av at dersom for eksempel gift og barn under 16 år er *for* høyt korrelert, kan det være vanskelig for modellen å identifisere de individuelle effektene av disse to variablene, når det er akkurat det vi prøver å gjøre.

Multikollinearitet blir i Wooldridge (2006) definert som høy (men ikke perfekt) korrelasjon mellom to eller flere uavhengige variabler. I ekstreme tilfeller, der en forklaringsvariabel er perfekt korrelert med en annen forklaringsvariabel, har vi *perfekt kollinearitet*, som bryter med forutsetning (A2) om at x-ene skal være uavhengig av hverandre. Multikollinearitet bryter imidlertid ikke med forutsetningene, og problemet med multikollinearitet er derfor ikke særlig godt definert. En ting er likevel klart: når vi estimerer β_i , gir det bedre resultater når vi har lite korrelasjon mellom x_i og de andre forklaringsvariablene, alt annet likt.

En løsning på problemet kan være å ta andre uavhengige variabler fra modellen for å på denne måten redusere multikollineariteten. Problemet med dette kan imidlertid være vi skaper en forventningsskjevhet, rett og slett på grunn av utelatte variabler. Ved å bruke korrelasjonsmatriser kan vi i stata enkelt teste for om forklaringsvariablene er sterkt korrelerte. Det vil også være et tegn på multikollinearitet dersom det å utelate enkelte uavhengige variabler fra modellen, endrer resultatene betydelig for andre forklaringsvariabler.

5.4.5 Ubalanserte panel og selekterte utvalg

I empiriske paneldatasett vil det ofte være problemer med manglende variabler. Dette kan komme av problemer med innsamlingen av data; etter et par perioder kan for eksempel individene i datasettet nekte å samarbeide, husstander kan flytte eller splittes opp, firma kan gå konkurs eller fusjonere med et annet firma osv. I mitt datasett vil det typisk være et problem med manglende verdier på utdanningsnivå, da det ikke er noen systematisk registrering av denne variabelen for personer som ikke kommer i kontakt med utdanningssystemet i Norge etter innflytting (se avsnitt 2.4). Dette fører til at den totale mengden av observasjoner ikke vil være komplett i datasettet.

En enkel løsning på problemet kan her være å slette alle observasjoner i datasettet som vi ikke har komplett informasjon om. Dette kan imidlertid være svært ineffektivt, da en mengde relevant informasjon ikke vil bli utnyttet. Dette kan løses ved å bruke alle observasjonene, inkludert for de individene som ikke er observert i alle periodene. Dette vil si å bruke et *ubalansert panel*. Det å bruke et ubalansert panel burde ikke medføre større problemer enn at antall observasjoner blir mindre. Dette gjelder imidlertid kun så lenge de manglende variablene er *tilfeldig* trukket fra utvalget. Et potensielt, og mer seriøst problem med å bruke paneldatasett med manglende variabler, er faren for *seleksjon i utvalget*. Dersom individer i datasettet *ikke* er observert for noen av forklaringsvariablene, på grunn av en eksogen årsak, kan både det å slette observasjonene, og å bruke et ubalansert panel, føre til forventningsskjeve estimatorer og ukorrekte tester.

6 Empirisk analyse

Jeg skal i dette kapittelet gå gjennom oppgavens empiriske analyse. Jeg vil begynne med å analysere hvordan forskjeller i landbakgrunn gir ulik innvirkning på utdanningsnivået. Deretter vil jeg gå videre til en analyse for hver av verdensregionene separert, og se på hvilke faktorer som påvirker utdanningsnivå. Jeg vil på slutten av kapittelet kommentere hvordan de økonometriske problemene jeg gikk gjennom i kapittel 5 har hatt innvirkning på analysen.

For å på best mulig måte kunne analysere forskjellene for de ulike verdensregionene, har jeg i den første analysen satt sammen de to datasettene. Jeg kjører her regresjoner både for kvinner og menn, for å se hvordan de ulike landbakgrunnene påvirker utdanningsnivå forskjellig for disse. For å få resultater som er så korrekt som mulig vil jeg her bruke tre modeller. I den første ønsker jeg å benytte meg av panelstrukturen i dataene mine. Jeg vil derfor bruke en tilfeldig effektmodell for paneldata. På grunn av at den avhengige variabelen, *utdanningsnivå*, varierer såpass lite over tid (se appendiks A) har jeg valgt å ”skrumpe inn” tidsdimensjonen i paneldatasettet. Jeg har derfor satt tidsdimensjonen til å være over 4 perioder, der periode 1 består av årene 1993-1995, periode 2 av 1996-1998, periode 3 av 1999-2001 og periode 4 består av årene 2002-2003. I hver av periodene er den avhengige variabelen, *utdanningsnivå*, satt til det høyeste utdanningsnivået individet hadde i løpet av perioden. Alle de andre variablene tar den verdien som var oppgitt for hvert individ i det siste året av perioden. Dersom verdien for det siste året er *'uoppgitt'*, gjelder den siste verdien som er oppgitt for individet i løpet av perioden. Siden *utdanningsnivå* er en variabel som kun øker med årene, og ikke vil kunne gå nedover, ser jeg også et poeng i å kjøre en tverrsnittsregresjon på den siste perioden i datasettet, altså periode 4 fra 2002-2003. I denne tverrsnittsanalysen vil jeg bruke en vanlig minste kvadrats metode modell, samt en probit modell for ordnede responsvariabler.

Innvandrerbefolkningen kan i stor grad sies å være en heterogen gruppe, og det vil derfor være lite hensiktsmessig og se på innvandrerbefolkningen under ett. Jeg vil på grunn av dette også kjøre en regresjon for hver av verdensregionene, for på best mulig måte å få frem det særegne ved hver region, for eksempel når det gjelder sivil status, barn og etterkommere versus førstegenerasjons innvandrere. Jeg skal her bruke en tilfeldig effekt modell, for å få utnyttet panelstrukturen i datasettet. Disse modellene vil bli referert til som regionsmodeller.

6.1 Resultater fra estimering med minste kvadrats metode, tilfeldig effekt og ordnet probit

Av tabell 11 kan vi se hvordan de ulike landbakgrunnene har påvirkning på utdanningsnivå. Alt annet likt, reflekterer variablene med andre ord hva som kun stammer fra den gitte regionen eller landbakgrunnen. Nordmenn, eller de uten innvandrerbakgrunn, er her referansegruppen. Disse regresjonene vil her utelukkende bli brukt til å analysere forskjellene i landbakgrunn, og resten av variablene utgjør dermed kun kontrollvariabler. Alle estimatene som ikke er signifikant på 5 % nivå er i tabellen satt i kursiv.

6.1.1 Tolkning av resultater, landbakgrunn

I mitt datasett har jeg dessverre ikke informasjon om hvilket land innenfor de forskjellige verdensregionene innvandrerne kommer fra. Selv om det kan være store forskjeller innenfor en region, har jeg altså kun mulighet til å se på forskjellene mellom regionene. Jeg har imidlertid i avsnitt 2.1.1 gått gjennom hvilke land innvandrerne hovedsakelig kom fra innen de forskjellige verdensdelene, før 1993. Fra Europa kom for eksempel den største andelen fra våre naboland, Danmark og Sverige. Jeg har heller ikke noen egen variabel for *innvandringsgrunn*, men i avsnitt 2.1.2 har jeg skrevet om typiske innvandrergrunner for de forskjellige regionene. Hovedvekten av innvandrere fra Europa har for eksempel kommet som følge av arbeidsinnvandring, mens størsteparten av de afrikanske innvandrerne har kommet som flyktninger.

Av tabell 11 ser vi at de med landbakgrunn fra Nord-Amerika, Vest-Europa og Øst-Europa ser ut til å ha et høyere utdanningsnivå enn de med landbakgrunn i Norge, alt annet likt. Dette gjelder både for kvinner og menn. De med landbakgrunn i de asiatiske regionene og Afrika har derimot lavere koeffisienter enn oss nordmenn for begge kjønn. For de med landbakgrunn i Sør-Amerika er imidlertid tallene litt motstridende for kvinner og menn. Kvinner har her en klart negativ koeffisient i forhold til kvinner uten innvandrerbakgrunn, mens menn ser ut til å ha en positiv koeffisient i forhold til norske menn, selv om denne ikke er særlig stor. Estimering med minste kvadrats metode (MKM), tilfeldig effekt (RE) og ordnet probit (OP) gir her liknende resultater for de fleste verdensregionene. Dette blir nærmere kommentert videre i teksten.

Tabell 11: Resultater fra estimering med MKM, Tilfeldig effekt og Oprobit, menn og kvinner hver for seg, med utdanningsnivå som avhengig variabel.

<i>Variabler</i>	<i>MKM</i>		<i>Random Effekt</i>		<i>Oprobit</i>	
	Kvinner	Menn	Kvinner	Menn	Kvinner	Menn
Vest-Europa	0.3101 (0.0357)	0.3255 (0.0339)	0.5408 (0.0254)	0.5001 (0.0253)	0.2219 (0.0261)	0.2059 (0.0233)
Øst-Europa	0.0063 (0.0334)	0.0508 (0.0337)	0.1866 (0.0298)	0.3024 (0.0289)	0.0318 (0.0239)	0.0434 (0.0225)
Afrika	-1.3392 (0.0452)	-0.3638 (0.0318)	-1.0159 (0.0389)	-0.0768 (0.0332)	-0.9519 (0.0339)	-0.2563 (0.0212)
Vest-Asia	-1.8081 (0.0350)	-0.7528 (0.0293)	-1.4816 (0.0298)	-0.4086 (0.0298)	-1.3586 (0.0266)	-0.5424 (0.0199)
Øst-Asia	-0.8764 (0.0277)	-0.6046 (0.0285)	-0.6394 (0.0243)	-0.2127 (0.0274)	-0.6432 (0.0205)	-0.4081 (0.0196)
Nord-Amerika	0.9917 (0.0530)	1.3930 (0.0613)	1.2671 (0.0386)	1.6293 (0.0448)	0.7193 (0.0387)	0.9050 (0.0423)
Sør-Amerika	-0.4093 (0.0453)	-0.0689 (0.0418)	-0.2218 (0.0385)	0.1423 (0.0380)	-0.2527 (0.0325)	-0.0375 (0.0277)
Etterkommere	-0.2434 (0.0546)	0.0949 (0.0557)	-0.2532 (0.0517)	0.0597 (0.0633)	-0.1859 (0.0365)	0.0553 (0.0363)
Barn u, 16 år	0.2073 (0.0101)	0.2309 (0.0102)	0.0089 (0.0016)	0.0223 (0.0013)	0.1429 (0.0070)	0.1552 (0.0067)
Barn u, 16 år ²	-0.0792 (0.0027)	-0.0521 (0.0031)	-0.0049 (0.0004)	-0.0030 (0.0003)	-0.0544 (0.0019)	-0.0339 (0.0021)
Botid	0.0646 (0.0036)	0.0037 (0.0036)	0.0689 (0.0030)	0.0074 (0.0032)	0.0405 (0.0026)	-0.0004 (0.0024)
Botid ²	-0.0021 (0.0001)	-0.0005 (0.0001)	-0.0023 (0.0001)	-0.0007 (0.0001)	-0.0013 (0.0001)	-0.0002 (0.0001)
Gift	-0.0307 (0.0086)	0.2543 (0.0090)	0.0201 (0.0020)	0.0274 (0.0016)	-0.0120 (0.0060)	0.1789 (0.0061)
Gift*Barn u, 16	0.1776 (0.0069)	0.0208 (0.0077)	-0.0008 (0.0013)	-0.0018 (0.0011)	0.1105 (0.0047)	0.0030 (0.0050)
Tettbygd strøk	0.2955 (0.0079)	0.5350 (0.0076)	0.0255 (0.0019)	0.0264 (0.0016)	0.1936 (0.0054)	0.3396 (0.0052)
Norsk statsborger	-0.1168 (0.0246)	-0.0268 (0.0248)	0.0445 (0.0042)	0.0479 (0.0033)	-0.0662 (0.0180)	-0.0164 (0.0170)
Alder 21-26 år	0.3864 (0.0134)	0.2637 (0.0119)	0.4777 (0.0134)	0.2157 (0.0137)	0.3061 (0.0090)	0.2083 (0.0077)
Alder 27-31 år	0.1893 (0.0133)	0.1016 (0.0120)	0.3343 (0.0119)	0.1284 (0.0116)	0.1751 (0.0091)	0.1036 (0.0079)
Alder 32-36 år	0.0751 (0.0124)	0.0097 (0.0119)	0.1748 (0.0118)	0.0525 (0.0111)	0.0771 (0.0085)	0.0293 (0.0079)
Alder 42-46 år	-0.1768 (0.0123)	-0.0510 (0.0125)	-0.2281 (0.0108)	-0.0812 (0.0102)	-0.1295 (0.0086)	-0.0473 (0.0084)
Alder 47-51 år	-0.3871 (0.0125)	-0.1600 (0.0133)	-0.4336 (0.0102)	-0.1992 (0.0100)	-0.2865 (0.0089)	-0.1298 (0.0089)
Alder 52-56 år	-0.6397 (0.0133)	-0.3529 (0.0148)	-0.6782 (0.0102)	-0.4085 (0.0103)	-0.4945 (0.0097)	-0.2749 (0.0102)
Periode 2			0.0697 (0.0009)	0.0607 (0.0007)		
Periode 3			0.1161 (0.0009)	0.0938 (0.0007)		

Periode 4			0.1432 (0.0009)	0.1084 (0.0008)		
Antall obs,	221282	229608	886536	926517	221282	229451
F-test	738,3	471,59				
Wald/LR-test			61371,3	36737,27	26163,57	18554,67
R ² /R ² overall/pseudo R ²	0,1222	0,0740	0,0893	0,0316	0,0436	0,0436

- Estimaterne som ikke er signifikante på et 5 % signifikansnivå er satt i *kursiv*, signifikante verdier er uthevet
- Fylkesdummyer er med i regresjonen, men ikke tatt med i tabellen pga mange insignifikante estimater
- Oprobit regresjonens cut (8) er ikke med i tabellen,
- Konstantleddet til regresjonene er ikke tatt med i tabellen
- Regresjonen er gjort med robuste standardavvik som korrigerer for heteroskedastisitet

Ser at de med landbakgrunn i **Vest-Asia** er de med det laveste utdanningsnivået, både for kvinner og menn. Dette stemmer overens med den deskriptive statistikken i kapittel 4, der jeg fant at over 40 prosent av de med landbakgrunn i Vest-Asia har ingen utdanning, kun grunnskoleutdanning eller uoppgitt utdanningsnivå, mens mindre enn 20 prosent hadde utdanning på universitets- og høyskolenivå. Vi ser at koeffisienten er spesielt lav for kvinner fra Vest-Asia i forhold til de uten innvandringsbakgrunn, og også i forhold til de andre regionene. Dette stemmer også overens med den deskriptive statistikken, der jeg fant at kvinner fra Vest-Asia har et betydelig lavere utdanningsnivå enn menn fra Asia.

