

Forord

Jeg ønsker å rette en stor takk til min veileder, professor Kåre Johansen, for glimrende veiledning og konstruktive kommentarer under arbeidet med oppgaven. Johansen introduserte meg for problemstillingen og har gjennom hele skriveprosessen gitt grundige og inspirerende tilbakemeldinger.

Jeg vil også takke mine foreldre som har støttet meg gjennom mange års skolegang, både økonomisk og faglig. Til slutt vil jeg takke mine medstudenter for faglige og ikke faglige samtaler.

Trondheim, 31. Mai 2012

Lars-Magnus Berg Eidsaune

Innhold

1	Innledning	1
2	Litteraturoversikt	3
2.1	Teoretisk bakgrunn	3
2.2	Empiriske studier av lønnskurven	6
2.3	Lønnskurven i Norge	7
2.4	Oppsummering	8
3	Empirisk tilnærming	9
3.1	Spesifikasjon av paneldatamodell	9
3.2	Datagrunnlag og inkluderte variabler	10
3.3	Estimeringsmetode	15
3.3.1	Fixed effects	15
3.3.2	Random effects-estimering	18
3.3.3	Random eller fixed effects?	18
3.3.4	Test av koeffisientforskjeller	19
4	Empiriske resultater	21
4.1	Innledning	21
4.2	Resultater fra fixed- og random effects-estimering	22
4.2.1	Grunnmodell	22
4.2.2	Utvidet modell	28
4.3	Diagnostiske tester	31
4.4	Oppsummering	32
4.5	Andre funksjonsformer	34
4.6	Oppsummering	40
5	Sammenlikning med tidligere studier	41
6	Konklusjon	47
	Refferanseliste	49
A	Appendiks	I
A.1	Andre funksjonsformer	I
A.2	Korrelasjonsmatrise.	III
A.3	Hausman test	IV

Figurer

3.1	Relativ lønn mellom GRU og HØY	11
3.2	Gruppespesifikk og aggregert regional ledighet	12
4.1	Lønnskurene for grunnskole, videregående skole og høyere utdanning . . .	23
4.2	Lønnskurene for alle grupper ved bruk av regional ledighet i M1 og M3. M1 er representert med hel linje, mens M3 har prikket linje.	39

Tabeller

1	Deskriptiv statistikk	11
2	Korrelasjonsmatrise	12
3	Resultater fra grunnmodell	22
4	Gruppesammenlikning av estimerte koeffisientverdier	24
5	Resultat fra grunnmodell med gruppebestemt ledighet	26
6	Resultater fra utvidet modell	28
7	Gruppesammenlikning av estimerte koeffisientverdier i utvidet modell . . .	30
8	Resultater fra ulike funksjonsformer	36
9	Resultater fra tidligere norske studier	45
10	Apendiks. Resultat av funksjonsform U	I
11	Apendiks. Resultat av funksjonsform U^{-1}	I
12	Apendiks. Resultat av funksjonsform U med gruppespesifikk ledighet . . .	II
13	Apendiks. Resultat av funksjonsform U^{-1} med gruppespesifikk ledighet . .	II
14	Apendiks. Fullstendig korrelasjonsmatrise	III
15	Apendiks. Hausman test	IV

1 Innledning

Historisk sett har Norge hatt relativ lav arbeidsledighet¹, og vi har ikke opplevd en sterk økning i ledigheten slik som mange andre europeiske land, noe som også var tilfelle under finanskrisen. Inntektsforskjellen er også lav, spesielt sammenliknet med USA og Storbritannia. Disse to faktorene gjør Norge til et ”annerledesland” i ulikhetssammenheng. I motsetning til andre land har Norge ikke opplevd en økning av lønnsforskjeller knyttet til utdanning i løpet av de siste tiårene. I Europa er det en trend med økende arbeidsledighet blant personer med lav utdanning². Uten å gå i detalj kan noe av årsaken være at den sentraliserte lønnsfastsettelsen bidrar til lave lønnsforskjeller i Norge. Naturlig nok vil dette også medføre at regionale lønnsforhandlinger får mindre betydning, i motsetning til land med lavere grad av sentrale lønnsforhandlinger³. Trygve Haavelmo (1992) trekker frem økonomiske ønskemål for den norske økonomien hvor inntekts- og lønnsolidaritet skal vektlegges, både i offentlig og privat sektor.

Innad i Norge er de regionale forskjellene i arbeidsledighet relativt små. For eksempel var gjennomsnittsledigheten høyest i Østfold i 2011 med 3,3 prosent og lavest i Sogn og Fjordane med 1,9 prosent⁴. Det tradisjonelle⁵ synet på sammenhengen mellom regional lønn og regional ledighet er at regioner med høy ledighetsrate må tilby høyere lønn enn en region med lav ledighet, for å tiltrekke arbeidere. Dette synet ble utfordret sent på 1980-tallet da forskere fikk tilgang på en rekke ny mikrodata som antydte at regional ledighet har en dempende lønnseffekt.

Bakgrunnen for denne studien er Blanchflowers og Oswalds (heretter BO) bok *”the wage curve”* (1994). De finner en empirisk sammenheng mellom lønn og arbeidsledighet som gir, på tvers av land, en negativ lønnsammenheng mellom lønn og ledighet, hvor økt ledighet gir en lønnsrespons på -0,1. Resultatene har vært til dels kontroversielle og etterprøvd av mange forskere. Også i Norge har litteraturen rundt tema økt voldsomt. Card (1995) gir i sin omtale av boken skarp kritikk av holdbarheten av resultatet. Han påpeker blant annet at resultatet kan være en mistolket Phillipskurve eller en arbeidstilbudsfunksjon. Videre poengterer Card; *”de konsentrerer seg utelukkende om effekten av total ledighet. Et interessant spørsmål er hvorvidt lønnen til en definert gruppe av arbeidere er sterkere knyttet til gruppens ledighet eller til den totale ledigheten”*⁶ Card (1995:794) og antyder med dette at ulik lønnsrespons mellom grupper kan oppstå som følger av en dårlig spesifisert modell. Modellspesifikasjon er altså en medvirkende årsak til variasjon i rapporterte

¹Se f.eks ”Labour market statistics” <http://dx.doi.org/10.1787/unemp-table-2011-1-en>

²Hægeland og Kirkebøen (2007).

³Se Calmfors og Driffill (1988).

⁴Kilde statbank.ssb.no/statistikkbanken

⁵Teorien om komenserende lønnsforskjeller strekker seg tilbake til Adam Smith. Nyere empiriske funn støtter hypotesen. Se f.eks Harris og Todaro (1970).

⁶Min oversettelse.

elastisiteter som er utført i etterkant.

I en empirisk meta-analytisk studie utført av Nijkamp og Poot (2005) ser de på 208 elastisiteter hentet fra den voksende litteraturen⁷ i et forsøk på å avdekke hvorfor elastisiteten varierer mellom studiene. De finner at bare 14,9 prosent tester andre funksjonsformer enn BOs konstante elastisitetsfunksjon⁸, og et fåtall ser på gruppeforskjeller. Som det ble nevnt innledningsvis, øker ledigheten mellom utdanningsgrupper, spesielt i Europa. Dette betyr at bruk av aggregert ledighet er et dårlig mål i et forsøk på å avdekke hvorvidt utdanningsgrupper opplever ulik lønnsrespons av endret ledighet.

Problemstillingen i denne studie er å teste om det finnes en norsk regional lønnskurve for tre ulike utdanningsnivå, henholdsvis grunnskole, videregående skole og høyere utdanning. Det undersøkes spesielt om gruppene responderer signifikant ulikt, både ved bruk av aggregert og gruppespesifikk regional arbeidsledighet. I en utvidet modell ser vi hvordan lønn responderer på ulike arbeidsmarkedsforhold og andre variabler, hvor hovedfokus vil være på langtidsvirkningen av lønn og arbeidsledighet. Det vil også bli undersøkt ulike funksjonsformer av lønnsrelasjonen, i tråd med BO.

Formålet med oppgaven er ved bruk av nytt datamateriale å fremlegge nye empiriske funn av lønnskurven i Norge for ulike utdanningsgrupper. Det som skiller denne studien fra andre, er bruk av gruppespesifikk regional ledighet som forklaringsvariabel.

Opgaven er delt inn i 6 kapitler. I kapittel 2 gis en bakgrunnsoversikt av ulike teorier som forklarer lønn og ledighetssammenhengen. Det blir også gitt en oversikt over internasjonale og norske empiriske funn tilknyttet lønnskurven. I kapittel 3 blir inkluderte variabler omtalt og spesifisering av modell vist. Det blir presentert to estimeringsmetoder, fixed og random effects. I kapittel 4 blir de empiriske resultatene fremlagt på bakgrunn av forgående kapittel. Først presenteres en basismodell hvor vi ser på gruppeforskjeller med og uten gruppespesifikk ledighet. Deretter blir det eksperimentert med ulike funksjonsformer av lønnskurven for å avdekke dens robusthet. Til slutt sammenliknes resultatene med tidligere studier. Siste kapittel gir en avsluttende konklusjon.

⁷De bruker 17 studier mellom 1990 og 2001.

⁸Blanchflower og Oswald tester for ulike funksjonsformer av ledighetsraten og konkluderer med at $LnW = \alpha + \beta LnU + \text{andre variabler}$ passer best.

2 Litteraturoversikt

Helt siden den store depresjonen på 1930-tallet har arbeidsledighet og lønn vært et av de fremste forskningstemaene i makroøkonomi. Det er gjort mange studier og fremlagt flere teorier om emnet, og i dette kapitlet vil jeg gi en oversikt over noen av dem. I nyere tid har flere forskere brukt en mikroøkonomisk tilnærming som grunnlag i sine studier. Formålet med kapitlet er å bli kjent med den viktigste litteraturen rundt dette tema, med fokus på den mikroøkonomiske litteraturen, for senere å bruke det i min empiriske analyse.

2.1 Teoretisk bakgrunn

En av de mest innflytelsesrike funnene ble gjort av Alban W. Phillips i 1958. Phillips bruker data over en 95-års periode og finner en negativ sammenheng mellom lønnsraten og arbeidsledighetsraten. Essensen i Phillipskurven er at det finnes en korttidsavveining mellom ledighet og inflasjon. Fall i ledigheten kan føre til økt inflasjon. Det finnes en stor mengde litteratur som omhandler Phillipskurven⁹. Milton Friedman (1968) kritiserte den originale Phillipskurven og tok i bruk begrepet NAIRU (non-accelerating-inflation rate of unemployment) for å skille mellom den kortsiktige og langsiktige Phillipskurven. På lang sikt ble kurven vertikal på grunn av forventninger folk satte seg. Hvis en opplevde høy inflasjon i dag, ville en forvente høyere gjennomsnittlig inflasjon i fremtidige perioder, som det måtte tas hensyn til. Kurven fikk en stabil inflasjonsrate og dermed ingen avveining mellom inflasjon og ledighet.

Siden Robert Lucas og Leonard Rappings (1969) studie av det aggregerte arbeidsmarkedet har flere argumentert for at ledighet bør tolkes som *arbeidstilbud*. Hovedideen er at husholdningen jobber flere timer når lønningene er midlertidig høye og færre timer når lønningene er midlertidig lave. Ved å estimere den intertemporære arbeidselastisiteten kan en modellere fluktuasjonen av arbeidskraft over en periode. Likevel er det ifølge Layard et al. (1991) ingen tilfredsstillende empiriske funn som støtter denne teorien. Et annet problem er at det ikke finnes ufrivillig ledighet i denne modellen.

På mikronivå finner vi flere teoretiske modeller som forsøker å forklare sammenhengen mellom lønn og ledighet. *Implisitt kontraktmodell* ble formulert av Azariadis i 1975. I utgangspunktet er slike modeller relativt komplekse, men lar seg sammenfatte ved noen enkle antakelser. En antar at det er usikkerhet om tilstanden i verden på grunn av for eksempel tilfeldige produktivitetssjokk. Husholdningen misliker usikkerhet og er risikoavers. Bedrifter på den annen side bryr seg ikke mye om usikkerhet og er risikonøytrale. Under disse omstendighetene kan en pareto optimal handel pågå mellom arbeiderne (husholdningen)

⁹se f.eks. Manning (1993)

og bedriftene. I bytte for en stabil reallønn må arbeiderne betale en forsikringspremie til arbeidsgiveren ved å akseptere å jobbe til en lavere reallønn. En slik teori forklarer lønnsstivhet over en periode, men ikke ledighet. Dette blir derimot behandlet i teorien om *effektivitetslønn*.

