

Forord

Arbeidet med oppgaven har vært både spennende og til tider krevende. Professor Kåre Johansen fortjener en stor takk for å ha veiledet meg gjennom de ulike fasene i oppgaveskrivingen. Veiledningen har vært grundig, inspirerende og humørfyllt. Ønsker også å si takk for god tilrettelegging i startfasen, noe som bidro til at tankeprosessen kom igang. Tilbakemeldingene har hele tiden vært konstruktive og dette har ført til at skriveprosessen har utviklet seg til å bli meget lærerik. Takk også til klassekameratene for et godt klassemiljø og for nyttige fagdiskusjoner underveis.

Trondheim 25.mai 2011

Stig Fiske Strand

Innholdsfortegnelse

1 INNLEDNING	1
1.1 Disponering av oppgaven	5
2 LITTERATUROVERSIKT	6
2.1 Internasjonal teori.....	6
2.2 Interregional migrasjon i Norge	9
2.3 Oppsummering.....	13
3 EMPIRISK TILNÆRMING	14
3.1 Spesifikasjon av statistisk paneldatamodell.....	14
3.1.1 Funksjonsform.....	15
3.2 Presentasjon av datamaterialet	16
3.2.1 Forklaring av inkluderte variabler	17
3.3. Estimeringsmetode	20
3.3.1 Fixed effects-estimering	21
3.3.2 Random effects-estimering.....	23
3.3.3 Test av koeffisientforskjeller	25
4 EMPIRISKE RESULTATER	26
4.1 Resultater fra fixed effects-estimering.....	27
4.1.1 Grunnmodell. Arbeidsledighet og vakanser.....	27
4.1.2 Utvidet modell. Den nasjonale konjunktursituasjonen.....	30
4.1.3 Utvidet modell. Kommunale inntekter	33
4.1.4 Fullstendig modell	35
4.1.5 Diagnostiske tester	38
4.1.6 Oppsummering.....	40
4.2 Sensitivitetsanalyse	41
4.2.1 Utvalgte tidsperioder	41
4.2.2 Boligpris.....	43
5 UTVIDELSER – dynamisk modellspesifikasjon og GMM	45
5.1 Re-spesifikasjon av fullstendig modell	45
5.1.1 Resultater fra dynamisk fixed effects-estimering	46
5.2 Generalisert momentmetode	49
5.2.1 Arellano og Bonds instrumentvariabelmetode.....	50
5.3 Resultater fra GMM-analysen	51
5.4 Oppsummering.....	56
6 AVSLUTTENDE DISKUSJON OG KONKLUSJON	57
REFERANSELISTE	61
APPENDIKS	64

1 Innledning

Hvordan vil regionale arbeidsmarkedsvariable påvirke flytting mellom norske regioner for ulike utdanningsgrupper? Vil den nasjonale konjunktursituasjonen påvirke flytteresponsen? Dette er sentrale problemstillinger i denne empiriske studien, der vi skal benytte paneldata for regionale arbeidsmarkedsvariabler i perioden 1986-2004 som et utgangspunkt for den økonometriske analysen. Vi estimerer både statiske og dynamiske modeller ved hjelp av faste og stokastiske effekter, og avslutningsvis utvides analysen med generalisert momentmetode.

I denne studien skal vi se på flyttestrømmer for tre ulike utdanningsgrupper som inneholder personer i alderen 16-66 år. Disse tre utdanningsgruppene er grunnskole, videregående skole og høyere utdanning. I den første gruppen, grunnskole, har personene kun utdanning fra grunnskolenivå. For gruppen videregående skole har de utdanning fra grunnskole og videregående skole, mens de i gruppen høyere utdanning har utdanning fra grunnskole, videregående og høyere utdanning.

En undersøkelse hentet fra statistisk sentralbyrå (SSB) i 2009 viser at omtrent 23 prosent av befolkningen i Norge som tilhører aldersgruppen 16 år og over har en utdanning på grunnskolenivå. For videregående skole og høyere utdanning har vi henholdsvis 43 og 34 prosent.¹ Elevtallet stiger både i grunnskolen og i videregående opplæring, og antallet som fortsetter i høyere utdanning øker for hvert år. Av de nordiske landene er det bare Finland som har større andeler med universitets- og høgskoleutdanning enn Norge, med 36 prosent².

Tall fra Statistisk sentralbyrå (SSB) viser at i 2009 flyttet 201 000 personer mellom kommuner i Norge og dette er en økning på 2000 ifra 2008.³ Mobiliteten var omtrent på samme nivå som året før, men et hovedtrekk er at flyttingene var mindre sentraliserte enn på flere år. Høsten 2008 begynte flyttingene å avta og det var i perioden 2008-2009 ventet at mobiliteten skulle bli lavere, særlig på grunn av finanskrisen. Man vet i ettertid at finanskrisen ikke fikk noen sterke realøkonomiske effekter i Norge, og en mulig dempende effekt på ledigheten kan være det økende antallet av unge under utdanning. Totalt sett flyttet vi like mye som i perioden 2003-2005. Tall fra begynnelsen av 1990-tallet, da det sist var en

Kilder:

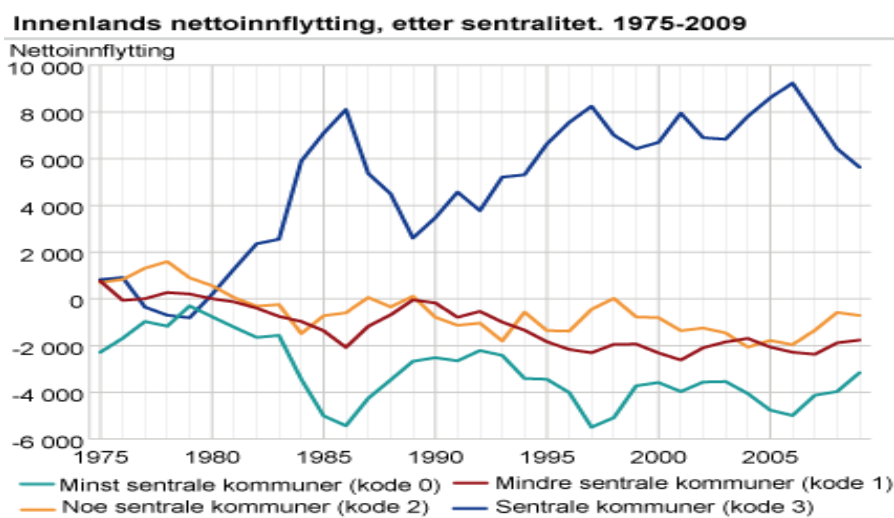
¹ <http://www.ssb.no/emner/04/01/utniv/>

² <http://www.ssb.no/emner/04/01/utniv/>

³ <http://www.ssb.no/emner/02/02/20/flytting/>

nedgang for norsk økonomi, viser at mobiliteten sank betydelig mer i denne perioden. Sentraliseringen har avtatt noe i de siste årene, likevel er det hvert eneste år en nettostrøm i befolkningsflyttingene til de mest sentrale delene av landet. Som et eksempel på dette kan det trekkes frem at i løpet av årene 1950-2009 har de innenlandske flyttingene gitt netto innflytting til Akershus og Oslo på ca 62 000, mens Sør-Østlandet har hatt en flyttegevinst på ca 90 000 i denne perioden. For Nord-Norge sin del har de hatt et netto flyttetap på ca 120 000, noe som tilsvarer ganske nær en fjerdedel av dagens befolkning på 464 649.⁴

Figur 1 viser at gjennom de siste 34 årene har innenlands nettoinnflytting vært størst i de sentrale regionene i Norge. Dette illustreres ved den mørkeblå linjen.



Kilde: <http://www.ssb.no/emner/02/02/20/flytting/>

Ifølge forsker Lars Østby (2002) ligger nordmenn på flyttetoppen sammenlignet med andre europeiske nasjoner. Danmark er det landet med høyest mobilitet, mens land som Norge, Sverige, Sveits og Finland følger deretter. Den høye mobiliteten i Norge fører til at de regionale forskjellene i arbeidsledigheten er relativt beskjedne sammenlignet med andre europeiske land, samtidig som den høye mobiliteten kan gi lav gjennomsnittlig arbeidsledighet på landsbasis. For Norge var ledigheten iløpet av 2007 lavest i Rogaland med 1,4 % og høyest i Finnmark med 3,7 %.⁵ Sammenlignet med land som Italia der ledigheten er ca 10 % høyere i Sør-Italia enn i Nord-Italia, ser man at de regionale forskjellene er mye mindre i Norge. Dette skyldes selvsagt at Norge har relativt høy mobiltet, men man kan heller ikke komme utenom den sterke økonomien som gjør det mulig å tilrettelegge tiltak for å

Kilder:

⁴ <http://www.ssb.no/emner/02/02/20/flytting/tab-2010-05-06-04.html>

⁵ <http://statbank.ssb.no/statistikkbanken/>

minske regionale forskjeller. Et eksempel som kan være med å bidra til lavere regionale forskjeller er at hvis man flytter til Nord-Troms eller Finnmark, kan de som har studielån få redusert studielånet med maksimalt med kr 25 000 per år⁶. Dette vil i størst grad være rettet mot nyutdannede, som i tillegg kan antas å ha høyere flyttetilbøyelighet enn resten av befolkningen. Det kan sies at Norge har en generelt aktiv regionalpolitikk, noe som kan bidra til lavere migrasjon mellom regionene og kan på flere måter virke som en dempende effekt på fraflytting fra ugunstige regioner.

Selv mener Østby at dette ikke nødvendigvis er gitt at høy mobilitet er et gode eller et onde, verken for den enkelte eller for samfunnet. Høy mobilitet i et samfunn vil ofte bety at vi har stor evne til å omstille oss, og det er nettopp dette som er et kjennetegn på den økonomiske veksten i Norge i etterkrigstida. Mye av veksten var da knyttet til flytting fra landsbygda. Høy mobilitet virker også positivt gjennom at det trekker i retning av lav likevektsledighet. På en annen side vil et samfunn med høy mobilitet bli mindre tett og føles mindre trygt. De sosiale kontrollene blir løsere, og kriminalitet og andre avvik kan få bedre grobunn. Lav mobilitet kan fort gi et samfunn uten dynamikk og forandring, men samtidig kan det bli preget av trygghet og oversiktighet.⁷

Et arbeidsmarked består av personer med både lav og høy utdanning. Dette vil være personer som har ulik produktivitet, og sammensetningen av flyttestrømmen for ulike kommuner og regioner vil dermed ikke være helt uten betydning. Hvis det er store forskjeller i den økonomiske utviklingen mellom regionene vil de med høyere utdanning mest sannsynlig flytte til regioner med størst økonomisk utvikling. Dette kan føre til at arbeidsledigheten vokser i ugunstige regioner og ikke minst vil den økonomiske utviklingen bli forverret i forhold til gunstige regioner. Man risikerer da å sitte igjen med få høyt utdannede og mange lavt utdannede i enkelte regioner, mens andre vil ha en betydelig større andel av høyt utdannede personer med høy produktivitet tilgjengelig. En utvikling som dette kalles ofte for "brain drain" og betyr at individer med høy utdanning og høy produktivitet migrerer til de mest attraktive regionene. Dette gir dermed grunnlag for at det blir interessant å undersøke hvordan og hvilke effekter som påvirker ulike utdanningsgrupper, som igjen vil ha betydning for flyttemønsteret mellom regionene.

Kilder:

⁶ <http://www.lanekassen.no>

⁷ <http://www.ssb.no/emner/02/02/20/>

Norske studier av nyere dato som for eksempel Carlsen, Johansen og Kaspersen (2007) og Carlsen, Johansen og Stambøl (2010) konkluderer med at ubalanser i det regionale arbeidsmarkedet har en signifikant effekt på interregional migrasjon. Med hovedvekten på hvordan flytting skjer mellom regioner, ønsker vi å se om det er noen forskjell mellom ulike utdanningsgrupper og samtidig se på muligheten for at den nasjonale konjunktursituasjonen kan påvirke effekten av regionale arbeidsmarkedsvariable. I perioden 1980 – 2000 har vi hatt to perioder med sterkt utflytting fra disktriktene. Mønsteret viser til å henge sammen med konjunktursituasjonen og en trend viste til at svingninger i økonomien skapte et sentraliserende flyttemønster⁸. Vi kan da spørre; Hvis det tas hensyn til regionale arbeidsmarkedsvariable, er det en forskjell mellom ulike utdanningsgrupper når det gjelder interregional migrasjon i Norge? Bli effektene av disse arbeidsmarkedsvariablene påvirket av den nasjonale konjunktursituasjonen?

Min analyse bygger i utgangspunktet på tre tidligere studier som tar for seg flytting mellom regioner. Disse studiene er Dahl (2007), Carlsen, Johansen og Kaspersen (2007) og Carlsen, Johansen og Stambøl (2010). Hovedpoengene i disse er følgende:

Dahl (2007) ser på tre forskjellige utdanningsgrupper og undersøker hvordan utdanningsnivået har betydning for den regionale migrasjonen i Norge. Ved hjelp av to estimeringsmetoder kontrolleres det for ulike arbeidsmarkedsvariabler og konklusjonen er at mobiliteten er stigende i utdanningsnivået. Disse to estimeringsmetodene forklares mer detaljert i kapittel 3. I artikkelen til Carlsen, Johansen og Kaspersen (2007) blir det sett på hvordan arbeidsmarkedsvariable påvirker flytting samtidig som det kontrolleres for effekten av konjunktursituasjonen. De ser ikke på forskjellen mellom ulike utdanningsgrupper, men har netto innflytting for hele befolkningen som avhengig variabel. Ved bruk av differensiering kontrolleres det her for tidsinvariante forhold som har betydning for flytting. De konkluderer med at Norge ser ut til å ha høy mobilitet i europeisk målestokk, i tillegg til at følsomheten for regional arbeidsledighet er betydelig større enn i EU.

I likhet med Dahl (2007) ser Carlsen, Johansen og Stambøl (2010) på hvordan effekter på det regionale arbeidsmarkedet påvirker migrasjonsstrømmene ved å ta hensyn til utdanningsnivået. Her legges det vekt på mobiliteten for de i aldersgruppen 20-66 og det undersøkes om det er forskjell mellom utdanningsgruppene ved å se på sjokk i markedet.

⁸ Kilde: www.regjeringen.no; NOU 2003: 24

Metoden fixed effects benyttes og det kommer frem at det er signifikante forskjeller blant de ulike utdanningsgruppene. De med høyere utdanning responderer mye sterkere på effekter av regional ledighet sammenlignet med personer med kun grunnskoleutdanning.

Min analyse vil fokusere på forskjeller mellom ulike utdanningsgrupper når vi ser på flytting mellom regioner og om den nasjonale konjunkturen påvirker effekten av regionale arbeidsmarkedsvariable. I de økonometriske modellene som benyttes i oppgaven vil netto innmigrasjonsrate for utdanningsgruppene være den avhengige variabelen. I oppgaven tas det i bruk paneldata der Norge er delt inn i 90 regioner for perioden 1986-2004. Paneldata gir oss muligheten til å se på utviklingen over tid for hver region og samtidig kan paneldatasettene gi informasjon om ulikheter mellom regionene i ønsket tidsperiode.

1.1 Disponering av oppgaven

Denne oppgaven består av 6 kapitler og er organisert på følgende måte: I *kapittel 2* gis en oversikt over tidligere undersøkelser og det presenteres empiriske funn som ser på sammenhengen mellom interregional migrasjon og arbeidsmarkedsvariable - både internasjonalt og i Norge. *Kapittel 3* gjør rede for den økonometriske modellen, beskriver datamaterialet og ser på metodebruk. I *kapittel 4* presenteres og tolkes empiriske estimeringsresultater av statiske modeller, samt diagnostiske tester og det utføres en sensitivitetsanalyse av utvalgte resultater. *Kapittel 5* utvider analysen ved å estimere en dynamisk modellspesifikasjon samtidig som generalisert momentmetode (GMM) benyttes. *Kapittel 6* gir diskusjon og konklusjon.

2 Litteraturoversikt

I dette kapitlet som tar for seg litteraturoversikten blir det fokusert på internasjonale studier som omhandler teori og empiriske undersøkelser rettet mot interregional migrasjon, samt migrasjonen sett i sammenheng med utdanningsnivået og konjunktursituasjonen. I Norge er det skrevet et begrenset antall undersøkelser om dette emnet, men internasjonalt er det en bredere litteratur der temaet migrasjon blir undersøkt og diskutert. Med migrasjon menes interregional migrasjon for videre lesing av denne oppgaven.

Denne studien tar sikte på å undersøke ulikheter i migrasjonsresponsen blant forskjellige utdanningsgrupper og da med tanke på effekten av migrasjonen mellom de regionale arbeidsmarkedene. De fleste europeiske studier finner at effekten av regionale arbeidsmarkedsforhold er svak eller at det ikke er noen effekt. De studiene som finner signifikante effekter på det regionale arbeidsmarkedet får at regioner med høy ledighet opplever netto utmigrasjon, samtidig som det viser seg at vakanseraten og lønnsnivået vil påvirke migrasjonen.

Resultatene fra de norske studiene tyder på at arbeidskraftmobiliteten i Norge er høyere sammenlignet med andre EU-land. Analyser viser at forskjeller i det regionale arbeidsmarkedet har signifikante effekter på netto innmigrasjon i norske regioner. Den nasjonale konjunktursituasjonen viser seg også å være en viktig faktor for påvirkningskraften på de ulike arbeidsmarkedsvariablene som relateres til flyttestrømmene. Studier påpeker også at det er signifikante forskjeller mellom ulike utdanningsgrupper når det gjelder responsen på de ulike arbeidsmarkedsvariablene.

2.1 Internasjonal litteratur

Det første delkapitlet ser nærmere på internasjonal empiri om migrasjon og skal derfor knytte sammen ulike teoretiske bakgrunner fra tidligere studier omkring dette temaet. Gjennom de siste tiårene har tilgangen til mikro- og paneldata økt. Samtidig har nye måter å gjennomføre regresjoner på bidratt til at flere interessante sammenhenger knyttet til migrasjonen har blitt avdekket. Det eksisterer flere teorier om arbeidskraftmobilitet og det blir sett nærmere på hvilke faktorer som har en utslagsgivende effekt på migrasjonen.

En av de enkleste fremstillingene av årsaken til flytting mellom regioner er formulert av Ravenstein (1989). Han ser på migrasjon som et resultat av personlig nyttemaksimering gitt en budsjettskranke, og ulike lønnsforskjeller kan dermed være en motivasjon for å flytte til en annen region. Andre forskere har poengtert at ikke alle beslutninger baserer seg på det pengemessige grunnlaget, men at individets individuelle preferanser til for eksempel familie og omgivelser kan påvirke flyttebeslutningen. En av de som poengterte dette tidlig var Sjaastad (1962) som hadde vinklingen om at migrasjon kan settes inn i et investeringsperspektiv. Individet veier nytteøkning versus kostnadene ved å foreta en flyttebeslutning. Det er akkurat dette som beskriver humankapitalteorien, dvs at et individ vil ønske å flytte til en annen region hvis nytten er større enn kostnadene. Ehreberg og Smith (2000) diskuterer også humankapitalteorien, men i tillegg forventer de at mobiliteten er høy blant unge mennesker som tar utdanning. Yngre vil dermed være mer mobile enn eldre og det skyldes både psykiske og geografiske forhold, som for eksempel at yngre er i startfasen av arbeidslivet samtidig som de antas å være mindre sensitive til lønn og bosted. Det blir dermed naturlig å anta at disse faktorene generelt vil kunne bidra til økt migrasjon.

Jackman og Savouri (1992) tar utgangspunkt i en enkel forklaring av migrasjon og sier at det handler om jobbmatching. Arbeidere ønsker i realiteten å flytte fra region_i til region_j om de mottar et bedre jobbtilbud, som derav gir økt nytte. I en studie fra Storbritannia som fokuserer på 10 regioner, kommer forskerne frem til at migrasjon er en mekanisme som i mange tilfeller kan skape ubalanser i arbeidsmarkedet. Denne studien sår tvil om hvor effektiv migrasjonen er når det gjelder å skape likevekt i markedet. Et annet element, som er en sentral problemstilling i min oppgave, er at ifølge forskerne viser migrasjonen tendenser til å følge nasjonale konjunkturer ved at migrasjonen er lav i dårlige tider og betydelig større i gode økonomiske tider. Dette har en sammenheng med andelen av ledige jobber og det faktum at bedrifter tilpasser seg markedsforholdene. De konkluderer derfor med at regionale forskjeller i ledigheten viser seg å være størst i lavkonjunkturer og dette gir at når behovet for migrasjon øker, er migrasjonen på sitt laveste for de ulike regionene. Dette er grunnen til at migrasjonens evne til å skape likevekt blir sett på som lite effektiv.