For kvinner er det de med landbakgrunn i **Afrika** som har det laveste utdanningsnivået, etter Vest-Asia. Dette resultatet gjelder imidlertid ikke for menn, da menn med landbakgrunn i Afrika ser ut til å ha et høyere utdanningsnivå enn menn fra **Øst-Asia**. Dette gjelder for begge tverrsnittsanalysene, MKM og OP. For estimering med RE ser vi imidlertid at menn med landbakgrunn fra Afrika ikke ser ut til å ha noen signifikant forskjell i utdanningsnivå i forhold til nordmenn. I den deskriptive statistikken i kapittel 4 så vi at i likhet med Vest-Asia, var det også store forskjeller blant afrikanerne når det gjaldt høyere utdanning for kvinner og menn. Her skilte kvinner med landbakgrunn fra Afrika seg ut med liten grad av høyere utdanning, i forhold til menn fra Afrika. Blant kvinner fra Øst-Asia så vi av den deskriptive statistikken at de i større grad enn menn hadde liten eller ingen utdanning, men også at de i større grad enn menn fra Øst-Asia hadde høyere utdanning. Dette kan være med på å gjøre at kvinnene fra Øst-Asia har et høyere utdanningsnivå enn de afrikanske kvinnene, mens menn fra Øst-Asia er under afrikanske menn på statistikken, gitt alt annet likt.

Utdanningsnivået for kvinner med landbakgrunn i **Sør-Amerika** er høyere enn utdanningsnivået til kvinner med landbakgrunn i de ikke-vestlige regionene Vest-Asia, Øst-Asia og Afrika, men lavere enn utdanningsnivået til kvinner med landbakgrunn i Nord-

Amerika, Vest-Europa, Øst-Europa og Norge. Det samme gjelder menn med landbakgrunn i Sør-Amerika, bortsett fra at disse som nevnt ser ut til å ha et utdanningsnivå høyere enn det nordmenn har. Både kvinner og menn med landbakgrunn fra Sør-Amerika har imidlertid koeffisienter som er veldig nære koeffisientene for nordmenn. Av regresjonsmodellen ser vi forskjellige resultater for tverrsnittsanalysen og paneldata analysen. Både MKM og OP rapporterer at det å ha landbakgrunn i Øst-Asia ikke gir noen signifikant forskjell i utdanningsnivå i forhold til de med norsk landbakgrunn. Ved estimering med RE ser vi imidlertid at menn og kvinner med landbakgrunn fra **Øst-Europa**, ser ut til å ha høyere utdanningsnivå enn menn og kvinner fra Norge, selv om disse koeffisientene ikke er særlig høye. Øst-Europa er dermed den ikke-vestlige regionen med det høyeste utdanningsnivået i utvalget mitt.

Som i den deskriptive statistikken, ser vi at **Nord-Amerika** er på topp når det gjelder utdanningsnivå, for både kvinner og menn. **Vest-Europa**, som er det landet som har det nest høyeste nivået, viser her langt lavere koeffisientestimer, og dette gjelder spesielt for menn. Da koeffisientene for menn og kvinner er relativt like for vesteuropeerne, i forhold til de uten innvandrerbakgrunn, er nordamerikanernes koeffisienter vesentlig høyere for menn enn for kvinner. Det er likevel ingen tvil om at kvinner fra Nord-Amerika har det høyeste utdanningsnivået av kvinnene i utvalget.

Som vi har sett, går det her et skille mellom de vestlige og de ikke-vestlige regionene når det gjelder utdanningsnivå. Dette skillet er imidlertid ikke helt klart, da begge de ikke-vestlige regionene, Øst-Europa og Sør-Amerika, har høyere koeffisienter enn de uten innvandrerbakgrunn i Norge. Det at Øst-Europa og Sør-Amerika¹⁰ er de vestlige landene med det høyeste utdanningsnivået stemmer med funn gjort av Olsen (2006), som fant at disse gruppene hadde det høyeste sysselsettingsnivået av de ikke-vestlige innvandrerne. Han antar at det høye sysselsettingsnivået blant personer fra Øst-Europa har en sammenheng med arbeidsinnvandringen fra de nye EU-landene, og det kan da være rimelig å anta at mange av disse innvandrerne har et visst utdanningsnivå innenfor fagfelt som er etterspurt på det norske markedet, som for eksempel byggebransjen. Når det gjelder innvandrere fra Sør-Amerika, vet vi at mange av disse kommer fra Chile, da innvandrere fra Chile er den tolvte største ikke-vestlige innvandrergruppen i Norge (Lie, 2004). Lie (2004) skriver at chilenerne har et

¹⁰ Dette gjelder imidlertid kun for menn fra denne regionen

gjennomsnittlig høyt utdanningsnivå, da over 26 prosent av innvandrerne fra Chile har utdanning på universitets- og høyskolenivå. Dette er nok medvirkende til at de med landbakgrunn fra Sør-Amerika gjør det såpass godt på utdanningsstatistikken.

De ikke-vestlige regionene Vest-Asia, Afrika og Øst-Asia har imidlertid svært lave koeffisienter, og vi kan i disse regionene se store forskjeller på størrelsen til koeffisientene for menn og kvinner. Menn i disse regionene har altså et klart høyere utdanningsnivå enn kvinnene, gitt alt annet likt. Innvandringsgrunn kan være med på å forklare disse forskjellene i utdanningsnivå mellom regionene. Fra Vest-Asia, som er den regionen med det laveste utdanningsnivået, har mange kommet som arbeidsimmigranter fra land som Tyrkia og Pakistan før innvandringsstoppen i 1975. I følge Lie (2003) har disse typisk et lavt utdanningsnivå, noe som også bekreftes av Barth, Bratsberg og Raaum (2004). Fra Vest-Asia er det imidlertid også en del av innvandrerne som har kommet som flyktninger fra land som Iran før 1993, og ifølge Lie (2003) har også disse ofte et lavt utdanningsnivå, eller i alle fall dårlig dokumentasjon på utdannelsen sin. Det samme gjelder innvandrere fra Afrika, som i hovedsak har immigrert til Norge av fluktgrunner. For innvandrerne fra Øst-Europa er det imidlertid et skille mellom innvandrere fra land som Vietnam og Sri Lanka, som for det meste har kommet til Norge av fluktgrunner, eller for å gjenforenes med familie som allerede har flyktet til Norge, og innvandrere fra land som Thailand og Filippinene som har kommet for å stifte familie i Norge. Da den første av disse gruppene kan tenkes å ha et lavt utdanningsnivå, vil ofte utdanningsnivået til de som kommer til Norge for å stifte familie avhenge av utdanningsnivået til den de kommer til¹¹ (Lie, 2003). En stor del av andelen av de som innvandrer fra Filippinene, samt den største andelen av de som innvandrer fra Kina, kommer nettopp til Norge for å ta utdanning her. Det er dermed sannsynlig at disse gruppene trekker opp utdanningsnivået for regionen. Det vil med andre ord være grunn til å tro at det er store forskjeller i utdanningsnivå også *innenfor* disse regionene, noe som gjør det vanskelig å trekke mer spesifikke konklusjoner på bakgrunn av disse funnene.

¹¹ Ifølge Lie (2003) skiller de med landbakgrunn fra Thailand seg ut som en gruppe med svært lavt utdanningsnivå, mens de med landbakgrunn fra Filippinene har et bemerkelsesverdig høyt utdanningsnivå i forhold til de uten innvandrerbakgrunn.

Mange av innvandrerne fra de vestlige regionene, Vest-Asia og Vest-Europa, samt mange av innvandrerne fra Øst-Europa har kommet til Norge av arbeidsgrunner. En av grunnene til at disse har et høyere utdanningsnivå enn de med landbakgrunn i Norge, kan være at de kommer til Norge nettopp *på grunn av* kvalifikasjonene sine. Som vi så i avsnitt 2.2 er behovet for kvalifiserte arbeidere i Norge stort, og innvandrere fra disse regionene med høy utdanning vil derfor ofte være etterspurt i Norge.

I følge Roys seleksjonsmodell (se avsnitt 3.2.2) vil arbeidsinnvandrere komme til Norge dersom de har sterke økonomiske insentiver til dette. Dersom avkastningen på utdanning i Norge er høy i forhold til det den er i for eksempel Nord-Amerika, vil dette typisk trekke innvandrere med høyt utdanningsnivå fra Nord-Amerika til Norge, da de kan tjene mer her enn hjemme. Dersom Norge har lav avkastning av utdanning på inntekt, altså større likhet i inntekt mellom utdanningsnivå, enn Nord-Amerika, ville imidlertid, i følge modellen, innvandrere med lavt utdanningsnivå fra Nord-Amerika immigrert til Norge. Det vil være rimelig å tro at Norge er et land med lav avkastning av utdanning på inntekt, i forhold til Nord-Amerika, da Norge som velferdsstat, er et land med relativt lave inntektsforskjeller. Det kan derfor argumenteres for at de med høy utdanning i Norge ville hatt økonomiske insentiver til å immigrere til USA, mens de med lav utdanning i USA ville hatt insentiver til å immigrere til Norge. På bakgrunn av resultatene i regresjonen, der nordamerikanerne er de med det høyeste utdanningsnivået i befolkningen, stemmer altså ikke intuisjonen i Roys seleksjonsmodell i akkurat dette tilfellet. Det kan her være rimelig å anta at Roy modellen ikke passer like godt til strukturen i det norske samfunnet som for den amerikanske.

6.2 Resultater fra estimering med regionsmodellene

Jeg vil nå se nærmere på forskjellene innad de forskjellige regionene, og tabell 12 gjengir resultatene fra estimering med tilfeldig effekt regionsmodeller. Jeg vil her gå gjennom forklaringsvariablene hver for seg, og også ta med hvordan eventuelle interaksjonsvariabler eller variabler som fanger opp 2. ordenseffekten påvirker koeffisientene.

Tabell 12: Resultater fra estimering med tilfeldig effekt regionsmodeller, utdanningsnivå som avhengig variabel

<i>Variable</i>	<i>Vest-Europa</i>	<i>Øst-Europa</i>	<i>Afrika</i>	<i>Vest-Asia</i>	<i>Øst-Asia</i>	<i>Nord-Amerika</i>	<i>Sør-Amerika</i>	<i>Norge</i>
Kvinne	0.2706 (0.0193)	<i>0.0667</i> (0.0369)	-0.6742 (0.0448)	-0.8141 (0.0326)	-0.1763 (0.0277)	-0.1262 (0.0587)	-0.1393 (0.0514)	-0.1545 (0.0051)
Kvinne*gift	<i>0.0112</i> (0.0081)	<i>-0.0195</i> (0.0140)	<i>-0.0175</i> (0.0191)	-0.0603 (0.0163)	<i>0.0054</i> (0.0114)	<i>0.0153</i> (0.0234)	<i>-0.0148</i> (0.0178)	-0.0145 (0.0024)
Kvinne*barn u.16 år	-0.0157 (0.0033)	<i>-0.0046</i> (0.0063)	-0.0296 (0.0064)	-0.0098 (0.0039)	<i>-0.0044</i> (0.0038)	-0.0230 (0.0090)	<i>0.0043</i> (0.0079)	-0.0197 (0.0010)
Etterkommere	0.5416 (0.0648)	0.7787 (0.1068)	0.9280 (0.3175)	1.5600 (0.6436)	<i>-0.2102</i> (0.2411)	0.3842 (0.1634)	<i>0.1648</i> (0.5544)	-
Etterkommere*Barn u. 16	<i>-0.0045</i> (0.0075)	<i>-0.0108</i> (0.0107)	<i>-0.0079</i> (0.0441)	<i>0.0408</i> (0.1085)	<i>-0.0052</i> (0.0288)	<i>0.0191</i> (0.0113)	<i>-0.0817</i> (0.0416)	-
Etterkommere*gift	0.0408 (0.0198)	<i>-0.0160</i> (0.0297)	<i>-0.0157</i> (0.0667)	<i>0.3866</i> (0.2275)	<i>0.1178</i> (0.0848)	-0.1036 (0.0438)	0.0413 (0.1472)	-
Barn under 16 år	0.0283 (0.0047)	<i>0.0115</i> (0.0084)	0.0323 (0.0071)	0.0029 (0.0060)	<i>0.0091</i> (0.0053)	<i>-0.0030</i> (0.0134)	<i>0.0111</i> (0.0091)	0.0226 (0.0012)
Barn under 16 år^2	<i>-0.0019</i> (0.0013)	<i>-0.0019</i> (0.0018)	<i>-0.0019</i> (0.0014)	-0.0030 (0.0008)	-0.0036 (0.0012)	<i>0.0014</i> (0.0019)	-0.0065 (0.0025)	-0.0037 (0.0003)
Botid	<i>-0.0048</i> (0.0043)	0.0488 (0.0075)	-0.0260 (0.0107)	-0.0210 (0.0098)	0.0480 (0.0066)	<i>0.0038</i> (0.0126)	<i>-0.0178</i> (0.0167)	-
Botid^2	-0.0010 (0.0001)	-0.0027 (0.0002)	<i>0.0006</i> (0.0004)	-0.0011 (0.0004)	<i>0.0002</i> (0.0002)	-0.0019 (0.0003)	<i>0.0003</i> (0.0006)	-
Gift	<i>-0.0027</i> (0.0093)	<i>-0.0110</i> (0.0116)	0.0285 (0.0132)	0.0337 (0.0126)	0.0298 (0.0108)	<i>-0.0031</i> (0.0218)	-0.0467 (0.0204)	0.0328 (0.0017)
Gift*barn under 16	<i>-0.0052</i> (0.0038)	<i>0.0032</i> (0.0069)	-0.0155 (0.0057)	<i>0.0097</i> (0.0054)	0.0099 (0.0048)	<i>0.0146</i> (0.0142)	0.0250 (0.0070)	<i>-0.0014</i> (0.0009)
Gift*Botid	<i>0.0007</i> (0.0004)	0.0017 (0.0007)	<i>-0.0020</i> (0.0012)	<i>-0.0021</i> (0.0011)	-0.0022 (0.0010)	<i>0.0008</i> (0.0011)	0.0037 (0.0018)	-
Tettbygd strøk	0.0104 (0.0046)	<i>0.0036</i> (0.0110)	0.0546 (0.0185)	<i>-0.0124</i> (0.0135)	<i>-0.0071</i> (0.0093)	<i>0.0291</i> (0.0208)	<i>-0.0071</i> (0.0128)	0.0284 (0.0013)
Norsk statsborger	<i>0.0104</i> (0.0090)	0.0738 (0.0072)	0.0568 (0.0076)	0.0735 (0.0052)	0.0239 (0.0044)	<i>0.0177</i> (0.0374)	0.0298 (0.0086)	-
Alder 21-26 år	-0.5802 (0.0618)	-0.6023 (0.0856)	-0.4266 (0.0793)	-0.3591 (0.0490)	<i>0.0524</i> (0.0489)	-1.4593 (0.1658)	-0.4124 (0.1128)	0.4671 (0.0101)
Alder 27-31 år	-0.3452 (0.0421)	-0.3373 (0.0620)	<i>-0.1053</i> (0.0649)	-0.1676 (0.0443)	0.1100 (0.0386)	-0.6840 (0.1116)	-0.1713 (0.0792)	0.3128 (0.0089)
Alder 32-36 år	-0.1693 (0.0355)	-0.1604 (0.0585)	<i>-0.0871</i> (0.0648)	-0.1362 (0.0440)	0.1622 (0.0384)	-0.2801 (0.0953)	<i>0.0017</i> (0.0681)	0.1451 (0.0089)
Alder 42-46 år	-0.0754 (0.0302)	<i>0.0223</i> (0.0570)	-0.1631 (0.0758)	-0.2096 (0.0476)	-0.1442 (0.0442)	0.2164 (0.0870)	-0.2822 (0.0779)	-0.1419 (0.0080)
Alder 47-51 år	-0.1763 (0.0315)	0.2141 (0.0560)	-0.5671 (0.1004)	-0.4857 (0.0523)	-0.2714 (0.0496)	0.2752 (0.0876)	-0.4840 (0.0943)	-0.2986 (0.0076)
Alder 52-56 år	-0.2362 (0.0326)	0.3430 (0.0644)	-0.9927 (0.1136)	-0.8221 (0.0659)	-0.7593 (0.0583)	<i>-0.0131</i> (0.1069)	-0.8488 (0.1179)	-0.1087 (0.0082)
Periode 2	0.0486 (0.0020)	0.0475 (0.0044)	0.0917 (0.0060)	0.0784 (0.0043)	0.0696 (0.0033)	0.0379 (0.0060)	0.0569 (0.0063)	0.0655 (0.0006)
Periode 3	0.0762 (0.0021)	0.0700 (0.0044)	0.1252 (0.0065)	0.1099 (0.0046)	0.1062 (0.0036)	0.0623 (0.0063)	0.0977 (0.0067)	0.1066 (0.0006)