Shapiro og Stiglitz (1984) tok utgangspunkt i denne teorien og fremla en modell som kunne forklare ufrivillig arbeidsledighet. Hypotesen argumenterer for at lønn ikke bare settes av tilbud og etterspørsel. Bedriften setter en høyere lønn enn markedet skulle tilsi for å øke produktiviteten (effektiviteten). Stiglitz (1986) viser at det er minst fem ulike forklaringer når det kommer til forholdet mellom lønn og en arbeiders produktivitet. For det første er det en direkte sammenheng mellom lønn og ernæringsinntak. Videre vil lav lønn gi større gjennomtrekk av arbeidere og derfor lavere produktivitet på grunn av at nye ansatte må gjennom en opplæringsfase. Når bedriften står overfor imperfekt informasjon om arbeideren, vil høy lønn tiltrekke seg kvalitetsarbeidere. Bedriften vil også ha imperfekt informasjon om arbeiderens handlinger i jobben. Derfor vil ledighet fungere som en disiplineringsmekanisme. Hvis arbeideren skulker jobben, kan bedriften sparke vedkommende. Høy lønn kan derimot medføre at arbeideren føler seg rettferdig behandlet og viser god moral. Ettersom høy lønn gir økt produktivitet, vil alle bedrifter i markedet tilby høyere lønn, som igjen fører til lønnspress som overstiger markedsklareringen, og ufrivillig arbeidsledighet oppstår.

Arbeidsledighet vil også skyldes en innsiders maktposisjon. Nettopp dette er hovedideen i den såkalte *insider-outsider* teorien til Lindbeck og Snower (1988). Ledighet oppstår fordi det er dyrt å bytte ut ansatte (innsidere) med ledige arbeidere (utsidere) på grunn av omsetningskostnader. Siden innsidere er med i lønnsprosessen, kan de få denne kostnaden i form av lønn fordi bedriften vil spare ansettelse-, oppsigelse- og opplæringskostnader. Det vil derfor oppstå ufrivillig ledighet på grunn av interessekonflikt mellom innsidere og utsidere siden innsiderne ikke bryr seg om utsiderne.

Mange forskere har pekt på *forhandlingsmodellen*¹⁰ som årsak til arbeidsledighet. I Hoel og Nymoen (1988) antas det at lønn fastsettes ved forhandlinger mellom fagforeningen og bedriften, der bedriften setter antall ansatte etter lønnsforhandlingene for å maksimere profitt. Fagforeningen er opptatt av forventet nytte til sine medlemmer. Velferden til den ansatte avhenger av reallønna og faktorer som bestemmer arbeidsforholdene. En arbeidsledig vil få ledighetstrygd, for så å finne en alternativ jobb. Det antas at å finne en ny jobb vil øke ens velferd, mens høy ledighet vil forlenge tiden en går arbeidsledig. Derfor vil forventet nytte for den enkelte arbeidende øke med reallønnen, forventet alternativ lønn og ledighetstrygd. Den vil avta med dårlige arbeidsforhold og høyt ledighetsnivå. Til slutt blir utfallet av lønnsforhandlingene løst med et asymmetrisk Nash problem¹¹. Uenighet

¹⁰Blant annet Layard et al (1991).

¹¹Se f.eks Moene, Wallerstein og Hoel (1993).

kan utløse streiker.

I Norge forhandles det om lønn både på sentralt og lokalt nivå hvor makroøkonomiske faktorer tas hensyn til ved nasjonal forhandling, mens de lokale forhandlingene tilpasser seg de regionspesifikke forholdene. En sentral lønnsforhandling vil ikke kunne fange opp alle faktorer som påvirker lønnen. For eksempel vil lokale forhold som arbeidsmiljø kunne påvirke lokal lønn. Derfor vil alle sentrale lønnsforhandlinger suppleres med lokale forhandlinger. Effekten av regionale lønnsforhandlinger vil variere med landets struktur, der land med høy sentralisering vil oppleve sterkere regional lønnsrigiditet enn land med mindre sentralisering. I Norge vil derfor betydningen av lokale forhandlinger være mindre sammenliknet med for eksempel USA.

I en omfattende studie gjort av BO (1994) finner de empirisk dokumentasjon på sammenhengen mellom ledighet og reallønn som antyder en universell *lønnskurve*. Studien baserer seg på mikrodata av 3.5 millioner mennesker i 12 land. Hovedresultatet er at den gjennomsnittlige ledighetselastisitet av lønn er nærmere lik $-0,1$. Dette betyr at en hypotetisk dobling i ledigheten, vil gi et lønnsfall på 10%. Som BO poengterer, kan denne negative sammenhengen mellom lokal arbeidsledighet og lønn forveksles med Phillipskurven og en arbeidstilbudsfunksjon. Likevel mener de at det finnes klare forskjeller og understreker at Phillipskurven var foreslått som en justeringsmekanisme mot likevekt, mens lønnskurven må sees på som en likevektsallokering som ikke beskriver et midlertidig fenomen. Videre argumenteres det for at Phillipskurven kobler endring i lønn med ledighetsraten, mens lønnskurven lenker lønnsnivået i seg selv til ledigheten. Til slutt er datagrunnlaget en vesensforskjell, hvor Phillipskurven er basert på makrodata, mens lønnskurven er estimert ved hjelp av mikrodata. Videre distanserer de lønnskurven fra arbeidstilbudstilnærming ved å vise til empiriske resultater. Resultatene fra USA og Storbritannia viser at det er lokal ledighet snarere enn lokal sysselsetting eller størrelsen av lokal arbeidsstyrke, som påvirker lønnen. I den grad en tror arbeidstilbudsfunksjonen skal uttrykke den totale mengden arbeidskraft som en funksjon av lønn, synes funnene å være uforenelig med arbeidstilbudtolkningen.

Funnene kan ikke desto mindre knyttes opp mot tidligere litteratur. BO trekker særlig frem forhandlingsmodellen, effektivitetslønnmodellen og implisittkontraktsmodellen som teoretisk forklaring av lønnskurven. I forhandlingsmodellen vil økt ledighet minke jobbalternativene og derfor svekke arbeidernes forhandlingsmakt, og en negativ sammenheng mellom ledighet og lønn oppstår. Intuisjonen for en lønnskurve i effektivitetslønnmodellen baserer seg på at kostnaden ved å skulke jobben øker ved økt ledighet fordi det er vanskeligere å finne seg jobb hvis en blir avslørt og oppsagt. Derfor kan bedriften gi lavere lønn ved høy ledighet. I implisittkontraktsmodellen antar BO at ulike områder har ulike kvaliteter, og disse ulikhetene skaper forskjell i lønn og forventet ledighet. Ved å anta lik ledighetstrygd i alle regionene viser de at regioner med gode kvaliteter vil tilby kontrakter

med lavere lønn (Card 1995).

2.2 Empiriske studier av lønnskurven

Etter BOs funn av en universell lønnskurve har flere forskere etterprøvd resultatene i ulike land.

Wim Groot et al. (1992) reproducerer BOs resultater fra 1990, men finner bare en svak bekreftelse på deres empiriske lov. Det er først når de redefinerer ledighetsraten, samt tar hensyn til at ledighet har to motstridende effekter, en ren lønnseffekt og en positiv effekt på lønn ved redusert arbeidstilbud, at de finner en lønnselastisitet relativt nær $-0,1$. Rudolf Winter-Ebmer (1996) analyserer i en tverrsnittstudie av østerriksk data forholdet mellom lønn og ledighet, og bekrefter en negativ lønnskurve. I hans funn er lønnselastisiteten lavere enn hos BO. Stefan Janssens og Jozef Konings (1998) finner også en negativ lønnskurve hvor elastisiteten for menn er nær $-0,1$. For kvinner derimot finner de en positiv insignifikant sammenheng. Deres forklaring på dette er at arbeidsmarkedet for kvinner er mer konkurransedyktig og mindre påvirket av fagforeninger enn arbeidsmarkedet for menn. Ikke uventet finner de også at mer utdanning gir høyere fortjeneste, både for kvinner og menn. I Tyskland viser en studie gjennomført av Joachim Wagner (1994) at lønnskurven er mest markant i første halvdel av 1980-årene. I estimat av data fra 1984 til 1990 er det ingen statistiske bevis for en negativ lønnskurve. Noe av forklaringen kan ligge i endringene av interregional variasjon i ledighetsraten mellom første og andre del av 1980-årene. Bell et al. (2002) bekrefter BO sine funn for Storbritannia. De konkluderer med at en langtidselastisitet for gjennomsnittlig regionale lønninger med hensyn på ledighet ligger mellom $-0,11$ og $-0,13$. Nicolaisen og Tranæs (1996) tar for seg en eventuell lønnskurve i Danmark og finner en signifikant lavere lønnselastisitet enn den for BO. Sammenhengen er fortsatt negativ og ligger mellom $-0,02$ og $-0,04$, avhengig av om de korrigerer for de ulike konjunktorene i regionene.

Som blant annet Card (1995) påpeker, vil lønnskurven være ulik for ulike grupper av arbeidere enten det er menn eller kvinner, ung eller gammel, høy eller lav utdanning, sterk eller svak fagforening og så videre.

Campbell (1997) tar for seg lønnsstivhet mellom ulike utdanningsgrupper i USA. Han bruker data innhentet fra 1975 til 1994. Han estimerer resultater for white collar og blue collar-ansatte hvor white collar-yrker er ledere, kontorpersonell og salgsarbeidere og blue collar er henholdsvis håndverkere, transportarbeidere og teknikere. Funnet i analysen gir en invers sammenheng mellom ferdighetsnivå i et yrke og i hvor stor grad aggregert ledighet påvirker lønnen i yrket. Videre konkluderes det at white collar-lønninger er mindre responsiv på aggregert ledighet enn blue collar-lønninger. Den inverse relasjonen mellom

ferdighetsnivå og responsen av lønn på aggregert ledighet kan være bevis for effektivitetslønn og implisitt kontraktsteori.

2.3 Lønnskurven i Norge

Det er gjort mange estimat av lønnskurven i Norge. BO (1994) var de første til å studere en mikrobasert lønnskurve i Norge. Som nevnt ble det funnet en lønnselastisitet på $-0,1$ for et datasett fra 1989 - 1991. Når de innfører 18 regionale dummyer, faller koeffisienten til $0,01$ og blir ubetydelig. Nøyer de seg med 4 dummyer, finner de en elastisitet på $-0,08$ som er i samsvar med funn i de øvrige landene de undersøker.

Johansen (1997) bruker aggregert tidsseriedata fra norsk industri i perioden 1964 - 1990. Hovedresultatet er at ledighet påvirker lønnsresponsen signifikant og har en langtidselastisitet på $-0,07$. Ved bruk av en stykkevis ("piecewise") regresjonsmodell som tillater ledighetskoeffisienten å endre seg for ulike nivåer av ledighet, finner han en langtidslønnselastisitet på $-0,16$ for lav arbeidsledighet, mens den er $-0,035$ for høy ledighet (over 3%). Konklusjonen er at lønnsresponsen på ledighet fungerer som en inflasjonsdrivende mekanisme i en het økonomi, snarere enn en likevektskraft i arbeidsmarkedet.

I Dyrstad og Johansen (2000) undersøker de hvor responsiv regionale lønninger er i forhold til regional ledighet. De bruker paneldata av industrilønninger i perioden 1970 - 1988, hvor de konkluderer med at den regionale lønnsresponsen på regional arbeidsmarkedsforhold er små og estimerer en langtidselastisitet av regional lønn på regional ledighet til $-0,02$. Responsen på aggregert ledighet er større og ligger på $-0,07$. Derfor konkluderer de med at aggregert ledighet har større betydning enn regional ledighet, og at ledighet i første rekke påvirker lønninger via sentrale lønnsoppgjør. Tolkningen av dette er at effekten for aggregert ledighet fungerer gjennom sentral lønnssetting. Til slutt påpeker de at summen av regional og aggregert ledighetselastisitet er $-0,09$ som er nært BOs funn på $-0,1$.

Albæk et. al. (1999) finner ingen tegn til en lønnskurve etter å ha kontrollert for fixed effects ved bruk av regional data. De konkluderer med at forbigående endring i regional arbeidsledighet ikke kan sees i de regionale lønningene. Barth et.al (2002) mener årsaken til dette er at studien baserer seg på tall fra 1989 til 1993, en periode karakterisert som høy grad av sentralisering og ganske lav nominell lønnsøkning. I deres egen studie inkluderer dataene også observasjoner fra 1993 - 1997, en periode hvor lokal lønnsforming utgjorde en betydelig større del av total lønnsvekst. Resultatet er det samme som hos Dyrstad og Johansen (2000) hvor det ble funnet en langtidselastisitet av lønn med hensyn til regional ledighet på $-0,02$.

I likhet med Campbell (1997) utfører Johansen (1999) en empirisk studie av lønnsrespons knyttet til arbeidsledighet hos høyt og lavt utdannede. Resultatene baserer seg på data

for gjennomsnittlig timeinntekt for faglærte og ufaglærte i den norske verkstedsindustrien. Analysen gir empirisk indikasjon på at responsen på ledighet er sterkere for lavt utdannede enn høyt utdannede. Langtidselastisiteten for "skilled" arbeidere blir estimert til -0,05 mens for "unskilled" arbeidere ligger den på -0,1. Stykkevis ("piecewise") regresjon gir for begge grupper en tilnærmet flat lønnskurve for ledighet over 1,7%. Langtidselastisiteten er -0,12 for faglærte og 0,18 for ufaglærte ved ledighet under 1,7%. Dette tolkes dit hen at ulike mekanismer er i sving ved ekstremt lav ledighet og ved høyere ledighet. Johansen anser at effektivitetslønnmekanismer slår til ved lav ledighet, og at bedrifter har sterke insentiver til å øke lønn for å rekruttere og beholde arbeidere. Med høy ledighet forsvinner insentivene til å øke lønnen, og fagforeningens innflytelse blir lønnsavgjørende siden forventet kostnad ved jobbavskjedigelse øker ved ledighet. Likevel kan lønnseffekten via forhandlingene være svak.