I en artikkel om ledighet og interregional mobilitet argumenterer Pissarides og Wadsworth (1989) for at arbeidsledigheten påvirker mobiliteten på tre måter. For det første er jobbsituasjonen for arbeiderne viktig. En arbeidsledig person kan være mer interessert i å flytte enn en som allerede er sysselsatt. Arbeidsledige antas å ha mindre goder å gi slipp på i

form av økonomiske insentiver, mens en sysselsatt må kompenseres gjennom økt nytte i forhold til å bli i regionen. Det andre er at betydelige forskjeller mellom regioner gjør at en arbeider med større sannsynlighet ønsker å flytte hvis han bor i en region med høy ledighet. Det kan forklares ut ifra at en region med høy ledighet i mange tilfeller vil oppleve netto utmigrasjon. Undersøkelser viser at en region med over gjennomsnittet høy ledighet har under gjennomsnittet lav ansettelsesprosent. Dette fører til ulik attraktivitet blant regionene. Det siste momentet er at høyere total ledighet på landsbasis fører til mindre migrasjon mellom regioner. Det tilbys færre jobber og de som er uten jobb finner de mindre interessant å flytte til en annen region hvis de med stor sannsynlighet forblir arbeidsledige. Arbeiderne blir mindre verdifulle på grunn av stort tilbud relativt til etterspørselen, samtidig som de som allerede har en jobb ser på den som mer verdifull enn i perioder med lav ledighet.

En ofte benyttet antagelse er at lønnen øker relativt til utdanningsnivået. Sysselsatte og da særlig høyt utdannede antas generelt å være opptatt av endringer i lønna, og dette gir da at regioner med lav ledighet og en stor andel av høyt utdannede i mange tilfeller vil oppleve netto innmigrasjon. Dette fordi arbeidere i regioner med høy lønn er lite tilbøyelig til å flytte, mens arbeidere i andre regioner med lavere lønn ønsker å flytte inn i regioner med gode økonomiske vilkår. Dette kommer også frem i undersøkelsen fra 1976-77 og 1983-84 som Pissarides og Wadsworth (1989) benyttet, der det slås fast at effekten er signifikant og at man får en positiv sammenheng mellom migrasjon og utdanningsnivå. Mobiliteten virker også å være stigende i graden av utdanningsnivået innenfor ulike grupperinger i arbeidsmarkedet.

McMaster og Pissarides (1990) finner i sin studie at regioner med høy ledighet vil ha netto utmigrasjon av arbeidskraft til regioner med lav ledighet. Denne migrasjonen vil fortsette helt til forskjellene mellom regionene utjevnes over tid. I og med at denne likevektsprosessen er treg, påvirkes av regionale sjokk og at lønninger ikke er konstante, vil ikke dette gi likevekt i kortere tidsperioder. I likhet med studiene fra Storbritannia 1961-1982 nevner McMaster og Pissarides at det på lengre sikt hadde vært optimalt at lønnen skulle ha vært den faktoren som i størst grad kompenserer for forskjeller i ledigheten, men på grunn av eksogene faktorer som konjunktursjokk blir ikke dette realistisk. Tregheter skaper kostnader, men for å unngå disse kostnadene mener de at regionalpolitikken kan tilpasses for å bidra til at jobber heller enn arbeidere bytter mellom regioner.

Resultatene i McCormick (1997) tyder på at ledigheten, både individuelt og totalt har en signifikant effekt på utмиграsjon for ulike regioner over en lengre periode. Ugunstige regionale sjokk gir i de fleste tilfeller økt ledighet som totalt gir effekter som vil påvirke både land og regioner. De finner også at regional mobilitet i mange tilfeller er korrelert med økonomiske sykluser og at dess høyere mobiliteten er, dess lavere blir ledigheten selv i perioder der man opplever lavkonjunkturer i arbeidsmarkedet. McCormick og Hughes (1994) ser også på forskjellen i ledighet mellom ulike utdanningsgrupper. Undersøkelsen basert på engelske data fra 1983-1986 viser at variasjon i ledigheten er større innad i grupper med lavere utdanning i forhold til grupper med høyt utdannede, men totalt sett er migrasjonen blant høyt utdannede større enn migrasjonen for lavt utdannede. For eksempel vil man ved et arbeidsmarkedssjokk oppleve at lavt utdannede ikke migrerer mye, mens høyt utdannede har netto utмиграsjon og dette kan dermed tyde på at høyt utdannede er mer mobile enn de med lavere utdanning. Dette kan ha en sammenheng med at høyt utdannede antas å få flere jobbtilbud, selv i perioder med økonomisk lavkonjunktur. McCormick & Hughes (1994) finner effektene til å være signifikante, men noe svake.

2.2 Interregional migrasjon i Norge

I det siste delkapitlet ser vi mer detaljert på norske studier som tar for seg hvordan regionale arbeidsmarkedsvariabler påvirker flyttestrømmen mellom regioner, og hvordan lav- eller høykonjunkturer igjen kan påvirke de ulike arbeidsmarkedsvariablene. Ved å se på flere europeiske studier fant vi at flere av effektene av regionale arbeidsmarkedsforhold på migrasjonen var til dels svake, mens andre forskere i større grad fant signifikante effekter for ledigheten på netto migrasjon. Vi skal se av de norske studiene at disse konkluderer med at flyttestrømmene i Norge er svært følsomme overfor regionale arbeidsmarkedsforskjeller, og at regionale arbeidsmarkedsvariabler har en relativ sterk og signifikant effekt på flyttestrømmene.

En studie av Carlsen og Johansen (2004) ser på 90 økonomiske regioner i perioden 1993-98. Der hvor flere av de engelske studiene ikke fant noen eller kun svake effekter av regionale arbeidsmarkedsvariabler på migrasjonen, konkluderer Carlsen og Johansen med at det er sterke og signifikante effekter på interregionale migrasjonsstrømmer i Norge. En ofte brukt forklaring er at arbeidskraftmobiliteten i Norge er høyere enn i de fleste andre europeiske land

og mye av dette skyldes den lave befolkningstettheten. En annen mulig forklaring kan også være at tidligere studier har brukt data om variabler som har inneholdt imperfekt informasjon om for eksempel ledigheten og jobbmuligheter, og at man dermed har fått feilestimerte resultater.

I en økonomisk analyse foretatt av Brunborg, Sørli og Texmon (2006) antyder de at mobiliteten mellom distriktene og sentrale strøk varierer i takt med svingningene i økonomien. De har tatt utgangspunkt i en kortere periode fra 2000-2004, da mobiliteten var relativt høy og man hadde en sterk grad av sentralisering. Studiene for denne perioden konkluderer med at flyttemønsteret i Norge er svært livsfasebetinget. Personer i ung alder, da særlig i gruppen med lavere utdanning, har en sterk tendens til å flytte inn til storbyene. Omtrent 60 % av disse flytter ut igjen innen de blir 35 år. (Sørli 2004). Det konkluderes med at effektene av konjunktursvingningene er signifikant med hensyn på mobiliteten, samtidig som at lavere utdanning har en positiv effekt på netto innmigrasjon til de største regionene.

Carlsen, Johansen og Kaspersen (2007) bruker i en annen studie paneldata til å se på sammenhengen mellom flyttestrømmer, regionale arbeidsmarkeder og konjunktursituasjonen for 90 norske regioner i perioden 1986-2004. Forskerne konkluderer med at netto innflytting til en region øker når arbeidsledigheten i regionen reduseres og antall vakanser i regionen øker. Det nye i denne studien er at effekten av regional arbeidsledighet er sterkest i oppgangstider, mens regionale vakanser betyr mest i nedgangstider, konjunktursituasjonen tatt i betraktning. Årsaken er at det skapes få nye jobber i nedgangstider slik at mobiliteten er generelt lav, men derimot forventes det at regionale forskjeller i vakanser betyr mest i dårlige tider på grunn av at det utlyses færre jobber i en lavkonjunktur. De finner signifikante effekter av regional ledighet og vakanserate, og at flyttestrømmene i Norge er svært følsomme ovenfor regionale arbeidsmarkedsforskjeller. Resultatene tilsier at arbeidskraftmobiliteten i Norge er relativt høy og kan dermed være med å bidra til forklaringen av hvorfor Norge har relativt små forskjeller i arbeidsledigheten. Bakgrunnen til denne høye mobiliteten kan komme av at Norge er et land med lav befolkningstetthet og jobbutvalget i de ulike regionene vil dermed være begrenset – noe som igjen kan føre til at man er nødt til å være på reisefot for å finne den jobben man er kvalifisert for.

Sammenlignet med europeiske studier ser det ut som Norge har langt høyere mobilitet. En skal være imidlertid være litt forsiktig med slike sammenligninger siden data som brukes i studiene for EU-landene baserer seg på langt større geografiske områder. Et eksempel på dette som det refereres til i artikkelen til Carlsen, Johansen og Kaspersen (2007) er Nahuis og Parikh (2004) som ser på regioner med over 2 millioner innbyggere i en paneldatastudie for flytting mellom regioner i EU-området. Gjennomsnittsstørrelsen for norske arbeidsmarkedsregioner er omtrent 50 000 innbyggere. I en av artiklene som er utgangspunktet for analysen sier Carlsen, Johansen og Kaspersen (2007) at de fleste norske kommuner er relativt små og at det er stor sannsynlighet for at arbeidsmarkedet utenfor bostedskommunen påvirker innbyggernes jobbmuligheter. Fra SSB får vi oppgitt at de deler inn landet i 90 økonomiske regioner med 4,8 kommuner i gjennomsnitt for hver region.

Når det kommer til ulike utdanningsgruppers flytterespons på arbeidsmarkedsvariablene konkluderer Geir A. Dahl (2007) i sin studie med at mobiliteten er stigende i utdanningsnivået. Dette selv om effekten i mange tilfeller er sterkest for gruppen med videregående utdanning. Effekten av den regionale ledighetsraten på netto innmigrasjon viser seg å være negativ for alle utdanningsgruppene, og for vakanseraten er det en positiv effekt som øker i takt med utdanningsnivået. Migrasjonsratene for de tre gruppene virker å være ulike og det tyder på at de vektlegger forklaringsvariablene på forskjellige måter. Dette kan vi se ut ifra hvor mye de ulike utdanningsgruppene responderer på to av variablene; arbeidsledighet og vakanserate. Ved endringer i de to forklaringsvariablene kommer ulikheten mellom gruppene betydelig frem. Dahl sier at de med høyest utdanning vil ha flere jobbmuligheter i et mer spredt geografisk område og at dette gir større valgfrihet ved valg av bosted. Disse faktorene vil da bidra til økt flyttetilbøyelighet. Av dette følger det også at de med høyere utdanning responderer mest i forhold til boligpriser, og det kan også tenkes at kommunale inntekter fanger opp ulike forhold som er av størst betydning for denne gruppen.

Motivert av analyser fra det europeiske arbeidsmarkedet ser Carlsen, Johansen og Stambøl (2010) på hvordan mobilitetsresponsen er mellom ulike utdanningsgrupper i Norge. I europeiske økonomier er det store variasjoner i ledigheten mellom regioner- og utdanningsgrupper. Lav grad av mobilitet fører til større regionale ulikheter og høyere ledighet. Resultatene som kommer frem i denne artikkelen viser at de regionale ulikhetene i arbeidsledigheten er synkende i utdanningsnivået, mens responsen av migrasjonsstrømmene på regionale ledighetssjokk øker med utdanningsnivået. Disse resultatene kan også sies å være

konsistente med resultatene fra Dahl (2007). Signifikante forskjeller mellom gruppene kan tyde på at de med lav utdanning har lav regional mobilitet, og at dette bidrar til større forskjeller mellom gruppene, samt at ledigheten øker blant de med lavere utdanning.

Videre i denne oppgaven vil resultatene fra de nevnte norske studiene om at det eksisterer signifikante effekter på migrasjonen av ulike arbeidsmarkedsvariabler, danne en plattform for ytterligere undersøkelser. Vi ønsker å knytte den nasjonale konjunktursituasjonen opp imot disse arbeidsmarkedsvariablene, og se hvordan dette påvirker migrasjonen for de tre utdanningsgruppene.

2.3 Oppsummering

Hvorvidt en person migrerer mellom regioner eller ikke virker å være avhengig av fremtidig nytte sammenlignet med kostnaden. En betingelse er at nytten blir større enn kostnaden. Arbeidsmarkedsforhold vil naturlig nok spille en viktig rolle for denne beslutningen, men man kan heller ikke utelukke at utdanningsnivået til denne personen vil bety mye for jobbmulighetene.

Europeiske studier har noe mer uklare konklusjoner. De forskerne som ser på sammenhengen mellom migrasjon, det regionale arbeidsmarkedet og konjunkturer ser ut til å mene at i de fleste tilfeller har økonomiske konjunkturer en signifikant effekt på regionale arbeidsmarkedsvariabler. Litt mer uklart er det med sammenhengen mellom migrasjon og arbeidsmarkedsvariabler. Noen finner ingen eller til dels svake effekter på migrasjonen, mens andre finner sterke og mer signifikante effekter. Selv om modellspesifikasjonene og metodene varierer mye for de ulike studiene, virker det til at de fleste funnene støtter undersøkelsene som viser til at mobiliteten er stigende i utdanningsnivået. Resultatene tyder også på at høyt utdannede personer påvirkes mer negativt av sjokk i arbeidsmarkedsvariablene enn de lavt utdannede. Dette kan vi se fra den estimerte effekten av den nasjonale konjunktursituasjonen på netto innmigrasjon, der de med høyere utdanning reduserer innmigrasjonen i større grad enn de med lavere utdanning.

Vi har sett flere eksempler på norske studier som konkluderer med at det eksisterer signifikante effekter av ulike markedsvariable på migrasjonen i Norge. Det kommer frem at arbeidskraftmobiliteten i Norge er høy sammenlignet med europeisk målestokk. Netto innflytting er følsom både ovenfor vakanseraten og for endringer i den regionale arbeidsledigheten. Når arbeidsmarkedet kommer inn i en periode med høykonjunktur finner de at regional ledighet betyr mest i oppgangstider, mens i en lavkonjunktur er det vakanseraten som er av størst betydning.

3 Empirisk tilnærming

For å undersøke sammenhengen mellom flyttestrømmer, regionale arbeidsmarkedsindikatorer for ulike utdanningsgrupper og om disse blir påvirket av den nasjonale konjunktursituasjonen, benytter vi et paneldatasett for perioden 1986-2004. Grunnen til at det benyttes paneldata er fordi vi ønsker å følge tverrsnittsenheter over tid, og de økonometriske metodene som brukes er tilpasset analyse av balansert paneldata⁹. I delkapittel 3.1 beskrives regresjonsligningen som utgjør studiets økonometriske modell. Delkapittel 3.2 presenterer deskriptiv statistikk og forklarer datamaterialet. Mens delkapittel 3.3 begrunner valget av estimeringsmetode.

3.1 Spesifikasjon av statistisk paneldatamodell

Denne analysen er en videreføring av artiklene til Carlsen, Johansen og Kaspersen (2007), Carlsen, Johansen og Stambøl (2010), og masteroppgaven til Geir A. Dahl (2007). Modellformuleringene bygger på flere av de samme inkluderte variablene som i studiene ovenfor. Det er en statistisk modell der det forutsettes at marginale endringer i en eller flere av de uavhengige forklaringsvariablene vil gi en momentan effekt i den avhengige variabelen, $Innmigrasjon[X]_{it}$, og vi ser bort ifra eventuelle tregheter. Basismodellen for den statistiske analysen gis ved følgende ligning:

$$(3.1) \quad Innmigrasjon[X]_{it} = \beta_1 \log(U_{it}) + \beta_2 \log(VakS_{it}) + \beta_3 \log(\mathbf{andre\ variabler}_{it}) + V_{it}$$

Den avhengige variabelen er $Innmigrasjon[X]_{it}$ som er definert som netto innflytting til en region, i prosent av befolkningen og innenfor de tre ulike utdanningsgruppene. Fotskrift i angir region og t angir tidspunktet. X 'en i denne variabelen inneholder de tre utdanningsgruppene som skal analyseres. Disse tre er; *grunnskole*, *videregående skole* og *høyere utdanning*. For enkelhets skyld forkortes disse til grunnskole = GRU, videregående skole = VGS og høyere utdanning = HØY. Ligning (3.1) kan her forstås som en felles avhengig variabel for netto innmigrasjon for alle de tre utdanningsgruppene, men i analysen vil de tre ulike utdanningsgruppene betraktes separat ved estimering. **Andre variabler**_{it} er en vektor av de andre inkluderte forklaringsvariablene. U_{it} er regional arbeidsledighet og $VakS_{it}$ er vakansestrøm på fylkesnivå. Alle forklaringsvariablene som brukes i analysen blir nærmere beskrevet i delkapittel 3.2.1.

⁹ Alle regresjoner i analysen er foretatt i programmet PcGive 12.10, OxPack for GiveWin 2.10

I modellen har vi valgt å inkludere et fullt sett av tidsdummyer for å kontrollere for effekten av alle aggregerte variable. Dette er fordi det er naturlig å tenke seg at flyttestrømmene påvirkes av nivået på de regionale variablene relativt til det nasjonale gjennomsnittet. Ligning (3.1) viser at forklaringsvariablene er log-transformert. Ved å inkludere tidsdummyer vil dette gi at effekten av de regionale variablene blir de samme enten man inkluderer variablene som i ligning (3.1), eller om man regner variablene relativt til sine respektive nasjonale gjennomsnitt. I tillegg inkluderer vi interaksjonsledd og tidsdummyer på aktuelle variabler for å kunne analysere hvordan forklaringsvariablene responderer på den nasjonale konjunktursituasjonen for en bestemt tidsperiode.

Restleddet V_{it} som inngår i modellen antas å bestå av to komponenter som er uavhengig normalfordelt og identisk på følgende vis; $V_{it} = \eta_i + \varepsilon_{it}$. Får da en tidsvarierende (idiosynkratisk) komponent ε_{it} og en regionspesifikk komponent η_i , som ikke varierer over tid. Det idiosynkratiske restleddet inneholder utelatte variabler som varierer både for hver region og over tid. Den regionspesifikke restleddskomponenten fanger opp stokastisk, uobserverbar heterogenitet som kan skyldes utelatte regionspesifikke variable. Da gjerne av variabler som antas konstante over tid, men som varierer mellom hver region. Det er også mulig å benytte tidsdummyer som kontrollvariable for å renske bort effekter av makroøkonomiske variable eller variable som endres over tid, men som er felles for individene.

3.1.1 Funksjonsform

Ligning (3.1) gir at de uavhengige høyresidevariablene er inkludert med (naturlig) logaritmisk form. Ved å gjøre dette tillater man for ikke-linære sammenhenger mellom høyre og venstre side i ligningen. Wooldridge(2006:191) forklarer flere grunner til at en logaritmisk spesifisering er å foretrekke. Wooldridge viser til at strengt positive variabler ofte har heteroskedastiske eller skjeve fordelinger. En transformasjon til en naturlig logaritmisk form av variabelen kan redusere problemet og gi normalfordeling, men dette vil normalt også innskrenke spennvidden i observasjonene i og med at spredningen i variabelen ”krymper”. Logaritmer kan dermed hjelpe til å redusere heterogenitetsproblemer eller i beste fall fjerne de helt.

Siden den avhengige variabelen $Innmigrasjon[X]_{it}$ kan anta negative verdier i modellen, vil ikke denne inngå med logaritmisk form. Som nevnt i Wooldridge(2006:46) får koeffisientverdiene i en modell med nivå-log-spesifikasjon en noe spesiell tolkning. Tolkningen er at man estimerer effekten av prosentvis endring i gjeldende variabler, ikke en endring i prosentpoeng. Hvis β_1 betraktes som en koeffisientverdi til en forklaringsvariabel, vil $(\beta_1/100)$ angi endring i den avhengige variabel $Innmigrasjon[X]_{it}$ ved at den nevnte forklaringsvariabelen endres med 1 prosent.¹⁰

3.2 Presentasjon av data

I denne analysen benyttes datamateriale som strekker seg fra 1986 til 2004. Gjennom denne perioden har vi aktuelle opplysninger om variabler som migrasjonsratene for de tre ulike utdanningsgruppene, regional arbeidsledighet, vakanser (på fylkesnivå) og kommunale inntekter (region). Det blir også benyttet data for den nasjonale konjunktursituasjonen til å teste om dette påvirker den estimerte effekten av arbeidsmarkedsvareblene. Grunnlaget til datamaterialet er basert på tall som er hentet fra NAV (Arbeids- og velferdsforvaltningen) og Statistisk sentralbyrå (SSB). For boligprisen har vi observasjoner for perioden 1991-2004, og på grunn av dette blir variabelen boligpris inkludert i en modell med delutvalgsanalyse.

Ifølge SSB bygger inndelingen av Norge inn i 90 økonomiske regioner med 4,8 kommuner i gjennomsnitt for hver region, i stor grad på hvilke kommuner som arbeidsmessig hører sammen - som igjen er basert på pendlingstabeller.¹¹ Dette fører til at hver kommune blir tilordnet den senterkommunen den har størst innpendling til, og at hver region representerer et tilnærmet "sant" arbeidsmarked for innbyggerne. Tallene for vakanseratene er imidlertid gitt på fylkesnivå. De tre forskerne Carlsen, Johansen og Kaspersen (2007) sier i artikkelen at en måte å nærme seg en tverrsnittsenhet som representerer et reelt lokalt arbeidsmarked for befolkningen, er ved å aggregere primærdataene til regionnivå. Dette fordi de fleste norske kommuner er små, og dermed er det sannsynlig at arbeidsmarkedet utenfor bostedskommunen har betydning for innbyggernes jobbmuligheter.

¹⁰ Illustrert i en modell ut ifra Wooldridge(2006:46). $Y = \beta_0 + \beta_1 \log(x)$ gir at $\Delta y = \beta_1 \Delta \log(x)$. Skriver dette som $\Delta y = (\beta_1/100)[100 * \Delta \log(x)]$. Får da at $\Delta y = (\beta_1/100)(\% \Delta x)$ ved å utnytte at $100 * \Delta \log(x) = \% \Delta x$. Dette viser dermed at $(\beta_1/100)$ er en enhets endring i y når x øker med 1 %.