Periode 4	0.0931 (0.0022)	0.0833 (0.0047)	0.1482 (0.0068)	0.1301 (0.0049)	0.1266 (0.0039)	0.0716 (0.0066)	0.1214 (0.0075)	0.1276 (0.0007)
Antall observasjoner	128465	32465	28382	56258	69002	10078	15596	1482562
R2 overall	0,0538	0,0440	0,0423	0,0742	0,0279	0,1389	0,0243	0,0534
Wald test	4995,45	1292,73	1581,09	3830,93	2895,56	694,34	652,12	81041,87

- Estimaterne som ikke er signifikante på et 5 % signifikansnivå er satt i *kursiv*, signifikante verdier er uthevet
- Fylkesdummyer er med i regresjonen, men ikke tatt med i tabellen pga mange insignifikante estimater
- Konstantleddet til regresjonene er ikke tatt med i tabellen
- Regresjonen er gjort med robuste standardavvik som korrigerer for heteroskedastisitet

6.2.1 Tolkning av resultater

Kvinne

For de med landbakgrunn i Vest-Europa og Øst-Europa¹² påvirker det å være kvinne utdanningsnivået positivt. Det vil si at kvinner i gjennomsnitt har høyere utdanning enn menn i disse regionene, når alt annet er likt. For verdensregionene Afrika, Vest-Asia, Øst-Asia, Nord-Amerika og Sør-Amerika er det imidlertid motsatt. Det å være kvinner i en av disse regionene påvirker utdanningsnivået negativt. Vest-Asia er det landet med størst forskjeller når det gjelder menn og kvinner, da vi ser at koeffisienten for kvinner i Vest-Asia er den absolutt laveste. Vi kan se at norske kvinner også har et lavere utdanningsnivå enn norske menn, og bare de ikke-vestlige regionene Vest-Asia, Øst-Asia og Afrika har lavere koeffisienter her.

Etterkommere

Vi ser av regresjonen at etterkommere, eller andregenerasjons innvandrere, har et høyere utdanningsnivå enn førstegenerasjonsinnvandrerne, og dette gjelder for alle verdensregionene. Dette stemmer overens med hva blant annet Lie (2003) har funnet om at etterkommere deltar i både videregående opplæring og høyere utdanning i høyere grad enn førstegenerasjonsinnvandrere.

Barn under 16 år

Verken Vest-Europa, Øst-Asia, Nord-Amerika eller Sør-Amerika har koeffisienter som er signifikant forskjellig fra null for denne variabelen, og det ser derfor ikke ut til at det å ha barn under 16 år har noen sammenheng med utdanningsnivå for disse. Det å ha barn under 16 år

¹² Denne variabelen er kun signifikant på 10 % nivå.

gir imidlertid positive koeffisientestimater for menn fra Vest-Europa, Afrika, Vest-Asia og Norge. Afrikanske menn har her den høyeste koeffisienten, mens vestasiatiske menn har den laveste. Dette positive utslaget er imidlertid signifikant marginalt avtakende for Vest-Asia og Norge. Det vil si at det å ha et barn under 16 kan påvirke utdanningsnivået positivt, gjerne fordi en da vil ha høyere utdanning for å kunne forsørge familien bedre. Dersom antallet barn øker vil imidlertid denne økningen i utdanningsnivå avta, og kanskje etter hvert begynne å synke. Vi ser at det å være kvinne og ha barn under 16 år gir en negativ sammenheng med utdanning i forhold til menn i samtlige regioner.

Jeg har her regnet ut toppunktet for hvor mange barn som gir det gjennomsnittlig høyeste utdanningsnivået for personer med landbakgrunn i Vest Asia og Norge. For menn fra Vest-Asia vil 0,5 barn gi det høyeste gjennomsnittlige utdanningsnivået, mens for norske menn vil 3,1 barn være det antallet barn under 16 år som gir det høyeste gjennomsnittlige utdanningsnivået. For kvinner fra Vest-Asia ser vi imidlertid at det å ha barn under 16 år gir et entydig negativt utslag på utdanningsnivået. For norske kvinner gir derimot det å ha barn under 16 år et positivt utslag på utdanningsnivået, men dette gjelder kun frem til et gjennomsnitt på 0,4 barn. Det ser altså ut som om det i størst grad er kvinners utdanningsnivå som vil være skadelidende av det å ha barn. Vi ser også at sammenhengen mellom å ha barn under 16 år og utdanningsnivå, er klart mer negativ enn for de med landbakgrunn fra Norge, for både menn og kvinner. Fra den deskriptive statistikken så vi at over 12 prosent av de med landbakgrunn fra Vest-Asia har mer enn tre barn under 16 år. Det er derfor ikke usannsynlig at dette kan gjøre det vanskeligere å komme seg ut i utdanning i Norge, da familien både vil trenge en forsørgerinntekt, samt noen til å passe barna.

Det å være gift i tillegg til å ha barn under 16 år ser ikke ut til å gi et særlig stort utslag på utdanningsnivå i forhold til det å ikke være gift. Gifte personer med landbakgrunn fra Afrika har imidlertid negativt koeffisientestimat på utdanningsnivå av det å ha barn under 16 år i forhold til de som ikke er gift. Dette resultatet er motsatt for de med landbakgrunn fra Øst-Asia og Sør-Amerika. For etterkommerne har ikke det å ha barn under 16 noen signifikant påvirkning på utdanningsnivået. Dette kan imidlertid komme av at de fleste etterkommerne ennå er såpass unge (Lie, 2003), og derfor ikke i særlig stor grad vil ha fått barn ennå.

Gift

Det å være gift har signifikant positiv innvirkning på utdanningsnivået for menn med landbakgrunn i Afrika, Vest-Asia, Øst-Asia og Norge. Denne variabelen har derimot ingen signifikant påvirkning på de med landbakgrunn fra de resterende regionene. Gifte kvinner har et lavere utdanningsnivå enn gifte menn, både i Vest-Asia og Norge. Forskjellene på disse koeffisientene er imidlertid store, og vi ser at kvinner med landbakgrunn i Vest-Asia har et negativt koeffisientestimat for det å være gift på utdanningsnivå, både relativt til menn og i absolutt forstand. Kvinner med norsk landbakgrunn har imidlertid et positivt koeffisientestimat av variabelen gift i absolutt forstand, selv om dette er negativt i forhold til estimatet for menn. Fra den deskriptive statistikken ser vi imidlertid at nesten 74 prosent av de med landbakgrunn fra Vest-Asia er gift, mens dette kun gjelder 54 prosent av de med norsk landbakgrunn. Det kan være nærliggende å tolke resultatene slik at kjønnsrollemønsteret er sterkere for de med landbakgrunn i Vest-Asia enn for de med norsk landbakgrunn, noe som kommer til uttrykk her ved at gifte kvinner fra Vest-Asia har lavere utdanningsnivå enn kvinner fra regionen som ikke er gift.

Det å være gift ser ikke ut til å ha noen entydig sammenheng med utdanningsnivå for etterkommerne i forhold til resten av innvandrerbefolkningen. For etterkommere med landbakgrunn fra Vest-Europa og Sør-Amerika ser vi et høyere utdanningsnivå for de som er gift, mens det er motsatt for etterkommere fra Nord-Amerika. Botiden i Norge ser ikke ut til å endre betydning av det å være gift i forhold til utdanningsnivå, og dette gjelder for alle regionene.

Tettbygd strøk

Det å bo i tettbygde strøk har en signifikant positiv innvirkning på utdanningsnivået til personer med landbakgrunn i Vest-Europa, Afrika og Norge. Dette gir mening, i og med at både utdanningsinstitusjoner og ofte de best betalte jobbene ligger i tettbygde strøk. For de andre regionene ser ikke det å bo i tettbygde strøk ut til å ha signifikant betydning her.

Norsk statsborger

Det å ha norsk statsborgerskap gir en signifikant positiv innvirkning på utdanningsnivå for alle de ikke-vestlige regionene. Denne variabelen har mest sannsynlig en sammenheng med

botiden i Norge, da man vanligvis må ha oppholdt seg i Norge i til sammen syv av de ti siste årene, med gyldig tillatelse, for å kunne få norsk statsborgerskap (Utlendingsdirektoratet, 2008). For å få statsborgerskap må man også ha til hensikt å fortsette å bo i Norge (ibid.), og variabelen gir da også en pekepinn på at vi har med permanente innvandrere å gjøre. Som nevnt i avsnitt 3.3 fant Kalena Cortes (2004) at flyktninger, som stort sett vil være en gruppe som ikke har mulighet til å returnere til hjemlandet sitt, vil investere mer i humankapital som språkopplæring og utdanning i vertslandet, enn andre typer innvandrere. Selv om resultatene hennes stammer fra USA, utelukker ikke mine resultater at dette også kan være tilfelle i Norge.

Norsk statsborgerskap har derimot ingen signifikant sammenheng med utdanningsnivået til de to vestlige innvandrergroppene, Vest-Asia og Nord-Amerika. Dette kan ha sammenheng med at svært få fra disse regionene skifter til norsk statsborgerskap etter ankomst til Norge (se avsnitt 4.2.2).

Botid

Vi ser at det er store forskjeller sammenhengen mellom botid og utdanningsnivå ettersom hvilken landbakgrunn vi ser på. Vi ser for eksempel at det er en negativ sammenheng mellom botid og utdanningsnivå for innvandrere fra Afrika og Vest-Asia, mens det er en positiv sammenheng mellom botid og utdanningsnivå for innvandrere fra Øst-Europa og Øst-Asia. Dette kan ha en sammenheng med at mange med landbakgrunn fra Øst-Europa og Øst-Asia har kommet til Norge nettopp for å ta utdanning (Statistisk sentralbyrå, 2008a). Vi ser at botid har en negativ marginalavkastning for de fleste regionene, noe som vil si at jo lenger innvandrere bor i Norge, jo mindre blir sammenhengen mellom botid og utdanningsnivå. Dette kan tyde på at de fleste innvandrere som tar utdanning i Norge, gjør det de første årene etter at de er kommet til Norge. For de med landbakgrunn i Øst-Europa gir en botid i Norge på ni år det gjennomsnittlig høyeste utdanningsnivået.

Når jeg kun undersøker for om botid har en innvirkning på utdanningsnivået, tar jeg imidlertid ikke hensyn til at det kan være såkalte kohort effekter mellom utdanningsnivået til de forskjellige innvandrergroppene (se avsnitt 3.2.1). Dersom innvandrere som kom i en tidlig innvandringsbølge typisk har høyt utdanningsnivå i forhold til de som har kommet senere til Norge, vil dette gi inntrykk av at botiden påvirker utdanningsnivået positivt, selv om dette

ikke nødvendigvis er tilfelle. Siden utdanningsnivå ikke kan synke med tiden, kan derfor ikke botid i Norge gi et negativt utslag på utdanningsnivå. For innvandrere fra Vest-Asia og Afrika, som får oppgitt en negativ effekt av botid, er det nok slik at de tidligere innvandrergruppene fra disse regionene har hatt et lavere utdanningsnivå enn de innvandrerne som har kommet i nyere tid. Fra Vest-Asia vet vi for eksempel at noen av de første innvandrerne til Norge fra denne regionen var arbeidsinnvandrere fra land som Tyrkia og Pakistan (se avsnitt 2.1.2), og i følge Bratsberg, Raaum og Røed (2007) hadde disse innvandrerne generelt lav utdanning.

Alder

Hos nordmenn ser vi en tendens til at utvalget har høyere utdanningsnivå jo lavere aldersgruppe vi er i. De i aldersgruppen 21 – 26 år er dermed de med den høyeste utdanningen her. For de med landbakgrunn i Vest-Europa, Afrika og Sør-Amerika ser vi at de i aldersgruppen 37 – 41 år er de med det høyeste utdanningsnivået, mens vi for de med landbakgrunn i Øst-Europa og Nord-Amerika ser en helt omvendt tendens fra den norske, nemlig at utdanningsnivået er høyere dess høyere aldersgruppe man er i. For de med østasiatisk landbakgrunn er det de i aldersgruppen 32 – 36 år som har det høyeste utdanningsnivået, men de som er yngre enn denne aldersgruppen har til gjengjeld høyere gjennomsnittlig utdanningsnivå enn de som er eldre. Resultatene her stemmer med funnene til Lie (2003) om at innvandrere blir ferdig med høyere utdanning på et senere tidspunkt enn nordmenn (se avsnitt 3.4).