2.4 Oppsummering

Det har i lengre tid vært en heftig diskusjon i fagmiljøet om hvorvidt det finnes en universell lønnskurve på -0,1. Det har også vært omstridt om BOs funn kan være en feilspesifisert Phillipskurve eller arbeidstilbudsfunksjon. I kjølvannet av deres bok er det en stadig voksende litteratur som etterprøver resultatene. Jevnt over ser det ut til at det finnes en lønnsrelasjon mellom lønn og ledighet, men resultatene er delvis sprikende.

Også i Norge er det gjort mange studier og de fleste finner en signifikant lønnsrespons av ledighet på lønn. For regional ledighet varierer elastisitetene mellom -0,08 og -0,02. Videre kan modellspesifikasjon og metodebruk være høyst varierende, noe som kan være noe av årsaken til ulike resultat.

Ulike teoretiske modeller er tatt i bruk som forsøk på å forklare sammenhengen. Internasjonalt er det ingen modell¹² som passer bedre enn andre, men for Norges del, med høy sentralisering, anses forhandlingsmodellen å gi god teoretisk forståelse. I fortsettelsen vil undertegnede gjøre en paneldatanalyse hvor utdanningsnivå mellom tre grupper, grunnskole, videregående skole og høyere utdanning blir sammenliknet. Det er så vidt meg bekjent ikke foretatt liknende norske undersøkelser hvor regional gruppespesifikk ledighet er brukt som forklaringsvariabel.

¹²Richard Townsend (2005) har utført en empirisk studie hvor han tar for seg ulike teoretiske modeller og prøver å finne ut hvilken modell som passer best.

3 Empirisk tilnærming

For å undersøke responsen av lønn på arbeidsledighet for ulike utdanningsgrupper benytter vi et paneldatasett for perioden 1995 - 2004. I paneldatanalysen ser vi på tverrsnittsenhet over tid, og det vil være utilstrekkelig å benytte vanlig MKM (minste kvadraters metode) for å oppnå forventningsrette estimatorer¹³.

3.1 Spesifikasjon av paneldatamodell

Vi tar utgangspunkt i den generelle lønnskurven gitt av BO hvor regional variasjon i ledighet og lønn viser effekten av ledighet på lønn. I tillegg inkluderer vi laggede verdier som forklaringsvariabler fordi effekten på lønn ved endring i forklaringsvariablene ikke inntreffer momentant. Målet med en slik dynamisk modell er å danne en mer realistisk lønnsrelasjon. Ved å sette opp det regionale lønnsnivået mot det regionale ledighetsnivået kan forholdet mellom lønn og arbeidsledighet estimeres. Basismodellen er gitt som en regresjonslikning på følgende form:

$$\begin{aligned} \logl\ddot{o}nn[X]_{it} = & \alpha_1 \log(W[X]_{it-1}) + \alpha_2 \log(U_{it}) + \alpha_3 \log(U_{it-1}) + \alpha_4 \log(VakS_{it}) \\ & + \alpha_5 \log(VakS_{it-1}) + \alpha_6 \log(Andre\ variabler)_{it} + \lambda_t + v_{it} \end{aligned} \quad (3.1)$$

Variabelen $\logl\ddot{o}nn[X]_{it}$ er definert som log reallønn innen en utdanningsgruppe X hvor fotskrift i angir region og t angir tidspunkt. X innebefatter de tre utdanningsgruppene grunnskole, videregående skole og høyere utdanning. Heretter blir de forkortet med GRU, VGS og HØY. Vi estimerer tre ulike lønnskurver for hver utdanningsgruppe, og modell 3.1 forstås som en kompakt formulering av disse. W_{it-1} er regional lønn innen utdanningsgruppe X under forrige periode som vi ser av fotskrift $t-1$. U_{it} og U_{it-1} er henholdsvis regional arbeidsledighet i periode t og $t-1$. $VakS_{it}$ og $VakS_{it-1}$ er vakanserestrøm på fylkesnivå i periode t , og $t-1$. *Andre variabler* er en vektor av andre forklaringsvariabler mens λ_t er et sett av tidsdummyer som blant annet fanger opp effekten av nasjonal arbeidsledighet, slik at effektene i modellen fremkommer av regionale forhold. Forklaringsvariablene vil bli nærmere beskrevet i neste delkapittel.

Restleddet i modellen, v_{it} , antas å bestå av to komponenter som er identiske og uavhengig normalfordelt på følgende vis: $v_{it} = \eta_i + \epsilon_{it}$. Her utgjør η_i en individspesifikk komponent som ikke varierer over tid. ϵ_{it} er en tidsvarierende (idiosynkratisk) komponent. Den individspesifikke komponenten fanger opp effekten av utelatte variabler som kan påvirke lønn og er tilnærmet konstant over tid, men varierer mellom regionene. Dette kan for eksempel være geografiske forhold¹⁴. I kapittel 3.3 går vi nærmere inn på antakelser som blir gjort

¹³Alle regresjoner er foretatt ved bruk av PcGive i OxMetrics 6.

¹⁴F.eks krafttilgang og mineraler.

i forbindelse med estimeringsmetode.

Vår høyresidevariabel er satt opp på naturlig logaritmisk form. Som Wooldridge (2006) påpeker, er det flere fordeler ved å bruke den naturlige logaritmen. Strengt positive variabler vil ofte ha fordelinger som er heteroskedastiske eller skjeve. Ved å ta log kan en redusere og i beste fall eliminere problemet. Videre vil log vanligvis begrense spennvidden i variabelen som gjør estimatet mindre sensitivt for ekstreme observasjoner. Siden både lønn og ledighet er på log form vil α tolkes som elastisiteten av lønn med hensyn til ledighet. Dette betyr at en hypotetisk dobling i ledigheten vil gi en lønnsendring på α prosent. Som vi vet, har BO i sin studie av lønnskurven kommet frem til et universelt estimat på $-0,1$. Dette vil gi et lønnsfall på 10 prosent ved en dobling i arbeidsledigheten.

Hovedpoenget med en modellspesifikasjon er å finne en avveining mellom en rikt spesifisert modell på bekostning av frihetsgrader, eller en enklere modell som i større grad ivaretar statistisk validitet. Problemet med enkle modeller kan være utelatt variabelskjevhet, der vi ikke har fanget opp signifikante forklaringsvariabler. I vårt datasett kan vi tillate en utvidet modellspesifikasjon fordi datagrunnlaget er av relativt god lengde. Poenget med den enkle grunnmodellen er å etterprøve tidligere resultat, for senere å kunne inkludere flere forklaringsvariabler og sammenlikne eventuelle forskjeller.

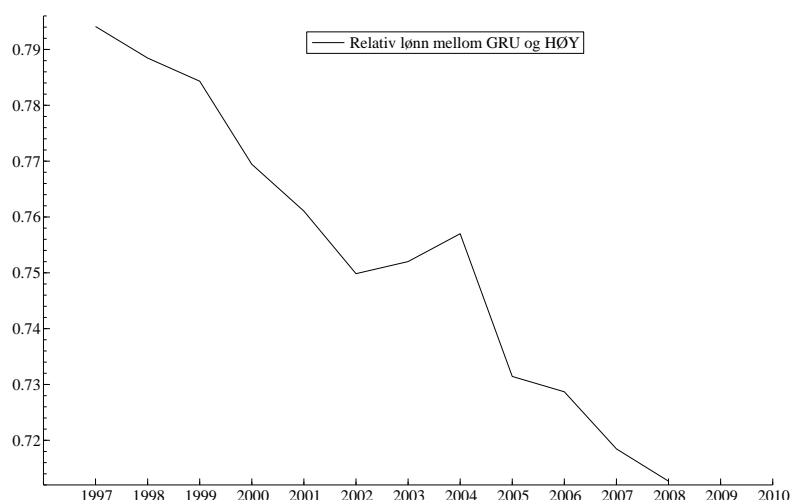
3.2 Datagrunnlag og inkluderte variabler

Datamaterialet som benyttes i analysen er hentet fra Statistisk sentralbyrå (SSB) og Arbeids- og velferdsforvaltningen (NAV). Vi har data i perioden 1995 til 2004 der vi bruker lønn, arbeidsledighet, vakanserate, ordinære tiltak, boligpris samt sommer- og vintertemperaturer (hentet fra meteorologisk institutt). I tillegg har vi separate variabler for lønn og arbeidsledighet delt inn i tre utdanningsgrupper; grunnskole, videregående skole og høyere utdanning.

Den regionale inndelingen som benyttes, består av 90 økonomiske regioner definert av SSB. Johansen (2003) påpeker at det faktiske jobbmarkedet for landets innbyggere ofte består av flere nærliggende kommuner. Dermed vil en kommunal inndeling være for liten, mens en fylkesinndeling vil være for stor. Grunnlaget for valg av økonomisk region satt av SSB, er pendlingstall, varehandelsstatistikk og befolkningstall i største tettsted. Dermed vil en region bestå av flere kommuner som har et felles arbeidsmarked.

Den avhengige variabelen $\log\text{l\o}nn[X]_{it}$ viser som nevnt reallønn i region i innen utdanningsgruppe X . I analysen er dataen justert for blant annet alder og kjønn, slik at vi kontrollerer for ulike sammensetninger mellom regioner og forskjeller over tid.

Figur 3.1: Relativ lønn mellom GRU og HØY



Når det kommer til de relative lønnsforskjellene mellom utdanningsgruppene, har de i Norge, i motsetning til andre land, tradisjonelt sett vært små, noe som kan sees av figur 3.1. Likevel ser vi en tendens til økte lønnsforskjeller, der høyere utdanning har størst lønnsvekst. Som Johansen (1999) poengterer, støtter dette hypotesen at lønnsforskjeller reduseres i høykonjunkturer og blir større i lavkonjunkturer.

I tabell 1 ser vi deskriptiv statistikk for variablene. Ved å se på standardavvikets størrelse relativt til variabelens gjennomsnittsverdi, kan vi sammenlikne variasjonen i de ulike variablene. Vi ser tydelig at økt utdanningsnivå gir økt gjennomsnittlig lønn samtidig som standardavviket øker tiltakende. Dette betyr at det er større spredning i lønnen for utdanningskategorien høyere utdanning enn for grunnskole. Dette er ingen overraskelse siden mangfoldet av jobber i høyere utdanning er større og strekker seg over flere sektorer med ulik lønnsstruktur.

Tabell 1: Deskriptiv statistikk for variabler som inngår i paneldataanalysen

Variabler	Forkortelse	Gjennomsnitt	Standardavvik
Lønn[GRU]		3109	527,67
Lønn[VGS]		3431,5	615,97
Lønn[HØY]		4468,3	952,36
Log Lønn[GRU]	lWG_{it}	8,02	0,171
Log Lønn[VGS]	lWV_{it}	8,125	0,180
Log Lønn[HØY]	lWH_{it}	8,382	0,212
Regional Ledighet	U_{it}	3,313	1,290
Regional Ledighet[GRU]	UG_{it}	5,619	2,147
Regional Ledighet[VGS]	UV_{it}	4,759	1,857
Regional Ledighet[HØY]	UH_{it}	1,972	0,741
Vaksansrate	$VakS_{it}$	15,431	6,224
Boligpris	$BPNY_{it}$	0,815	0,327
Ordinære tiltak	OT_{it}	0,934	0,671
Sommertemperatur	$sommertemp_{it}$	14,436	2,17
Vintertemperatur	$vintertemp_{it}$	-2,05	3,791

Tabell 2: Korrelasjonsmatrise

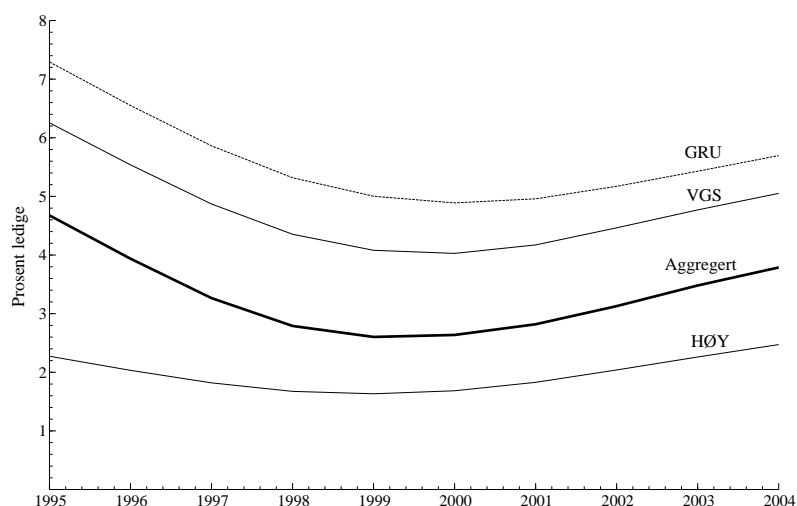
	LogLønn[GRU]	LogLønn[VGS]	LogLønn[HØY]
LogLønn[GRU]	1.0000		
LogLønn[VGS]	0.9875	1.0000	
LogLønn[HØY]	0.9750	0.9847	1.0000
LogU	-0.1670	-0.1761	-0.1504

Vi ser videre en negativ sammenheng mellom gjennomsnittlig lønn og arbeidsledighet, hvor lav utdanning gir lav gjennomsnittlig lønn og høy arbeidsledighet. Det avgjørende for en god analyse av denne typen, er at vi har tilstrekkelig tidsvariasjon i data innenfor den enkelte region.