¹¹ Se SSB (2000): NOS C 16 "Standard for økonomiske regioner".

3.2.1 Forklaring av inkluderte variabler

Avhengig variabel er Innмиграsjon $[X]_{it}$, som viser netto innflytting til region i , år t , i prosent av befolkningen ved begynnelsen av året. Migrasjonen baserer seg på tall hentet fra de tre ulike utdanningsgruppene GRU, VGS og HØY. Netto migrasjon baserer seg på tall hentet fra opplysninger om flytting mellom kommuner, innen samme kommune, fra utlandet til Norge og fra Norge til utlandet. Flyttestrømmene registreres i hver kommune og dette gjør det mulig å beregne netto flytting mellom hver region.

I tabell 3.1 vises deskriptiv statistikk i form av standardavvik og gjennomsnittsverdier. Mer detaljerte definisjoner av variablene finnes i tabell A1 i appendikset. Grunnen til at vi undersøker standardavvikets størrelse relativt til variablenes gjennomsnittsverdi, er fordi det gir oss muligheten til å si noe om forskjellen i variasjonen i de ulike variablene. I vår empiriske analyse kontrollerer vi for regionspesifikke forhold som ikke varierer over tid, og estimering ved hjelp av FE-modellene fører til at vi benytter en transformasjon som reduserer utvalgsvariasjonen. Det som da blir viktig for denne typen regresjoner er at vi har tilstrekkelig variasjon i data innenfor hver enkelt av regionene. Som vi ser av tabell 3.1 benyttes standardavvik som er basert på variasjon i tverrsnitts- og tidsdimensjonen (st.avvik₁). I tillegg til standardavvik som er basert på ren tidsserievariasjon (st.avvik₂) som kommer til nytte når resultatene i kapittel 4 og 5 skal tolkes. Av tabellen ser vi at standardavvikenes størrelse relativt til gjennomsnittsverdiene er størst for de tre migrasjonsratene.

Tabell 3.1. Deskriptiv statistikk for variabler i paneldataanalysen.

Variabel	Forkortelse	Gj.snitt	St.avvik ₁	St.avvik ₂
Innмиграsjon[GRU]	Innмиграsjon $[X]_{it}$; X= GRU	0,107	0,542	0,481
Innмиграsjon[VGS]	Innмиграsjon $[X]_{it}$; X= VGS	-0,253	0,705	0,531
Innмиграsjon[HØY]	Innмиграsjon $[X]_{it}$; X= HØY	-1,173	1,750	1,300
Regional	U_{it}	3,703	1,607	1,282
arbeidsledighetsrate				
Vakanserate (strøm)	$VakS_{it}$	14,400	5,432	4,900
Kommunale inntekter	$KommInn_{it}$	1,385	0,263	0,152
Boligpris	$Boligpris_{it}$	0,648	0,279	0,222
Msysvekst	ΔN_t			1,661

Merk: Senket skrift i angir regio, t angir år. Gj.snitt og St.avvik₁ er basert på variasjon både i tids- og tverrsnittsdimensjonen. St.avvik₂ er basert på ren tidsserievariasjon innenfor den enkelte regionen, og finnes ved å definere variabelen vi betrakter som avhengig variabel, for så å kjøre MKM-regresjonen med 89 regiondummyer. Standardavviket til denne regresjonen kan tolkes som tidsvariasjonen i variabelen.

I denne oppgaven blir det sett på hvordan migrasjonsratene for de tre ulike utdanningsgruppene blir påvirket av stramheten i arbeidsmarkedet. Ved å inkludere vakanseraten på fylkesnivå, $VakS_{it}$ og regional arbeidsledighet, U_{it} kan dette testes¹². Vakanser forklares som "strøm" av vakanser, det vil si antall nye ledige stillinger i jobbmarkedet, mens arbeidsledighet vil være den del av arbeidsstyrken som er i stand til, og villig til å arbeide, men som ikke finner arbeid. Disse variablene er oppgitt i prosent av beregnet arbeidsstyrke, der arbeidsstyrken er definert som summen av antall sysselsatte og arbeidsledige.¹³

Ved å inkludere regional ledighet og vakanserate i analysen, ønsker vi å kontrollere for ulike effekter av utdanningsnivå og påvirkning fra den nasjonale konjunktursituasjonen. Det forventes at høy regional ledighet gir en negativ effekt på netto innmigrasjon og at høy vakanserate har en positiv effekt på netto innmigrasjon. En annen rimelig antagelse er at mobiliteten vil være stigende i utdanningsnivået og at dette antas å gi individet flere muligheter i arbeidsmarkedet både når det gjelder yrke- og bostedsalternativer.

For å ta hensyn til konjunkturreffektene definerer vi den makroøkonomiske indikatoren gitt ved nasjonal sysselsettingsvekst, ΔN_t . I appendikset viser figur 3A2 hvordan denne indikatoren varierer over estimeringsperioden. Regresjonsligningen (3.1) utvides med interaksjonsledd og med denne indikatoren har vi muligheten til å undersøke hvordan de partielle effektene av de regionale arbeidsmarkedsvARIABLENE påvirkes av den nasjonale konjunktursituasjonen. I delkapittel 4.1.2 benyttes metoden med interaksjonsledd ved estimering. Interaksjonsleddet vil mer konkret fortelle oss hvordan migrasjonsresponsene på regional arbeidsledighet og vakanser varierer over den nasjonale konjunktursyklusen, og om variasjonen er forskjellig for ulike utdanningsgrupper. For å gjøre tolkningen av koeffisienten enklere er den makroøkonomiske indikatoren målt i prosent og som avvik fra gjennomsnittet for hele perioden (1986-2004).¹⁴

¹² "Johansen, Carlsen og Kaspersen (2007:28) om fylkesvis vakanserate: " På begynnelsen av 80-tallet omfattet statistikken for antall ledige stillinger kun stillinger som var meldt til Aetat. I løpet av 80- og 90-tallet økte andel offentliggjorte stillinger som ble omfattet av statistikken, og i dag antar NAV (tidligere Aetat) at deres tall omfatter praktisk talt alle stillinger som er utlyst i aviser, tidsskrifter, etc. Vår vakansrate er derfor strengt tatt ikke sammenlignbar over tid. Imidlertid tror vi ikke dette er noe stort problem i praksis fordi vi inkluderer tidsdummier i spesifikasjonene slik at det vil være differensen mellom fylkesvis og nasjonale vakansrater som driver resultatene.

¹³ Arbeidsstyrken er definert som summen av antall sysselsatte og arbeidsledige. Statistisk sentralbyrå publiserer kommunedata om sysselsatte for perioden 2000-2004, mens det foreligger kommunetall for ansatte fra 2000 og tilbake til 1986. Ved å benytte forholdstallet mellom sysselsatte og ansatte i år 2000 har vi estimert antall sysselsatte i hver kommune i perioden 1986-2000, og anvender denne variabelen ved beregning av regionale og fylkesvis tidsserier for arbeidsstyrken".

¹⁴ Se Johansen, Carlsen og Kaspersen (2007:28) om konjunkturreffekter.

I tillegg til å se på stramheten i arbeidsmarkedet og om konjunktursituasjonen skaper ulikheter for de tre utdanningsgruppene, skal vi også undersøke noen utvalgte variabler som kan være relevant for flyttebeslutningene til de tre utdanningsgruppene. Disse variablene er kommunale inntekter og boligpriser.

Ved å inkludere kommunale inntekter, $KommInn_{it}$ som en forklaringsvariabel, kan man fange opp eventuelle effekter av tjenestetilbudet i regionen. Denne variabelen omfatter skatt og rammetilskudd per innbygger i 10 000 kr deflatert med en aggregert prisindeks for kommunalt konsum. En region vil dermed være mer attraktiv jo lavere skattenivået er ut ifra et gitt tjenestetilbud. Tilfellet med Norge er at det er moderate geografiske forskjeller i skattebyrde på grunn av at inntekt og formue skattelegges med samme skattesats i nesten alle landets kommuner, bortsett fra Finnmark og Troms. Forskjellene i brukerbetaling og eiendomsskatt på tvers av kommuner er relativt små i forhold til kommunale inntekter, og dette kommer av at kommunene selv bestemmer om de ønsker å benytte eiendomsskatt. Dersom man finner noe effekt av variabelen kommunale inntekter er det naturlig å forvente en positiv effekt på netto innflytting. Dette vil generelt gjelde uavhengig av hvilken utdanningsgruppe man ser på, fordi kommuner med høye inntekter vil normalt være i stand til å tilby sine innbyggere et bedre velferds- og tjenestetilbud i motsetning til mindre pengesterke kommuner. Dette vil naturlig øke attraktiviteten til kommunens omkrets, og samtidig kunne bidra til økt innflytting for den tilhørende regionen.

Vi forventer at økt boligpris, $Boligpris_{it}$, partielt sett bidrar til redusert netto innflytting. Carlsen og Johansen (2004) samt Carlsen, Johansen og Kaspersen (2007) finner en signifikant negativ effekt av denne forklaringsvariabelen på netto innflytting, der data er hentet fra perioden 1994-1998. Dahl (2007) rapporterer også lav respons på boligprisvariabelen for lavt utdannede, mens effekten er sterkere og signifikant for de to andre gruppene. Konstruksjonen av denne variabelen er foretatt ved hjelp av hedoniske regresjoner på data om boligomsetninger der kvadratmeterprisen estimeres som en funksjon av boligkarakteristika og regiondummyer.¹⁵ Variabelen kan tolkes som en gjennomsnittlig kvadratmeterpris for en standardisert type leilighet i vedkommende region og år. Boligprisen er også deflatert med konsumprisindeksen i årene 1991-2004.

¹⁵ Se Carlsen, Johansen og Kaspersen (2007)

Tabell 3.2. Korrelasjonsmatrise for den avhengige variabelen $Innmigrasjon[X]_{it}$ for de tre utdanningsgruppene.

	Innmigrasjon[GRU]	Innmigrasjon[VGS]	Innmigrasjon[HØY]
Innmigrasjon[GRU]	1,000		
Innmigrasjon[VGS]	0,42416	1,000	
Innmigrasjon[HØY]	0,12106	0,53665	1,000

Korrelasjonen brukes for å se om det er forskjeller i migrasjonen mellom de tre utdanningsgruppene. Av tabell 3.2 ser vi at korrelasjonen mellom gruppene grunnskole og høyere utdanning er meget lav, mens den er størst (men generelt lav) mellom gruppene videregående skole og høyere utdanning. Dess høyere korrelasjonen er mellom to grupper, dess mer likt vil de respondere på variablene i modellen. I korrelasjonsmatrisen observeres alt fra lav til relativt moderat korrelasjon og det gir oss en indikasjon på at det eksisterer forskjeller i gruppenes flyttebeslutninger. Den lave korrelasjonen mellom grunnskole- og høyere utdanning kan tolkes som at det er liten forskjell mellom gruppene, men at det innenfor gruppen med høyere utdanning eksisterer betydelige forskjeller. En måte å forklare dette på er at gruppene vektlegger variablene ulikt og at variabler som er av ingen eller liten betydning for en gruppe, kan ha en betydelig innvirkning på en annen gruppe. I appendikset vil figur 2 vise en grafisk framstilling av korrelasjonen i form av et kryssplott, mens tabell A2 og A3 viser korrelasjonen mellom samtlige forklaringsvariable i paneldatanalysen.

3.3 Estimeringsmetode

Når man har regresjoner som er av rene tverrsnittsdata vil bruk av minste kvadraters metode (MKM) gi forventningsrette og konsistente estimatorer, gitt visse standardforutsetninger om restleddet. Det er dette man kjenner som Gauss-Markov-Teoremet (Wooldridge, 2006:kap.3). I vår oppgave tar vi i bruk et paneldatamateriale, som igjen vil si at man har informasjon om den enkelte tverrsnittsenhet over tid ut ifra flere sammenhengende tidspunkt eller for flere sammenhengende perioder. Ved bruk av tverrsnittsdata brukes ofte metoden ”pooled-MKM” for å oppnå forventningsrette estimatorer, men behovet for å gjøre andre forutsetninger da spesielt om variabler som ikke varierer over tid, fører til at denne metoden blir utilstrekkelig. Vår analyse benytter årlige observasjoner som er gjort over et visst tidsrom for en rekke regioner. Vi skal se nærmere på to metoder for estimering av paneldata der forskjellen ligger i hvordan restleddet tolkes.

3.3.1. Fixed effects-estimering

Om man ønsker å bruke "pooled-MKM" som estimeringsmetode må en forutsetning ligge til grunn, og det er at utelatte variabler som kun varierer mellom regionene er uavhengige av de inkluderte forklaringsvariablene. Dette er en sterk forutsetning og siden det er grunn til å anta at det eksisterer korrelasjon mellom de inkluderte forklaringsvariablene og det regionspesifikke restleddet på grunn av betydelige strukturelle forskjeller mellom regionene, kan dette løses ved å bruke fixed effects-estimering. FE benytter noe man kjenner som "within-group"-transformasjon og det betyr at variablene måles som avvik fra sine respektive regionspesifikke gjennomsnitt.

Brukes "pooled-MKM" som estimeringsmetode uten at det tas hensyn til korrelasjon mellom forklaringsvariablene og det regionspesifikke restleddet η_i , vil dette gi skjeve estimater. FE transformerer bort det regionspesifikke restleddet og dermed unngår man dette problemet i modellen. Fixed effects-estimering identifiserer effektene av forklaringsvariablene ved at en over tid utnytter tidsvariasjonen i alle variablene innen hver enkelt region. For at FE-estimatoren skal være konsistent er en forutsetning at utvalget er tilfeldig trukket ut ifra populasjonen, og dermed vil forutsetningene for FE-estimatoren være mer restriktiv en for "pooled-MKM". Dette medfører også at mindre variasjon i datamaterialet utnyttes siden kun variasjon over tid benyttes, og der all ren tverrsnittsvariasjon transformeres bort. Datasettet som vi skal bruke i analysen kan ikke sies direkte å inneholde informasjon fra et tilfeldig utvalg. Vårt argument for at kriteriet er oppfylt, er at variablene må kunne tolkes tilfeldig siden man på forhånd ikke kan forutsi nivået på variablene.

Basert på Wooldridge (2006:503) finner vi følgende antagelser om sammenhengen mellom restleddet og variablene.

$$i) E(\varepsilon_{it} | x_{it}) = 0 \quad \text{der: } i = 1, 2, \dots, J \text{ og } t = 1, 2, \dots, T$$

Forutsetningen gir at forventningsverdien til det idiosynkratiske restleddet skal være lik null, gitt forklaringsvariablene x_{it} . Dette betyr at forklaringsvariablene er eksogene i forhold til restleddskomponenten ε_{it} , som igjen vil si at det ikke er noe korrelasjon med den delen av restleddet som varierer over tid.

ii) Alle forklaringsvariable endrer seg over tid og det eksisterer ingen perfekte lineære sammenhenger mellom forklaringsvariablene. Dette er forutsetningen om ingen perfekt multikollinearitet. Definisjonen av multikollinearitet sier oss at det er høy, men ikke perfekt korrelasjon mellom to eller flere uavhengige variabler.¹⁶ I dette tilfellet kan man nesten skrive variablene som en lineær kombinasjon av hverandre – og estimatoren vil få en høy varians. Standardavvikene er store, men koeffisientene vil være konsistente.

Selv om man har multikollinearitet, vil estimatoren fortsatt være forventningsrett. Gitt andre faktorer er det bedre jo lavere varians man har mellom x_i og de uavhengige variablene når en skal estimere. En måte å redusere denne korrelasjonen på, kan være ved å fjerne en eller flere av de uavhengige variablene fra modellen. Problemer oppstår særlig om man utelater variabler som hører med i populasjonsmodellen grunnet insignifikans basert på lav t-verdi. Konsekvensen blir at man får inkonsistente estimat på grunn av korrelasjon mellom utelatte variable og inkluderte forklaringsvariable.

iii) $\text{Var}(\varepsilon_{it} | x_{it}, \eta_i) = \sigma_\varepsilon^2$

Variansen til ε_{it} er konstant for alle regioner i og tidsperioder t , gitt x_{it} og η_i . Dermed vil det idiosynkratiske restleddet være homoskedastisk. Skulle det oppstå et motsatt tilfelle der restleddet er heteroskedastisk vil parameterestimaten fortsatt være konsistente, men standardavvikene vil bli feilberegnet – påvirker inferens og testing. Konsekvensen er at ”de vanlige MKM-formlene” for varians og standardavvik ikke lenger er gyldige. Videre er både F-testen og LM-testen basert på forutsetning om konstant restleddsvariens, slik at heller ikke disse testene er gyldige ved heteroskedastiske restledd. En mulighet for å gjøre F-, LM- og t-testen gyldig under heteroskedastisitet er at det konstrueres standardavvik som er heteroskedastisk-robuste. Dette estimeres så ved vanlig MKM der det tas hensyn til de nye standardavvikene.

iv) $\text{Cov}(\varepsilon_{it}, \varepsilon_{is} | x_{it}, \eta_i) = 0$

Her eksisterer det ingen korrelasjon mellom restleddene – og ikke over tid. Gitt ε_{it} og η_i er kovariansen mellom to restledd for samme enhet i ulik tidsperiode lik null for de idiosynkratiske restleddene. Et potensielt problem kan være at modellen blir feilspesifisert ved at utelatte variabler er seriekorrelerte- eller at man opplever ren seriekorrelasjon, og dette

¹⁶ Se Wooldridge (2006:96)

kan være en medvirkende årsak til forventningskjevne estimater hvis det er seriekorrelasjon i restleddet. MKM-estimatorene vil ikke lenger være effisiente og de vanlige variansformlene blir ikke lenger gyldig.

$$v) \text{Var}(\varepsilon_{it} | x_{it}, \eta_i) = \sigma_\varepsilon^2 \text{ og } \text{Var}(\eta_i | x_{it}, \varepsilon_{it}) = \sigma_\eta^2$$

Den siste forutsetningen går ut på at restleddet ε_{it} er uavhengig på tvers av regioner og identisk normalfordelt, gitt det regionspesifikke restleddet og forklaringsvariablene. Det samme gjelder for restleddet η_i . Disse forutsetningene blir viktig om vi ønsker å foreta tester på regresjonskoeffisientene siden vanlige t- og F-tester bygger på normalfordeling.

3.3.2 Random effects-estimering

Random effects (RE) skiller seg fra fixed effects-modeller i hvordan den behandler seriekorrelasjon, men begge modellene tar hensyn til regionspesifikke effekter. Metoden blir også kalt for generalisert minste kvadraters metode (GLS). Denne metoden velges foran fixed effects (FE) om man tror at det regionspesifikke restleddet η_i , er ukorrelert med de enkelte forklaringsvariablene. Metoden har de samme forutsetningene som FE, men bygger på ytterligere en antakelse om at det regionspesifikke restleddet er uavhengig av forklaringsvariablene for alle gjeldende perioder:

$$vi) E(\eta_i | x_{it}) = 0$$

Gitt forklaringsvariablene x_{it} , er forventningsverdien til det regionspesifikke restleddet lik null. Restriksjonen pålegger at restleddskomponenten η_i og de inkluderte forklaringsvariablene er ukorrelert for at RE skal gi forventningsrette estimater.

$E(\eta_i | x_{it}) \neq 0$ gir at forutsetningen brytes og η_i fanger dermed opp effekter av utelatte variabler som er korrelert med x_{it} . Antas videre at variansen til η_i er konstant, gitt x_{it} .

I denne oppgaven skal vi ved estimering kun benytte FE siden man kan forvente at det regionspesifikke restleddet er korrelert med inkluderte forklaringsvariable i modellen. Fordelen med FE versus RE er at denne metoden gir konsistente estimater under mindre restriktive forutsetninger. Dette på grunn av at metoden er konsistent uansett om η_i er korrelert med forklaringsvariablene eller ikke - samtidig som det kan kontrolleres for mer og dermed redusere sannsynligheten for skjeve estimater. En annen begrunnelse for denne

beslutningen om å fokusere på FE sammenlignet med RE er at Geir A. Dahl (2007:34) i sin konklusjon sier at FE er den metoden som gir flest pålitelige resultater, og dette relateres til estimeringsteknikker fordi det som nevnt godtas at man tillater korrelasjon mellom forklaringsvariablene og regionspesifikke effekter. Wooldridge (2006:493) benytter samme forklaring på hvorfor FE er en mer robust estimeringsmetode enn RE og i mange tilfeller vil man oppleve at de regionspesifikke effektene i restleddet er korrelerte med de inkluderte forklaringsvariablene.

I analysen vil en utfordring ved fixed effects-estimering som vist i forutsetning (ii), være at det kan oppstå høy grad av seriekorrelasjon mellom to eller flere av de inkluderte forklaringsvariablene. Dette kan føre til høye standardavvik og feiltolket signifikans. Ved å se på de to korrelasjonsmatrisene fra tabell A2 og A3 i appendikset, kan vi sjekke for multikollinearitet for de variablene som blir brukt i estimering av ligning (3.1) og i de andre modellspesifikasjonene. Wooldridge (2006) har ingen klar definisjon av ”høy” korrelasjon, men ingen av våre modeller har variabler som viser til en korrelasjon nær 1 eller -1. I økonomiske analyser er det sannsynlig at man opplever brudd på forutsetningen om eksogenitet i forklaringsvariablene, $[E(\epsilon_{it}|x_{it}) = 0]$. Endogenitet kan da oppstå i en eller flere av forklaringsvariablene på grunn av utelatt variabel eller simultanitet. I kapittel 5 kommer vi nærmere inn på dette da vi skal se på noen måter å håndtere endogenitet ved å inkludere lag og/eller instrumentvariabler i ligningen.