Periode

Vi ser at utdanningsnivået øker for hver periode for alle regionene, selv om denne økningen ikke er særlig stor. Dette viser at utdanningsvariabelen varierer over tid, og dermed at noen av individene i utvalget fortsatt holder på med utdanningen sin.

6.3 Diskusjon av resultatene

Jeg har i denne oppgaven gått gjennom forskjeller i utdanningsnivå mellom innvandrere i Norge fra ulike verdensregioner, og faktorer som kan ha innvirkning på disse forskjellene. Mangelen på tilgjengelige data har imidlertid gjort at resultatene blir veldig generelle, selv om vi vet at utdanningsnivå nok også varierer mye *innen* de forskjellige verdensregionene i

utvalget. Dersom jeg hadde hatt tilgang på mer spesifikke kategorier for landbakgrunn, som for eksempel enkeltland, eller landgrupper, kunne jeg også ha fanget opp deler av disse forskjellene, og dermed fått et bedre innblikk når det gjelder også disse forskjellene. Siden jeg ikke har tilgang på en variabel som forklarer innvandrergrunn har jeg heller ingen mulighet til å undersøke om det er forskjeller i utdanningsnivå for innvandrere med flyktningbakgrunn i forhold til for eksempel arbeidsinnvandrere. Det vil være nærliggende å tro at det kan være forskjeller innad i disse gruppene, da arbeidsinnvandrere kommer til Norge av frivillige årsaker, mens flyktninger ofte "tvinges" til å komme til Norge, på grunn av frykt for eget liv i hjemlandet. Variabelen innvandrergrunn kunne også til en viss grad ha erstattet en mer spesifikk variabel for landbakgrunn, da innvandringsgrunn også kan fortelle oss noe om hvilke kår personen som innvandrer kommer fra, og hva som motiverte innvandrerens til å immigrere til Norge.

Datasettet mitt inneholder som sagt observasjoner av alle innvandrerne i Norge i 1993, og følger disse frem til 2003. Jeg har med andre ord ingen informasjon om innvandrere som kom til Norge *etter* 1993. I kapittel 2 kan vi se at innvandrerbefolkningen i Norge nesten er fordoblet siden 1993 og frem til i dag, noe som gjør at mine resultater ikke automatisk kan overføres til dagens innvandrerbefolkning. Innvandrerbefolkningen har også endret form, fra i hovedvekt å bestå av flyktninger og immigranter som kommer for familiegjening med disse, til nå å i hovedsak bestå av arbeidsinnvandring (se avsnitt 2.1.2). Dette gjør at utdanningsnivået til dagens innvandrere i Norge kan være vidt forskjellig fra hva det var i 1993, også med tanke på landbakgrunn. Det kunne derfor ha vært interessant å gjort samme analyse med nyere variabler, og sammenlignet funnene med resultatet i denne oppgaven.

Et potensielt problem for oppgaven min er den store andelen individer som ikke har oppgitt utdanningsnivå i utvalget for innvandrere. Som vi så i den deskriptive statistikken, i avsnitt 4.2.1, har over 19 prosent av utvalget for innvandrere uoppgitt utdanningsnivå. Andelen var her størst for de vestlige innvandrerne, og nesten 40 prosent av innvandrerne fra Nord-Amerika har ikke oppgitt utdanningsnivå. Det at vi ikke observerer den avhengige variabelen *utdanningsnivå* for alle individene over alle periodene, vil si at vi har å gjøre med et ubalansert panel. Som nevnt i avsnitt 5.4.5 burde ikke dette medføre andre problemer enn at antallet observasjoner synker. Dette gjelder imidlertid kun så lenge de individene vi ikke observerer utdanningsnivå for, er tilfeldig trukket fra utvalget. Dersom størsteparten av de vi ikke observerer utdanningsnivå for, eksempelvis er personer som har et svært lavt utdannings-

nivå, og derfor ikke har kommet i kontakt med det norske utdanningssystemet eller fått registrert utdanningen sin i Norge, vil det å estimere med det ubalanserte panelet kunne føre til forventningsskjevne koeffisientestimer og ukorrekte tester.

På bakgrunn av dette har jeg kjørt en probit analyse, der jeg har sammenlignet gruppen med uoppgitt utdanningsnivå, med hver av de andre utdanningsnivågruppene. Tabell E.1 i appendiks E viser resultatene av disse analysene. Jeg kunne i prinsippet ha gjort disse for hver verdensdel, men belyser dette heller ved å se på grupper med dummyer for de forskjellige verdensdelene. På bakgrunn av verdiene til Likelihood ratio testene, og verdiene til R^2 , er det ingen av utdanningsnivågruppene som skiller seg ut som spesielt like gruppen med uoppgitt utdanningsnivå. Dette tyder på at vi ikke har et selektert utvalg, og at gruppen som det ikke er oppgitt utdanningsnivå for dermed fordeler seg ganske likt over alle utdanningsnivåene.

Som gjennomgått i kapittel 5.4 er endogenitet et generelt problem i økonometrisk analyse. En av årsakene til endogenitet kan være simultanitet, som vil si at en eller flere av forklaringsvariablene bestemmes samtidig som den avhengige variabelen. Man vil da ikke lenger kunne tolke de rene kausale effektene av forklaringsvariablene, og estimatene vil bli forventningsskjevne. I min analyse vet vi for eksempel ikke om innvandrere har fullført utdanningen sin før de kommer til Norge, eller om de tar utdanning etter at de har innvandret til Norge. Dette gjør at vi ikke kan skille ut den rene kausale effekten landbakgrunn har på utdanning, da vi ikke vet om det er faktorer fra landet de kommer fra som fører til at de har tatt den utdanningen de har, eller om dette kommer på bakgrunn av innvandringen til Norge.

Problemet med simultanitet vil også være til stede når det gjelder variablene sivilstand og antall barn. Verdiene for disse variablene vil ikke nødvendigvis være bestemt slik de er i datasettet, på det tidspunktet man tar utdanning. Det er for eksempel minst like sannsynlig at man tar utdanning først, for så å gifte seg og få barn etterpå, som at man gifter seg og får barn før man tar utdanning. Dersom utdanning kommer først, kan ikke disse variablene sies å ha noen direkte effekt på utdanningsnivå. Utdanningsnivå kan imidlertid også ha en innvirkning på antall barn og sivilstand, da såkalte karrieremennesker med høyt utdanningsnivå kan tenkes å få færre barn enn personer med et lavere utdanningsnivå. Dette vil kanskje i størst grad gjelde kvinner, da det å få barn ofte vil kunne ha en effekt på kvinners karriere. Når effekten går i forskjellige retninger på denne måten har vi såkalt forskjøvet simultanitet i

forklaringsvariablene. Det å estimere effekten av antall barn på utdanning, vil da gi forventningsskjevne koeffisientestimer.

Som vi har sett i kapittel 3 har utdanningsnivå stor påvirkning på inntekt, og disse variablene er dermed sterkt positivt korrelerte. Jeg har i appendiks 3 vist at dette er tilfelle, da inntekten øker med utdanningsnivå for hele utvalget, selv om størrelsen på inntektsøkningen varierer med landbakgrunn. Det kan imidlertid være feil å anta at utdanningen kommer først, og så inntekt, da det kanskje spesielt for innvandrere kan være aktuelt at de tar utdanning på bakgrunn av tidligere lav inntekt i Norge, og et ønske om å øke denne. Av resultatene mine kan det se ut som om innvandrere i gjennomsnitt er eldre enn nordmenn når de er ferdig med sin utdanning, noe som kan tyde på at dette er tilfelle. I så fall burde også inntekt vært med som en av forklaringsvariablene. Dersom inntekt også er korrelert med en av forklaringsvariablene, som for eksempel bosted¹³, vil dette innebære at vi har et endogenitetsproblem i modellen, på grunn av utelatte variabler. Dette vil med andre ord også føre til at vi ikke lenger kan måle den rene kausale effekten av bosted på utdanningsnivå. Det at inntekt har en innvirkning på hvilket fylke man bor i virker sannsynlig, i og med at mange velger å flytte dit de kan få den høyeste inntekten, noe som ofte er i de store byene. Det vil altså være store sjanser for at de med høyest inntekt er selektert her. For å korrigere for endogenitet kreves instrumentvariabler, men disse kan i tilfellet med paneldatamodeller være svært krevende å oppdrive.

I tidsserieanalyser og paneldatanalyser kan det som nevnt i avsnitt 5.4 også være problemer med autokorrelasjon og multikollinearitet. Siden jeg i analysen min kun ser på effekter på utdanningsnivå over fire perioder, reduserer dette sjansene for autokorrelasjon, da tidsserieaspektet i oppgaven er veldig lite. Jeg har likevel kjørt en t-test for autokorrelasjon for å kontrollere for om dette forekommer i min analyse. Da denne hadde en t-verdi som ikke var signifikant forskjellig fra null, er det ikke autokorrelasjon av betydning i regresjonen. Jeg har også testet for multikollinearitet ved bruk av korrelasjonsmatriser i stata. Funnene her er at selv om enkelte variabler, som for eksempel alder og botid er korrelert med hverandre, er ikke dette nok til at vi har et problem med multikollinearitet. Det kan imidlertid nevnes her at paneldata dimensjonen ofte visker bort mye av problemene med autokorrelasjon og multikollinearitet som ofte kan oppstå i tidsserieanalyser.

¹³ Forklaringsvariabelen 'bosted' er tilstede i regresjonen min i form av fylker og variabelen for tettbygde strøk.

7 Konklusjon

Jeg har gjennom denne oppgaven gjort rede for forskjeller i utdanningsnivå for innvandrere i Norge fra forskjellige verdensregioner, og sammenlignet disse med befolkningen uten innvandrerbakgrunn. Analysen er basert på to paneldatasett, ett over innvandrerbefolkningen i Norge og ett over de uten innvandrerbakgrunn, over perioden 1993 til 2003. Datasettene inneholder variabler for *utdanningsnivå*¹⁴ og *landbakgrunn*, i tillegg til forklaringsvariabler som *antall barn*, *sivilstand*, *botid*, *alder* og *statsborgerskap*. I oppgaven har jeg valgt å hovedsakelig fokusere på antall barn og sivilstand, da det i liten grad er kartlagt hvordan innvirkning disse forklaringsvariablene har på utdanningsnivået. I den empiriske analysen har jeg brukt tilfeldig effekt modellen for paneldata som grunnleggende hovedmodell, men jeg supplerer analysen ved bruk av minste kvadrats metode og en ordnet probitmodell. Jeg skal her presentere oppgavens hovedfunn, og forslag til videre forskning på emnet. Det må imidlertid presiseres at resultatene mine ikke automatisk kan overføres til dagens innvandrerbefolkning, da denne har gjennomgått store endringer siden 1993.

I den deskriptive statistikken finner jeg at de med landbakgrunn i Vest-Asia skiller seg ut ved å ha en spesielt stor andel personer med ingen utdanning, eller kun grunnskoleutdanning. Dette gjelder spesielt for kvinner fra Vest-Asia, der nærmere 35 prosent av disse ikke har utdanning over grunnskolenivå. De med landbakgrunn i Nord-Amerika skiller seg derimot ut som den gruppen som i desidert størst grad har tatt høyere utdanning. Selv om kvinner fra Nord-Amerika er øverst på denne statistikken, kan vi i mindre grad skille mellom kvinner og menn her.

Dette stemmer godt overens med den empiriske analysen i oppgaven, der jeg finner at de med landbakgrunn i de vestlige regionene, Nord-Amerika og Vest-Europa, har et svært høyt utdanningsnivå i forhold til de uten innvandrerbakgrunn. Mer overraskende er det kanskje at de med landbakgrunn fra Øst-Europa, samt menn med landbakgrunn fra Sør-Amerika, også har et høyere utdanningsnivå enn de uten innvandrerbakgrunn i Norge, gitt alt annet likt. Vest-Asia, etterfulgt av Afrika, er imidlertid den landbakgrunnen med det laveste

¹⁴ Utdanningsnivåene er hentet fra Norsk standard for utdanningsgruppering (NUS2000).

utdanningsnivået. I disse regionene ser vi også store forskjeller i utdanningsnivå når det gjelder kvinner og menn, og kvinner ligger her langt under menn på utdanningsstatistikken.

For menn ser det ut til å være en positiv sammenheng mellom det å ha barn under 16 år, og utdanningsnivået. Denne blir imidlertid mindre når antall barn øker. Kvinner med barn under 16 år ser ut til å ha et lavere utdanningsnivå enn menn med barn under 16 år, og kvinner fra Vest-Asia skiller seg her ut ved å ha en negativ sammenheng mellom barn under 16 år og utdanningsnivå. Lignende resultater finner jeg også når det gjelder det å være gift. Denne faktoren viser en positiv sammenheng med utdanningsnivå for menn fra flere av regionene, men har faktisk en negativ sammenheng med menn fra Sør-Amerika. Gifte kvinner har også et lavere utdanningsnivå enn gifte menn, og dette gjelder både for de uten innvandrerbakgrunn, og for de med landbakgrunn fra Vest-Asia. For kvinner fra Vest-Asia har imidlertid det å være gift en absolutt negativ sammenheng med utdanningsnivået.

Det kan videre trekkes frem at etterkommer har et høyere utdanningsnivå enn førstegenerasjonsinnvandrere, for alle verdensregionene. For disse gir imidlertid det å ha barn under 16 år ingen signifikant forskjell i utdanningsnivå, og vi kan heller ikke se noen signifikant skille mellom de som er gift og de som ikke er gift. Den største andelen etterkommere i Norge er imidlertid under ti år (Lie, 2003), noe som kan tyde på at det kan være for tidlig å se på slike sammenhenger for etterkommerne på nåværende tidspunkt.

Hovedresultatet mitt viser altså at innvandrere med landbakgrunn i Vest-Asia og Afrika er de med det laveste utdanningsnivået i Norge, samt at menn har et gjennomgående høyere utdanningsnivå enn kvinner i disse regionene. Disse resultatene stemmer overens med funnene til Lie (2003), selv om hennes analyse ikke gjør et skille i landbakgrunn mellom Vest-Asia og Øst-Asia, men heller ser direkte på forskjeller mellom ulike land. Gjennom min oppgave har jeg imidlertid funnet en negativ sammenheng mellom det å ha barn under 16 år og utdanningsnivå, samt det å være gift og utdanningsnivå, for kvinner med landbakgrunn i Vest-Asia, noe som i liten grad har blitt undersøkt tidligere. Da de med landbakgrunn fra denne regionen er de som i størst grad er gift, og har den største andelen barn under 16 år, samtidig som kvinner fra denne regionen ligger nederst på utdanningsstatistikken, vil det være sannsynlig at disse faktorene kan være et hinder for at vestasiatiske kvinner kommer seg ut i utdanning i Norge. Dette kan i så fall medføre dårligere tilegning av det norske språk, samt

dårligere integrering i det norske samfunnet for denne gruppen, og kan derfor være et viktig resultat å undersøke videre.