I tabell 2 over ser vi korrelasjonen av lønn og arbeidsledighet mellom utdanningsgruppene. Ikke overraskende er det høy korrelasjon i lønnen hvor GRU og VGS korrelerer mer enn GRU og HØY. Dette betyr at lønnen i utdanningsgruppene følger hverandre nært. Vi ser også en negativ korrelasjon mellom ledighet og lønn. Selv om korrelasjonen er lik mellom alle gruppene, er det uansett interessant å studere lønnsresponsen ved endring i forklaringsvariablene. Tilsynelatende små forskjeller kan likevel gi store endringer i lønn ved endring i arbeidsledigheten. Tidligere studier har vist at lavt utdannede har en sterkere lønnsrespons ved endring i ledighet, spesielt når arbeidsledigheten er generell lav (Johansen 1999).

Figur 3.2 viser tidsutviklingen for de gruppespesifikke arbeidsledighetene samt aggregert regional ledighet. Utviklingen over tid er nokså lik for gruppene. Lavest arbeidsledighet finner vi rundt årtusenskiftet før den igjen tiltar. De tre gruppenes ledighet korrelerer ulikt med aggregert regional ledighet som kan sees i korrelasjonsmatrise i apendiks A.2.

Figur 3.2: Gruppenspesifikk og aggregert regional ledighet



Videregåendegruppen korrelerer sterkest med 0,951. De to øvrige gruppene korrelerer med 0,848 og 0,703 for henholdsvis grunnskole og høyere utdanning. På bakgrunn av dette vil vi forvente et justert bilde av sammenhengen mellom lønn og ledighet ved bruk av gruppespesifikk ledighet. Siden videregåendegruppen korrelerer sterkest, vil vi se minst utslag her, mens resultatene for høyere utdanning vil endres mye på grunn av lavere korrelasjon med aggregert regional ledighet.

Når vi ser på de inkluderte variablene, er lagget lønn, W_{it-1} , inkludert for å fange opp hvordan lønn blir påvirket av forrige periodes lønn. Det kan også argumenteres for at regional lønn blir påvirket, ikke bare av den aktuelle regionens arbeidsmarked, men også av hele det aggregerte arbeidsmarkedet. Dette kan skje fordi noen lønninger blir satt, eller var satt, nasjonalt (Bell et al. 2002). Denne effekten fanges opp i inkluderte dummyvariabler og vi ser derfor bare på effekter på regionalt nivå.

Ved å inkludere regional ledighet og vakanseraten i analysen, kan vi kontrollere for ulike effekter av utdanningsnivå og påvirkning fra den nasjonale konjunktursituasjonen. Vi vil forvente at den regionale arbeidsledigheten og vakanseraten vil påvirke lønn i hver sin retning. Økt regional ledighet vil gi lavere lønn, mens økt vakanserate vil øke lønnen. Det er tenkelig at disse variablene vil påvirke utdanningsgruppenes lønn i ulik grad. For eksempel vil utdanning normalt sett utvide et individs yrkesmuligheter. Disse variablene er oppgitt i prosent av beregnet arbeidsstyrke, der arbeidsstyrken er definert som summen av antall sysselsatte og arbeidsledige.

BO tester i sin studie fem ulike lønnskurver ved å bruke ulike spesifikasjoner av ledigheten. Spesifikasjonene i regresjonen er som følger:

1. Naturlige logaritmen av ledighetsraten, $\log U$
2. Naturlige logaritmen av ledighetsraten, $\log U$, og dens kube $(\log U)^3$
3. Ledighetsraten U
4. Ledighetsraten U og dens kvadrat U^2
5. Den inverse av ledighetsraten U^{-1}

Alle spesifikasjonene resulterer i en fallende lønnskurve, men BO finner at den første spesifikasjonen passer best og bruker den i sin analyse. Albæk et. al (1999) diskuterer hvordan ulike arbeidsmarkedsforhold påvirker lønnen. Fra et teoretisk ståsted er det derfor ikke klart hvilken arbeidsledighet vi bør inkludere i vår regresjon. De tester empirisk forskjellen mellom lang- og korttidsledighet hvor det antas at langtidsledige ikke vil konkurrerer i like stor grad som korttidsledige om arbeid, fordi ledighet over lang tid vil depresierte ens arbeidsferdigheter. Arbeidsgivere kan være skeptiske til å ansette arbeidere som har vært arbeidsledige over lang tid. Hvis langtidsledige ikke har den samme dempende effekten

på lønn som korttidsledige, vil det være viktig å tilrettelegge arbeidsmarkedsforhold som reduserer langtidsledighet. Dette vil styrke avhengigheten av endringer i ledighet på lønn. Albæk et. al konkluderer med at langtidsledige ikke demper lønnspresset like mye som korttidsledige, samt at arbeidsmarkedstiltak øker effektiviteten i arbeidsmarkedet i Norge. Dette tatt i betraktning, bruker jeg den generelle regionale arbeidsledigheten for hver utdanningsgruppe og ser bort fra eventuelle forskjeller mellom lang- og korttidsledige. Jeg vil først ta for meg funksjonsformen $\log U$, deretter undersøke andre funksjonsformer hvor ledighetsraten U og den inverse av ledighetsraten U^{-1} brukes som forklaringsvariabel.

Når det kommer til andre variabler, er det i den generelle litteraturen inkludert kontrollvariabler for blant annet alder, kjønn, fagforeninger, utdanning, yrke og boligpris. Siden mitt datasett ikke inkluderer alle disse variablene, må jeg konsentrere meg om de jeg har tilgjengelig.

Boligpris, $BPNY_{it}$, inkluderes for å fange opp eventuelle effekter av boligprisen i regionen. Som Bell et. al (2002) poengterer, er boligprisen en av de mest signifikante forskjellene som gjør en region mer attraktiv enn andre. Boligprisen er for eksempel mye høyere på Sørøstlandet enn i nord og gir Nord-Norge en signifikant fordel. Vermeulen og Ommeren (2009) bruker data fra Nederland og 12 EU-land, og viser i sin undersøkelse at regioner med høy ledighet blir kompensert med lavere boligpriser. Dersom vi finner en effekt av variabelen, kan vi forvente en positiv sammenheng mellom boligpris og lønn. Dette er interessant da vi på den ene siden har sterke empiriske belegg som antyder at folk i regioner med høy ledighet ikke blir kompensert med høyere lønn¹⁵. På den annen side er regionale arbeidsledighetsforskjeller, spesielt i Europa, vedvarende og justeringen via migrasjon er treg. Arbeidere vil avstå fra migrasjon til andre områder fordi en vil pådra seg høyere bokostnader. Videre vil lave boligpriser innebære en reel fordel hvis fordelene ikke er justert til regional kjøpekraft. Dette kan redusere insentivene for jobbsøk i andre regioner (Vermeulen og Ommeren 2009).

Ordinære arbeidsmarkedstiltak, OT_{it} , vil som nevnt kunne påvirke lønnsresponsen. Johansen (2002) analyserer empirisk hvordan norske regionale lønninger er påvirket av regionale sjokk i ledigheten. Sterk regional lønnsrespons gir en lav ledighetslikevekt, mens en regional lønnsstivhet impliserer at regional ledighet er standhaftig. Resultatene viser at lønnsresponsen er liten. Det argumenteres derfor for en aktiv tiltakspolitikk, og Johansen finner en insignifikant effekt av arbeidsmarkedstiltak som er lønnsdempende (Albæk et. al 1999 finner at andelen som er langtidsledige, har en signifikant positiv effekt på lønn i Norge), som implisitt reduserer antall langtidsledige. Resultatene av arbeidsmarkedstiltak er ikke entydige når vi sammenlikner med andre land som for eksempel Sverige. Flere studier¹⁶ argumenterer for at arbeidsmarkedstiltak har en tendens til å øke lønnspresset i

¹⁵Blanchflower og Oswald (1994), Card (1995).

¹⁶Se f.eks Calmfors (1995), Calmfors og Nymoen (1990).

økonomien.

De to siste variablene, $sommertemp_{it}$ og $vintertemp_{it}$, viser temperaturen for henholdsvis sommer og vinter. Slike variabler vil antyde en regions bekvemmelighet. Det kan være tenkelig at nordmenn foretrekker et varmt og behagelig klima fremfor et kaldt klima. Derfor er det også tenkelig at temperaturen (klima) påvirker arbeidsledighet og lønn. For USA er det gjort empiriske funn som støtter hypotesen at arbeidere godtar dårligere arbeidsmarkedsvilkår, hvis regionen har et godt klima (Blomquist et al., 1988).

3.3 Estimeringsmetode

Utgangspunktet for estimeringen er vanlig MKM, men i min analyse har vi et paneldata-materiale, med informasjon bestående av både tidsserier og tversnittselementer hvor vi følger samme individ (region) over tid. MKM vil gi skjeve estimatorer, blant annet fordi vi ikke kan anta at observasjonene er uavhengig fordelt over tid. Jeg vil fokusere på to estimeringsmetoder, de såkalte fixed og random effects-metodene. Under gitte antakelser oppnår vi forventningsrette og effisiente estimat.

3.3.1 Fixed effects

I min enkle modell er det lett å tenke seg at vi har et utelatt variabel problem. Pooled-MKM metoden¹⁷ forutsetter at utelatte variabler som bare varierer mellom regionene, er uavhengige av de inkluderte forklaringsvariablene¹⁸. Dette er en sterk forutsetning som ikke holder i min modellspesifikasjon. Derfor vil bruk av denne metoden gi skjeve estimatorer. Eksempler som er karakteristiske for en region, kan være næringsstrukturer, alderssammensetning, klima og andre forhold¹⁹. Det kan tenkes at slike variabler vil korrelere med tidsvarierende variabler som den regionale arbeidsledighetsraten. I likning 3.2 under, vil slike effekter fanges opp i den regionspesifikke restleddskomponenten η_i .

Løsningen på dette problemet er å transformere bort det regionspesifikke restleddet slik at det ikke lenger inngår i modellen. Fixed effects-metoden gjør dette, og det underliggende problemet med en individspesifikk restleddskomponent som er korrelert med inkluderte forklaringsvariabler, er løst. For å se hva metoden innebærer, kan vi tenke oss en modell

¹⁷Pooled-MKM er tilfeldig innsamlet data fra en stor befolkning der poenget er å samle de tilfeldige observasjonene slik at utvalgsstørrelsen øker, og vi oppnår mer presise estimatorer og statistisk mer slagkraft.

¹⁸Forutsetningen innebærer at det sammensatte restleddet $v_{it} = \eta_i + \epsilon_{it}$ ikke korrelerer med utelatte variabler slik at $E(\eta_i|x_{it}) = 0$ og $E(\epsilon_{it}|x_{it}) = 0$.

¹⁹Enkelte faktorer er ikke nøyaktig konstant, men relativt konstant over en 10-årsperiode.

med en forklaringsvariabel for hver i (region)

$$y_{it} = \alpha_1 x_{it} + \eta_i + \epsilon_{it} \quad t = 1, 2, \dots, T \quad (3.2)$$

for hver i (region) vil gjennomsnittet over tid være:

$$\bar{y}_i = \alpha_1 \bar{x}_i + \eta_i + \bar{\epsilon}_i \quad (3.3)$$

hvor $\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$, og så videre. Siden η_i ikke varierer over tid, vil den forsvinne ved å trekke 3.3 fra 3.2. Vi ender opp med:

$$\ddot{y}_{it} = \alpha_1 \ddot{x}_{it} + \ddot{\epsilon}_{it} \quad (3.4)$$

hvor $\ddot{y}_{it} = y_{it} - \bar{y}_i$ og så videre.