3.3.3 Test av koeffisientforskjeller

Resultatene fra delkapittel 2 kan tyde på at det eksisterer ulikheter i flytteresponsen mellom utdanningsgruppene. Ved å gjennomføre en test av koeffisientforskjellene kan det testes om disse ulikhetene er signifikante. Dette gjøres ved at to av ligningene i regresjonen settes sammen til én, slik at det er mulig å teste for parameterlikhet. Metoden kan hentes ut ifra kapittel 13 i Wooldridge (2006). Dette illustreres nedenfor;

$$(3.2) \quad \text{Innmigrasjon}[\text{GRU}]_{it} = \alpha_1 \log(U_{it}) + \alpha_2 \log(\text{VakS}_{it}) + v_{1it}$$

$$(3.3) \quad \text{Innmigrasjon}[\text{VGS}]_{it} = \gamma_1 \log(U_{it}) + \gamma_2 \log(\text{VakS}_{it}) + v_{2it}$$

Ved å subtrahere (2) ifra (1), får vi:

$$(3.4) \quad \text{Innmigrasjon}[\text{GRU}]_{it} - \text{Innmigrasjon}[\text{VGS}]_{it} = (\alpha_1 - \gamma_1) \log(U_{it}) + (\alpha_2 - \gamma_2) \log(\text{VakS}_{it}) + (v_{1it} - v_{2it})$$

Skrevet på en mer kompakt form:

$$(3.5) \quad \text{Innmigrasjon}_{it}^* = \Theta_1 \log(U_{it}) + \Theta_2 \log(\text{VakS}_{it}) + v_{it}^*$$

Her er $\text{Innmigrasjon}^* = \text{Innmigrasjon}[\text{GRU}] - \text{Innmigrasjon}[\text{VGS}]$, $\Theta_1 = (\alpha_1 - \gamma_1)$, $\Theta_2 = (\alpha_2 - \gamma_2)$ og $v^* = (v_1 - v_2)$. Denne transformerte modellen gir oss da muligheten til å teste for parameterlikhet ved å estimere Θ_1 og Θ_2 . Deretter kan t-verdien benyttes til å se om det er signifikant forskjell mellom α_1 og γ_1 , samt α_2 og γ_2 . Man kan også teste hypoteser om ingen ulikhet mellom utdanningsgruppene, dvs $\Theta_1 = 0$ og $\Theta_2 = 0$.

4 Empiriske resultater

I dette kapitlet estimeres ulike modellspesifikasjoner for å avdekke eventuelle forskjeller mellom utdanningsgruppenes innmigrasjonsrespons for ulike variabler i arbeidsmarkedet. Ved å benytte den regionale ledighetsraten og vakanseraten på fylkesnivå har vi mulighet til å undersøke effekten av stramheten i arbeidsmarkedet. Basismodellen som ble omtalt i delkapittel 3.1, fremstår slik:

$$(4.1) \quad \text{Innmigrasjon}[X]_{it} = \beta_1 \log(U_{it}) + \beta_2 \log(\text{VakS}_{it}) + \beta_3 \log(\text{andre variabler}_{it}) + V_{it}$$

Metodene som benyttes i den statiske regresjonsanalysen er fixed effects. Hver av modellene vil estimeres for seg selv med hensyn på de tre utdanningsgruppene. Grunnmodellen vil være lik ligning (4.1) der kun arbeidsmarkedsvariablene U_{it} og VakS_{it} inkluderes, mens kommunale inntekter vil inngå i en utvidet modellen. I noen modeller blir det tatt hensyn til den nasjonale konjunktursituasjonen ved at det inkluderes et interaksjonsledd, N_t , som fanger opp effekten av nasjonal sysselsettingsvekst. En dynamisk modellspesifikasjon med ett års lag benyttes i et tilfelle for å se om dette gir endrede resultater for innmigrasjonsraten. Disse regresjonene utføres ved bruk av data fra perioden 1986-2004. I analysene basert på delutvalg og generalisert momentmetode (GMM) kontrollerer vi også for boligpris, men her er tidsserien noe kortere. Delutvalgsanalysen brukes til å undersøke variablenes stabilitet, mens det tas hensyn til potensiell simultanitet i GMM-analysen.

Estimeringsmetoden som vil bli brukt i dette kapitlet er fixed effects(FE). Dette kom frem under diskusjonen i kapittel 3.3 som konkluderer med at fixed effects er den metoden som høyst sannsynlig gir mest pålitelige estimater i analysen. Dette på grunn av at det kontrolleres for eventuell korrelasjon mellom inkluderte forklaringsvariabler og utelatte eller uobserverbare regionale forhold som ikke varierer over tid. I modellen inkluderes tidsdummyer for alle år t unntatt det første året i tidsserien, og derav kan det kontrolleres for faktorer som påvirker tilpasningen over tid. Tidsdummyene vil ikke rapporteres i tabellene.

De estimerte koeffisientene fra modellene vil rapporteres i delkapittel 4.1. Resultatene fra fixed effects-estimeringene tolkes, samt at det gis en oversikt over diagnostiske tester. I delkapittel 4.2 foretas en sensitivitetsanalyse der modeller fra delkapittel 4.1 re-estimeres med utgangspunkt i en delutvalgsanalyse.

4.1 Resultater fra fixed effects-estimering

4.1.1 Grunnmodell. Arbeidsledighet og vakanser

Tabell 4.1 gir oss de estimerte koeffisientene for variablene i grunnmodellen med regional arbeidsledighet og vakanser på fylkesnivå. Resultatene er basert på fixed effects-estimering for hver av de tre utdanningsgruppene.

Tabell 4.1. Effekt av regional arbeidsledighet og vakansestrøm på netto innmigrasjon for ulike utdanningsgrupper. Avhengig variabel: $\text{Innmigrasjon}[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
	FE(I)	FE(I)	FE(I)
Log (U_{it})	-0,235 (-3,78)	-0,744 (-7,44)	-0,612 (-3,29)
Log ($VakS_{it}$)	0,130 (1,53)	0,228 (2,44)	0,498 (1,77)
Observasjoner	1710	1710	1710
R^2	0,077	0,181	0,132
σ	0,465	0,483	1,218
Wald (joint)	16,32**	59,82**	19,07**
AR(1)-test	-0,445	4,194**	2,378*
AR(2)-test	-0,231	3,145**	2,008*

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

Resultatene viser at fortegnene på arbeidsmarkedsvariablene er som forventet. Den regionale arbeidsledigheten har en negativ effekt på netto innmigrasjon for alle utdanningsgruppene. Variabelen er høyst signifikant og med en absolutt t-verdi på over 3 vil det bety at den partielle effekten er signifikant ved 1 prosents nivå. For vakanseraten er effekten på netto innmigrasjon positiv for alle utdanningsgruppene. Jackman & Savouri (1992) og Carlsen, Johansen og Kaspersen (2007) finner tilsvarende effekter for disse variablene i sine studier. Resultatene gir at vakanseraten for gruppen videregående skole er signifikant med en t-verdi lik 2,44, mens de to andre gruppene har en lavere t-verdi. En forklaring til at det observeres lav t-verdi for vakanseraten kan komme av at variabelen er gitt på fylkesnivå i datamaterialet. Dette medfører at antall observasjoner er betraktelig færre for vakanseraten i forhold til de regionale forklaringsvariablene.

Ved å se på den simultane effekten av de to arbeidsmarkedsvariablene kan de se ut mobiliteten blant grunnskolegruppen er lavere enn for de to andre gruppene. Dette kan illustreres ved samtidig å øke ledigheten og redusere vakanseraten med ett standardavvik fra gjennomsnittene.¹⁷ Estimaten fra FE(I) gir da en reduksjon i netto innmigrasjonrate for grunnskolegruppen med 0,12 prosentpoeng, 0,32 prosentpoeng for videregående skolegruppen og 0,39 prosentpoeng for gruppen med høyt utdannede.

Resultatene tyder på at gruppene vektlegger de to arbeidsmarkedsvariablene forskjellig når det kommer til flyttebeslutninger. Forskjellene mellom gruppene kan testes for å finne ut om effektene er signifikante ved hjelp av metoden beskrevet i delkapittel 3.3.3. Gruppene sammenlignes parvis ved å sette sammen to ligninger før en ny estimering gir ny t-verdi. Denne kan brukes til å se om effekten mellom gruppene er signifikant. T-verdi større eller lik 1,96 for en forklaringsvariabel betyr at responsen er signifikant forskjellig mellom gruppene ved 5 prosents nivå. For en t-verdi større eller lik 1,282, vil dette bety signifikante forskjeller ved en 10 prosents nivå ensidig hypotese.

Tabell 4.1.1. Gruppesammenligning av estimerte koeffisientforskjeller fra FE(I)-modellene. Avhengige variabler: $\text{Innmigrasjon}[X_1]_{it} - \text{Innmigrasjon}[X_2]_{it}$, der X_1 og X_2 står for utdanningsgruppene. X vil være GRU, VGS og HØY.

Variabler	Gruppesammenligning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(I)	D-FE(I)	D-FE(I)
Log (U_{it})	0,509 (6,26)	0,376 (2,21)	-0,133 (-0,972)
Log (VakS_{it})	-0,098 (-1,04)	-0,368 (-1,24)	-0,270 (-1,05)

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik.

Ved å sammenligne grunnskolegruppen med de to andre gruppene ser vi at forskjellen i effekten av ledighet er signifikant, mens samme effekt mellom gruppene videregående skole og høyere utdanning er insignifikant. Resultatene fra D-FE(I) i tabell 4.1.1 viser dermed at gruppene responderer ulikt på ledighetsraten. Hvis vi øker ledighetsraten med ett standardavvik fra gjennomsnittet blir reduksjonen i innmigrasjonsraten for grunnskolegruppen 0,07 prosentpoeng. For de to andre gruppene er responsen større med 0,22 og 0,18 prosentpoeng for henholdsvis videregående skolegruppen og gruppen for høyere utdanning.

¹⁷ Basert på Std.avvik₂ i tabell 3.1 (standardavvikene til tidsserievariasjonen i forklaringsvariablene).

Vi ønsker nå å undersøke hvor stor del av standardavvikene til gruppens innmigrasjonsrate som kan forklares av den regionale arbeidsledighetsraten. Dette gjøres ved å øke ledighetsraten med ett standardavvik fra gjennomsnittet som vist over. Størst effekt finnes for videregående skolegruppen der dette vil resultere i en nedgang i gruppens innmigrasjonsrate tilsvarende 41 prosent av innmigrasjonsratens standardavvik. Tilsvarende får vi for grunnskolegruppen 14 prosent, mens det blir 15 prosent for gruppen med høyere utdanning.

For vakanseraten er forskjellen størst mellom gruppene grunnskole og høyere utdanning, men denne effekten er som de andre, ikke signifikant. Selv om effektene ikke er signifikante, tyder det på at gruppen med videregående skole-utdanning responderer mest på ledigheten og gruppen med høyere utdanning mest på vakanseraten i henhold til effekten på migrasjonen. Dette kan skyldes at de med høyere utdanning har bedre oversikt i jobbmarkedet og at de derfor vil respondere sterkere på denne variabelen.

En økning av vakanseraten med ett standardavvik fra gjennomsnittet gir i FE(I)-modellen en økning i netto innmigrasjon for grunnskolegruppen med 0,04 prosentpoeng, 0,07 prosentpoeng for videregående skolegruppen og 0,15 prosentpoeng for gruppen med høyere utdanning. En sjekk av vakanseratens forklaringsgrad viser at den er relativt lik for de tre gruppene. De tre utdanningsgruppene innmigrasjonsrate endres fra 8 til 13 prosent av innmigrasjonsratens standardavvik når variabelen for vakanseraten økes med ett standardavvik fra gjennomsnittet.

4.1.2 Utvidet modell. Den nasjonale konjunktursituasjonen

Vi utvider nå grunnmodellen og kontrollerer for hvordan den nasjonale konjunktursituasjonen vil påvirke den regionale arbeidsledigheten og vakanseraten i arbeidsmarkedet. Den nye modellspesifikasjonen vil da være definert ved (4.2) der vi har inkludert interaksjonsledd for å fange opp konjunktoreffekten. $\text{Log}(U_{it}) * \Delta N_t$ er interaksjonsleddet mellom regional arbeidsledighet og nasjonal sysselsettingsvekst, mens $\text{Log}(\text{VakS}_{it}) * \Delta N_t$ er tilsvarende for vakanseraten. Interaksjonsleddet vil vise endring i regional ledighet og vakanserate når vi får endringer i den nasjonale sysselsettingsveksten. Grunnmodellen som inkluderer interaksjonsledd for den nasjonale konjunktursituasjonen, er da gitt ved:

$$(4.2) \text{Innmigrasjon}[X]_{it} = \beta_1 \log(U_{it}) + \beta_{11} \log(U_{it}) * \Delta N_t + \beta_2 \log(\text{VakS}_{it}) + \beta_{21} \log(\text{VakS}_{it}) * \Delta N_t + V_{it}$$

Margineffekten på netto innmigrasjon av økt ledighetsrate er i (4.2) gitt ved $\beta_1 + \beta_{11} * \Delta N_t$, mens margineffekten av økt vakanserate er gitt ved $\beta_2 + \beta_{21} * \Delta N_t$. I appendikset viser figur 3A2 hvordan aggregert sysselsettingsvekst varierer over tidsperioden 1986 – 2004, mens figur 4A3 viser margineffekten av regional ledighetsrate for hver enkelt utdanningsgruppe.

Tabell 4.2. Effekten av regional arbeidsledighet og vakansestrøm på netto innmigrasjonsrate for ulike utdanningsgrupper når vi kontrollerer for den nasjonale konjunktursituasjonen. Avhengig variabel: $\text{Innmigrasjon}[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
	FE(II)	FE(II)	FE(II)
Log (U_{it})	-0,237 (-3,72)	-0,722 (-7,34)	-0,462 (-2,35)
Log (U_{it})* ΔN_t	-0,0021 (-0,132)	-0,038 (-2,13)	-0,204 (-2,76)
Log (VakS_{it})	0,098 (1,16)	0,139 (1,55)	0,269 (1,05)
Log (VakS_{it})* ΔN_t	-0,071 (-2,01)	-0,181 (-3,11)	-0,424 (-3,39)
Observasjoner	1710	1710	1710
R ²	0,080	0,202	0,162
σ	0,464	0,478	1,19
Wald (joint)	17,56**	64,97**	25,83**
AR(1)-test	-0,583	4,006**	2,153*
AR(2)-test	-0,304	2,841**	1,770

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

Ved å kontrollere for den nasjonale konjunktursituasjonen på de to arbeidsmarkedsvariablene gir dette at effekten på migrasjonen blir negativt påvirket for alle de tre utdanningsgruppene. Tolkningen blir at en lavkonjunktur gir økt ledighet samtidig som det tilbys færre stillinger i arbeidsmarkedet, slik at det blir mindre gunstig å skifte jobb og derav reduseres migrasjonen. Effektene er signifikante for gruppene videregående skole og høyere utdanning med t-verdier over 1,96. Responsen viser seg å være stigende i utdanningsnivået.

Den makroøkonomiske konjunkturindikatoren er målt som avvik fra sitt respektive gjennomsnitt. Marginaleffektene av regional ledighet og vakanser vil henholdsvis være lik β_1 og β_2 i en "normal" konjunktursituasjon. Som forventet har økt regional arbeidsledighet negativ effekt på netto innmigrasjon, mens økt regional vakanserate vil øke netto innmigrasjon. Den estimerte effekten av regional arbeidsledighet er statistisk signifikant for alle de tre utdanningsgruppene, mens estimert effekt av fylkesvis vakanserate er tilsvarende insignifikant. Videre ser vi at estimatene til interaksjonsleddene er negative og statistisk signifikant i alle tilfeller, bortsett fra ledighetseffekten for gruppen grunnskoleutdanning. Dette bekrefter igjen at effektene av regionale arbeidsmarkedsvariabler påvirkes av konjunktursituasjonen; ledighet betyr mest i oppgangstider, mens vakanser betyr mest i nedgangstider. Ved å kontrollere for effekten av den nasjonale konjunktursituasjonen, fant Carlsen, Johansen og Kasperen (2007) tilsvarende effekter for de regionale arbeidsmarkedsvariablene på netto innmigrasjon.

En mulighet for å illustrere betydningen for netto innmigrasjon av de regionale arbeidsmarkedsvariablene kan gjøres ved og først ta utgangspunkt i en konjunktursituasjon der makroindikatoren er lik sitt respektive gjennomsnitt, det vil si at interaksjonsleddene blir lik null. Estimert marginaleffekt av regional ledighet blir -0,237 for gruppen med grunnskoleutdanning som betyr at en økning i den regionale ledigheten med ett prosentpoeng, alt likt, vil redusere netto innmigrasjon med 0,237 prosentpoeng. Tilsvarende for gruppene videregående skole og høyere utdanning med henholdsvis -0,722 og -0,462. I artikkelen til Carlsen, Johansen og Kaspersen (2007) er den marginale effekten lik -0,184, der den avhengige variabelen $Innflytting_{it}$ ser på hele befolkningen samlet under ett. Ved å dele befolkningen opp i ulike utdanningsgrupper ser vi at de marginale effektene blir større. Dette på grunn av at forskjellene mellom og innad i hver gruppe kommer sterkere frem ved en økning i regional ledighet og ikke bare den totale effekten.

Margineffekten av regional arbeidsledighet er sterkere ved høykonjunktur enn ved lavkonjunktur, mens margineffekten av fylkesvis vakanserate er sterkest ved lavkonjunktur. Fra tabell 4.2 finner vi at margineffekten av regional arbeidsledighet for gruppen lavere utdanning endres fra -0,23 når nasjonal sysselsettingsvekst er ett prosentpoeng under gjennomsnittet til -0,24 når sysselsettingsveksten er ett prosentpoeng over gjennomsnittet. Tilsvarende fremgangsmåte gir for utdanningsgruppen videregående skole at regional arbeidsledighet endres fra -0,68 til ca -0,76, mens for gruppen med høyere utdanning blir endring i regional arbeidsledighet fra -0,26 til om lag -0,67. Carlsen, Johansen og Kaspersen (2007) finner at den samlede endringen blir -0,15 til -0,22. Sammenlignet med våre resultater ser man at forskjellen er størst sett i forhold til gruppen høyere utdanning, og at det kan sies at forskjellen er økende i utdanningsnivået. Det samme kan sies om endringen i fylkesvis vakanserate som de finner til å endres fra 0,056 til 0,025. Margineffekten av fylkesvis vakanserate avhenger både av aggregert sysselsettingsvekst og aggregert vakanserate. I vårt tilfelle inkluderes ikke aggregerte vakanser i regresjonen. Når sysselsettingsveksten endres fra ett prosentpoeng under gjennomsnittet til ett prosentpoeng over gjennomsnittet gir dette en endring i margineffekten av fylkesvis vakanserate fra 0,17 til 0,03 for gruppen med grunnskoleutdanning. Tilsvarende for gruppene videregående skole-utdanning og høyere utdanning med henholdsvis fra 0,32 til -0,04 og fra 0,70 til -0,16.

Ved å tilstramme arbeidsmarkedsforholdene ved samtidig å øke ledigheten og redusere vakanseraten med ett standardavvik fra gjennomsnittene finner vi at netto innmigrasjon reduseres med 0,25 prosentpoeng for gruppen høyere utdanning. Det er 0,14 prosentpoeng lavere enn i tabell 4.1. Resultatene ved å øke ledigheten og vakanseraten hver for seg blir lik som i tabell 4.1, men en forskjell som kan nevnes er at vakanseraten for gruppen høyere utdanning her får en positiv påvirkning med 0,08 prosentpoeng imot 0,15 prosentpoeng tidligere. Dette medfører en reduksjon i utdanningsgruppens innmigrasjonsrate fra 12 % til 6 % av innmigrasjonsratens standardavvik. I en lavkonjunktur blir den positive effekten av vakanseraten på netto innmigrasjon lavest for de med høyere utdanning når vi tar hensyn til den samlede effekten. Dette ser vi ut ifra tabell 4.2 der effekten av den nasjonale konjunktursituasjon viser seg å være størst for denne utdanningsgruppen. En tolkning av dette blir at i en lavkonjunktur vil de med høyere utdanning redusere sin netto innmigrasjon i større grad sammenlignet med de to andre gruppene. Dette kan ses i sammenheng med at høyt utdannede antas å ha flere jobbtillbud samtidig som mobiliteten antas å være størst for denne gruppen.

Tabell 4.2.1. Gruppesammenligning av estimerte koeffisientforskjeller fra FE(II)-modellene. Avhengige variabler: $\text{Innmigrasjon}[X_1]_{it} - \text{Innmigrasjon}[X_2]_{it}$, der X_1 og X_2 står for utdanningsgruppene. X vil være GRU, VGS og HØY.

Variabler	Gruppesammenligning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(II)	D-FE(II)	D-FE(II)
Log (U_{it})	0,485 (6,23)	0,225 (1,24)	-0,26 (-1,70)
Log (U_{it})* ΔN_t	0,036 (1,66)	0,202 (2,66)	0,166 (2,43)
Log ($VakS_{it}$)	-0,042 (-0,458)	-0,172 (-0,613)	-0,12 (-0,525)
Log ($VakS_{it}$)* ΔN_t	0,11 (2,82)	0,353 (3,03)	0,243 (2,06)

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik.