Selv om jeg har inkludert fylkesdummyer i regresjonen min, har jeg ikke gått nærmere inn på disse i oppgaven min. Da vi vet at størsteparten av innvandrerne i Norge holder til i Oslo, og at faktisk omkring 22 prosent av befolkningen i Oslo består av innvandrere (Djuve og Friberg, 2004), er det sannsynlig at utdanningsnivået til innvandrere her avviker fra utdanningsnivået til innvandrere i andre deler av landet. En nærmere analyse av dette kunne derfor ha vært et interessant utgangspunkt for videre forskning.

Jeg har gjennom mine data ikke hatt tilgang til opplysninger om hvilke innvandrere som har tatt utdanning i Norge etter ankomst, og hvilke innvandrere som har med seg utdanning i "bagasjen" fra hjemlandet. Internasjonale studier har vist at utdanning som er tilegnet etter ankomst til vertslandet ofte gir en bedre avkastning på arbeidsmarkedet enn utdanning tatt i hjemlandet (Friedberg, 2000). Det kunne derfor vært hensiktsmessig å også gjøre en studie av dette i Norge, da resultatene kan ha en betydning for hvordan den norske integreringspolitikken bør utformes. Dersom innvandrere som tar utdanning i Norge gjør det bedre på arbeidsmarkedet enn innvandrere som har tatt utdanning i hjemlandet, vil utfordringene i like stor grad bestå i å få innvandrerne til å ta utdanning i Norge, som å innføre tiltak for å få opp statusen på utdanning tilegnet i andre verdensregioner.

Vi kan ut i fra appendiks C se at alle innvandrergруппene har en positiv avkastning på inntekt som følge av økt utdanning, men at størrelsen på denne inntektsøkningen avhenger sterkt av landbakgrunn. Mens nordmenn uten innvandrerbakgrunn har den høyeste inntekten etter utdanningsnivå, med de vestlige regionene Nord-Amerika og Vest-Europa like bak, har de med landbakgrunn fra Vest-Asia og Afrika den laveste den laveste inntekten etter utdanningsnivå av alle regionene i utvalget. Det vil derfor være nødvendig å gjøre analyser av hvilke faktorer som spiller inn her. Er problemet at disse innvandrergруппene ikke får godkjent utdanningen sin fra hjemlandet i Norge, eller lider denne gruppen i særlig grad av diskriminering på arbeidsmarkedet?

Statistisk sentralbyrå gjennomførte i 1999/2000 en spørreskjemaundersøkelse for å kartlegge utdanningsnivået til innvandrere i Norge. Dette var et bra og viktig tiltak, men initiativet må ikke stopper her. Da innvandrerbefolkningen i Norge har økt kraftig det siste tiåret, er nok

andelen vi ikke har oppgitt utdanningsnivå for i dag betydelig. For at myndighetene skal kunne komme med tiltak for å forbedre utdanningsnivået for innvandrere, er det ekstremt viktig å ha tilstrekkelig informasjon om emnet. Da utdanningsnivå i tillegg har en så sterk effekt på lønn, vil kartlegging av utdanningsnivå også være en viktig faktor for å kunne likestille innvandrere og nordmenn på arbeidsmarkedet, i større grad enn det som er tilfellet i dag.

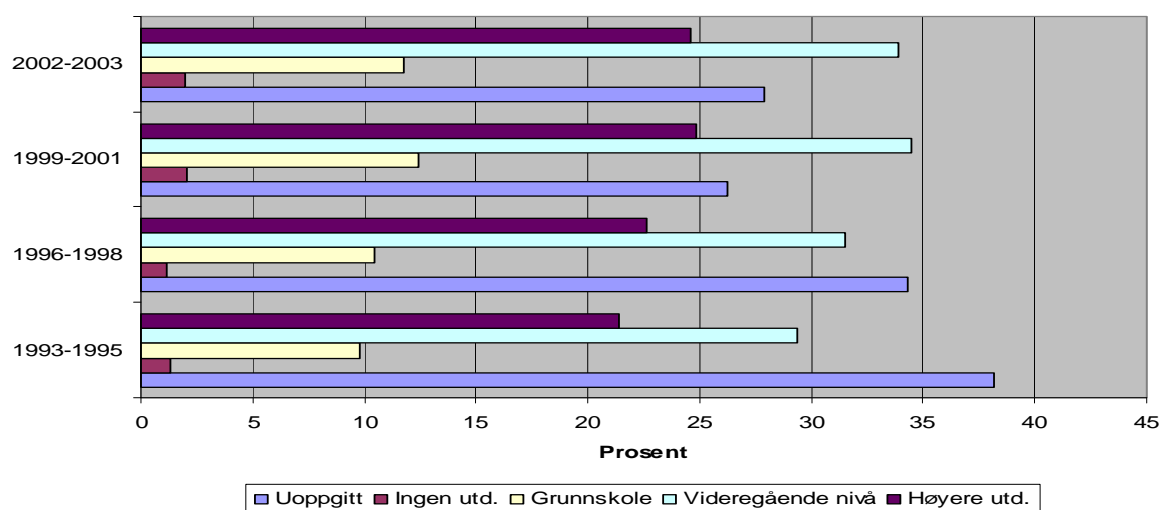
Mer oppsummerende kan det sies at de fleste innvandrergруппene i Norge har et relativt høyt utdanningsnivå, i forhold til de uten innvandrerbakgrunn. Nødvendige politikkendringer her vil altså ikke være å få disse gruppene ut i utdanning, men heller å få innvandrerne opp på samme lønnsnivå i arbeidsmarkedet som de uten innvandrerbakgrunn. Innvandrere fra Vest-Asia, og da spesielt kvinner fra denne regionen, utgjør imidlertid et særlig unntak her. Dette er en gruppe med særdeles lavt utdanningsnivå, der mange ikke har utdanning i det hele tatt, eller kun grunnskoleutdanning. For kvinner fra denne regionen ser det dessuten ut til at det å være gift og ha barn kan være et hinder for i det hele tatt å ta utdanning. Innvandrere fra Vest-Asia er også den innvandrergруппen som har den laveste lønnen på nesten alle utdanningsnivå. Det vil derfor være viktig at myndighetene tar spesielt hensyn til denne innvandrergруппen når utdanningspolitikk for innvandrere utformes.

Appendiks A

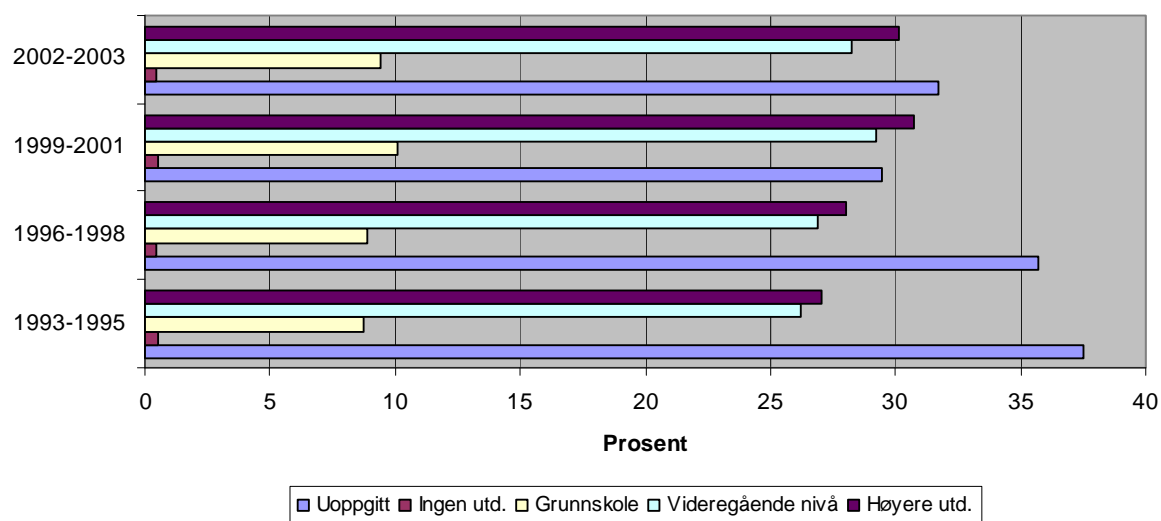
Utdanningsnivå over tid

Ser av figurene at det er en markant nedgang i andelen med uoppgitt utdanningsnivå i perioden 1999-2001, og av figur 4 ser vi at dette gjelder for alle regionene. Dette tyder på resultatene fra spørreskjemaundersøkelsen i 1999/2000 har blitt registrert i datasettet. Ser av figurene at alle utdanningsnivåene øker, for alle gruppene, når andelen uoppgitt synker.

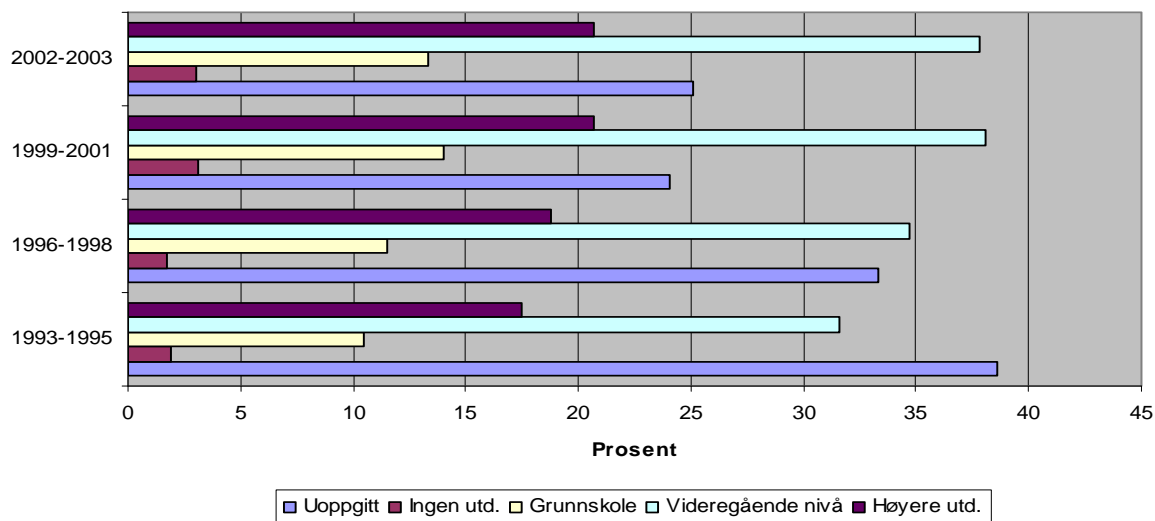
Figur A. 1: Utdanningsnivå blant innvandrere basert på periode



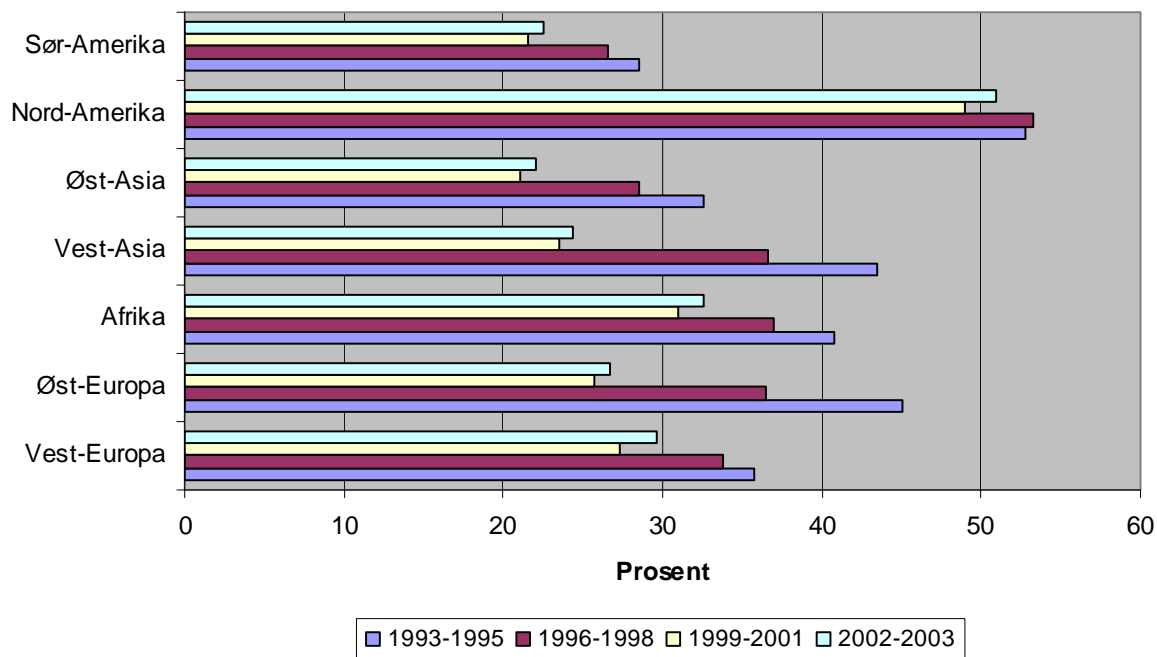
Figur A. 2: Utdanningsnivå hos vestlige innvandrere basert på perioder



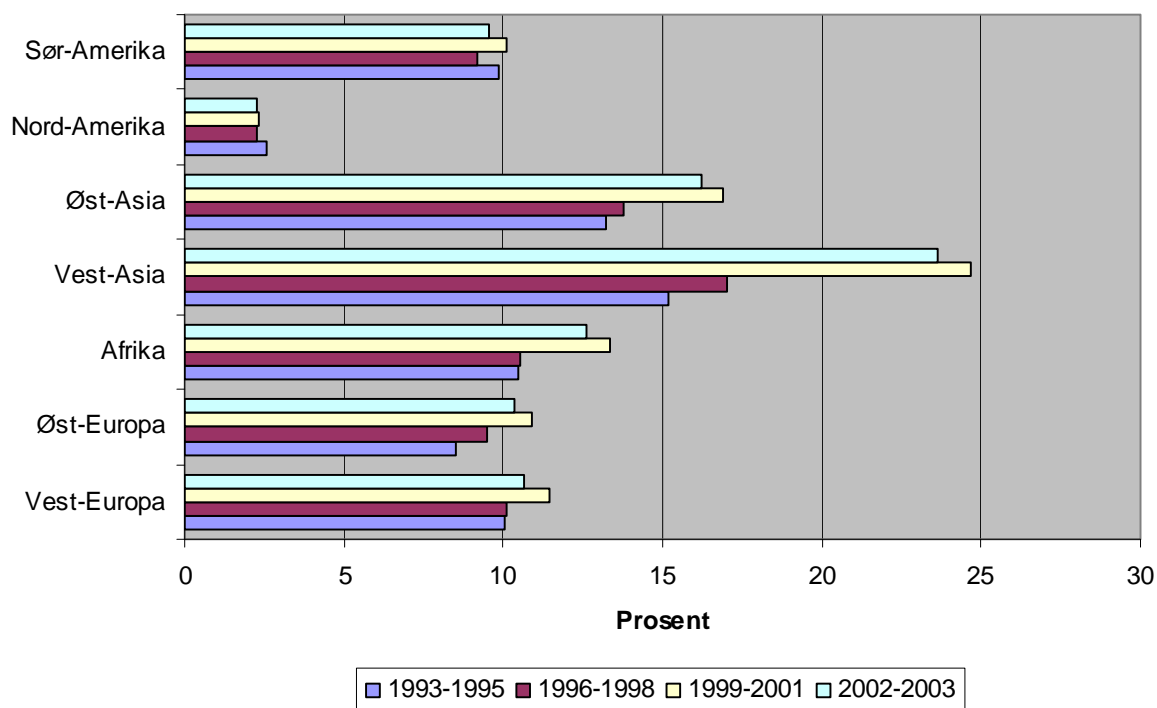
Figur A. 3: Utdanningsnivå hos ikke-vestlige innvandrere basert på periode