En forutsetning for at FE-estimatoren skal være konsistent, er at utvalget er tilfeldig trukket fra populasjonen. Ulempen med denne transformasjonen er at estimatoren bare utnytter variasjon over tid, og vi er ikke i stand til å estimere eller indentifisere effekten av eventuelle variabler som kun varierer mellom regionene. Variabler som varierer lite i tid, vil også bli upresist estimert. FE-metoden utnytter altså mindre variasjon (informasjon) i data sammenliknet med pooled MKM. I vårt datasett kan vi ikke forutsi nivået på variablene, og dette kan tolkes dit hen at kriteriet er oppfylt. Når det kommer til de resterende forutsetningene for restleddet, er de basert på Wooldridge (2006:503).

i) $E(\epsilon_{it}|x_{it}) = 0$, $i = 1, 2, \dots, J$ og $t = 1, 2, \dots, T$

Forutsetningen sier at restleddet skal ha forventningsverdi lik null, gitt forklaringsvariablene x . Når denne holder, har vi ingen seriekorrelasjon mellom ϵ_{it} og forklaringsvariablene.

ii) $R_j^2 \neq 1$ hvor R_j^2 finnes ved regresjon av x_j på x_i , $j \neq i$

Alle forklaringsvariablene endres over tid, og det er ingen perfekte lineære sammenhenger mellom forklaringsvariablene. Dette benevnes som forutsetningen om ingen perfekt multikollinearitet. Uformelt blir multikollinearitet definert som "høy" korrelasjon mellom forklaringsvariabler. Konsekvensen av multikollinearitet er altså upresise estimatorer som resulterer i for høye eller lave t -verdier, avhengig av om det rapporteres for lave eller for høye estimat. Dette fører til at vi forkaster eller beholder nullhypotesen for ofte eller for sjeldent.

Tidsforskjøvne forklaringsvariabler kan også føre til multikollinearitet. For eksempel kan vi forvente at lønn i nåværende periode delvis vil være bestemt av lønnen i forrige periode. I appendiks A.2 presenteres en korrelasjonsmatrise. De inkluderte forklaringsvariablene korrelerer i liten grad, og vi kan med trygghet anse forutsetning ii) som oppfylt.

iii) $Var(\epsilon_{it}|x_{it}, \eta_i) = \sigma_\epsilon^2$

Denne forutsetningen sier at restleddet må ha konstant varians for alle regioner i og tidsperioder t , gitt x og η . Når dette er oppfylt, omtales det idiosynkratiske restleddet som homoskedastisk. Brudd på denne forutsetningen gjør at vi får det som kalles heteroskedastisitet. I dette tilfelle vil parameterestimaterne fortsatt være konsistente, men ikke variansminimale. Konsekvensene er at vi systematisk feilberegner standardavvikene som inngår i testobservatoren som vil påvirke t- og F-testene. De vanlige MKM-formlene for varians og standardavvik er ikke lengre gyldige, og F- og LM-tester som baserer seg på konstant restleddsvariens, blir også ugyldige. En løsning er å konstruere standardavvik, samt t-, F- og LM-verdier som er gyldige under heteroskedastisitet. Disse kalles da for heteroskedastisk-robuste standardavvik. Metoden som benyttes tilegnes ofte White (1980).

I vår oppgave kan det være vanskelig å oppfylle kravet om homoskedastisitet for regioner med for eksempel ulik næringsstruktur. I fixed effects-modellen tas dette hensyn til ved å anta at slike ulikheter er konstante over tid. Vi kan således utelukke problemet som forklart ovenfor.

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

Denne antakelsen impliserer at de idiosynkratiske restleddene ikke er seriekorrelert. Det eksisterer ingen korrelasjon mellom restledd på ulike tidspunkt. Hvis restleddet i to påfølgende perioder avhenger av hverandre, har vi seriekorrelasjon. Årsaken kan være utelatte variabler som er seriekorrelerte, eller at modellen er for enkelt spesifisert. Problemet med seriekorrelasjon er at det kan være en medvirkende årsak til forventningskjevne estimater. Modellen vil ikke være velfungerende, og de vanlige variansformlene blir ikke lengre gyldige.

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

Den siste forutsetningen innebærer at restleddene ϵ og η er uavhengig på tvers av regioner og identisk normalfordelt, gitt det motsatte restleddet og forklaringsvariablene X . Forutsetningen er nødvendig ved t- og F-tester av hypoteser om regresjonskoeffisientene siden disse bygger på en normalfordeling.

Under disse forutsetningene vil FE-estimering gi forventningsrette og effisiente koeffisienter. Det viktigste vi bør merke oss er at faste forskjeller i lønnsnivået mellom regionene fanges opp i restleddet η_i , som vi transformerer bort. Inkludering av tidsdummier impliserer at effekten av aggregert nasjonal ledighet også tas hensyn til²⁰.

²⁰Anta at tidsdummen kan skrives som $\lambda_t = A + D\hat{u}_t + B(\text{relevante tidsspesifikke variabler})$ hvor A er en konstant, \hat{u}_t er logaritmen til gjennomsnittlig ledighet på tvers av regionene ved tidspunkt t og til slutt inkluderes relevante tidsspesifikke variabler. Vi kan da tenke på D som aggregert lønnskurveeffekt i Norge som kommer av sentrale forhandlinger, hvor det tas hensyn til gjennomsnittlig ledighet i forhandlingsprosessen. Fordi vi inkluderer tidsdummi vil ikke effekten av D fanges opp i regresjonslikning 3.1. Estimaterne i vår fixed effect-modell vil derfor være ulik rene tidsseriestudier av lønnskurven.

3.3.2 Random effects-estimering

Som nevnt er målet med en fixed effects-modell å eliminere det korrelerte restleddet η_i fra regresjonen, fordi det vil gi skjeve estimater. Dersom vi tenker oss at η_i *ikke* korrelerer med forklaringsvariablene, vil en eliminering av η_i resultere i ineffisiente estimater. I slike tilfeller foretrekkes en random effects-modell. Metoden blir også kalt for generalisert minste kvadraters metode (GLS). Metoden har de samme forutsetningene som i fixed effects, og i tillegg legger vi til en forutsetning:

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

Forutsetningen sier at forventningsverdien til restleddet η_i gitt forklaringsvariablene x_{it} , skal være lik null. Denne restriksjonen er lik forutsetning i) som sier at restleddet er ukorrelert med de inkluderte forklaringsvariablene. Betydningen av dette er at eventuelle utelatte variabler som bare varierer mellom enhetene (regionene), er uavhengig av de inkluderte forklaringsvariablene. Vi kan dermed innbefatte forklaringsvariabler som ikke har tidsvariasjon.

Under disse forutsetningene kan vi bruke både pooled MKM- og FE-estimering. Begge metodene vil gi forventningsrette estimater, men i dette tilfellet er pooled MKM å foretrekke siden det bruker større variasjon i datamaterialet. Vi bør likevel merke oss at da η_i er i hvert restledd i hver periode, vil v_{it} være seriekorrelert over tid²¹. Denne positive seriekorrelasjonen kan være betydelig, og fordi standardavvikene i vanlig pooled MKM ignorerer denne korrelasjonen, vil regresjonsestimatene være skjeve, samt vil vi ikke oppnå vanlig t normalfordeling. Problemet løses ved å ta i bruk random effects-estimering som transformerer det sammensatte restleddet slik at det ikke seriekorrelerer (Wooldridge 2006:491).

3.3.3 Random eller fixed effects?

Siden fixed effects tillater korrelasjon mellom η_i og forklaringsvariablene x_{it} , mens random effects ikke gjør det, anses ofte fixed effects-metoden som en bedre metode. Videre er det ofte sagt at random effects-estimering er mer passende når enhetene i utvalget kan sees på som tilfeldig valgt. En annen fordel er at det er færre parametere som skal estimeres i en random effects modell (ingen dummy variabler eller "within-" transformasjon skal utføres) og vi vil derfor spare frihetsgrader som gjør at modellen produserer mer effisiente estimat enn en fixed effects-modell.

I en random effects-modell antas det at det sammensatte restleddet v_{it} ikke må korrelere med forklaringsvariablene. Denne forutsetningen er mye strengere enn den tilsvarende i

²¹Under random effects-antakelsene er $Corr(v_{it}, v_{is}) = \sigma_a^2 / (\sigma_a^2 + \sigma_u^2)$, $t \neq s$ hvor $\sigma_a^2 = Var(\eta_i)$ og $\sigma_u^2 = Var(\epsilon_{it})$.

fixed effects-modellen og kan generelt anses som en svakhet. Hausman (1978) foreslo en formell test som kan avgjøre hvilken av de to metodene som gir best estimat. Hvis vi ikke kan forkaste testen, betyr dette i praksis at enten er estimatene i de to modellene veldig nære slik at det ikke spiller noen rolle hvilken vi bruker, eller så er utvalgsvariasjonen så stor i fixed effects-estimatet at vi ikke kan finne statistisk signifikante forskjeller. I det sistnevnte tilfellet må vi se om vi har nok informasjon i data til å produsere presise estimat. Altså, hvis vi kan forkaste Hausman testen, betyr det at den essensielle forutsetningen i random effects-modellen er feil,²² og vi bør bruke fixed effects-estimering.

I vår analyse er det grunn til å tro at fixed effects-modellering gir de mest robuste estimatene fordi det regionsspesifikke restleddet korrelerer med de inkluderte forklaringsvariablene. Bratsberg og Turunen (1996) samt Sanz-de-Galdeano og Turunen (2006) estimerer både fixed effects og random effects og konkluderer med at fixed effects-metoden er best egnet i sine studier.

3.3.4 Test av koeffisientforskjeller

Tidligere analyser kan tyde på at det eksisterer ulikheter i lønnsresponsen mellom utdanningsgruppene. Det er derfor av interesse å finne ut hvorvidt disse forskjellene er signifikante. Dette kan gjøres ved å sette regresjonslikningen til to utdanningsgrupper sammen til en likning. Metoden er hentet fra Wooldridge (2006) og kan sees nedenfor:

$$\begin{aligned} \text{logløn}[GRU]_{it} &= \beta_1 \log(U_{it}) + \beta_2 \log(\text{VakS}_{it}) + v_{1it} \\ \text{logløn}[VGS]_{it} &= \alpha_1 \log(U_{it}) + \alpha_2 \log(\text{VakS}_{it}) + v_{2it} \end{aligned}$$

Ved å subtrahere den siste likningen med den første, får vi:

$$\text{logløn}[GRU]_{it} - \text{logløn}[VGS]_{it} = (\beta_1 - \alpha_1) \log(U_{it}) + (\beta_2 - \alpha_2) \log(\text{VakS}_{it}) + (v_{1it} - v_{2it})$$

På kompakt form kan dette skrives:

$$\text{logløn}_{it}^* = \theta_1 \log(U_{it}) + \theta_2 \log(\text{VakS}_{it}) + v_{it}^*$$

hvor $\text{logløn}_{it}^* = \text{logløn}[GRU] - \text{logløn}[VGS]$, $\theta_1 = (\beta_1 - \alpha_1)$, $\theta_2 = (\beta_2 - \alpha_2)$ og $v^* = (v_{1it} - v_{2it})$. Vi kan nå estimere θ_1 og θ_2 . T-verdiene til disse koeffisientene vil fortelle oss hvorvidt det er en signifikant forskjell mellom β_1 og α_1 samt β_2 og α_2 . Vi kan også teste hypoteser om ingen ulikheter, det vil si $\theta_1 = 0$ og $\theta_2 = 0$.

²²Anta vi bare har en forklaringsvariabel x_{it} som varierer positivt både med y_{it} og restleddet v_{it} . Estimatoren vil tilskrive enhver økning i y til x når det i virkeligheten også oppstår av restleddet v_{it} , som resulterer i skjeve koeffisienter.

4 Empiriske resultater

4.1 Innledning

I denne delen av oppgaven presenteres mine egne empiriske resultater hvor hovedmålet er å avdekke hvorvidt det eksisterer ulike lønnskurver for de ulike utdanningsgruppene. Vi starter med å se på den generelle modellen, som omtalt i kapittel 3, gitt som:

$$\begin{aligned} \log\text{lønn}[X]_{it} = & \alpha_1 \log(W[X]_{it-1}) + \alpha_2 \log(U_{it}) + \alpha_3 \log(U_{it-1}) + \alpha_4 \log(VakS_{it}) + \\ & \alpha_5 \log(VakS_{it-1}) + \alpha_6 \log(\text{Andre variabler})_{it} + \lambda_t + v_{it} \end{aligned} \quad (4.1)$$

Vi presenterer to ulike estimat, henholdsvis fixed effects, som fjerner regionsspesifikke effekter, og random effects, som antar ingen seriekorrelasjon mellom restleddet og forklaringsvariablene. Til å begynne med inkluderes bare variablene W_{it-1} , U_{it} , U_{it-1} , $VakS_{it}$ og $VakS_{it-1}$. Hensikten er å se hvordan disse variablene påvirker lønn for utdanningsgruppene i ulik grad.

Det fremgår av diskusjonen over at fixed effects høyst sannsynlig gir de beste estimatene i denne analysen, siden den kontrollerer for eventuell korrelasjon mellom inkluderte forklaringsvariabler og utelatte eller uobserverbare regionale forhold som ikke varierer over tid. Tidsdummyer er inkludert i modellen for samtlige år, og derav kan det kontrolleres for aggregerte variabler som påvirker tilpasningen over tid. Dummyene vil derfor fange opp nasjonale lønnsforhandlinger slik at vår modell ser på effekten av regionale lønnsforhandlinger. Tidsdummyene vil ikke rapporteres i tabellene.