Ved å inkludere effekten av konjunktursituasjonen vil det fortsatt være en signifikant forskjell mellom gruppene, men effekten er noe redusert. De estimerte konjunktoreffektene er signifikante og vi legger merke til at forskjellen i arbeidsledigheten blir mindre mellom de to første gruppene, mens den øker mellom gruppene videregående skole og høyere utdanning sammenlignet med tabell 4.1.1. Dette fordi netto innmigrasjon for de tre gruppene reduseres i størst grad for gruppen høyere utdanning i en lavkonjunktur. Effekten av vakanseraten vil også reduseres mellom gruppene, men t-verdier lavere enn 0,6 gir at disse effektene ikke er signifikante. Gruppesammenligningen viser at forskjellen med hensyn på konjunktursituasjonen støtter resultatene om at responsen øker med utdanningsnivået siden koeffisientforskjellene er størst mellom gruppene grunnskole og høyere utdanning. Disse effektene er signifikante og viser til t-verdier lik 2,66 og 3,03.

4.1.3 Utvidet modell. Kommunale inntekter

Her vil modellen utvides til å se på hvordan inkludering av kommunale inntekter vil påvirke flytteresponsen. Sammenlignet med grunnmodellen i tabell 4.1 vil estimering av utvidet modell gi at vakanseraten og den regionale ledigheten endres relativt lite. Det er kun for gruppen høyere utdanning at vi får en merkbar reduksjon i effekten av regional ledighet på netto innmigrasjon. Vi kan dermed si at disse to estimerte effektene er robuste. Estimaten for de to variablene ligger for FE(III) mellom 17 til 24 prosent lavere enn i FE(I). Resultatene fra fixed-effects-regresjonen er rapportert i tabell 4.3.

Tabell 4.3. Effekt av regional arbeidsledighet, vakanserate og kommunale inntekter på netto innmigrasjonsrate. Avhengig variabel: Innmigrasjon $[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
	FE(III)	FE(III)	FE(III)
Log (U_{it})	-0,213 (-3,28)	-0,708 (-7,10)	-0,492 (-2,47)
Log ($VakS_{it}$)	0,133 (1,59)	0,234 (2,51)	0,513 (1,85)
Log ($KommInn_{it}$)	0,413 (1,93)	0,698 (1,95)	2,28 (3,00)
Observasjoner	1710	1710	1710
R^2	0,0786	0,186	0,139
σ	0,464	0,482	1,213
Wald (joint)	19,02**	55,42**	21,03**
AR(1)-test	-0,5012	4,107**	2,234*
AR(2)-test	-0,3325	3,102**	1,827

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

En økning i kommunale inntekter vil gi grunnlag for et bedre kommunalt tjenestetilbud. Forventningen til denne variabelen er at den er positiv for alle utdanningsgruppene fordi den kan antas å gjøre hver region mer attraktiv og kan partielt sett bidra til økt innflytting. Et mulig problem er at kommunale inntekter kan antas å være korrelert med regionale priser, noe som kan gi skjevhet i estimatene. En positiv korrelasjon mellom regionale priser og kommunale inntekter kan dermed gi underestimering for effekten av kommunale inntekter, da det forventes at økte regionale priser partielt sett vil ha en negativ effekt på netto innmigrasjon. Resultatene i våre FE-estimeringer viser at effektene på kommunale inntekter er positive for utdanningsgruppene og det tyder da på at within group-transformasjonen kan korrigere for en antatt korrelasjon med utelatte regionale priser.

Vi ser av resultatene i tabell 4.3 at økte kommunale inntekter bidrar til økt netto innmigrasjon for alle gruppene, men at variabelen har klart størst effekt for gruppen høyere utdanning. For denne gruppen vil den statistiske koeffisientverdien være signifikant og den ligger langt over estimatet for de to andre gruppene. Ved å se på estimerte koeffisientforskjeller i tabell 4.3.1 bekreftes det at estimert effekt for gruppen høyere utdanning er signifikant høyere enn tilsvarende estimat for de to andre gruppene.

Tabell 4.3.1. Gruppesammenligning av estimerte koeffisientforskjeller fra FE(III)-modellene. Avhengige variabler: $\text{Innmigrasjon}[X_1]_{it} - \text{Innmigrasjon}[X_2]_{it}$, der X_1 og X_2 står for utdanningsgruppene. X vil være GRU, VGS og HØY.

Variabler	Gruppesammenligning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(III)	D-FE(III)	D-FE(III)
Log (U_{it})	0,494 (6,10)	0,278 (1,53)	-0,216 (-1,43)
Log ($VakS_{it}$)	-0,10 (-1,06)	-0,381 (-1,28)	-0,281 (-1,09)
Log ($KommInn_{it}$)	-0,285 (-0,774)	-1,86 (-2,20)	-1,58 (-2,38)

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik.

Ved å øke kommunale inntekter med ett standardavvik fra gjennomsnittet øker netto innmigrasjonsrate med 0,24 prosentpoeng for gruppen med høyere utdanning. Denne variabelen er signifikant med en t-verdi på 3. For de to andre gruppene er også variabelen tilnærmet signifikant ved 5 prosents signifikansnivå ved at de får estimert en t-verdi lik 1,93 og 1,95. For grunnskolegruppen og videregående skolegruppen vil en økning med ett standardavvik gi økt netto innmigrasjon med 0,04 og 0,07 prosentpoeng. Vi merker oss at de to siste estimatene ligger nærme hverandre, men av resultatene så langt har vi grunn til å tro at kommunale inntekter gir økt flytterespons ved stigende utdanningsnivå. Økning med ett standardavvik fra gjennomsnittet vil medføre en økning i utdanningsgruppens innmigrasjonsrate med 8 prosent av innmigrasjonsratens standardavvik for grunnskolegruppen, 13 prosent for videregående skolegruppen og 18 prosent for gruppen med høyere utdanning.

4.1.4. Fullstendig modell

I tabell 4.4 finner vi de estimerte koeffisientene for variablene i den fullstendige modellen. Her inngår regional arbeidsledighet, vakanser på fylkesnivå og kommunale inntekter, samtidig som det kontrolleres for hvordan den nasjonale konjunktursituasjonen vil påvirke arbeidsmarkedsvariablene. Resultatene ved å inkludere kommunale inntekter gir små endringer for estimerte effekter av arbeidsmarkedsvariablene og de tilhørende marginaleffektene når man sammenligner med tabell 4.2. Effekten av regional ledighet på netto innmigrasjon blir noe redusert for de tre gruppene, mens effekten av vakanserate blir bortimot uendret.

Tabell 4.4. Effekten av regional arbeidsledighet, vakansestrøm og kommunale inntekter på netto innmigrasjonsrate for ulike utdanningsgrupper når vi kontrollerer for den nasjonale konjunktursituasjonen. Avhengig variabel: $\text{Innmigrasjon}[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
	FE(IV)	FE(IV)	FE(IV)
Log (U_{it})	-0,218 (-3,27)	-0,691 (-6,97)	-0,362 (-1,69)
Log (U_{it})* ΔN_t	-0,00075 (-0,0465)	-0,0359 (-2,03)	-0,197 (-2,72)
Log (VakS _{it})	0,102 (1,22)	0,146 (1,62)	0,289 (1,13)
Log (VakS _{it})* ΔN_t	-0,069 (-1,95)	-0,178 (-3,06)	-0,414 (-3,40)
Log (KommInn _{it})	0,389 (1,81)	0,611 (1,96)	1,994 (2,92)
Observasjoner	1710	1710	1710
R ²	0,082	0,205	0,168
σ	0,464	0,477	1,194
Wald (joint)	20,87**	60,42**	31,51**
AR(1)-test	-0,6346	3,923**	2,007*
AR(2)-test	-0,3986	2,801**	1,614

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

Estimert effekt av den nasjonale konjunktursituasjonen på regional ledighet er signifikant for gruppene videregående skole og høyere utdanning, mens effekten på vakanseraten er signifikant for alle tre gruppene. Ved å kontrollere for den nasjonale konjunktursituasjonen ser vi at det er gruppen høyere utdanning som responderer mest negativ på netto innmigrasjon. De estimerte effektene er signifikant med absolutte t-verdier lik 2,72 og 3,40.

Fra tabell 4.4 får vi at estimert marginaleffekt av regional ledighet blir -0,218 for gruppen grunnskoleutdanning når makroindikatoren er lik sitt respektive gjennomsnitt. Interaksjonsleddene faller da bort. Dette betyr at en økning i den regionale ledigheten med ett prosentpoeng, alt likt, vil redusere netto innmigrasjon med 0,218 prosentpoeng. Tilsvarende for gruppene videregående skole- og høyere utdanning med henholdsvis 0,691 og 0,362 prosentpoeng. Samme metode som i delkapittel 4.1.2 gir her at marginaleffekten av regional arbeidsledighet for gruppen lavere utdanning endres fra -0,21 når nasjonal sysselsettingsvekst er ett prosentpoeng under gjennomsnittet til -0,22 når sysselsettingsveksten er ett

prosentpoeng over gjennomsnittet. Tilsvarende fremgangsmåte gir for gruppen videregående skole at regional arbeidsledighet endres fra -0,65 til ca -0,73, mens for gruppen med høyere utdanning blir endringen i regional arbeidsledighet fra -0,16 til om lag -0,56. For fylkesvis vakanserate vil en endring i sysselsettingsveksten fra ett prosentpoeng under gjennomsnittet til ett prosentpoeng over gjennomsnittet gi en endring i marginaleffekten fra 0,17 til 0,03 for gruppen med grunnskoleutdanning. Tilsvarende for gruppene videregående skole og høyere utdanning fra henholdsvis 0,33 til -0,03 og fra 0,71 til -0,13. En inkludering av kommunale inntekter bidrar dermed ikke til noen nevneverdige endringer for marginaleffektene sammenlignet med resultatene i avsnitt 4.1.2.

Vi kan illustrere en tilstramming av arbeidsmarkedsforholdene ved samtidig å øke ledighetsraten og redusere vakanseraten med ett standardavvik fra gjennomsnittene. Brukes resultatene fra FE(IV) vil dette redusere netto innmigrasjonrate med 0,11 prosentpoeng for grunnskolegruppen. For gruppen videregående skole og gruppen for høyere utdanning blir den totale effekten av arbeidsmarkedsvariablene over to ganger sterkere, da netto innmigrasjonsrate for gruppene reduseres med henholdsvis 0,27 og 0,23 prosentpoeng.

Tabell 4.4.1. Gruppesammenligning av estimerte koeffisientforskjeller fra FE(IV)-modellene. Avhengige variabler: Innmigrasjon $[X_1]_{it}$ – Innmigrasjon $[X_2]_{it}$, der X_1 og X_2 står for utdanningsgruppene. X vil være GRU, VGS og HØY.

Variabler	Gruppesammenligning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(IV)	D-FE(IV)	D-FE(IV)
Log (U_{it})	0,474 (6,12)	0,144 (0,74)	-0,33 (-1,95)
Log (U_{it})* ΔN_t	0,035 (1,61)	0,196 (2,63)	0,16 (2,40)
Log (VakS $_{it}$)	-0,044 (-0,48)	-0,188 (-0,671)	-0,14 (-0,582)
Log (VakS $_{it}$)* ΔN_t	0,109 (2,82)	0,345 (3,05)	0,24 (2,04)
Log (KommInn $_{it}$)	-0,22 (-0,653)	-1,61 (-2,03)	-1,38 (-2,14)

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik.

Gruppesammenligningen i D-FE(IV) forsterker resultatene fra tabell 4.4 ved å vise at gruppen videregående skole responderer sterkest på arbeidsledigheten, mens gruppen for høyere utdanning responderer sterkest på vakanseraten og kommunale inntekter når vi inkluderer

kontrollvariable for den nasjonale konjunktursituasjonen. Koeffisientverdiene viser oss at gruppene reagerer ulikt og at effekten av konjunktursituasjonen får større betydning for migrasjonen når utdanningsnivået øker. For eksempel får vi at en reduksjon i netto innmigrasjon forsterkes av økt utdanningsnivå når vi foretar en tilstramming av arbeidsmarkedsforholdene.

4.1.5 Diagnostiske tester

I paneldata har R^2 en noe uklar tolkning, men kan i utgangspunktet sies å gi en indikasjon på hvor mye av tidsvariasjonen i den avhengige variabelen $Innmigrasjon_{it}$ som forklares av modellens inkluderte høyresidevariabler. Uavhengig av hvilken FE-modell vi ser på i den statiske modellspesifikasjonen ligger R^2 for utdanningsgruppen grunnskole på rundt 8 prosent, mens den er 19 prosent for videregående skole. For gruppen høyere utdanning ligger den rundt 13 prosent. I regresjonen er nok flere viktige variable som påvirker flytteresponsen og arbeidsmarkedsituasjonen utelatt og det er nok en av grunnene til at R^2 ikke har større forklaringsgrad i FE-modellene.

Wald (joint) forkaster i alle modellene en nullhypotese ved 1 prosents signifikansnivå om at de uavhengige variablene ikke har forklaringskraft. Wald (time) som ikke er rapportert, viser at tidsdummyene er signifikant ved 1 prosents signifikansnivå i alle modellene.

Sigma viser restleddets estimerte standardavvik. Størst restleddsvariasjon finner vi i grunnmodellen for høyere utdanning med en verdi lik 1,22. Grunnskole og videregående skole ligger i størrelsesorden rundt 0,46 til 0,48. Siden estimerte standardavvik til restleddet er større for gruppen høyere utdanning i alle modellene, kan det være nærliggende å tenke seg at flyttebeslutningene til personer med høyere utdanning påvirkes av flere variabler enn hva tilfellet er for personer med lavere utdanning. Vi kan anta at egenskaper og interesser som fører til at en person velger høyere utdanning er forskjellig fra en annen person som velger et lavere utdanningsnivå.

AR(1) og AR(2) er tester for første- og andreordens seriekorrelasjon i restleddet. Restleddet i regresjonen vil per definisjon være seriekorrelert på grunn av det regionspesifikke, tidsinvariante restleddet. Basert på forutsetningene til FE-estimering vil ikke det transformerte

restleddet være seriekorrelert, som beskrevet under forutsetningen iv) i kapittel 3.3.1. Dette er ikke et problem for grunnskolegruppen siden AR(1) ikke forkastes i noen tilfeller. For de to øvrige gruppene forkaster AR(1) og AR(2) en nullhypotese om ingen seriekorrelasjon ved 1 prosents signifikansnivå for videregående skole, og ved 5 prosents signifikansnivå for høyere utdanning.

Problemet med seriekorrelasjon i restleddene kan oppstå som følge av umodellert dynamikk. Som for eksempel misspesifikasjon av modellen, forventningstregheter eller fra spatial korrelasjon, det vil si korrelasjon som følge av påvirkning fra nærliggende observasjoner i tverrsnittsenheten. Konsekvensene blir at vi får korrelasjon mellom inkluderte, eller mellom utelatte, variabler *over tid* som vil påvirke standardavvik og inferens, som vi igjen må ta høyde for i våre resultater. I kapittel 5 utvider vi analysen med en dynamisk spesifisering og kan da undersøke om problemet med seriekorrelasjon opphører.

4.1.6 Oppsummering

Til nå har vi sett på sammenhengen mellom de tre utdanningsgruppene netto innmigrasjonsrate og følgende forklaringsvariabler: Regional arbeidsledighetsrate, fylkesvis vakanserate og kommunale inntekter, samtidig som vi har kontrollert for den nasjonale konjunktursituasjonen. Vi har estimert modellene med fixed-effects og i tillegg har vi testet for koeffisientforskjeller mellom gruppene.

Uavhengig av utdanningsgruppe er estimatene for den regionale arbeidsledighetsraten negative og signifikant. Ved å se på effekten av den regionale arbeidsledighetsraten på netto innmigrasjon framgår det av estimatene at grunnskolegruppen skiller seg fra de to andre gruppene. For gruppene videregående skole og høyere utdanning er effekten to til tre ganger sterkere. Estimaterne for vakanseraten viser seg å være positiv og i de aller fleste tilfellene signifikant forskjellig fra null i ensidige tester. Estimaterne blir også mer signifikant i de regresjonene der vi ikke tar hensyn til den nasjonale konjunktursituasjonen. Effekten av vakanseraten virker å være stigende i utdanningsnivået og gruppene koeffisienter er ikke signifikant forskjellig i de statiske modellene.

Estimert effekt av den nasjonale konjunktursituasjonen på vakanseraten er signifikant for alle gruppene, mens samme effekt på den regionale arbeidsledighetsraten er signifikant for gruppene videregående skole og høyere utdanning i den statiske modellen. Koeffisientverdiene viser dermed at effekten av konjunkturvariablene på netto innmigrasjon er stigende med utdanningsnivået.

Estimerte effekter av kommunale inntekter er positive og signifikant for alle gruppene. Responsen for gruppen høyere utdanning skiller seg med ut med en koeffisientverdi som er 3 til 4 ganger høyere enn de andre gruppene, samtidig som disse verdiene er signifikant forskjellig gruppene imellom.

4.2 Sensitivitetsanalyse

For å teste resultatenes stabilitet med hensyn til utvalgsvariasjon i tids- og tverrsnittsdimensjonen, deles tidsserien i to kortere perioder. Her vil tabell 4.4 bli re-estimert basert på underutvalg av datamaterialet. Den fulle tidsserien i datamaterialet strekker seg fra 1986 til 2004, men modellen vil her estimeres for periodene 1986-1995 og 1996-2004. I tillegg benytter vi i dette delkapitlet muligheten for å undersøke effekten av boligpris der tidsserien (1991-2004) er kortere enn for de øvrige variablene.

4.2.1 Utvalgte tidsperioder

Ved å dele tidsserien i to, står vi igjen med 900 og 810 observasjoner i datamaterialet for de to periodene. Dette gjør at variasjonen reduseres betraktelig og vi bør derfor være ekstra oppmerksomme på begrensningene FE-estimeringen nå innebærer, diskutert i kapittel 3. Resultater fra regresjonen av den fullstendige modellen er gjengitt i tabell 4.5 og viser den estimerte effekten av arbeidsmarkedsvariablene på netto innmigrasjon for de tre utdanningsgruppene.

Tabell 4.5. Delutvalgsestimeringer for periodene 1986-1995 og 1996-2004. Avhengig variabel: Innmigrasjon $[X]_{it}$. Transformasjon: Within groups.

Tidsperiode →	Gruppe					
	GRU		VGS		HØY	
	86-95	96-04	86-95	96-04	86-95	96-04
Variabler	FE(V)	FE(VI)	FE(V)	FE(VI)	FE(V)	FE(VI)
Log (U_{it})	-0,311 (-3,24)	-0,161 (-1,08)	-0,784 (-6,18)	-0,701 (-4,72)	-0,170 (-0,754)	-1,04 (-3,32)
Log (U_{it})* ΔN_t	0,002 (0,097)	-0,058 (-1,73)	-0,0038 (-0,191)	-0,258 (-4,18)	-0,206 (-2,80)	-0,252 (-1,78)
Log (Vak S_{it})	-0,081 (-0,59)	0,049 (0,384)	0,0034 (0,023)	0,044 (0,397)	0,510 (1,51)	-0,549 (-1,86)
Log(Vak S_{it})* ΔN_t	-0,071 (-1,91)	0,113 (1,54)	-0,167 (-2,86)	0,018 (0,192)	-0,370 (-2,80)	-0,299 (-1,54)
Log(KommInn $_{it}$)	-0,415 (-1,15)	-0,175 (-0,349)	-0,254 (-0,584)	0,491 (0,806)	0,925 (1,09)	4,552 (4,28)
Observasjoner	900	810	900	810	900	810
R ²	0,048	0,094	0,242	0,183	0,153	0,206
σ	0,42	0,497	0,469	0,428	1,294	0,977
Wald (joint)	19,34**	9,922	56,19**	52,03**	31,80**	44,71**
AR(1)-test	-0,345	-4,332**	2,496*	0,231	-1,339	0,211
AR(2)-test	-1,346	-2,396*	-0,584	-0,575	-1,213	-2,511*

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

For gruppene grunnskole og videregående skole viser estimatene for ledighetsraten seg å være relativt like for de to tidsperiodene. Resultatene er også stabile sammenlignet med resultatene for hele perioden 86-04, gjengitt i modellene FE(III) og FE(IV) i tabell 4.3 og 4.4. Signifikansen er ikke uventet lavere sammenlignet med estimatene for hele perioden, og da spesielt for gruppene grunnskole og videregående skole i den siste tidsperioden. For gruppen høyere utdanning er estimatet for perioden 86-95 noe lavere enn for hele perioden, mens estimatet for perioden 96-04, er tre ganger så høyt.

Estimatene for vakanseraten har blitt enda lavere sammenlignet med hele perioden og i tillegg er estimatene insignifikant for alle gruppene med unntak av gruppen høyere utdanning. Fra FE(VI) ser man at t-verdien og koeffisientestimatet ligger noe høyere. En forklaring til de lave estimatene kan være på grunn av korrelasjon med den regionale ledighetsraten, samtidig som at vi benytter within-groups transformasjonen og dette fører til at noe av variasjonen i data forsvinner. I tillegg er data for vakanseraten gitt på fylkesnivå og dette gir at variasjonen i tverrsnittsdimensjonen er lav i utgangspunktet. Den resterende variasjonen kan dermed bli utilstrekkelig for å oppnå presise estimater og det blir ikke bedre av at man benytter delperiodeestimeringer.