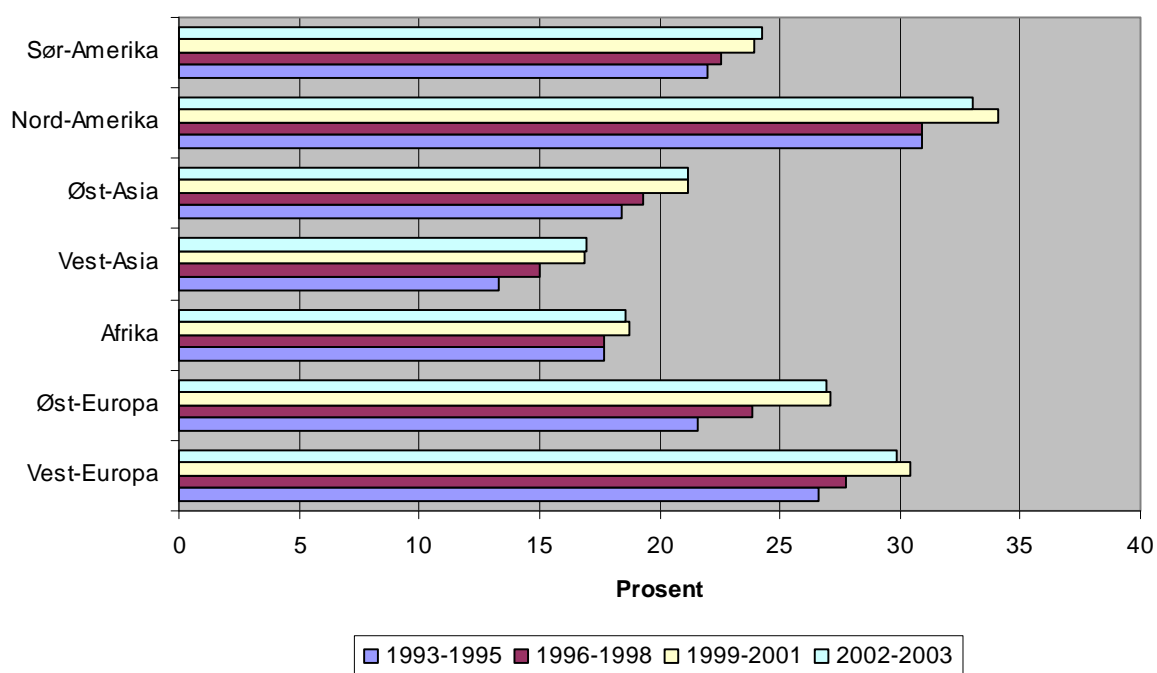
Figur A. 4: Andelen med uoppgitt utdanningsnivå basert på verdensregion og periode



Figur A. 5: Andelen med ingen utdanning eller kun grunnskoleutdanning basert på landbakgrunn og periode



Figur A.6: Andelen med høyere utdanning basert på verdensregioner og periode

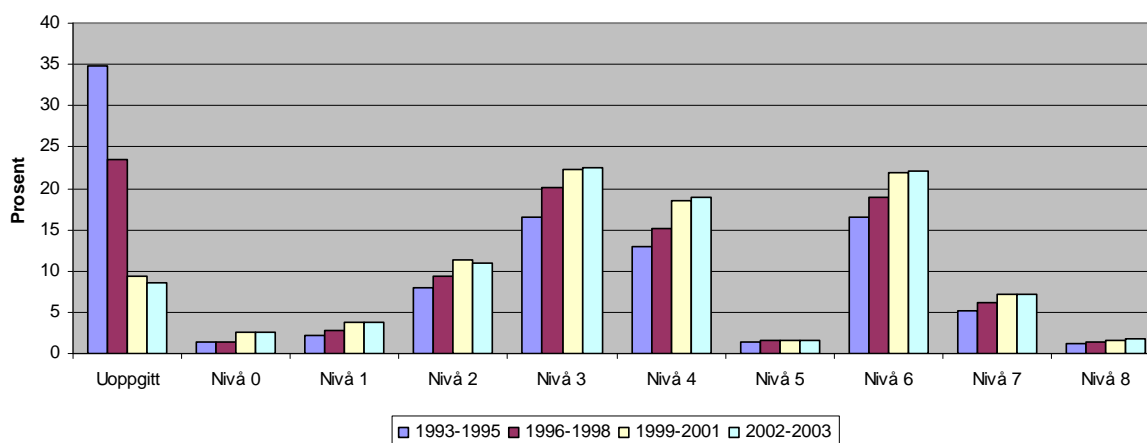


Appendiks B

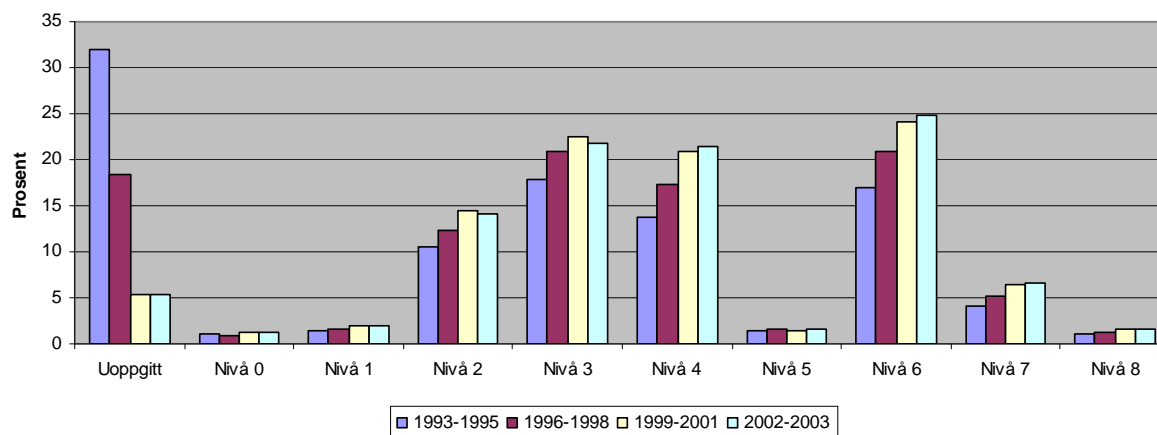
Hvor bor innvandrerne med høy utdanning?

Når vi skal se på hvor innvandrere med høy utdanning bor i Norge kan vi både sjekke for om innvandrere med høy utdanning typisk bor i tettbygde eller spredtbygde strøk, og hvilket fylke flest med høy utdanning bor i. I statistikken for spredtbygde/ tettbygde strøk ser jeg her på utdanningsnivå over tid, mens jeg ser på det maksimale utdanningsnivået i statistikken for fylkene.

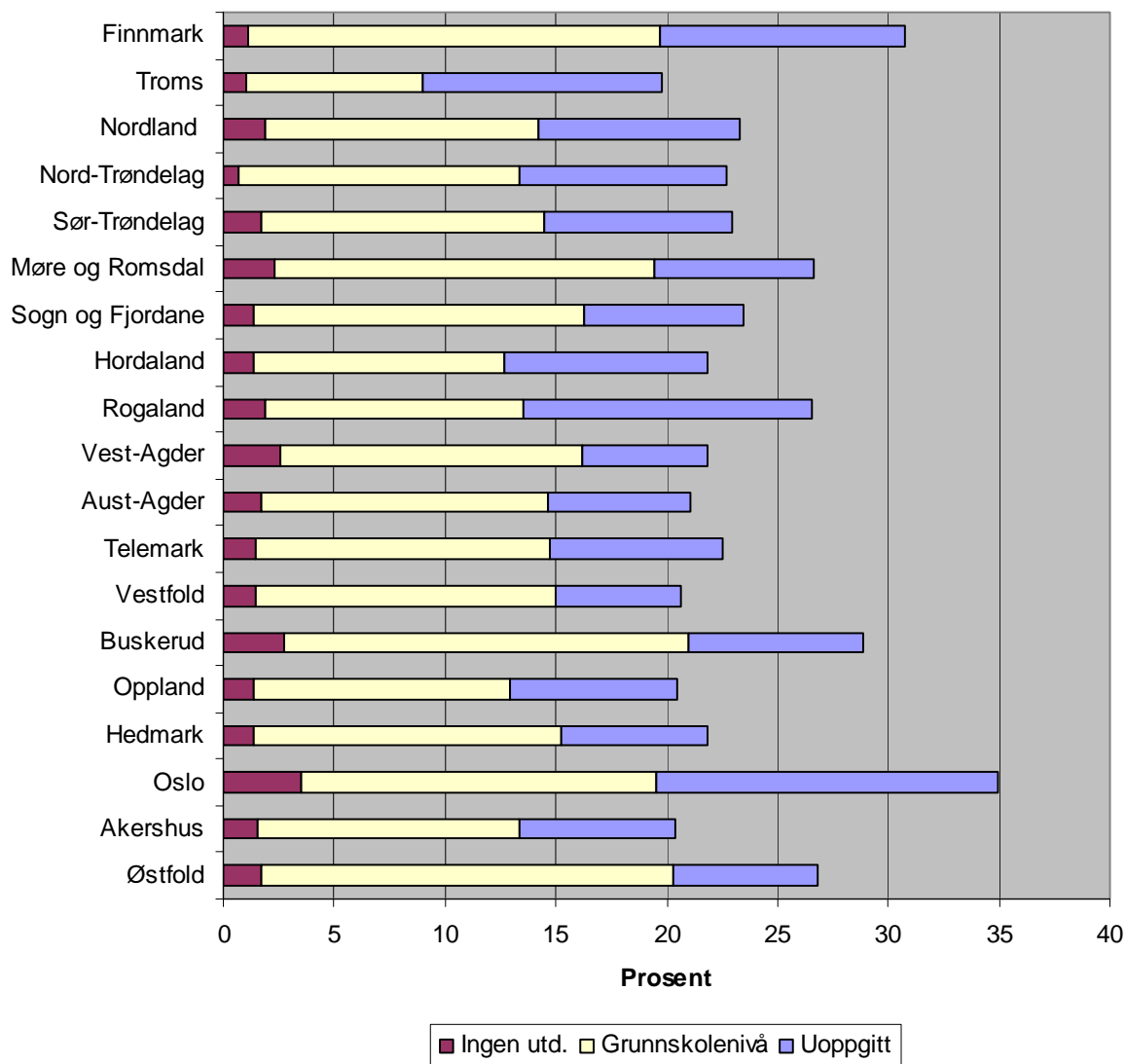
Figur B.1: Forskjeller i utdanningsnivå i tettbygde strøk basert på årsintervaller



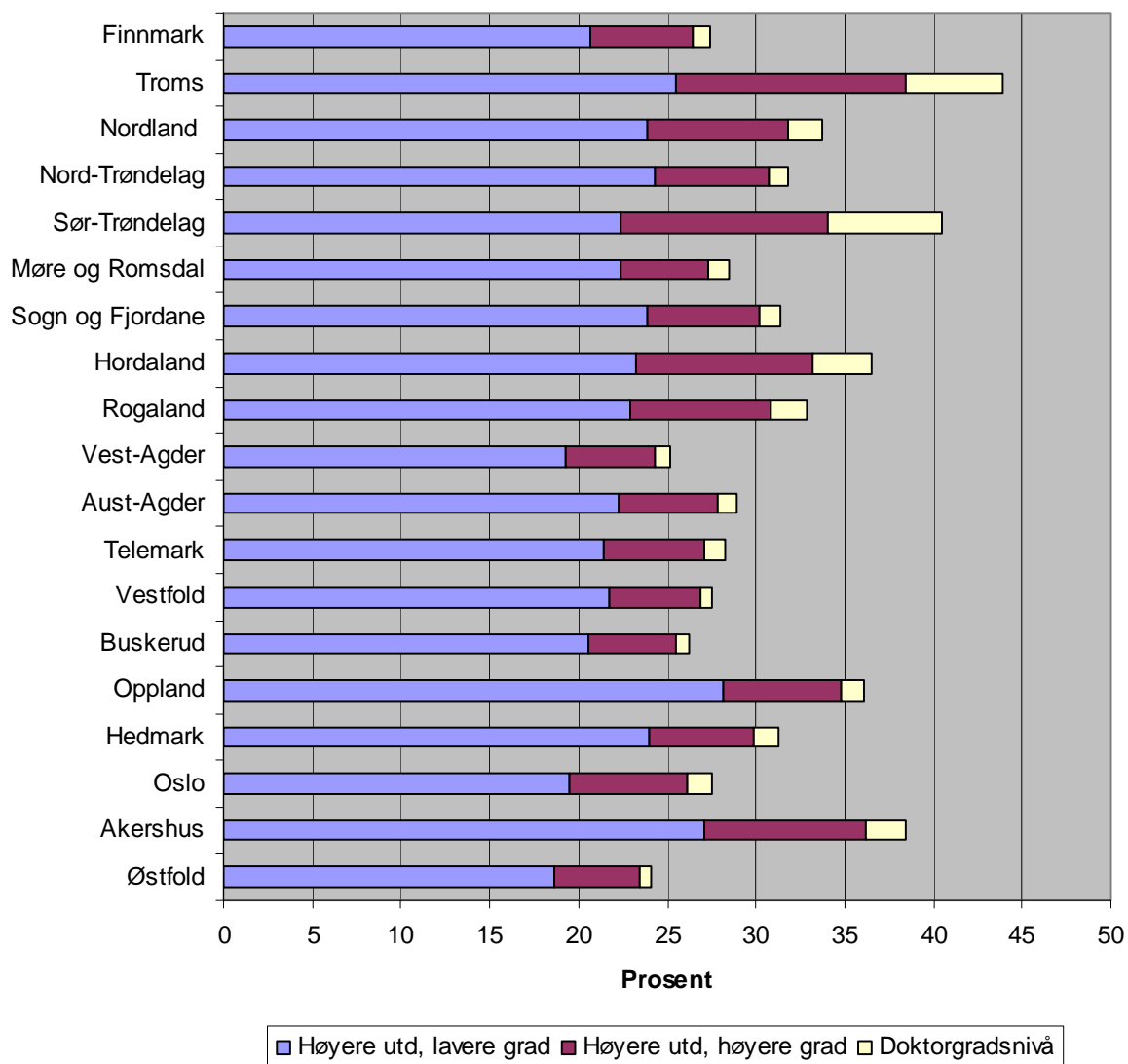
Figur B.2: Forskjeller i utdanningsnivå i spredtbygde strøk basert på årsintervaller



Figur B.3: Innvandrerbefolkningen med ingen utdanning, grunnskoleutdanning eller uoppgitt utdanningsnivå basert på fylker



Figur B. 4: Innvandrerbefolkningen med høyere utdanning basert på fylker



Tabell B. 1: Oppføring av antall personer i hvert fylke med uoppgitt utdanningsnivå.