I delkapittel 4.2 rapporteres de estimerte koeffisientene og resultatene tolkes. I delkapittel 4.3 gis det en oversikt over diagnostiske tester. I siste delkapittel foretas en sensitivitetsanalyse der ulike funksjonsformer av basismodellen fra delkapittel 4.1 reestimeres.

4.2 Resultater fra fixed- og random effects-estimering

4.2.1 Grunnmodell

De estimerte koeffisientene for variablene i grunnmodellen rapporteres i tabell 3. Resultater for både fixed- og random effects-regresjonene er inkludert.

Tabell 3: Effekt av logaritmen til regional arbeidsledighet og vakanseratestrøm på lønn for ulike utdanningsgrupper. Avhengig variabel $\log\text{lønn}[X]_{it}$.

Variabler	Gruppe					
	GRU		VGS		HØY	
	FE(I)	RE	FE(I)	RE	FE(I)	RE
$\log(W_{it-1})$	0,773 (28,7)	0,988 (157)	0,757 (31,3)	1,009 (275)	0,76 (30)	1,024 (389)
$\log(U_{it})$	-0,026 (-8,95)	-0,024 (-9,90)	-0,02 (-10,1)	-0,019 (-11,7)	-0,011 (-4,90)	-0,012 (-6,85)
$\log(U_{it-1})$	0,012 (2,89)	0,023 (9,52)	0,011 (4,76)	0,019 (11,5)	0,006 (2,46)	0,011 (6,21)
$\log(\text{Vak}S_{it})$	-0,001 (-0,69)	-0,001 (-0,5)	-0 (-0,02)	0,001 (0,52)	0,003 (2,06)	0,003 (2,22)
$\log(\text{Vak}S_{it-1})$	0,002 (0,8)	0,003 (1,3)	-0,002 (-1,00)	-0 (-0,30)	-0,002 (-1,38)	-0,003 (-1,77)
Oservasjoner	810	810	810	810	810	810
R^2	0,998	0,998	0,999	0,999	0,999	0,999
σ	0,007	0,008	0,005	0,005	0,005	0,006
Wald (joint)	912,7**	26030**	1499**	79110**	926,7**	154300**
AR(1)-test	-0,633	-0,243	-1,601	-0,17	-0,098	3,197**
AR(2)-test	-1,398	0,215	-2,79**	-2,004*	-2,642**	-0,86
	Langtidsledighetskoeffisient					
	-0,062 (-4,19)	-0,064	-0,038 (-5,12)	-0,018	-0,021 (-3,22)	-0,025

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 ved 1% signifikansnivå er angitt med **, samt ved 5% signifikansnivå med*.

Resultatene viser at fortegnene på forklaringsvariablene er stort sett som forventet. *Den laggede lønnsvariabelen* har en positiv effekt for alle utdanningsgruppene i begge modellene. Variabelen er høyst signifikant, spesielt i RE-modellen. Koeffisientestimatene er noe høyere ved bruk av random effects-estimering. Som BO poengterer, vil en verdi nær 1 antyde en Phillipskurve. I RE-modellen får vi en verdi på 1, og vi kunne omskrive likningen slik at venstreside blir på endringsform. På den annen side kan vi med god margin forkaste $H_0: \alpha_1 = 1$ i FE(I) modellen for samtlige grupper, og vi kan fastslå at det finnes en langtidssammenheng mellom lønn og ledighetsraten²³.

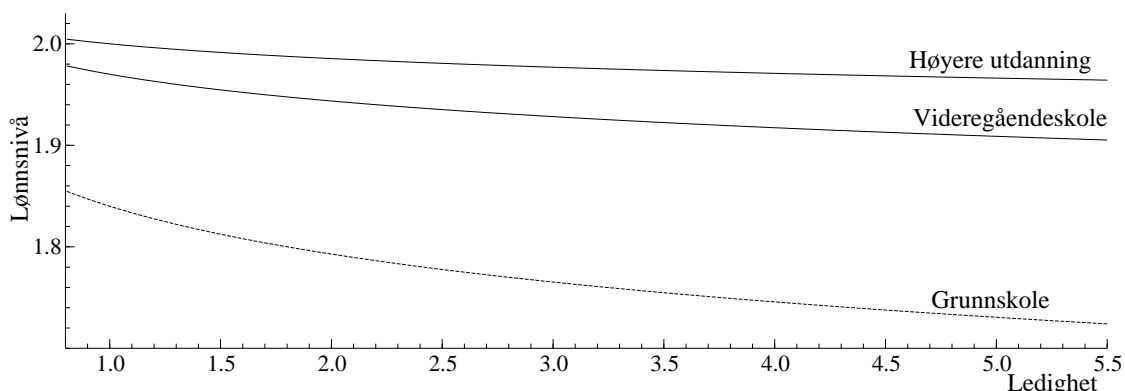
²³Som Blanchflower og Oswald (2005) poengterer antyder en høy verdi av lagget lønn en betydelig persistens i lønn slik at Phillips-liknende effekter i regional lønn eksisterer. Albæk et al. (2000) finner liknende effekter i sin studie av nordiske lønnskurver.

Den regionale arbeidsledigheten har en negativ effekt for alle utdanningsgruppene og er signifikant ved 1 prosentnivå, siden t-verdiene er over 3 i absoluttverdi. Vi antyder en noe svakere effekt av ledighet for gruppen høyere utdanning enn for grunnskolegruppen. Når det kommer til *den laggede ledighetsvariabelen*, er denne positiv som betyr at økt ledighet i forrige periode faktisk øker lønnen i nåværende periode. Variabelen har en signifikant betydning som sees av t-verdiene. Til slutt er *den langsiktige ledighetskoeffisienten* rapportert. Vi ser en klar sammenheng mellom utdanningsnivå og lønnsresponsen på arbeidsledighet. Langtidsresponsen er totaleffekten av ledighet på lønn og består av både den direkte effekten og den indirekte effekten. Effekten finnes ved å kalkulere $(\alpha_2 + \alpha_3)/(1 - \alpha_1)$. Lønnskurven er -0,062, -0,038 og -0,021 for henholdsvis grunnskole, videregående skole og høyere utdanning og kan sees i figur 4.1. I tråd med tidligere studier finner vi at lavt utdannede har en sterkere lønnsrespons enn høyt utdannede ved endring i ledigheten.

Våre to siste forklaringsvariabler, *vakanseraten* i nåværende og forrige periode, er insignifikante for grunnskole- og videregående skolegruppene. De vil derfor ikke påvirke lønn av noen betydning. For høyere utdanning er variabelen signifikant ved 5 prosentnivå. Det betyr at ledige stillinger kun har signifikant påvirkning på lønnen til de med høyere utdanning. Så langt ser vi en motsatt effekt mellom gruppene. Arbeidsledighet har en større betydning på lønnen for gruppene med lavest utdanning, mens vakanseraten bare har en effekt hos høyere utdanning. Dette kan forklares med gruppenes ulike migrasjonsmønster, hvor gruppen høyere utdanning er klart mer mobil enn for eksempel grunnskolegruppen²⁴. Derfor vil flere ledige stillinger ha større innvirkning på lønnen for gruppen med høyest migrasjon, altså høyere utdanning. Generelt er estimatene for RE-modellene høyere enn for FE-modellene.

Hvis vi konsentrerer oss om arbeidsledigheten, ser vi at effekten av økt arbeidsledighet antyder at grunnskolegruppen responderer sterkere enn de andre to gruppene. Vi kan illustrere dette ved en forverring i arbeidsmarkedet ved å øke ledighetsraten med ett standardavvik fra gjennomsnittet.

Figur 4.1: Lønnskurvene for grunnskole, videregående skole og høyere utdanning



²⁴Se Carlsen og Johansen (2004).

Estimatene fra FE(I) gir da en reduksjon i lønnen for grunnskolegruppen med 0,02 prosentpoeng. For videregående og høyere utdanning er responsen endel lavere med verdier på henholdsvis -0,012 og -0,007. For å finne en like sterk respons hos høyere utdanning som hos grunnskole må vi øke ledighetsraten med to standardavvik fra gjennomsnittet.

En økning av arbeidsledigheten med ett standardavvik tilsvarer en økning av arbeidsledigheten med 39 prosent. Dette utgjør en nedgang i lønn for henholdsvis grunnskole, videregående og høyere utdanning med 2,4, 1,5 og 0,8 prosent. Ved å bruke gjennomsnittslønnen i tabell 1 ser vi at økt arbeidsledighet på ett standardavvik utgjør en reduksjon på 74,6 kroner for grunnskolegruppen, 51,5 kroner for videregående skolegruppen og 35,7 kroner for høyere utdanningsgruppen.

Vi kan videre undersøke i hvor stor grad standardavvikene til gruppenes lønn som kan forklares av den regionale arbeidsledigheten. Å øke ledigheten med ett standardavvik fra gjennomsnittet vil for grunnskolegruppen resultere i en nedgang i gruppens lønn tilsvarende 12 prosent av lønnens standardavvik. For videregående- og høyere utdanningsgruppene forklarer ledighetsraten enda mindre variasjon, da ett standardavviks økning utgjør henholdsvis 6,5 og 3,5 prosent. Som vi ser, kan ikke den regionale arbeidsledigheten forklare mye av variasjonen til gruppenes lønn. Her vil nok andre variabler kunne forklare spredningen bedre, som for eksempel yrke, ansiennitet, alder og så videre. Likevel er det interessant å påpeke at arbeidsledighet har større forklaringskraft i lønn[GRU]s variasjon enn i lønn[HØY]s variasjon.

Resultatene tyder altså på at lønnsresponsen på arbeidsledigheten er ulik i de tre ulike utdanningsgruppene hvor grunnskole responderer mest og høyere utdanning minst. Hvorvidt disse ulikhetene er signifikant forskjellige, kan vi teste ved å bruke fremgangsmåten beskrevet i delkapittel 3.3.4. I korte trekk går metoden ut på å parvis sammenlikne gruppene ved å sette regresjonslikningene til to grupper sammen til en likning. Vi utfører en ny estimering som gir nye t-verdier. Disse t-verdiene forteller oss om forklaringsvariablene er signifikant forskjellige mellom gruppene. En t-verdi større eller lik 1,96 betyr signifikant forskjell ved 5 prosentnivå, og en t-verdi større eller lik 1,282 betyr signifikant forskjell ved 20 prosentnivå. Resultatene fra estimeringen er rapportert i tabell 4, og er basert på langtidsløsningene.

Tabell 4: Gruppesammenlikning av estimerte koeffisientverdier fra FE(I)-modellene. Avhengig variabel: $\log\text{lønn}[X_1]_{it} - \log\text{lønn}[X_2]_{it}$, der X_1 og X_2 står for utdanningsgrupper.

Variabler	Gruppesammenlikning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(I)	D-FE(I)	D-FE(I)
$\log(U)$	-0,0191 (-2,43)	0,4488 (3,33)	0,3355 (3,77)
$\log(VakS)$	0,0071 (0,93)	0,0509 (0,58)	0,1122 (1,57)

I tabell 4 har vi testet om endret ledighet har signifikant langsiktig effekt på relativ lønn mellom gruppene. Vi ser at sammenlikningen mellom alle gruppene gir signifikante forskjeller i effekten av ledighet. Resultatene fra D-FE(I) viser dermed at gruppene responderer ulikt på ledighetsraten. Som vi forventet, er ikke effekten av vakanseraten signifikant ulik mellom gruppene.

Funnene insinuerer at det finnes ulike lønnskurver for ulike utdanningsgrupper i Norge. Empirien antyder at lønnen øker i utdanningsnivået, mens lønnsresponsen med hensyn til ledighet avtar. Mer utdanning har derfor dobbel positiv effekt: det øker inntekten relativt til antall år i skolen, og en blir skjermet mot lønnsreduksjon i perioden med økt ledighet. Til sammenlikning har lavt utdannede generelt lavere lønn og vil være utsatt for ytterligere lønnsreduksjon i perioder hvor arbeidsledigheten øker.

I fortsettelsen av dette delkapitlet vil vi estimere basismodellen ved bruk av den gruppebestemte ledighetsraten som forklaringvariabel og kan undersøke om lønnsresponsen mellom utdanningsgruppene er ulik. Vi kan forvente at estimatet for grunnskole- og høyere utdanningsgruppen får lavere lønnsrepons med hensyn til ledighet, på bakgrunn av diskusjonen i kapittel 3.2. Den nye regresjonen tar følgende form:

$$\begin{aligned} \log l\ddot{o}nn[X]_{it} = & \alpha_1 \log(W[X]_{it-1}) + \alpha_2 \log(U[X]_{it}) + \alpha_3 \log(U[X]_{it-1}) + \alpha_4 \log(VakS_{it}) + \\ & \alpha_5 \log(VakS_{it-1}) + \alpha_6 \log(Andre\ variabler)_{it} + \lambda_t + v_{it} \end{aligned} \quad (4.2)$$

hvor $[X]$ angir gruppe. Estimaten er rapportert i tabell 5.