For estimatene av den nasjonale konjunktursituasjonen på regional ledighet er disse signifikant for de to første gruppene i den siste perioden, mens estimatene for gruppen høyere utdanning er signifikant i begge periodene. Resultatene ved å sammenligne den siste perioden i FE(VII) for de to siste utdanningsgruppene med hele perioden, gjengitt som FE(IV) i tabell 4.4, viser til å være stabile. Tilsvarende finner vi at den nasjonale konjunktursituasjonen på vakanseraten gir signifikante estimater for perioden 86-95 for alle de tre utdanningsgruppene. Sammenligner vi estimatene i denne perioden med hele perioden, gjengitt i tabell 4.4, vil også resultatene her være stabile. Estimaten for den siste perioden har varierende fortegn for de to første gruppene samtidig som signifikansen er lav for alle de tre utdanningsgruppene. Den nasjonale konjunktursituasjonen har ulik påvirkningskraft på de to arbeidsmarkedsvariablene. Estimert effekt på regional ledighet ser ut til å være sterkest i perioden 96-04, mens effekten på vakanseraten er sterkest i perioden 86-95.

Estimatene for kommunale inntekter har lav signifikans for alle periodene med unntak av den siste perioden, 96-04, som gjelder gruppen høyere utdanning. Responsen er her markant høyere sammenlignet med den første perioden. Koeffisientverdiene er positive for denne

gruppen, men man ser av t-verdiene at estimatet er mindre presist for den første perioden, 86-95. Vi merker oss også at estimatene har varierende fortegn for de to første utdanningsgruppene.

4.2.2 Boligpris

Her utvides modellen med boligpris. Tidsserien som for denne variabelen strekker seg fra 1991-2004 er dermed kortere enn for de øvrige variablene i regresjonen. Carlsen og Johansen (2004) og Dahl (2007) finner signifikante negative effekter av boligpris på netto innmigrasjon. I tillegg kommer Dahl (2007) frem til at boligprisen har sterkest effekt på gruppene videregående skole og høyere utdanning. Tabell 4.6 rapporterer resultatene for den utvidede modellen.

Tabell 4.6. Estimer fra regresjonsmodeller som inkluderer boligpris. Avhengig variabel: Innmigrasjon $[X]_{it}$.

	Gruppe		
	GRU	VGS	HØY
Tidsperiode →	91-04	91-04	91-04
Variabler	FE(VII)	FE(VII)	FE(VII)
Log (U_{it})	-0,415 (-4,04)	-0,981 (-7,26)	-1,16 (-4,72)
Log (U_{it}) * ΔN_t	0,007 (0,229)	-0,135 (-2,97)	-0,271 (-2,16)
Log ($VakS_{it}$)	0,147 (1,39)	0,129 (1,19)	-0,305 (-0,89)
Log ($VakS_{it}$) * ΔN_t	-0,005 (-0,099)	-0,181 (-2,29)	-0,448 (-2,47)
Log ($KommInn_{it}$)	0,913 (2,45)	1,605 (3,91)	4,19 (4,31)
Log ($Boligpris_{it}$)	-0,215 (-1,60)	-0,063 (-0,33)	0,275 (0,734)
Observasjoner	1260	1260	1260
R^2	0,109	0,212	0,167
σ	0,465	0,446	1,059
Wald (joint)	60,03**	81,23**	52,17**
AR(1)-test	-1,986*	2,089*	2,077*
AR(2)-test	-1,553	1,907	0,2520

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

Resultatene fra tabell 4.6 viser at responsen på variabelen boligpris er lav for alle de tre utdanningsgruppene. FE(VIII) viser at boligprisen har to negative og ett positivt koeffisientestimat, samt at den estimerte effekten for gruppene videregående skole og høyere utdanning er insignifikant. Man kan tenke seg at alle forklaringsvariablene som inngår i analysen vil være beheftet med potensiell simultanitetsskjevheter av ulik grad, men på grunn av de lave estimatene kan dette fremstå som et synlig problem for variabelen boligpris. Demografiske faktorer som flyttemønster og utdanning blant befolkningen vil kunne påvirke etterspørselen etter bolig og boligprisnivået, gjennom ulike behov og hvilke boligtyper som etterspørres. Man kan også forvente at den nasjonale konjunktursituasjonen vil sette sitt preg på boliggetterspørselen, og en studie av Andersen (2001) viser til at førstegangsetablerere helst investerer i boligmarkedet i en oppgangskonjunktur, og at den største andelen av førstegangsetablerere følger konjunktorene. Det er rimelig å anta at oppgangskonjunkturer gir en økning i boligprisene, som igjen gir forventninger om enda høyere fremtidig pris. Ut ifra dette vil det være nærliggende å betrakte boligprisen som endogen på grunn av simultanitet. I kapittel 5 vil simultanitet drøftes nærmere. Som vi skal se vil estimering ved hjelp av instrumentvariabelmetoden gi oss mer pålitelige resultater.

5 Utvidelser – dynamisk modellspesifikasjon og GMM

Så langt i analysen har vi benyttet statiske modeller ved estimering og disse ser på *momentane* effekter av arbeidsmarkedsvariablene på flytteresponsen. I dette kapitlet ønsker vi å ta hensyn til at flyttebeslutninger er en prosess som gjerne bygger på fremtidige forventninger, og man kan dermed anta at en slik avgjørelse tar tid å beslutte. En fordel ved å benytte paneldata er at dette tillater for undersøkelse av den dynamiske tilpasningsprosessen.¹⁸ På bakgrunn av dette vil analysen bli utvidet med en dynamisk modellspesifikasjon og en *generalisert momentmetode* (GMM) for å redusere potensielle simultanitetsproblemer.

Delkapittel 5.1 re-spesifiserer modellen fra tabell 4.4 ved å inkludere både løpende og laggede verdier på variablene. WG-transformasjon benyttes for å komme frem til estimeringsresultatene og disse sammenlignes med den statiske analysen. Delkapittel 5.2 presenterer kort prinsippene bak Arellano & Bonds metode, mens delkapittel 5.3 gir empiriske resultater fra regresjoner der forklaringsvariablene er instrumenterte. Del 5.4 gir en oppsummering.

5.1 Re-spesifikasjon av fullstendig modell

Utvider nå den statiske modellspesifikasjonen fra tabell 4.4 til en dynamisk modell, som estimeres ved hjelp av WG-transformasjonen. Dette gir oss muligheten til å undersøke en eventuell effekt på innmigrasjonen på tidspunkt t , av verdier for de uavhengige variablene fra tidspunkt $t-1$. Forklaringsvariablene fremgår i tabellen som både løpende og laggede verdier (datert tilbake ett år). Den dynamiske modellspesifikasjonen blir gitt ved:

$$(5.1) \text{Innmigrasjon}[X]_{it} = \beta_1 \log(U_{it}) + \beta_{11} \log(U_{it-1}) + \beta_{12} \log(U_{it}) * \Delta N_t + \beta_{13} \log(U_{it-1}) * \Delta N_{t-1} + \beta_2 \log(\text{VakS}_{it}) + \beta_{21} \log(\text{VakS}_{it-1}) + \beta_{22} \log(\text{VakS}_{it}) * \Delta N_t + \beta_{23} \log(\text{VakS}_{it-1}) * \Delta N_{t-1} + V_{it}$$

Restleddet består fortsatt av to komponenter:

$$V_{it} = \eta_i + \varepsilon_{it}$$

Som beskrevet i kapittel 3, fanger η_i opp regionspesifikke effekter som er konstante over tid, mens det idiosynkratiske restleddet, ε_{it} vil inneholde uobserverbare variabler som varierer både over tid og for hver region.

¹⁸ Se Baltagi (2001: kap 8): Oversikt over dynamiske paneldatamodeller.

5.1.1. Resultater fra dynamisk fixed effects-estimering

Koeffisientene fra estimering av den dynamiske modellspesifikasjonen er gitt i tabell 5.1, modell FE(VIII). I de to påfølgende modellene (se tabell 5.1 og 5.1.1) er variabler som ikke viste seg å ha en sterk eller signifikant effekt ved ulike signifikansnivå, utelatt ved estimering. Fullstendig modell kan ses i tabell A4 og A5 i appendikset. Vi finner at resultatene fra den dynamiske FE-analysen i liten grad skiller seg fra de statiske modellene, og dette kan da bidra til å styrke valget av en statistisk modellspesifikasjon i kapittel 4. Den statiske modellen kan dermed betraktes som en langsiktig løsning på en bakenforliggende dynamisk modell.

Tabell 5.1. Dynamisk modellspesifikasjon med Within Groups-transformasjon. Avhengig variabel: $Innmigrasjon_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
	FE(VIII)	FE(VIII)	FE(VIII)
Log (U_{it})	-0,192 (-2,78)	-0,629 (-5,66)	-1,07 (-3,50)
Log (U_{it-1})	-	-	0,744 (2,60)
Log (U_{it})* ΔN_t	0,028 (0,136)	-0,050 (-2,30)	-0,232 (-2,45)
Log ($VakS_{it}$)	0,099 (1,21)	0,266 (2,85)	-
Log ($VakS_{it-1}$)	-	-	0,480 (1,73)
Log ($VakS_{it}$)* ΔN_t	-	-	-0,391 (-2,83)
Log ($VakS_{it-1}$)* ΔN_{t-1}	-0,132 (-3,25)	-0,179 (-3,85)	-
Log (KommInn $_{it}$)	0,378 (1,67)	-	2,09 (2,99)
Log (KommInn $_{it-1}$)	-	0,957 (2,71)	-
Observasjoner	1710	1710	1710
R ²	0,091	0,193	0,183
σ	0,463	0,468	1,176
Wald (joint)	25,62**	56,35**	56,54**
AR(1)-test	-1,110	3,374**	1,856
AR(2)-test	-1,012	2,093*	1,726

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

De estimerte effektene av løpende verdi på regional arbeidsledighet inngår med forventet fortegn for de tre gruppene og estimatene er signifikant ved 1 % nivå. For gruppen høyere utdanning er langtidseffekten av ledigheten på netto innmigrasjon lik -0,326. Det kan dermed tyde på at migrasjonsbeslutninger tas relativt raskt og dette gir grunnlag til å forkaste en eventuell hypotese om at det er treghet i tilpasningen. Sammenlignet med tidligere resultater får vi at for gruppen videregående skole vil den laggede verdien gi et sterkere og mer signifikant estimat for kommunale inntekter. Estimaterne på konjunkturvariablene gir fortsatt at gruppen høyere utdanning responderer sterkest på konjunkturvariablene.

En tilstramning av arbeidsmarkedsforholdene illustreres ved samtidig å øke ledigheten og redusere vakanseraten med ett standardavvik fra gjennomsnittene. Estimaterne fra FE(V) gir da en reduksjon i netto innmigrasjonsrate for grunnskolegruppen med 0,06 prosentpoeng. For gruppene videregående skole og høyere utdanning vil reduksjonen være lik 0,30 prosentpoeng. Sammenlignet med tabell 4.4 vil en tilstramning av arbeidsmarkedsforholdene her gi at effekten på netto innmigrasjon blir noe mer negativ for gruppen høyere utdanning, endres fra 0,23 til 0,30 prosentpoeng. For de to andre gruppene blir det liten eller ingen effekt.

Tabell 5.1.1. Gruppesammenligning av estimerte koeffisientforskjeller fra FE(V)-modellene. Avhengige variabler: $\text{Innmigrasjon}[X_1]_{it} - \text{Innmigrasjon}[X_2]_{it}$, der X_1 og X_2 står for utdanningsgruppene. X vil være GRU, VGS og HØY.

Variabler	Gruppesammenligning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(VIII)	D-FE(VIII)	D-FE(VIII)
Log (U_{it})	0,426 (5,18)	0,834 (2,82)	-
Log (U_{it-1})	-	-0,716 (-2,50)	-0,403 (-2,55)
Log (U_{it})* ΔN_t	0,049 (2,26)	0,234 (2,37)	0,207 (2,29)
Log ($VakS_{it-1}$)	-0,216 (-2,40)	-5,05 (-1,77)	-0,299 (-1,16)
Log ($VakS_{it}$)* ΔN_t	0,082 (1,89)	0,394 (2,83)	0,345 (2,66)
Log ($VakS_{it-1}$)* ΔN_t	-	-0,141 (-1,80)	-0,144 (-1,83)
Log (KommInn $_{it}$)	-	-1,72 (-1,80)	-1,456 (-2,23)
Log (KommInn $_{it-1}$)	-0,689 (-1,68)	-	-

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik.

Den største forskjellen ved å øke ledigheten og vakanseraten hver for seg, er at gruppen med høyere utdanning får en økt positiv effekt på vakanseraten med 0,06 prosentpoeng. Det vil si en økning fra 0,08 til 0,14 prosentpoeng. Dette gir oss muligheten til å undersøke hvor stor del av standardavvikene til gruppenes innmigrasjonsrate som kan forklares av den regionale arbeidsledighetsraten og den fylkesvise vakanseraten. Ved sammenligning av tabell 5.1 og de andre modellspesifikasjonene i analysen får vi at gruppen videregående skole vil få en lavere forklaringsgrad for regional arbeidsledighet enn tidligere. Samme gruppe vil få en reduksjon i forklaringsgraden for vakanseraten fra 15 til 8 prosent sammenlignet med tabell 4.4, men sett i forhold til grunnmodellen i tabell 4.1 vil denne endringen være marginal.

En økning av kommunale inntekter med ett standardavvik fra gjennomsnittet vil øke effekten på netto innmigrasjon for gruppen videregående skole-utdanning fra 0,06 til 0,09 prosentpoeng. Dette medfører at utdanningsgruppens innmigrasjonsrate forklarer 17 prosent av innmigrasjonsratens standardavvik, noe som er en økning på 6 prosent sammenlignet med tabell 4.4. Effektene av kommunale inntekter for de to resterende gruppene er lik de vi tidligere har funnet.

Resultatene fra gruppesammenligningen i tabell 5.1.1 viser at gruppen med utdanning fra videregående skole responderer sterkest på regional arbeidsledighet, mens gruppen med høyere utdanning responderer sterkest på variablene vakanserate og kommunale inntekter. Sammenlignet med tabell 4.4.1 tyder det på at man oppnår de samme resultatene selv om vi lagrer forklaringsvariablene i regresjonen med ett år for å ta hensyn til at flyttebeslutninger til tid. Dette forsterker bare antagelsen om at et høyere utdanningsnivå fører til økende migrasjon når vi kontrollerer for effekten av den nasjonale konjunktursituasjonen, og samtidig lagrer forklaringsvariablene med en periode.

Når det gjelder de diagnostiske testene i tabell 5.1 får vi at AR(1) i modell FE(VIII) for gruppen videregående skole er positiv og signifikant ved 1 % nivå. Disse resultatene viser også at for gruppen høyere utdanning vil verken AR(1) eller AR(2) forkastes ved den dynamiske modellspesifikasjonen. Seriekorrelasjon virker da å opphøre for denne gruppen sammenlignet med funnene fra de statiske modellene.

5.2 Generalisert momentmetode

I delkapittel 3.3.1 så vi at en ulempe ved å bruke WG-transformasjon i estimering er at denne metoden utnytter kun den delen av datamaterialet som varierer over tid, og dermed mister man mye av variasjonen i tverrsnittet. Siden datamaterialet i analysen strekker seg over en relativ kort tidsperiode (1986-2004) kan dette gi opphav til skjevhet i estimatorene ved at man får over- eller underestimerte estimater (Baltagi, 2001). Som tidligere nevnt i kapittel 4 vil forklaringsvariablene som inngår i regresjonene i større eller mindre grad kunne betraktes som endogene, og dette kan være et potensielt problem. Problemer med multikollinearitet som følge av sterk korrelasjon mellom forklaringsvariablene, målefeil og simultanitet er alle eksempler på endogenitet som kan være med på å gi upresise estimater.

En mulig løsning på problemet med simultanitet og utelatt variabelskjevhet kan være å instrumentere de endogene forklaringsvariablene. Arellano & Bond har foreslått en instrumentvariabelmetode innenfor rammeverket av «Generalized Method of Moments» (GMM) – på norsk *generalisert momentmetode*. Metoden er spesielt tilpasset for dynamiske paneldatamodeller og GMM-estimatoren produserer både første- og andrestegs estimatorer. Mer detaljert beskrivelse av metoden finnes i Hansen (1982).

Under delutvalgsanalyse i delkapittel 4.2.2 ble det kort diskutert at boligpris antas å være den variabelen som i størst grad er beheftet med potensiell simultanskjevhet i analysen. Det er naturlig å tenke seg at flere variabler kan være preget av denne typen endogenitet, og ved å benytte *generalisert momentmetode* kan vi ta hensyn til det som kalles *toveis kausalitet*. Et eksempel på *toveis kausalitet* kan illustreres gjennom å se på forholdet mellom regionale boligpriser og netto innмиграsjon: Man kan anta at høyere boligpris for en region har en *negativ* effekt på netto innмиграsjon, men samtidig vil det være naturlig å forvente at en reduksjon i netto innмиграsjon fremgår som en *dempende* effekt på boligprisen. Ved estimering av boligpris under delutvalgsanalysen fant vi at effekt på netto innмиграsjon ble positiv for gruppen høyere utdanning, mens forventningen er at effekten skal være negativ for alle de tre gruppene. Det blir dermed av interesse å forsøke og instrumentere boligprisen for å se om dette bidrar til at den estimerte effekten på netto innмиграsjon også blir negativ for gruppen høyere utdanning.

Det vil fremgå av estimeringene med *generalisert momentmetode* at det ikke nødvendigvis vil gi mer pålitelige resultater enn andre metoder, selv om man med denne metoden håndterer simultanitetproblemet. Det kan likevel være hensiktsmessig å benytte generalisert momentmetode som en kontroll og et supplement for FE-estimering i kapittel 4 og den dynamiske modellspesifikasjonen i delkapittel 5.1.

5.2.1 Arellano & Bonds instrumentvariabelmetode

Ved å benytte en instrumentvariabel kan man løse problemer med endogenitet for en eller flere forklaringsvariable. Som ved vanlig minste kvadraters metode er det generelle kravet til instrumentvariablene i simultane ligningssystemer at de skal være korrelert med den endogene variabelen man ønsker å instrumentere, men samtidig ukorrelert med restleddet (Wooldridge, 2006:508). Grunntanken i Arellano & Bonds instrumentvariabelmetode er å benytte en rekke momentrestriksjoner for å konstruere en estimator, der kovariansen mellom instrumentet og det transformerte restleddet er lik null [$E(\Delta\varepsilon_{it}|X_{it-s})=0$], der $s \geq 2$, (Verbeek, 2009:kap.10). Ifølge Arellano & Bond kan man oppnå konsistente estimatører ved å differensiere modellen, for deretter å benytte laggede verdier av de endogene forklaringsvariablene som instrumenter. Ved å ta utgangspunkt i grunnmodellen kan vi forklare metoden beskrevet over.

$$(5.2) \text{Innmigrasjon}[X]_{it} = \beta_1 \log(U)_{it} + \beta_2 \log(\text{VakS})_{it} + \eta_i + \varepsilon_{it}$$

der U_{it} og VakS_{it} representerer forklaringsvariablene ledighet og vakanserate, η_i er det regionspesifikke restleddet og ε_{it} er det idiosynkratiske restleddet. X er GRU, VGS og HØY. Førstedifferensiering av modellen i (5.2) vil eliminere regionspesifikke effekter:

$$(5.3) \Delta \text{Innmigrasjon}[X]_{it} = \beta_1 \Delta \log(U)_{it} + \beta_2 \Delta \log(\text{VakS})_{it} + \Delta \varepsilon_{it}$$

hvor

$$\Delta \log(U)_{it} = \log(U)_{it} - \log(U)_{it-1},$$

$$\Delta \log(\text{VakS})_{it} = \log(\text{VakS})_{it} - \log(\text{VakS})_{it-1}$$

$$\Delta \varepsilon_{it} = \varepsilon_{it} - \varepsilon_{it-1}.$$

For at vi skal kunne oppnå gyldige estimater for de to endogene variablene i ligning (5.3), må $(U)_{it-2}$ være korrelert med $\log \Delta \log(U)_{it}$, men samtidig ukorrelert med $\Delta \varepsilon_{it}$ så lenge ε_{it} ikke er

seriekorrelert (Verbeek, 2009:kap.10). Med andre ord betyr det at kravene for at de endogene variablene skal være gyldige instrumenter er oppfylt når variabelen på *nivåform* er korrelert med variabelen på *differensiert* form, og samtidig ukorrelert med det differensierte restleddet. I dette tilfellet kan vi under den nevnte forutsetningen benytte U_{it-s} og $VakS_{it-s}$ istedenfor X_{it-s} fra den generelle formelen¹⁹, der $s \geq 2$, som instrumenter for U_{it} og $VakS_{it}$. Det vil si at variablene U_{it-2} og $VakS_{it-2}$ lagget to eller flere perioder tilbake i tid, kan i analysen benyttes som gyldige instrumenter.