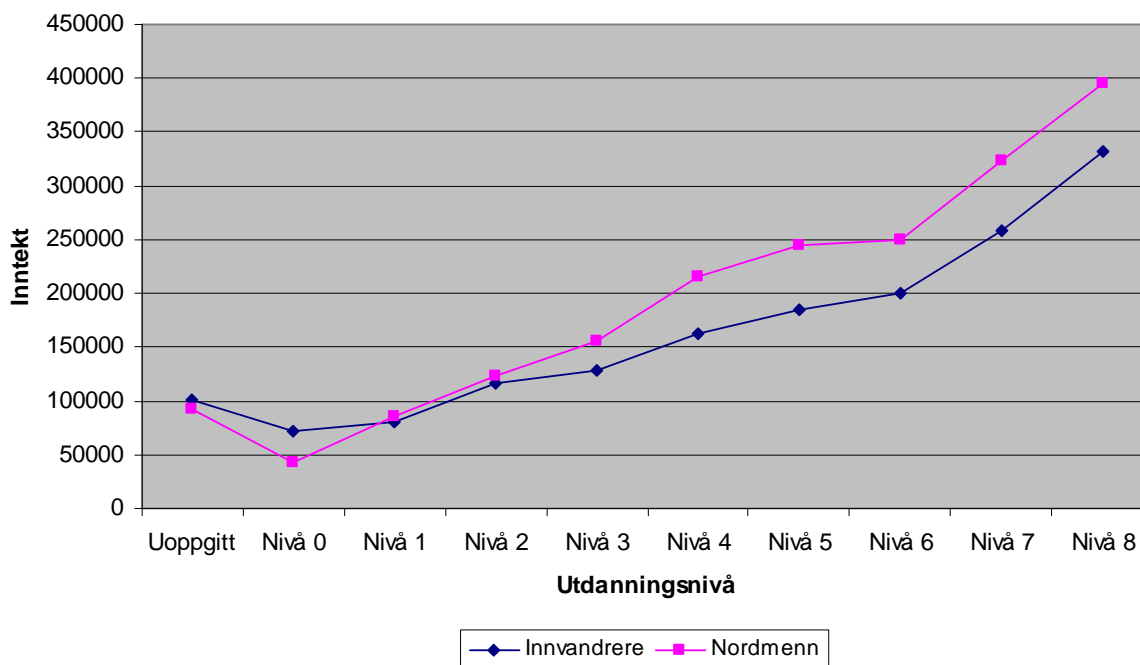
<i>Fylke</i>	<i>Hyppighet</i>	<i>Prosent</i>
Østfold	2,245	3.55
Akershus	5,729	9.05
Oslo	29,802	47.06
Hedmark	1,048	1.66
Oppland	912	1.44
Buskerud	2,849	4.50
Vestfold	1,376	2.17
Telemark	1,277	2.02
Aust-Agder	668	1.05
Vest-Agder	1,258	1.99
Rogaland	5,371	8.48
Hordaland	3,931	6.21
Sogn og Fjordane	576	0.91
Møre og Romsdal	1,028	1.62
Sør-Trøndelag	1,806	2.85
Nord-Trøndelag	511	0.81
Nordland	1,049	1.66
Troms	1,032	1.63
Finnmark	853	1.35
Total	63,321	100

Appendiks C

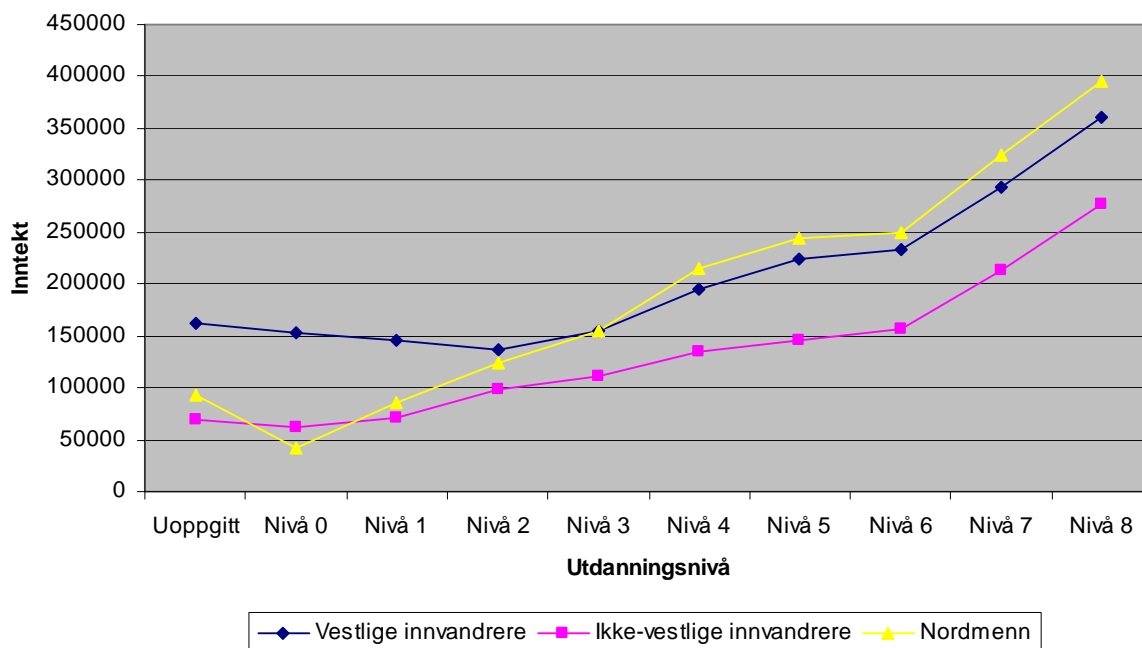
Sammenhengen mellom inntekt og utdanning

Jeg har her sammenlignet gjennomsnittlig inntekt på alle utdanningsnivå for nordmenn og for innvandrere, og for nordmenn, vestlige- og ikke-vestlige innvandrere. Videre har jeg gjort samme sammenligning for både kvinner og menn, med og uten innvandrerbakgrunn, og til slutt for de ulike landbakgrunnene jeg har brukt i datasettet. Utdanningsnivået som er oppgitt i figurene er det maksimale utdanningsnivået som er oppgitt i perioden 1993-2003.

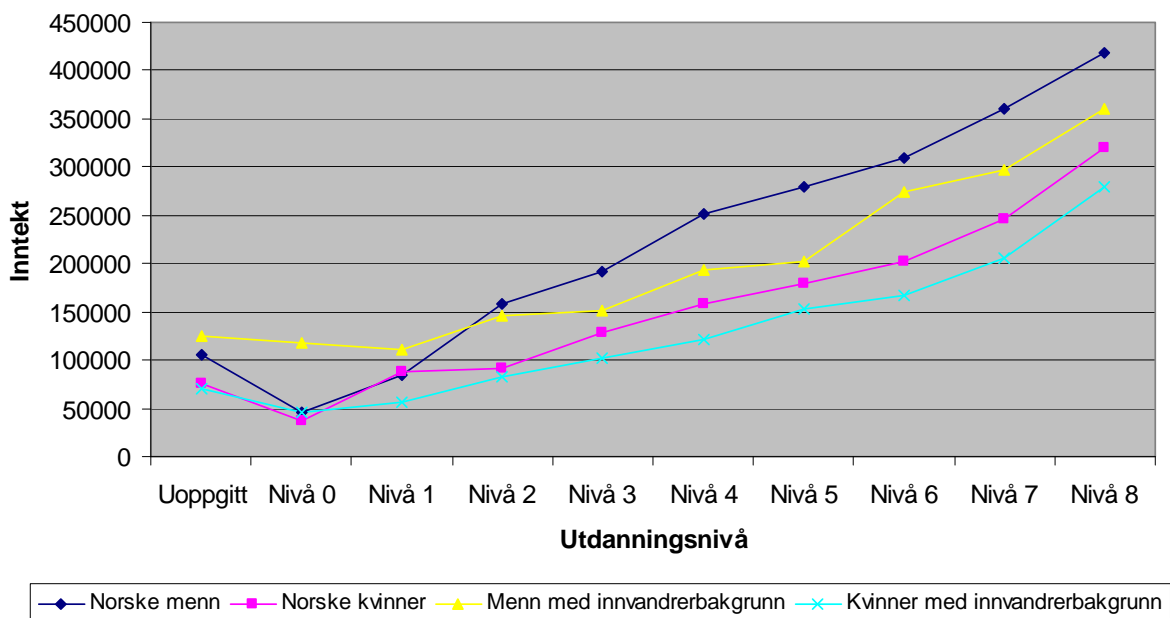
Figur C. 1: Gjennomsnittlig inntekt basert på utdanningsnivå



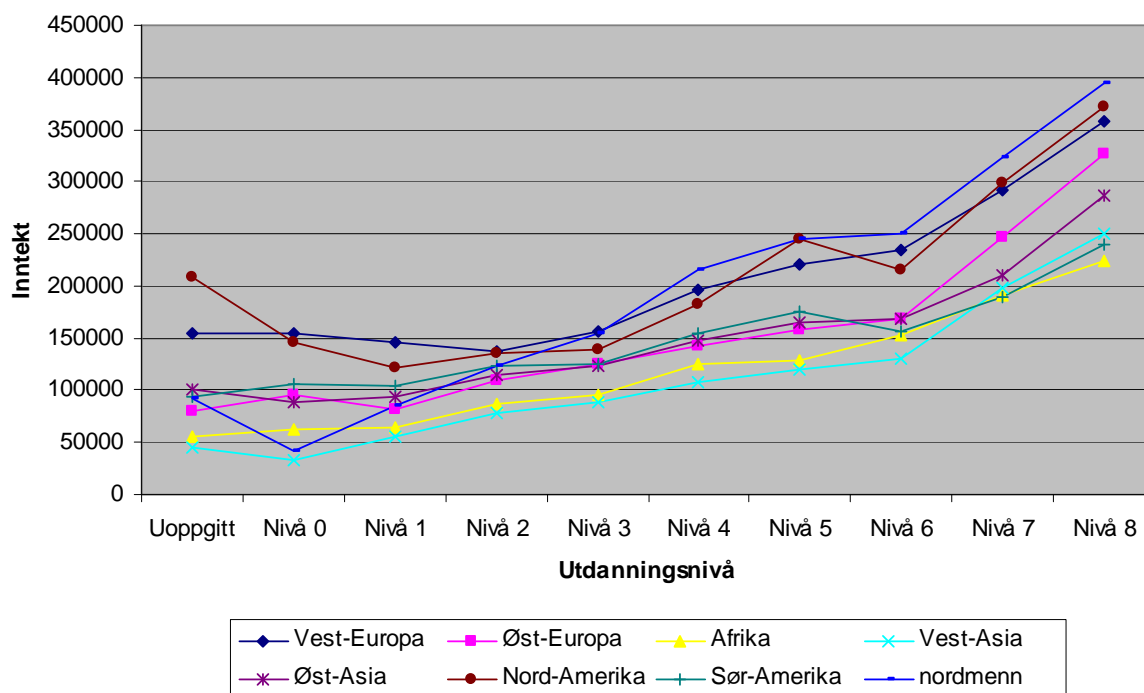
Figur C. 2: Gjennomsnittlig inntekt basert på utdanningsnivå, skille mellom vestlige- og ikke-vestlige innvandrere



Figur C. 3: Gjennomsnittlig inntekt basert på utdanningsnivå og kjønn



Figur C. 4: Gjennomsnittlig inntekt basert på utdanningsnivå og landbakgrunn



Appendiks D

Tabell D. 1: Resultater fra regresjon med probitmodell, positiv utdanning som avhengig variabel (i forhold til ingen utdanning og uoppgitt utdanningsnivå). Resultatene er oppgitt i marginaleffekter.

<i>Variabler</i>	<i>Probit kvinner</i>	<i>Probit menn</i>
<i>Vest-Europa</i>	0.1094 (0.0041)	0.0644 (0.0037)
<i>Øst-Europa</i>	0.0578 (0.0027)	0.0086 (0.0045)
<i>Afrika</i>	0.0113 (0.0046)	0.0167 (0.0038)
<i>Øst-Asia</i>	0.0553 (0.0028)	0.0127 (0.0036)
<i>Nord-Amerika</i>	0.0741 (0.0024)	0.0459 (0.0052)
<i>Sør-Amerika</i>	0.0594 (0.0030)	0.0355 (0.0045)
<i>Etterkommere</i>	-0.1124 (0.0543)	-0.0419 (0.0369)
<i>Etterkommere*gift</i>	0.0496 (0.0177)	0.0160 (0.0372)
<i>Barn under 16 år</i>	0.0137 (0.0061)	0.0030 (0.0056)
<i>Barn under 16 år²</i>	-0.0044 (0.0028)	-0.0026 (0.0023)
<i>Botid</i>	0.0047 (0.0006)	0.0056 (0.0006)
<i>Botid²</i>	0.0000 (0.0000)	-0.0001 (0.0000)
<i>Gift</i>	0.0102 (0.0036)	-0.0032 (0.0038)
<i>Tettbygd strøk</i>	0.0018 (0.0055)	0.0054 (0.0059)
<i>Norsk statsborger</i>	0.0451 (0.0036)	0.0630 (0.0040)
<i>Alder 21-26 år</i>	0.0336 (0.0039)	0.0299 (0.0039)
<i>Alder 27-31 år</i>	0.0233 (0.0039)	0.0266 (0.0037)
<i>Alder 32-36 år</i>	0.0234	0.0136

	(0.0038)	(0.0039)
<i>Alder 42-46 år</i>	-0.0293	-0.0101
	(0.0057)	(0.0049)
<i>Alder 47-51 år</i>	-0.0555	-0.0206
	(0.0073)	(0.0059)
<i>Alder 52-56 år</i>	-0.0855	-0.0387
	(0.0101)	(0.0079)
Antall observasjoner	44178	48917
LR Chi2 test	3638,44	1916,84
Pseudo R ²	0,1214	0,0617

- Estimaterne som ikke er signifikante på et 5 % signifikansnivå er satt i *kursiv*
- Fylkesdummyer er med i regresjonen, men ikke tatt med i tabellen pga mange insignifikante estimater
- Konstantleddet til regresjonene er ikke tatt med i tabellen
- Vest-Asia er sammenligningsgruppen
- Utvalget i analysen er et tverrsnittsutvalg fra periode 4

Appendiks E

Hvor passer de med uoppgitt utdanningsnivå best inn?