Resultatene er ikke veldig forskjellig fra tabell 3. Den største forskjellen er at estimatet av langtidsledighetskoeffisienten blir generelt lavere i samtlige grupper. I grunnskolegruppen ser vi nesten en halvering av estimatet til -0,032. Endringen i videregående skolegruppen er ikke like stor. Her faller estimatet til -0,03. Den største endringen finner vi i høyere utdanningsgruppen hvor estimatet blir tre ganger så lavt og får en verdi på -0,007. T-verdien er -1,37 som betyr at lønnskurven for høyt utdannede ikke er signifikant ved 5% nivå. Disse resultatene er som forventet fordi gruppene korrelerer ulikt med aggregert regional ledighet, hvor høyere utdanning korrelerer minst. Siden vi nå bruker gruppespesifikk arbeidsledighet, kan vi ikke sammenlikne og teste om gruppene responderer ulikt på ledigheten basert på metoden i delkapittel 3.3.4²⁵. Likevel ser vi en numerisk likhet mellom grunnskole- og videregående skolegruppene, mens høyere utdanningsgruppen skiller seg ut. Det kan også utføres en t-test av tallverdier som gir en indikasjon om det finnes en

²⁵En formell test er $t = \frac{\hat{\beta}_1 - \hat{\beta}_2}{se(\hat{\beta}_1 - \hat{\beta}_2)}$ hvor $se(\hat{\beta}_1 - \hat{\beta}_2) = \left\{ [se(\hat{\beta}_1)]^2 + [se(\hat{\beta}_2)]^2 - 2Cov(\hat{\beta}_1, \hat{\beta}_2) \right\}^{1/2}$. Programvaren som er benyttet gjør det vanskelig å utføre en slik test ved bruk av fixed effects-estimering.

lineær sammenheng mellom estimatene²⁶. Vi finner da en signifikant sammenheng mellom grunnskole og videregående skole, mens gruppen for høyere utdanning er signifikant ulik de andre gruppene.

Tabell 5: Effekt av logaritmen til gruppebestemt regional arbeidsledighet og vakanseratestrøm på lønn for ulike utdanningsgrupper. Avhengig variabel $\log\text{l\o}nn[X]_{it}$.

Variabler	Gruppe					
	GRU		VGS		HØY	
	FE	RE	FE	RE	FE	RE
$\log(W_{it-1})$	0,759 (27,1)	0,985 (159)	0,731 (27,1)	1,009 (282)	0,756 (32)	1,025 (377)
$\log(U_{it})$	-0,01 (-4,44)	-0,007 (-4,72)	-0,011 (-6,87)	-0,01 (-6,91)	-0,003 (-2,95)	-0,0026 (-3,35)
$\log(U_{it-1})$	0,002 (0,781)	-0,007 (-4,53)	0,003 (1,91)	(0,010) (7,18)	0,001 (0,823)	0,002 (2,59)
$\log(\text{Vak}S_{it})$	0,001 (0,509)	0,0004 (0,198)	0,001 (0,925)	0,002 (1,3)	0,004 (2,68)	0,004 (2,58)
$\log(\text{Vak}S_{it-1})$	0,002 (0,755)	0,001 (0,468)	-0,002 (-1,05)	-0,002 (-1,15)	-0,002 (-1,47)	-0,003 (-2,29)
Oservasjoner	810	810	810	810	810	810
R^2	0,998	0,997	0,999	0,998	0,999	0,999
σ	0,008	0,003	0,005	0,005	0,005	0,006
Wald (joint)	894,9**	2593**	1329**	8147**	1067**	14940**
AR(1)-test	0,775	0,952	-1,091	0,208	0,826	4,169**
AR(2)-test	-2,197*	-1,326	-3,161**	-2,677**	-2,755**	-0,983
	Langtidsledighetskoeffisient					
	-0,0319 (-2,88)	-0,0139	-0,0302 (-5,17)	-0,0540	-0,007 (-1,37)	-0,0184

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 ved 1% signifikansnivå er angitt med **, samt ved 5% signifikansnivå med*.

En forverring i arbeidsmarkedet ved å øke den gruppespesifikke ledighetsraten med ett standardavvik fra gjennomsnittet gir nå en reduksjon på 0,01, 0,01 og 0,002 prosentpoeng for henholdsvis grunnskole, videregående skole og høyere utdanning. Som vi ser, vil en økning med ett standardavvik fra gjennomsnittet forklare enda mindre i gruppenes lønnsvariasjon. For grunnskolegruppen vil dette tilsvare 6 prosent av lønnens standardavvik, mens for gruppene videregående og høyere utdanning forklarer en økning av ledigheten med ett standardavvik henholdsvis 5,5 og 1 prosent. Til sammenlikning med aggregert

²⁶Testen er $t = \frac{\hat{\beta} - a_j}{se(\hat{\beta})}$ hvor a_j er en kjent tallverdi. Vi tester en lineær sammenheng mellom estimatet, β og ønsket tallverdi a_j .

regional ledighet har forklaringskraften minnet betraktelig for gruppene grunnskole og høgere utdanning.

4.2.2 Utvidet modell

I dette delkapitlet ser vi på en utvidet modell hvor vi inkluderer forklaringsvariablene $BPNY_{it}$, OT_{it} , samt klimavariablene $sommertemp_{it}$ og $vintertemp_{it}$. Estimatenes er gjort ved bruk av aggregert regional ledighet. Som resultatene viser, ser det ut som estimatene endres lite. Det ble også estimert effekt av lagget boligpris og ordinære tiltak, men verdiene var insignifikante og vil derfor utelates. Effekten av den regionale arbeidsledigheten er negativ, og den kortsiktige elastisiteten er noe lavere enn estimatene i basismodellen. Vi ser en endring på henholdsvis 7 og 9 prosent for grunnskole- og høyere utdanningsgruppen, mens for videregående er nedgangen 5 prosent. Resultater fra fixed- og random effects-estimeringen er rapportert i tabell 6.

Tabell 6: Estimer fra regresjonsmodeller som inkluderer boligpris, ordinære tiltak, sommer og vinter-temperatur.

Variabler	Gruppe					
	GRU		VGS		HØY	
	FE(II)	RE	FE(II)	RE	FE(II)	RE
$\log(W_{it-1})$	0,77 (29,7)	0,987 (110)	0,744 (30,2)	1,004 (188)	0,762 (31,9)	1,019 (241)
$\log(U_{it})$	-0,024 (-8,87)	-0,023 (-9,41)	-0,019 (-9,84)	-0,019 (-11)	-0,01 (-4,78)	-0,011 (-6,25)
$\log(U_{it-1})$	0,011 (2,7)	0,023 (9,44)	0,011 (4,74)	0,019 (11,4)	0,006 (2,43)	0,011 (6,19)
$\log(VakS_{it})$	-0,00003 (-0,019)	-0,001 (-0,45)	-0,0005 (-0,035)	0,001 (0,642)	0,003 (2,36)	0,003 (2,31)
$\log(VakS_{it-1})$	0,002 (1,11)	0,003 (1,32)	-0,001 (-0,92)	-0,0002 (-0,145)	-0,002 (-1,34)	-0,002 (-1,66)
$\log BPNY_{it}$	-0,001 (-0,45)	0,0004 (0,27)	0,004 (2,31)	0,001 (0,841)	0,0002 (0,109)	0,001 (1,23)
OT_{it}	-0,003 (-1,83)	-0,001 (-0,405)	-0,002 (-1,29)	-0,001 (-0,649)	-0,001 (-0,863)	-0,001 (-1,16)
$sommertemp_{it}$	-0,0001 (-0,504)	-0,0007 (-0,417)	0,0005 (2,74)	0,0002 (1,46)	-0,0002 (-1,07)	0,0006 (0,486)
$vintertemp_{it}$	-0,0007 (-3,28)	-0,0001 (-1,21)	-0,0004 (-3,02)	-0,0008 (-1,29)	-0,0002 (-0,973)	-0,0001 (-2,07)
Oservasjoner	810	810	810	810	810	810
R^2	0,998	0,998	0,999	0,999	0,999	0,999
σ	0,007	0,008	0,005	0,005	0,005	0,005
Wald (joint)	1122**	25540**	1686**	1506**	1235**	158600**
AR(1)-test	-0,834	-0,21	-1,547	0,039	-0,115	3,326**
AR(2)-test	-1,673	0,213	-2,572*	-1,779	-2,67**	-0,772
	Langtidsledighetskoeffisient					
	-0,057 (-3,90)	-0,030	-0,033 (-4,22)	-0,066	-0,020 (-3,05)	-0,016

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 ved 1% signifikansnivå er angitt med **, samt ved 5% signifikansnivå med*.

I likhet med basismodellen er det ingen store forskjeller mellom RE- og FE-estimatene. De gir samme fortegn og spriker ikke mye i verdi. Bare den laggede lønnsvariabelen skiller seg ut, i likhet med basismodellen, med en del høyere verdier i RE-modellen. Vi kan også nå forkaste Phillipskurven i FE(II) for samtlige grupper. Estimaten for ledighet og vakanseraten er nesten uendret i forhold til basismodellen. Vi kan se at vakanseraten nå får negativ verdi både for grunnskole- og videregåendegruppen, men er meget insignifikant med t-verdier nær null. Effekten for videregående får en litt høyere t-verdi nå enn i grunnmodellen.

For boligprisen viser FE(II)-modellen et negativt koeffisientestimat for grunnskolegruppen. Resultatet er ikke signifikant, men vil bety at økt boligpris vil gi insignifikant nedgang i lønn. Resultatene til de to andre gruppene er som forventet. Økt boligpris gir en økt lønnsrespons. Verdiene er henholdsvis 0,004 og 0,0002 for de videregående- og høyere utdanningsgruppene hvor estimatet for videregående er signifikant med en t-verdi på 2,31. Videregående-skolegruppen opplever en signifikant økning i lønnen ved en økning i boligprisen og blir derfor kompensert. En region med høy ledighet som opplever økt boligpris, vil bli kompensert med økt lønn med unntak av grunnskolegruppen. Videre vet vi at denne gruppen har lavest gjennomsnittlig lønn, og er gruppen med høyest gjennomsnittlig arbeidsledighet i regionen. Begge disse faktorene peker i retning av lavere kjøpekraft. De to andre gruppene opplever derimot økt lønn, som kan presse boligprisen enda høyere, noe som forverrer situasjonen til grunnskolegruppen. Vi kan merke oss at vi kan ha et mulig simultanitetsproblem. Boligprisen blir åpenbart påvirket av folks kjøpekraft, og økt lønn vil forsterke kjøpekraften.

For ordinære tiltak viser FE(II)-estimeringen negative koeffisientverdier for alle gruppene. Grunnskolegruppens koeffisientverdi er tre ganger så stor som for gruppen høyere utdanning og halvannen gang så stor som koeffisientverdien for gruppen videregående skole. T-verdiene er relativt lave og bare grunnskole- og videregående-skolegruppens verdier er signifikante på henholdsvis 10 og 20 prosentnivå. Vi finner altså en svak lønnsdempende effekt.

Til slutt har vi utvidet modellen med sommer- og vintertemperaturer som forklaringsvariabler. Alle koeffisientene har negative fortegn med unntak av sommertemperatur for videregående-skolegruppen i FE(II)-modellen som også er den eneste signifikante verdien av sommertemperatur med en t-verdi på 2,74. Estimaten for vintertemperatur er signifikante bortsett fra i høyere utdanningsgruppen.

De nye inkluderte variablene påvirker effekten av den kortsiktige arbeidsledigheten, og den får svakere forklaringskraft. Dette gjenspeiles i den langsiktige arbeidsledighetselastisiteten. Vi ser fortsatt at lønnsresponsen på arbeidsledighet er sterkest hos grunnskolegruppen med en verdi på -0,057. Videregående-skolegruppens lønnskurve faller med 13 prosent og

får en verdi på -0,033. Endringen for høyere utdannede er mindre, og lønnsresponsen faller med 4 prosent til -0,02. I tabell 7 har vi igjen sammenliknet gruppene mot hverandre for å teste hvorvidt de estimerte koeffisientene gir signifikante forskjeller. Prosedyren er den samme som i delkapittel 3.3.4.

Tabell 7: Gruppesammenlikning av estimerte koeffisientverdier fra FE(II)-modellene. Avhengig variabel: $\log\text{lønn}[X_1]_{it} - \log\text{lønn}[X_2]_{it}$, der X_1 og X_2 står for utdanningsgrupper.