I denne sammenhengen blir det viktig å se på hva diagnosetestene sier om det transformerte restleddet. Arellano & Bond mener at en forutsetning for å oppnå konsistente estimatører er ved fravær av andreordens seriekorrelasjon i det transformerte restleddet. Det betyr at AR(2) ikke kan være signifikant for at instrumentene skal være gyldige. Samtidig som det transformerte restleddet viser til en negativ førsteordens seriekorrelasjon og at AR(1) er signifikant. En kombinasjon av negativ førsteordens seriekorrelasjon og fravær av andreordens seriekorrelasjon i de transformerte restleddene indikerer dermed at restleddene på nivåform er hvit støy (Wooldridge, 2006, kap. 10).

I tillegg til de diagnostiske testene benytter vi Sargans test som er en generell test av gyldigheten til instrumentene. Her blir det testet en simultan hypotese om at ingen av de inkluderte instrumentene er korrelert med restleddet. I delkapittel 5.3 vil vi bare rapportere resultatene fra andrestegs-testen, fordi det i praksis har vist seg at Sargans førstestegs-test nesten alltid forkaster den nevnte hypotesen om ingen korrelasjon med restleddet.

5.3 Resultater fra GMM-analysen

Som i studiene til Kaspersen (2006) og Dahl (2007) vil GMM-analysen basere seg på førstestegsestimater. I tabell 5.2 rapporteres førstestegsestimater fra GMM-analysen av en enkel grunnmodell og den fullstendige modellen fra kapittel 4, mens tabell 5.3 viser gruppesammenligningen fra GMM(I) og GMM(II)-modellene. Ved å benytte førstedifferensiering av variablene vil dette medføre at tidsseriene blir et år kortere og datamaterialet består nå av paneldata for perioden 1987-2004.

¹⁹ Se Kaspersen (2006:44)

Tabell 5.2. GMM-estimer basert på instrumentvariabelmetoden til Arellano & Bond (1991). Avhengig variabel: $Innmigrasjon[X]_{it}$. Transformasjon: Førstedifferensiering.

Variabler	Gruppe					
	GRU		VGS		HØY	
	GMM(I)	GMM(II)	GMM(I)	GMM(II)	GMM(I)	GMM(II)
Log (U_{it})	-0,365 (-2,81)	-0,369 (-2,83)	-0,824 (-4,91)	-0,855 (-4,79)	-0,223 (-0,86)	-0,050 (-0,177)
Log (U_{it})* ΔN_t		-0,0039 (-0,19)		-0,033 (-1,53)		-0,25 (-2,63)
Log (VakS $_{it}$)	0,047 (0,26)	0,053 (0,278)	0,577 (3,24)	0,399 (1,97)	2,372 (4,89)	1,54 (3,95)
Log (VakS $_{it}$)* ΔN_t		-0,029 (-0,61)		-0,078 (-1,18)		-0,311 (-2,12)
Log (KommInn $_{it}$)		0,365 (0,762)		0,361 (0,56)		3,47 (2,40)
Observasjoner	1620	1620	1620	1620	1620	1620
σ	0,644	0,644	0,581	0,579	1,603	1,587
AR(1)-test	-6,403**	-6,416**	-5,914**	-5,89**	-5,417**	-5,459**
AR(2)-test	0,241	0,229	0,074	-0,023	0,250	0,178
Sargan (p-verdi)	66,80	69,38	70,87	73,90	74,55	74,20

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå. Sargan er en χ -fordelt test for instrumentenes gyldighet, og dens "two-step" p-verdi rapporteres her.

Hvis vi starter med å se på de diagnostiske testene så viser resultatene for AR(1) at det transformerte restleddet har signifikant førsteordens seriekorrelasjon, og dette impliserer at restleddet på nivåform ikke er seriekorrelert. Samtidig viser AR(2) at vi har fravær av andreordens seriekorrelasjon. I FE-analysen i kapittel 4 var det ved flere anledninger andreordens seriekorrelasjon og dette kan være en forklaring bak problemene ved upresise estimer. I tillegg viser resultatene at Sargans andrestegs-test for gyldigheten av overidentifiserende restriksjoner ikke forkastes for noen av de tre gruppene og dette kan tyde på at instrumentene som vi har brukt er gyldige.

Ved å benytte instrumentvariabelmetoden blir den største forskjellen fra FE-modellene at vi får en betydelig endring i betydningen av arbeidsmarkedsvareblene på netto innmigrasjon for gruppen høyere utdanning. Her vil estimatene til den regionale ledigheten være svært lave og insignifikant, mens vakanseraten får høye og signifikante estimer. Ser vi på den simultane effekten av de to arbeidsmarkedsvareblene ved å øke ledigheten og redusere vakanseraten med ett standardavvik samtidig, vil reduksjonen på netto innmigrasjonsrate bli betydelig høyere enn i kapittel 4. GMM(I) viser størst endring, og reduksjonen i netto innmigrasjon for

gruppen videregående skole blir -0,48 prosentpoeng, mens for gruppen høyere utdanning får vi -1,05 prosentpoeng. Til sammenligning er reduksjonen størst i FE(I)-modellen med -0,32 for gruppen videregående skole og -0,39 for gruppen høyere utdanning. Ved å øke den fylkesvise vakanseraten med ett standardavvik vil GMM(I) og GMM(II) vise til en positiv effekt på henholdsvis 0,65 prosentpoeng og 0,45 prosentpoeng. Disse estimatene ligger mye høyere enn de gjorde i FE-modellene i kapittel 4, og dermed understreker dette bare at instrumentvariabelmetoden synes å tillegge betydningen av arbeidsmarkedet i langt større grad. For gruppen høyere utdanning peker vakanseraten seg ut som den avgjørende variabelen og ikke regional ledighet som tidligere hadde klart størst effekt på netto innmigrasjonsrate.

En ytterligere sammenligning av GMM-regresjonene og FE-modellene viser at gruppene grunnskole og videregående skole fortsatt responderer mer på regional ledighet enn på fylkesvis vakanserate. Koeffisientverdiene til variabelen regional ledighet er her høyere (mer negativ) enn i FE-modellene. Ved å se på estimatene til konjunktoreffektene i GMM-regresjonene vil koeffisientverdiene endres lite, men variablene vil kun være signifikant for gruppen høyere utdanning i forhold til FE-modellene der de også hadde en signifikant effekt for gruppen videregående skole. I likhet med vakanseraten vil kommunale inntekter også få en større betydning for gruppen høyere utdanning enn tidligere på netto innmigrasjonsrate, mens den samme effekten blir lavere og insignifikant for de to andre gruppene.

Tabell 5.3. Gruppesammenligning av estimerte koeffisientforskjeller fra GMM(I) og (II)-modellene. Avhengig variabel: $\text{Innmigrasjon}[X_1]_{it} - \text{Innmigrasjon}[X_2]_{it}$, der X_1 og X_2 står for utdanningsgruppene. X vil være GRU, VGS og HØY. Metode: GMM.

Variabler	Gruppesammenligning					
	GRU vs VGS		GRU vs HØY		VGS vs HØY	
	D-GMM(I)	D-GMM(II)	D-GMM(I)	D-GMM(II)	D-GMM(I)	D-GMM(II)
Log (U_{it})	0,459 (3,62)	0,487 (3,51)	-0,142 (-0,53)	-0,319 (-1,08)	-0,601 (-2,28)	-0,806 (-2,73)
Log (U_{it})* ΔN_t		0,029 (1,22)		0,246 (2,57)		0,217 (2,39)
Log (Vak S_{it})	-0,529 (-2,63)	-0,347 (-1,72)	-2,324 (-4,74)	-1,49 (-3,73)	-1,794 (-3,94)	-1,14 (-2,88)
Log (Vak S_{it})* ΔN_t		0,048 (0,867)		0,28 (2,13)		0,233 (1,89)
Log (KommInn $_{it}$)		0,004 (0,006)		-3,10 (-2,08)		-3,11 (-2,34)

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik.

Ved å benytte samme fremgangsmåte som for testene i kapittel 4 kan vi teste koeffisientforskjellene mellom utdanningsgruppene som rapporteres i tabell 5.3. Mer detaljert beskrivelse finnes i delkapittel 3.3.3, men det viktigste er at en t-verdi større eller lik 1,96 for en bestemt forklaringsvariabel viser at gruppene responderer signifikant forskjellig ved 5 prosents nivå. En sammenligning med FE-modellene viser at vakanseraten er signifikant forskjellig på tvers av utdanningsgruppene, mens forskjellene mellom utdanningsgruppene for de resterende variablene viser seg å være tilsvarende lik de vi fant i kapittel 4. En indikasjon fra GMM-regresjonen er dermed at responsen på netto innmigrasjon av vakanseraten er stigende i utdanningsnivået.

Vi har dermed sett at GMM-estimatene generelt gir høyere koeffisientestimer enn resultatene i FE-modellene. Dette kan da være med på å underbygge antagelsene våre om at de foreligger skjevhet i FE-estimatene på grunn av simultanitet og utelatt variabelskjevhet, og dette har da bidratt til at koeffisientverdiene er blitt underestimert.

Tabell 5.4. GMM-estimer fra regresjonsmodellen som inkluderer boligpris. Avhengig variabel: $Innmigrasjon[X]_{it}$. Transformasjon: Førstedifferensiering.

Tidsperiode →	Gruppe		
	GRU	VGS	HØY
	92-04	92-04	92-04
Variabler	GMM(III)	GMM(III)	GMM(III)
Log (U_{it})	-0,759 (-2,79)	-1,89 (-7,68)	-2,35 (-4,19)
Log (U_{it})* ΔN_t	-0,020 (-0,578)	-0,088 (-1,89)	-0,15 (-1,18)
Log ($VakS_{it}$)	0,0719 (0,282)	-0,138 (-0,618)	0,349 (0,87)
Log ($VakS_{it}$)* ΔN_t	0,068 (1,10)	-0,167 (-1,91)	-0,350 (-2,16)
Log ($KommInn_{it}$)	1,14 (1,85)	2,39 (2,20)	5,37 (2,86)
Log ($Boligpris_{it}$)	-0,173 (-0,522)	-0,647 (-1,19)	-1,26 (-1,41)
Observasjoner	1170	1170	1170
σ	0,660	0,579	1,388
AR(1)-test	-6,349**	-5,335**	-5,811**
AR(2)-test	0,2919	0,6594	-1,026
Sargan (p-verdi)	68,91	70,64	78,97

Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå. Sargan er en χ -fordelt test for instrumentenes gyldighet, og dens "two-step" p-verdi rapporteres her.

Resultater fra tabell 5.4 viser at estimering ved bruk av instrumentvariabelmetoden gir negative koeffisientverdier for boligprisen. I motsetning til delutvalgsanalysen vil variabelen boligpris få en forventet negativ effekt på netto innmigrasjonsrate for alle de tre utdanningsgruppene. Koeffisientverdien til boligprisen for gruppen videregående skole er mye sterkere i GMM-regresjonen enn i delutvalgsanalysen, men likevel insignifikant. For gruppen høyere utdanning er den estimerte effekten av boligpris negativ på netto innmigrasjon og signifikant ved 10 prosents nivå i en ensidig nullhypotese. Hvis vi estimerer regresjonen i tabell 5.4 uten å ta hensyn til den nasjonale konjunktursituasjonen og kommunale inntekter, vil effekten av boligprisen bli negativ og signifikant for alle de tre gruppene²⁰. Dermed ser vi at inkludering av disse to variablene vil dempe den negative effekten av boligprisen på netto innmigrasjonsrate. For boligprisen vil koeffisientforskjellene mellom gruppene være insignifikant uavhengig av valgt signifikansnivå, men resultatene tyder på at grunnskolegruppen skiller seg fra de to andre gruppene.²¹

²⁰ Koeffisient- og t-verdier (i parentes) for boligprisen: GRU: -0,51 (1,53), VGS: -1,36 (2,54), HØY: -2,45 (2,34)

²¹ T-verdier ved test av koeffisientforskjeller: GRU vs VGS: 0,82, GRU vs HØY: 1,21, VGS vs HØY: 0,55

5.4 Oppsummering

I kapittel 5 har analysen blitt utvidet med en dynamisk fixed effects-modell og en GMM-analyse. Variabler lagget ett år tilbake ble inkludert i regresjonen for å ta hensyn til at flyttebeslutninger tar tid. Dette viste seg at laggede variabler hadde liten påvirkningskraft, og dermed ble langtidseffektene relativt like resultatene vi fikk i de statiske analysene. Vi kan derav konkludere med at det ikke er så mye å hente fra en dynamisk modellspesifikasjon.

I GMM-analysen ble endogene instrumenter for arbeidsmarkedsvariablene inkludert i regresjonen ved at variabelen på nivåform ble lagget to og tre ganger. Hovedgrunnen til at endogene instrumenter inngår i regresjonen under GMM, er fordi vi ønsker å ta hensyn til potensiell skjevhet i FE-estimatorene fra kapittel 4 som følge av simultanitet. Resultatene fra GMM-analysen forsterker det vi fant fra FE-modellene: I likhet med tidligere responderer gruppen videregående skole mest på regional ledighet. Når det gjelder vakanseraten er den insignificant for grunnskolegruppen, mens for gruppen høyere utdanning er estimer for denne variabelen mangedoblet sammenlignet med FE-modellene. I tillegg viser det seg at effekten av vakanseraten er stigende i utdanningsnivået. Den største endringen i GMM-estimatene er imidlertid at den estimerte effekten av regional ledighet er svært lav og insignificant for gruppen høyere utdanning.

Estimatene av kommunale inntekter når vi ikke inkluderer boligpris er til forskjell fra FE-modellene kun signifikant for gruppen høyere utdanning, samtidig som den estimerte effekten er sterkere. Inkludering av boligprisen som i tabell 5.4 gir at variabelen blir positiv og signifikant for alle utdanningsgruppene. Effekten på netto innmigrasjon viser også til å være stigende i utdanningsnivået.

Estimering ved bruk av instrumentvariabelmetoden gir at resultatene for boligprisen blir mer robuste enn i FE-modellene. Den estimerte effekten på netto innmigrasjon blir negativ for alle de tre gruppene, men signifikant kun for gruppen høyere utdanning. I en test av koeffisientforskjellene skiller grunnskolegruppen seg mest fra de to andre gruppene, men testene viser seg å være insignificant. Estimering av den nasjonale konjunktursituasjonen i GMM-regresjonen skiller seg ikke nevneverdig fra FE-modellene. Estimatene er negativ på netto innmigrasjon og gruppen høyere utdanning responderer sterkest på variabelen. Dette blir underbygget ved å foreta en test av koeffisientforskjellene.

6 Avsluttende diskusjon og konklusjon

I denne studien har vi undersøkt betydningen av ulike utdanningsnivå for effekt på interregional migrasjon, og da med spesiell oppmerksomhet rettet mot variabler som knyttes opp imot arbeidsmarkedsmotivert flytting. Analysen er en paneldatastudie og baserer seg på tall fra 90 økonomiske regioner i Norge for perioden 1986-2004. Utgangspunktet for denne studien ligger i funnene fra Carlsen, Johansen og Kaspersen (2007), Dahl (2007) og Carlsen, Johansen og Stambøl (2010), som alle konkluderer med at ubalanser i det regionale arbeidsmarkedet har signifikante effekter på netto innflyttingsrate i norske regioner. I tillegg viser kontroll for den nasjonale konjunktursituasjonen til å forsterke effekten av arbeidsmarkedsvariablene. Disse resultatene samsvarer ikke med den internasjonale empirien der det påpekes at mobiliteten i Europa er lav og at insentivene til å flytte mellom regioner påvirkes i liten grad av arbeidsmarkedsvariablene. Selv om forskningsmetodene i internasjonale studier inneholder mye variasjon tyder funnene på at mobiliteten er stigende i utdanningsnivået, og en kontroll for konjunktursituasjonen viser til å ha signifikant effekt på migrasjonen.

I kapittel 3 under estimeringsmetode argumenterte vi for at fixed effects-estimering ville bli brukt i denne empiriske analysen. Dette på grunn av at tidligere studier påpeker at denne metoden gir de mest troverdige resultatene samtidig som metoden tillater for korrelasjon mellom forklaringsvariable og regionspesifikke effekter. Videre benyttet vi en sensitivetsanalyse for å teste resultatenes stabilitet, der tidsserien ble delt i to kortere perioder. For å løse potensielle endogenitetsproblemer blir resultatene ytterligere kontrollert ved å ta i bruk en dynamisk modellspesifikasjon. Til slutt foretok vi en estimering ved hjelp av Arellano & Bond sin foreslåtte instrumentvariabelmetode innenfor rammeverket generalisert momentmetode (GMM).

Estimering i den statiske modellen viste at estimatene for den regionale arbeidsledigheten var negativ og signifikant - uavhengig av utdanningsgruppe. For effekten på netto innmigrasjonsrate fremgår det at gruppen grunnskole skiller seg fra de to andre gruppene, mens gruppen videregående skole viser til den sterkeste effekten. Resultatene fra den dynamiske modellespesifikasjonen lagget ett år tilbake ga oss tilsvarende resultater som i de statiske modellene, og konklusjonen blir da at flyttebeslutninger tas relativt raskt. Sammenlignet med FE-modellene blir den største endringen i GMM-modellene at effekten av

regional ledighet viser seg å være svært lav og insignifikant for gruppen høyere utdanning. Vi kan likevel si at estimatene for utdanningsgruppens flytterespons av den totale effekten på arbeidsmarkedsvariablene viser til at mobiliteten er stigende i utdanningsnivået – uavhengig av modellespesifikasjon.

For den fylkesvise vakanseraten viser estimatene til å ha en positiv effekt på netto innmigrasjon og i de fleste tilfeller være signifikant forskjellig fra null i ensidige tester. Resultatene tyder også på at betydningen av vakanseraten øker i utdanningsnivået. Som for regional ledighet ga en dynamisk spesifikasjon liten endring for vakanseraten, men ved GMM-modellene førte estimering til at vakanseraten blir insignifikant for gruppen grunnskole. Den samme variabelen er signifikant og koeffisientverdien mangedobles for gruppen høyere utdanning.

Resultatene i vår empiriske analyse viser dermed at effektene av regional ledighet og fylkesvis vakanserate stemmer godt overens med forventningene om at effektene er stigende i utdanningsnivået, basert på funnene i nasjonal og internasjonal empiri. Tolkningen av vakanseraten i GMM-modellene kan tyde på at antallet ledige stillinger er det mest sentrale bak flyttetilbøyeligheten til de med høyere utdanning, og ikke regional ledighet som i de statiske modellene. Ifølge McCormick og Hughes (1994) har vi grunn til å tro at høyt utdannede får flere jobbtillbud, selv i perioder med økonomisk lavkonjunktur, og dette kan påvirke utdanningsgruppens forventninger til arbeidsmarkedet og gi utslag i flyttetilbøyeligheten.

De øvrige variablene som ble brukt ved estimering for perioden 1986-2004 var kommunale inntekter og interaksjonsledd som kontrollerte for den nasjonale konjunktursituasjonen. Som for regional ledighet og fylkesvis vakanserate viser effekten av konjunkturer til å være stigende i utdanningsnivået. Dette gir at mobiliteten til de med høyere utdanning vil bli mer påvirket av den nasjonale konjunktursituasjonen enn de med lavere utdanning. GMM-modellene gir tilsvarende resultater. Disse resultatene bekrefter det Carlsen, Johansen og Kaspersen (2007) fant i sin artikkel om at regionale forskjeller i arbeidsledighet betyr mest for mobiliteten i en høykonjunktur, mens regionale forskjeller i vakanseraten betyr mest i en lavkonjunktur. Dette på grunn av at det skapes færre jobber i nedgangstider.

Vi fant at den estimerte effekten på netto innflytting for variabelen kommunale inntekter er langt sterkere for gruppen høyere utdanning enn for de andre gruppene. Resultatene viser seg å være entydige for denne variabelen ved at effekten er signifikant for alle gruppene i FE-modellene. I GMM-modellene blir responsen ytterligere forsterket for gruppen høyere utdanning og er kun signifikant for denne gruppen. Inkludering av boligpris vil imidlertid gjøre koeffisientverdiene av kommunale inntekter signifikant for alle utdanningsgruppene, men effekten er fortsatt størst for de med høyere utdanning. Det tyder dermed på at høyt utdannede legger mer vekt på tjenestetilbudene, noe som kan ses i sammenheng med at denne gruppen antas å være bedre kvalifisert til et bredere spekter av jobber i et større geografisk område. Dette medfører et økt utvalg av bosted- og jobbmuligheter, og det kan være naturlig å velge bosted i en kommune med gode velferds- og tjenestetilbud. Dette kan igjen relateres til at utdanningsgruppene vektlegger variabler ulikt og det som er av betydning for en gruppe, kan ha ingen eller liten betydning for en annen gruppe. Kommunale inntekter virker dermed til å fange opp effekter som betyr mest for de med høy utdanning.

I sensitivitetsanalysen så vi på effekten av boligpris, som har en kortere tidsserie. De øvrige forklaringsvariablene viste til effekter som var relativt varierende i styrke og signifikans, mens boligprisen har en lav, men signifikant koeffisientverdi for gruppen grunnskole. Siden boligprisen er den variabelen vi med størst sannsynlig tror er beheftet med endogenitet i form av toveis kausalitet, og samtidig som resultatene fra FE-modellene er svake, gir en GMM-modellering de mest troverdige resultatene. Ved å benytte GMM på boligprisen ble estimatene som forventet mer robuste og i tillegg negativ for alle utdanningsgruppene, men kun signifikant for de med høyere utdanning. Effekten av endringer i boligprisen ser dermed ut til å påvirke kun de med høyere utdanning. Siden boligprisen og kommunale inntekter påvirker de med høyere utdanning i størst grad, vil denne gruppen dra fordel av å være mest attraktiv i arbeidsmarkedet fordi det som kjent gjør det enklere å tilpasse valg av bolig og bosted.