I tabell A.2 ser vi av LR-testen og pseudo R^2 (nederst i tabellen) at disse har en rimelig sterk verdi for alle utdanningsnivåene, og vi får derfor ikke noe sikkert svar på hvilket utdanningsnivå som er mest likt de med uoppgitt utdanningsnivå. Dette kan tyde på at de med uoppgitt utdanningsnivå i stor grad fordeler seg over alle nivåene. Jeg kunne i prinsippet ha gjort disse for hver verdensdel, men belyser dette heller ved å se på grupper med dummyer for de forskjellige verdensdelene. Regresjonen er kun gjort med datasettet for innvandrere, da det er i dette problemet med manglende utdanningsnivå ligger.

Tabell E. 1: Resultater fra regresjon med probit modell, uoppgitt utdanningsnivå som avhengig variabel i forhold til hvert av utdanningsnivåene

Variabler	Nivå 0	Nivå 1	Nivå 2	Nivå 3	Nivå 4	Nivå 5	Nivå 6	Nivå 7	Nivå 8
Vest-Europa	-0.5426 (0.0553)	-0.5445 (0.0491)	0.4339 (0.0352)	0.3711 (0.0307)	0.7057 (0.0318)	0.3968 (0.0631)	0.9385 (0.0310)	0.9758 (0.0403)	1.0628 (0.0621)
Øst-Europa	-0.8016 (0.0669)	-1.0200 (0.0604)	-0.0117 (0.0382)	0.0106 (0.0323)	0.3568 (0.0326)	-0.1697 (0.0723)	0.3040 (0.0329)	0.6856 (0.0413)	0.5197 (0.0699)
Afrika	0.0669 (0.0487)	-0.3544 (0.0473)	-0.0872 (0.0395)	0.2014 (0.0312)	0.2831 (0.0333)	0.3228 (0.0614)	0.2820 (0.0335)	0.2152 (0.0463)	0.1772 (0.0798)
Øst-Asia	0.1449 (0.0426)	-0.0191 (0.0380)	0.3177 (0.0322)	0.4103 (0.0262)	0.4413 (0.0283)	0.2762 (0.0544)	0.4707 (0.0281)	0.4400 (0.0386)	0.3531 (0.0656)
Sør-Amerika	-0.1769 (0.0836)	-0.3548 (0.0760)	0.1327 (0.0594)	0.3889 (0.0458)	0.7198 (0.0450)	0.3688 (0.0924)	0.6062 (0.0462)	0.5550 (0.0625)	0.3641 (0.1033)
Etterkommere	0.0280 (0.2679)	-1.3531 (0.4993)	-0.2501 (0.1108)	-0.1179 (0.0946)	0.0209 (0.0927)	0.0185 (0.1530)	0.0763 (0.0937)	0.1717 (0.1165)	0.4005 (0.1940)
Barn u. 16 år	-0.0903 (0.0205)	-0.0743 (0.0187)	-0.0611 (0.0153)	-0.0648 (0.0133)	-0.0827 (0.0138)	-0.1121 (0.0290)	-0.0606 (0.0135)	-0.0771 (0.0180)	-0.0286 (0.0278)
Kvinne	0.3681 (0.0314)	0.2636 (0.0280)	-0.0058 (0.0208)	-0.1087 (0.0177)	-0.0803 (0.0183)	-0.3132 (0.0381)	0.0818 (0.0179)	-0.0820 (0.0236)	-0.2505 (0.0370)
Botid	0.0007 (0.0026)	0.0026 (0.0024)	0.0489 (0.0015)	0.0374 (0.0014)	0.0325 (0.0015)	0.0550 (0.0030)	0.0242 (0.0015)	0.0228 (0.0019)	-0.0020 (0.0029)
Gift	0.0037 (0.0399)	0.1462 (0.0365)	0.1751 (0.0272)	0.0386 (0.0227)	0.1033 (0.0234)	0.1215 (0.0472)	0.1577 (0.0231)	0.2150 (0.0303)	0.3707 (0.0479)
Tettbygd strøk	0.0065 (0.0789)	-0.0790 (0.0680)	-0.0929 (0.0432)	0.0001 (0.0381)	-0.0036 (0.0378)	0.0404 (0.0762)	0.0558 (0.0366)	0.1348 (0.0493)	0.2308 (0.0745)
Norsk statsb.	0.3216 (0.0375)	0.4753 (0.0343)	0.3978 (0.0260)	0.5315 (0.0222)	0.4734 (0.0235)	0.5594 (0.0469)	0.4410 (0.0234)	0.3772 (0.0300)	0.0718 (0.0457)
Alder 21-26 år	-0.6216 (0.0621)	-0.4091 (0.0547)	0.3265 (0.0387)	0.4636 (0.0318)	0.4335 (0.0331)	0.9469 (0.0681)	-0.0278 (0.0334)	-0.0323 (0.0447)	-0.6651 (0.0802)
Alder 27-31 år	-0.3451 (0.0512)	-0.1198 (0.0455)	0.1444 (0.0354)	0.3692 (0.0291)	0.3688 (0.0301)	0.8041 (0.0633)	0.1264 (0.0290)	0.0928 (0.0387)	-0.4666 (0.0610)
Alder 32-36 år	-0.1705 (0.0479)	-0.0275 (0.0432)	0.0909 (0.0341)	0.2638 (0.0285)	0.2786 (0.0294)	0.5757 (0.0620)	0.1753 (0.0279)	0.1929 (0.0369)	-0.1718 (0.0552)
Alder 42-46 år	0.0217 (0.0501)	0.0465 (0.0458)	-0.0287 (0.0345)	-0.0665 (0.0310)	-0.0495 (0.0317)	-0.1567 (0.0695)	-0.0273 (0.0296)	-0.0245 (0.0387)	-0.0084 (0.0553)
Alder 47-51 år	0.1832 (0.0539)	0.0165 (0.0523)	0.0406 (0.0370)	-0.0965 (0.0344)	-0.1130 (0.0355)	-0.1780 (0.0748)	-0.1487 (0.0337)	-0.0677 (0.0433)	0.0151 (0.0617)
LR Chi2 test	1252.99	1559.82	3081.79	3255.18	3048.12	1356.24	3298.5	2137.22	1524.68
Antall obs	9269	10325	17419	27695	24718	8468	27017	13289	8509
Pseudo R ²	0.1221	0.1205	0.1312	0.1038	0.1033	0.1753	0.1065	0.1163	0.1935

- Estimatene som ikke er signifikante på et 5 % signifikansnivå er satt i kursiv
- Fylkesdummyer er med i regresjonen, men ikke tatt med i tabellen pga mange insignifikante estimater
- Konstantleddet til regresjonene er ikke tatt med i tabellen
- Vest-Asia er sammenligningsgruppen

Referanser

Akselsen, Anders, Lien, Sandra og Sivertstøl, Øyvind (2007) FD-trygd Variabelliste. *Notater 2007/17*, Statistisk sentralbyrå.

Arbeids- og inkluderingsdepartementet (2003) Introduksjonsloven. *Lov om introduksjonsordning og norskopplæring for nyankomne innvandrere*. URL: <http://www.lovdatab.no/all/hl-20030704-080.html>.

Arbeids- og inkluderingsdepartementet (2006) Lov om introduksjonsordning og norskopplæring for nyankomne innvandrere (introduksjonsloven). URL: <http://www.utdanningsforbundet.no/upload/51154/AID%20Introduksjonsloven%20og%20finansieringsordninger%20mars%202006.ppt#257,1,Lov>

Barne- og likestillingsdepartementet (2006) Registrert partnerskap - lov 30. april 1993 nr 40 om registrert partnerskap. URL: http://www.regjeringen.no/upload/kilde/bfd/bro/2001/0020/ddd/pdfv/284780-partnerskap_2006.pdf.

Barth, Erling, Bratsberg, Bernt og Raaum, Oddbjørn (2004) Mulighetenes land? *Tidsskrift for samfunnsforskning*, Vol 45, 637-658.

Borjas, George J. (1987) Self-Selection and the Earnings of Immigrants. *The American Economic Review*, Vol. 77, 531-553

Borjas, George J. (1991) Immigration Policy, National Origin, and Immigrant Skills: A Comparison of Canada and The United States. *NBER Working Papers Series*, No. 3691.

Borjas, George J. (1994) The Economics of Immigration. *Journal of Economic Literature*, Vol 32, 1667-1717.

Borjas, George J. (1996) *Labour Economics*, New York, McGraw Hill/ Irwin.

Borjas, George J. (1999) The Economic Analysis of Immigration. *Handbook of Labour Economics*, Volume 3, 1697-1760.

Bratberg, Espen (2004) *Econometric models for discrete outcomes, Lecture note*, Department of Economics, University of Bergen.

Bratsberg, Bernt, Raaum, Oddbjørn og Røed, Knut (2006) The Rise and Fall of Immigrant Employment: A Lifecycle Study of Labor Migrants to Norway. *The Ragnar Frisch Centre for Economic Research, Frisch project 1391*.

Bratsberg, Bernt, Raaum, Oddbjørn og Røed, Knut (2007) Opp og ned: Yrkesaktivitet og trygd over livsløpet for tidlige arbeidsinnvandrere i Norge *IMDI-rapport nr. 9*, Intergreringskart 2007; Arbeidsinnvandring - En kunnskapsstatus.

- Card, David (1995) Earnings, Schooling, and Ability Revisited. *Research in Labour Economics*, Vol. 14, 23-48.
- Chriswick, Barry R. (1978) The Effect of Americanization on the Earnings of Foreign-Born Men. *The Journal of Political Economy*, Vol. 86, 897-921.
- Cortes, Kalena E. (2004) Are refugees different from economic immigrants? Some empirical evidence on the heterogeneity of immigrant groups in the United States. *The Review of Economics and Statistics*, 86(2).
- Dalheim, Elisabeth (2001) Med utdanning i bagasjen. *Samfunnsspeilet nr. 2*, Statistisk sentralbyrå.
- Dalheim, Elisabeth (2002) En skjemabasert komplettering av registeret over befolkningens høyeste utdanning - Opplysninger om opplæring, skolegang og utdanning 1999. *Notater 2002/53*, Statistisk sentralbyrå.
- Daugstad, Gunnlaug (2006) Innvandring og innvandrere 2006. *Statistiske analyser*, Statistisk sentralbyrå.
- Djuve, Anne Britt og Friberg, Jon Horgen (2004) Innvandring og det flerkulturelle samfunn. *Fafo-notat 2004:32*.
- Dæhlen, Marianne (2001) Sosial bakgrunn betyr mer enn innvandrerbakgrunn. *Samfunnsspeilet nr. 2*, Statistisk sentralbyrå.
- Friedberg, Rachel M. (2000) You Can't Take It With You? Immigrant Assimilation and the Portability of Human Capital. *Journal of Labour Economics*, Vol. 18, 221-251.
- Jørgensen, Tor (2006) Nye definisjoner av utdanningsnivåer. URL: http://www.ssb.no/vis/magasinet/slik_lever_vi/art-2006-09-14-01.html.
- Kunnskapsdepartementet (2007) Likeverdig opplæring i praksis! Strategi for bedre læring og større deltakelse av minoritetsspråklige elever i barnehage, skole og utdanning 2007-2009, Revidert utgave. *Strategiplan*.
- Lie, Benedicte (2003) Store forskjeller i innvandreres utdanningsnivå. *Samfunnsspeilet nr. 3*, Statistisk sentralbyrå.
- Lie, Benedicte (2004) Fakta om ti innvandrergrupper i Norge. *Rapporter*, 2004/14.
- Nygård, Geir og Daugstad, Gunnlaug (2006) Store forskjeller i utdanningsnivået blant innvandrere i Norge. URL: <http://www.ssb.no/vis/magasinet/blandet/art-2006-05-09-01.html>.
- Olsen, Bjørn (2006) Er innvandrerungdom en marginalisert gruppe? *Samfunnsspeilet nr. 4*, Statistisk sentralbyrå.
- Raaum, Oddbjørn (1999) Inntektseffekter av utdanning i Norge - en litteraturoversikt. *Arbeidsnotat 2/1999*. Frischsenteret.

- Roy, A. D. (1951) Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers, New Series*, Vol 3, 135-146.
- Statistisk sentralbyrå (2000) Norsk standard for utdanningsgruppering, Revidert 2002. *Norges offisielle statistikk*.
- Statistisk sentralbyrå (2001) Mange innvandreregrupper mer utdannet enn resten av folket. *Utdanningsstatistikk*, URL: <http://www.ssb.no/vis/emner/04/01/utinnv/main.html>.
- Statistisk sentralbyrå (2004) Økonomi og levekår for ulike grupper 2003. *Rapporter 2004/2*. Oslo-Kongsvinger.
- Statistisk sentralbyrå (2007) Innvandring1 etter innvandringsgrunn og innvandringsår. 1990-2007. URL: <http://www.ssb.no/innvgrunn/tab-2008-10-02-01.html>.
- Statistisk sentralbyrå (2008a) Innvandring etter innvandringsgrunn og statsborgerskap. 1990-2007. URL: <http://www.ssb.no/emner/02/01/10/innvgrunn/tab-2008-10-02-02.html>.
- Statistisk sentralbyrå (2008b) Statistikkbanken. URL: <http://statbank.ssb.no/statistikkbanken/>.
- United Nations (2006) Press Release; Secretary-General sees UN as site for exchange of views on migration, in new report out today. URL: www.un.org/esa/population/migration/hld/Text/Migration_press%20release.pdf
- Utlendingsdirektoratet (2008) Statsborgerskap. URL: <http://www.udi.no/templates/Tema.aspx?id=2044>.
- Verbeek, Marno (2004) *A Guide to Modern Econometrics*, Chichester, West Sussex, John Wiley & Sons Ltd.
- Wooldrige, Jeffrey M. (2006) *Introductory Econometrics - A Modern Approach*, Mason, Ohio, USA, Thompson South-Western.
- Østby, Lars (2004) Innvandrere i Norge - Hvem er de , og hvordan går det med dem? *Notater 2004/66*, Statistisk sentralbyrå.