Utvidelser av modellene fra kapittel 4 til en dynamisk modellspesifikasjon førte ikke til nevneverdige endringer i resultatene. Siden de laggede variablene hadde liten påvirkningskraft på netto innmigrasjonsrate utover resultatene fra de statiske modellene, understreker dette at ubalanser i arbeidsmarkedet påvirker migrasjonsbeslutningene relativt raskt. Inkludering av instrumenter i GMM-analysen ga oss koeffisientverdier som generelt var høyere enn i FE-modellene, og dermed kan vi på mange måter si at instrumentene til de

endogene forklaringsvariablene korrigerer for simultanitetsproblemene. Dette gir grunnlag for å si at koeffisientverdiene i FE-modellene var underestimerte.

For videre analyser av interregional migrasjon kunne det vært interessant å inkludere flere indikatorer for utdanningsspesifikke arbeidsmarkedsvariabler. Dette kan være variabler som tar hensyn til lønnsnivået for hver av de tre utdanningsgruppene eller variabler som tar for seg ledighetsraten for hver enkelt gruppe, og ikke en samlet regional ledighetsrate. Vi har grunn til å tro at personer som har kvalifikasjoner og egenskaper som hører til i gruppen høyere utdanning, vil sammenligne seg med arbeidstakere fra samme utdanningsgruppe. Utdanningsspesifikke arbeidsmarkedsvariabler kunne dermed bidratt til at man hadde fått et mer presist bilde av flytteresponsen blant de ulike utdanningsgruppene, mye på grunnlag av at flytteresponsen for de med høyere utdanning ikke forventes å bli påvirket i stor grad av ledigheten blant de med lavere utdanning. I tillegg kunne det vært interessant å se ytterligere på variabler som karakteriserer regionene, som for eksempel kriminalitet og klimaindikatorer.

De data for mål på stramheten i arbeidsmarkedet som Carlsen og Johansen (2004) benytter kan være relevant for vår studie. Survey-undersøkelsene kan avdekke nyttig informasjon om befolkningens oppfatning av det regionale arbeidsmarkedet. En slik indikasjon kan hjelpe oss til å få mer presise estimater hvis vi er i stand til å kontrollere for flere karakteristikk ved regionene.

I likhet med norske studier av nyere dato fant vår studie at resultatene skiller seg fra internasjonal empiri og vi konkluderer med at Norge generelt har høyere arbeidskraftmobilitet enn andre europeiske land. I teorien kan vi dermed forvente at vi observerer kun beskjedne forskjeller i regional ledighet, men som Østby (2002) sier er ikke høy mobilitet nødvendigvis et samfunnsgode.

Referanseliste

Andersen, Arne (2001): "Høykonjunktur på boligmarkedet: Unge er ikke blitt presset ut.", Økonomiske analyser 5/2001, (s 23 – 27), Statistisk Sentralbyrå

Baltagi, B.H. (2001): «Econometric Analysis of Panel Data». 2. utgave, John Wiley & Sons, Ltd.

Brunborg, H., Sørli, K. og Texmon, I. (2005): «Innenlandske flyttinger». *Økonomiske analyser*, 6/2005: 39-42, Statistisk sentralbyrå.

Carlsen, F. og Johansen, K. (2004): «Subjective Measures of Employment Opportunities and Interregional Migration». *Labour* 18 (4): 563-589

Carlsen, F., Johansen, K. og Røed, K. (2006): «Wage Formation, Regional Migration and Local Labour Market Tightness». *Oxford Bulletin of Economics and Statistics* 68, 4: 423-444.

Carlsen, F., Johansen, K. og Kaspersen, S. (2007): «Flytting mellom norske regioner: Betydningen av regionale arbeidsmarkeder og konjunktursituasjonen». *Publisert i Norsk Tidsskrift nr. 1-2007*.

Dahl, G.A. (2007): «Arbeidsmarkedsmotivert flytting og betydningen av utdanningsnivå». Norges teknisk-naturvitenskapelige universitet, Trondheim.

Ehreberg, R. G. og Smith, R. S. (2000): «Modern Labor Economics». Addison Wesley Longman, Inc.

Hansen, L.P. (1982): «Large Sample Properties of Generalized Method of Moments Estimators». *Econometrica*, 50: 1029-1054

Hughes, G. og McCormick, B. (1994): «Did Migration in the 1980s Narrow the North-South Divide?». *Economica* 61: 509-527

Jackman, R. og Savouri, S. (1992): «Regional Migration in Britain: An Analysis of Gross Flows Using NHS Central Register Data». *The Economic Journal*, 102 (1992): 1433-1450.

Kaspersen, S.L. (2006): «Løs og ledig? Arbeidsmarkedsmotivert flytting i Norge». Norges teknisk-naturvitenskapelige universitet, Trondheim.

Lånekassen (2011): <http://www.lanekassen.no/Hovedmeny/Tilbakebetaling/Finnmark-og-Nord-Troms/>

McCormick, Barry. (1997): «Regional Unemployment and Labour Mobility in the UK». *European Economic Review*, 41 (1997): 581-589.

Nahuis, R. og A. Parikh (2004): «Factor mobility and regional disparities. East, west, home's best? ». Working paper 26, European Network of Economic Policy Research Institutes.

NAV (2006): «Historisk arbeidsmarkedsstatistikk»

Pissarides, Christopher A. og Wadsworth, Jonathan (1989): «Unemployment and the Inter-regional Mobility of Labour». *The Economic Journal*, 99 (1989): 739-755.

Pissarides, Christopher A. og McMaster, Ian. (1990): «Regional Migration, Wages and Unemployment: Empirical Evidence and Implication For Policy». *Oxford Economic Papers*, 42. (1990): 812-831.

Sjaastad, L. (1962): «The Costs and Returns of Human Migration». *The Journal of Political Economy*, Vol. 70, No. 5, Part 2: Investments in Human Beings: 80-93.

Statistisk sentralbyrå (2000): Norges offisielle statistikk C 616 «Standard for økonomiske regioner».

Sørli, K. (2004): «Flyttehistoriske hovedtrekk for årskullene født på 1950- og 1960-tallet». *Regionale trender* 2004:2, NIBR.

Strøm, B. (2010): «Forelesningsnotater SV SØK 3515: Mikro- og paneldataøkonometri», Institutt for samfunnsøkonomi, Norges teknisk-naturvitenskapelige universitet.

Verbeek, M. (2009): «*A Guide to Modern Econometrics*». 3rd edition. John Wiley & Sons, Ltd.

Wooldridge, J.M. (2006): «*Introductory Econometrics – A Modern Approach*». Thomson South-Western, Ohio.

Østby, L. (2002): «Nordmenn på flyttetoppen». *Samfunnsspeilet nr. 2*: 55-57, Statistisk sentralbyrå.

Internettressurser:

<http://www.aetat.no>

<http://www.ssb.no>

<http://www.lanekassen.no>

<http://www.regjeringen.no>

Appendiks – oversikt over tabeller og figurer

Tabell A1. Definisjoner av variabler som inngår i studiets økonometriske modeller.

Variabel	Forkortelse	Definisjon
Innmigrasjon[GRU]	Innmigrasjon[X] _{it} X = GRU	Netto innflyttingsrate for personer med kun grunnskole i aldersgruppen 16-66 år.
Innmigrasjon[VGS]	Innmigrasjon[X] _{it} X = VGS	Netto innflyttingsrate for personer med videregående skole i aldersgruppen 16-66 år.
Innmigrasjon[HØY]	Innmigrasjon[X] _{it} X = HØY	Netto innflyttingsrate for personer med høyere utdanning i aldersgruppen 16-66 år.
Regional arbeidsledighetsrate	U _{it}	Regional ledighet (prosent av arbeidsstyrken)
Vakanserate	VakS _{it}	Vakansestrøm. Tilgang vakanser (prosent av arbeidsstyrken, fylkesnivå).
Kommunale inntekter	KommInn _{it}	Kommunale inntekter (skatt og rammetilskudd) per innbygger i 10 000 kr, skalert med prisindeks for kommunalt konsum. Basisår = 1986.
Msyssvekst	ΔN _t	Aggregert sysselsettingsvekst i % og målt som avvik fra eget gjennomsnitt.
Boligpris	Boligpris _{it}	Pris på standardisert bolig (i millioner).

Merk: Senket skrift *i* angir region, *t* angir år.

Tabell A2. Korrelasjonsmatrise. Tidsperiode: 1986-2004

	Innmigrasjon [GRU]	Innmigrasjon [VGS]	Innmigrasjon [HØY]	U	U*ΔN_t	VakS	VakS*ΔN_t	KommInn
Innmigrasjon[GRU]	1,0000							
Innmigrasjon[VGS]	0,42416	1,0000						
Innmigrasjon[HØY]	0,12106	0,53666	1,0000					
U	-0,15682	-0,19602	-0,12394	1,0000				
U*ΔN_t	0,020928	-0,061159	-0,22956	-0,22684	1,0000			
VakS	-0,013983	-0,036564	-0,028608	-0,21785	0,23352	1,0000		
VakS*ΔN_t	0,0058755	-0,074794	-0,18868	-0,27390	0,85699	0,25877	1,0000	
KommInn	-0,17131	-0,23970	-0,16751	0,095082	-0,10272	0,19980	-0,18829	1,0000

Tabell A3. Korrelasjonsmatrise. Tidsperiode: 1991-2004

	Innmigrasjon [GRU]	Innmigrasjon [VGS]	Innmigrasjon [HØY]	U	U*ΔN_t	VakS	VakS*ΔN_t	KommInn	Boligpris
Innmigrasjon[GRU]	1,0000								
Innmigrasjon[VGS]	0,54296	1,0000							
Innmigrasjon[HØY]	0,43044	0,61575	1,0000						
U	-0,66352	-0,69676	-0,62693	1,0000					
U*ΔN_t	-0,054842	0,040793	0,36469	-0,15446	1,0000				
VakS	0,73449	0,57364	0,64964	-0,80530	0,13605	1,0000			
VakS*ΔN_t	-0,025664	-0,16220	0,35551	-0,096548	0,9253	0,089529	1,0000		
KommInn	0,41318	0,66437	0,25221	-0,71764	0,05979	0,40131	-0,092464	1,0000	
Boligpris	0,33499	0,59053	0,10617	-0,69284	-0,17717	0,28194	-0,29895	0,92430	1,0000

Tabell A4. Dynamisk modellspesifikasjon med Within Groups-transformasjon. Avhengig variabel: $Innmigrasjon_{it}$

Variabler	Gruppe		
	GRU	VGS	HØY
	FE(V)	FE(V)	FE(V)
Log (U_{it})	-0,202 (-1,49)	-0,752 (-4,62)	-1,172 (-3,59)
Log (U_{it-1})	0,008 (0,07)	0,099 (0,91)	0,743 (2,63)
Log (U_{it})* ΔN_t	0,006 (0,201)	-0,062 (-2,08)	-0,293 (-2,83)
Log (U_{it-1})* ΔN_{t-1}	-0,0036 (-0,12)	0,028 (0,879)	0,080 (1,23)
Log (VakS $_{it}$)	0,249 (2,50)	0,164 (1,46)	-0,22 (-0,722)
Log (VakS $_{it-1}$)	-0,178 (-1,51)	0,094 (0,818)	0,616 (2,12)
Log (VakS $_{it}$)* ΔN_t	0,041 (1,16)	-0,048 (-0,821)	-0,44 (-2,58)
Log (VakS $_{it-1}$)* ΔN_{t-1}	-0,155 (-3,71)	-0,16 (-3,45)	0,014 (0,155)
Log (KommInn $_{it}$)	0,314 (0,932)	0,043 (0,124)	1,685 (2,15)
Log (KommInn $_{it-1}$)	0,066 (0,161)	0,95 (2,24)	0,58 (0,609)
Observasjoner	1710	1710	1710
R ²	0,093	0,196	0,185
σ	0,463	0,468	1,18
Wald (joint)	41,94**	95,15**	59,64**
AR(1)-test	-1,167	3,370**	1,819
AR(2)-test	-1,060	2,028*	1,721

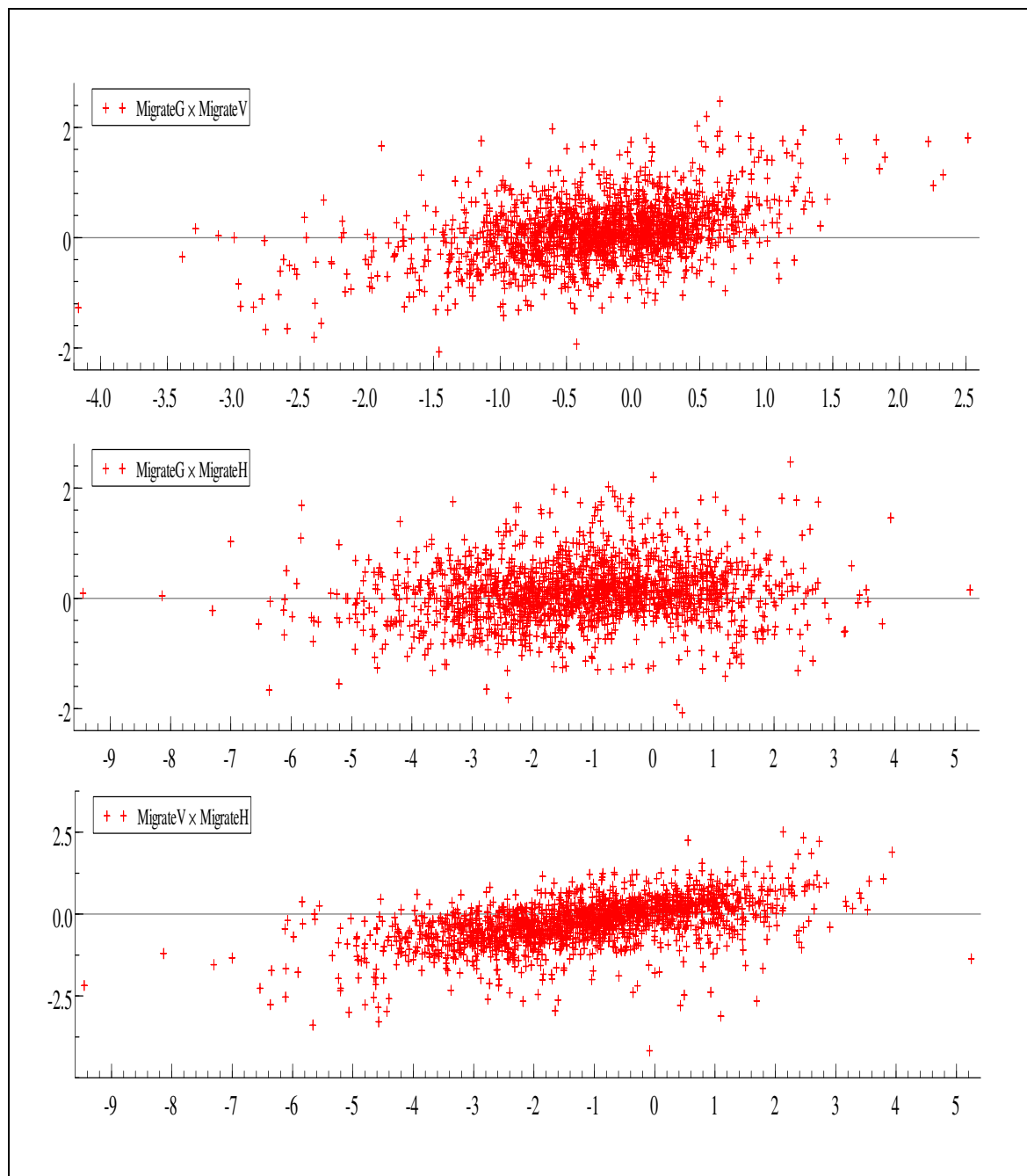
Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik. σ viser estimert standardavvik for restleddet. Wald(joint) er en kjikvadratfordelt test der H_0 : «Ingen av forklaringsvariablene har forklaringskraft». ** betyr forkastning av H_0 ved 1 % signifikansnivå. AR(1) og AR(2) er tester for henholdsvis første- og andreordens seriekorrelasjon i restledd, med nullhypotese H_0 : «Ingen seriekorrelasjon». Forkastning av H_0 indikeres med ** ved 1 % signifikansnivå, og * ved 5 % nivå.

Tabell A5. Gruppesammenligning av estimerte koeffisientforskjeller fra FE(V)-modellene. Avhengige variabler: Innмиграsjon $[X_1]_{it}$ – Innмиграsjon $[X_2]_{it}$, der X_1 og X_2 står for utdanningsgruppene. X vil være GRU, VGS og HØY.

Variabler	Gruppesammenligning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(V)	D-FE(V)	D-FE(V)
Log (U_{it})	0,552 (3,35)	0,970 (3,18)	0,418 (1,40)
Log (U_{it-1})	-0,092 (-0,700)	-0,735 (-2,57)	-0,643 (-2,39)
Log (U_{it})* ΔN_t	0,068 (2,23)	0,299 (2,78)	0,231 (2,11)
Log (U_{it-1})* ΔN_{t-1}	-0,031 (-1,22)	-0,083 (-1,31)	-0,052 (-0,859)
Log (VakS $_{it}$)	0,085 (0,552)	0,472 (1,38)	0,387 (1,41)
Log (VakS $_{it-1}$)	-0,273 (-1,98)	-0,794 (-2,61)	-0,522 (-2,09)
Log (VakS $_{it}$)* ΔN_t	0,09 (1,75)	0,482 (2,96)	0,392 (2,58)
Log (VakS $_{it-1}$)* ΔN_{t-1}	0,00065 (0,0192)	-0,168 (-1,97)	-0,169 (-2,03)
Log (KommInn $_{it}$)	0,272 (0,56)	-1,37 (-1,61)	-1,64 (-2,45)
Log (KommInn $_{it-1}$)	-0,894 (-1,70)	-0,516 (-0,487)	0,379 (0,451)

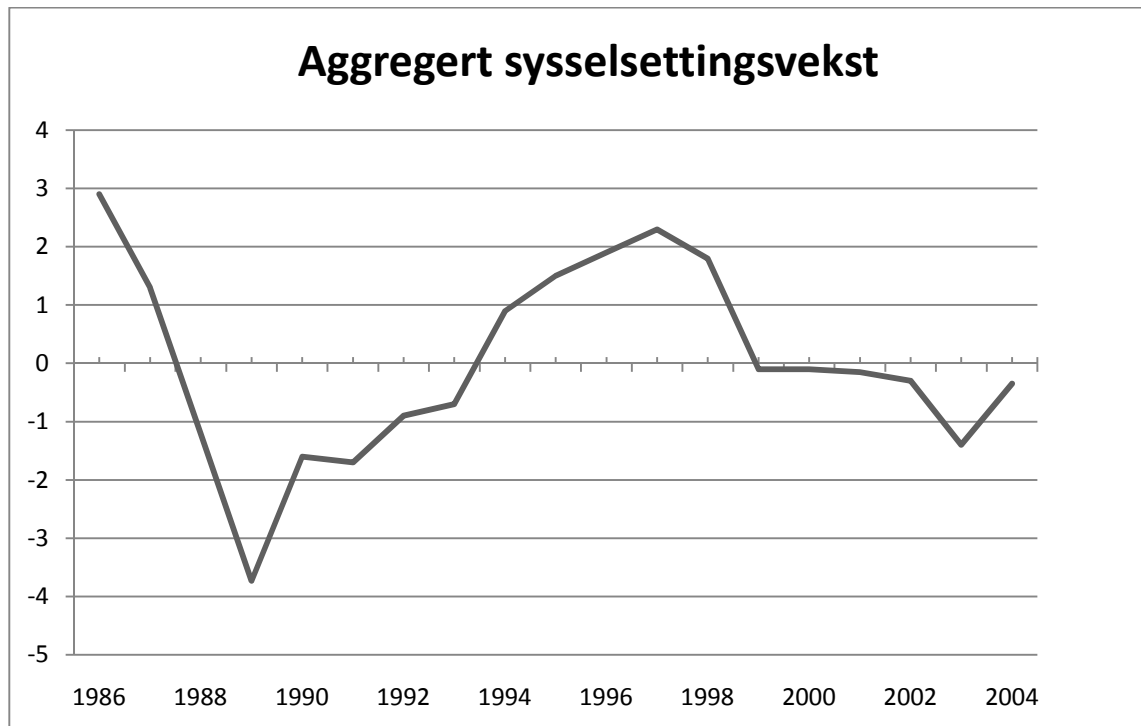
Merk: t-verdier er oppgitt i parentes, og er basert på heteroskedastisk robuste standardavvik.

Figur 2A1. Grafisk framstilling av korrelasjon mellom utdanningsgruppenes netto innmigrasjonsrater.



Merk: MigrateG/V/H tilsvarer her Innmigrasjon[GRU/VGS/HØY].

Figur 3A2. Utvikling i aggregerte konjunkturindikatorer 1986 - 2004.



Figur 4A3. Marginaleffekt av regional ledighetsrate for alle utdanningsgruppene. Sammenlignet med figur 3, gir figur 4 et mer detaljert inntrykk av hvordan marginaleffekten av regional ledighet varierer over tid som følge av variasjon i aggregert sysselsettingsvekst.