Variabler	Gruppesammenlikning		
	GRU vs VGS	GRU vs HØY	VGS vs HØY
	D-FE(II)	D-FE(II)	D-FE(II)
$\log(U)$	-0,019 (-2,65)	0,473 (3,35)	0,362 (3,87)
$\log(VakS)$	0,013 (1,63)	0,040 (0,44)	0,109 (1,48)
$\log(BPNY)$	-0,190 (-2,19)	-0,0189 (-0,12)	0,08 (0,80)
OT	-0,006 (-0,54)	-0,070 (-0,97)	-0,049 (-0,92)
$sommertemp$	-0,002 (-2,34)	0,021 (1,64)	0,009 (0,94)
$vintertemp$	-0,0009 (-1,41)	-0,0005 (-0,05)	-0,003 (-0,39)

De langsiktige arbeidsledighetselastisitetene er fortsatt signifikant ulike mellom gruppene. I denne modellspesifikasjonen ble også vakanseraten signifikant ulik mellom videregående og høyere utdanning. Øvrige variabler har derimot ikke signifikante forskjellige koeffisientestimater. Bare boligpris og sommertemperatur er signifikant ulik i sammenlikningen mellom grunnskole- og videregåendegruppen ved 5 prosentnivå²⁷.

Resultatene i fixed og random effects gir delvis ulike resultat spesielt med tanke på den laggede lønnsvariabelen. Dette er en god indikasjon på at random effects-estimering gir skjeve resultat. Vi har utført en Hausman-test²⁸ og resultatene er rapportert i appendiks A.3. Nullhypotesen er at begge estimeringsmetodene er akseptable og gir omtrent samme koeffisientverdier. Alternativhypotesen er at fixed effects-estimeringen er den beste, og at random effects gir skjeve resultat. Hvis dette er tilfelle, bør vi forvente en signifikant differanse mellom de to metodene. Som vi kan se av resultatene, finner vi en stor og signifikant Hausman-verdi, som betyr vi kan forkaste nullhypotesen med god margin. Her, i likhet med andre studier, finner vi at fixed effects estimeringen gir de beste resultatene.

²⁷En utvidet modell ble forsøkt estimert med gruppespesifikk ledighet. Resultatene ga relativt små utslag i langtidsløsningene sammenliknet med basismodellen.

²⁸Testen er gjort i STATA IC 12

4.3 Diagnostiske tester

Sigma, σ , er regresjonens standardavvik og er relativt lik for FE- og RE-modellene. Størst variasjon finner vi for grunnskolegruppen med en verdi på rundt 0,008. Verdiene for videregående- og høyere utdanningsgruppene ligger relativt stabilt på 0,005. Tolkningen er at lønnen til grunnskolegruppen påvirkes av flere (utelatte) variabler som gir utslag i høyere standardavvik. En mulig forklaring kan ligge i gruppenes ulike migrasjonsrate. For eksempel kan migrasjonsraten mellom gruppene påvirke lønnen, hvor grunnskolegruppen er mer stasjonær enn høyere utdanningsgruppen.

Wald-testen er motstykket til en F-test ved MKM-estimering der H_0 er at de uavhengige variablene ikke har forklaringskraft. Vi har oppgitt Wald(joint) som kontrollerer for tidsdummyer og fixed effects, og tester signifikansen for de økonomiske variablene. Nullhypotesen er forkastet ved 1 prosents signifikansnivå i samtlige modeller. Wald(time) er ikke rapportert, men viser at tidsdummyene er signifikante ved 1 prosents signifikansnivå i alle modeller.

Vår R^2 verdi er meget høy og tilnærmet lik 1 for alle gruppene. Grunnen til dette er antakeligvis fordi vi bruker tidsdummyer og fixed effects med laggede verdier. I utgangspunktet er R^2 et mål på hvor mye av variasjonen i avhengig variabel som forklares av modellens inkluderte høyresidevariabler. Dette betyr at modellens forklaringskraft ligger på 99 prosent for gruppene.

AR(1) og AR(2) er tester for første- og andreordens seriekorrelasjon i restleddet. Under forutsetning (iv) antar vi at de idiosynkratiske restleddene ikke er seriekorrelert.

$$v_{it} = \rho_1 v_{it-1} + \rho_2 v_{it-2} + \eta_{it}$$

hvor v_{it} er normalfordelt med forventningsverdi lik null og konstant varians. Vår hypotese blir da

$$\rho_1 = \rho_2 = 0$$

I våre resultat har vi ingen førsteordens seriekorrelasjon i noen av gruppene. For andreordens seriekorrelasjon forkaster vi nullhypotesen ved 1 prosents signifikansnivå for videregående- og høyere utdanningsgruppen. Dette kan muligens gi opphav til en viss grad av seriekorrelasjon i restleddet som vil forårsake skjeve resultat. Grunnen til seriekorrelasjon kan skyldes utelatte variabler som varierer i både tverrsnitts- og tidsdimensjonen.

Test for parameterstabilitet er ikke oppgitt i tabellen, men undersøkes her. Parameterstabilitet tester om estimatene er stabile over tidsperioden. Ved bruk av Chow test deler vi utvalget inn i to tidsperioder, en periode med data for 1995 - 1999 og en periode med data for 2000 - 2004. Vi estimerer så disse utvalgene og finner SSR (sum of squared residuals). Selve testen er formulert på følgende måte:

$$F = \frac{[SSR_p - (SSR_1 + SSR_2)]}{SSR_1 + SSR_2} \times \frac{[n - 2(k + 1)]}{k + 1}$$

hvor SSR_1 er for tidsperiode 1995 - 1999, SSR_2 er for tidsperiode 2000 - 2004 og SSR_p er for hele tidsperioden 1995 - 2004, n er antall observasjoner mens k er antall forklaringsvariabler. H_0 er «ingen forskjeller mellom tidsperiodene». F-verdien for våre grupper gir $F_{GRU} = 4,55$, $F_{VGS} = 3,64$ og $F_{HØY} = 3,12$. På grunn av mange frihetsgrader kan vi forkaste H_0 og vi har *ikke* stabile parameter over utvalgsperioden. Det ble også utført test ved bruk av gruppespesifikk ledighet, men også her forkaster vi parameterstabilitet. Resultatene kan være tilfeldig fordi i grunnskolegruppen påvirker lagget lønn mer i siste delperiode (økt persistens), mens for videregående- og høyere utdanningsgruppene er den estimerte verdien på lagget lønn størst i første delperiode.

Til slutt er det testet for normalfordelte restledd. H_0 er «restleddet er normalfordelt». Kvikvadratverdiene for gruppene gir $\chi_{GRU}^2 = 5,8$, $\chi_{VGS}^2 = 7,5$ og $\chi_{HØY}^2 = 0,2$. Dette betyr vi kan forkaste H_0 for videregående skolegruppen ved 5 prosent signifikansnivå. Dette kan påvirke t- og F-tester, og antakelse v) om normalfordelte restledd brytes.

4.4 Oppsummering

I dette delavsnittet har vi undersøkt sammenhengen mellom lønn og arbeidsledighet i tre utdanningsgrupper. Vi har først estimert en grunnmodell med log lønn som venstresidevariabel og lagget lønn, aggregert regional ledighet, lagget aggregert regional ledighet, vakanserate og lagget vakanserate som høyresidevariabel. Vi har estimert både fixed og random effects-modeller, og det viser seg at fixed effects-modellene gir de beste resultatene fordi de tillater korrelasjon mellom forklaringsvariablene og tidsinvariante regionspesifikke effekter. Også vår formelle Hausman-test tilsier at fixed effects gir de beste estimatene.

Resultatene viser en negativ langsiktig lønnselastisitet som er signifikant for samtlige grupper. Effekten er størst for grunnskolegruppen og minst for høyere utdanningsgruppen. Estimaten av vakanseraten gir insignifikante verdier for grunnskole- og videregående skolegruppen, men signifikant positive verdier for høyere utdanning. Noe av forklaringen ligger i gruppenes ulike migrasjon. Det er også testet for langsiktige koeffisientforskjeller mellom gruppene, og vi finner signifikante forskjeller på effekten av ledighet på lønn. Grunnskolegruppen reagerer 3 ganger kraftigere på endret ledighet enn høyere utdanningsgruppen. For vakanseraten finner vi ingen signifikant forskjell mellom gruppene.

Basismodellen ble så estimert med gruppespesifikk arbeidsledighet som forklaringsvariabel. Langtidsledighetskoeffisientene for høyere utdanning fikk størst reduksjon med 66,7 prosent. For grunnskole- og videregående gruppen var reduksjonen på henholdsvis 48,3 og 21 prosent. Årsaken til dette er at gjennomsnittsledigheten for høyere utdanning korrelerer mindre med det aggregerte gjennomsnittet og vil derfor feilestimere effekten av ledighet

på lønn. Den langsiktige lønnselastisiteten for høyere utdanning ble også insignifikant ved 5 prosentsnivå. Det er heller ikke grunn til å tro det eksisterer en signifikant forskjell mellom grunnskole og videregående skole.

I en utvidet modell ble det inkludert logaritmen av boligpris, ordinære tiltak, sommer- og vintertemperatur. Boligpris har en positiv signifikant effekt på lønn for videregåendeskelegruppen, mens estimatene var insignifikante for de to andre gruppene. Ordinære tiltak har en negativ effekt på lønn, men er insignifikant ved 5 prosentsnivå for alle gruppene. Sommertemperatur har en signifikant positiv effekt på lønn for videregående gruppen, mens estimatet gir insignifikante negative verdier for de to andre gruppene. Vintertemperatur har en signifikant negativ effekt på lønn for grunnskole- og videregåendeskelegruppene, mens verdien for høyere utdanning er insignifikant. Hovedfunnet er at langtidsledighetskoeffisientene nå får en lavere verdi og er signifikant forskjellig mellom gruppene.

4.5 Andre funksjonsformer

I likhet med BO og flere²⁹ vil jeg eksperimentere med ulike funksjonsformer av lønnskurven for å undersøke om lønnskurven er mer eller mindre ikke-lineær enn hva den logaritmiske basisspesifikasjonen innebærer. Resultatene til BO ga en robust, negativ ikke-lineær sammenheng mellom ledighet og lønn (Blanchard og Oswald 1994:102). Videre konkluderer de at log-funksjonsformen sannsynligvis er den best egnete på grunn av dens enkelhet og presisjon. Imidlertid betyr det ikke at log-spesifikasjonen er den beste for alle land. Videre kan vi sammenlikne de ulike modellene og fastslå hvilken som gir best estimat i vårt datamateriale. Den nye modellen tar følgende form³⁰:

$$\begin{aligned} \logl\ddot{o}nn[X]_{it} = & \alpha_1 \log(W_{it-1}) + \alpha_2(U_{it}) + \alpha_3(U_{it-1}) + \alpha_4 \log(VakS_{it}) + \\ & \alpha_5 \log(VakS_{it-1}) + \alpha_6 \log(Andre\ variabler)_{it} + \lambda_t + v_{it} \end{aligned} \quad (4.3)$$

Med log lønn som avhengig variabel vil likningen tolkes som en semilogaritmisk funksjon. Fordelen med en semilogaritmisk funksjonsform er at den ikke er veldig restriktiv i motsetning til en ren logaritmisk form. For eksempel vil lønnsrelasjonen på logaritmisk form være log-lineær over hele utvalget og impliserer konstant ledighetselastisitet. Det vil også bety at relasjonen mellom lønn og ledighet er enten tiltakende eller avtakende uten grenser. En semilogaritmisk spesifisering vil være asymptotisk og vil aldri nå x-aksen, noe som er mer realistisk siden lønn aldri vil være lik null.

I neste regresjon bruker vi den inverse av ledighetsraten som forklaringsvariabel og tar følgende form:

$$\begin{aligned} \logl\ddot{o}nn[X]_{it} = & \alpha_1 \log(W_{it-1}) + \alpha_2(U_{it}^{-1}) + \alpha_3(U_{it-1}^{-1}) + \alpha_4 \log(VakS_{it}) + \\ & \alpha_5 \log(VakS_{it-1}) + \alpha_6 \log(Andre\ variabler)_{it} + \lambda_t + v_{it} \end{aligned} \quad (4.4)$$

Også denne regresjonen vil være log - level form som betyr at en enhets endring i arbeidsledigheten (ett prosentpoeng) gir endring i lønn. De semielastiske verdiene for samtlige funksjonsformer finnes på følgende måte:

M1:

$$\log W[X] = \beta \log(U) \Rightarrow \frac{\Delta W}{W} / \Delta U = \beta \frac{1}{U} \quad (4.5)$$

M2:

$$\log W[X] = \beta(U) \Rightarrow \frac{\Delta W}{W} / \Delta U = \beta \quad (4.6)$$

M3:

$$\log W[X] = \beta(U^{-1}) \Rightarrow \frac{\Delta W}{W} / \Delta U = -\beta \frac{1}{U^2} \quad (4.7)$$

²⁹Groot, Mekkeholt og Oosterbeek (1992), Wagner (1994), Johansen (1997).

³⁰Det blir ikke estimert en utvidet modell da langtidsestimaterne ikke endres mye.

hvor $M1$ er vår basismodell på log - log form fra likning 4.1, $M2$ er modell med aggregert ledighet som forklaringsvariabel, likning 4.3, mens $M3$ er modell med den inverse av aggregert ledighet som forklaringsvariabel, likning 4.4.

Resultatene for likning 4.3 er rapportert i appendiks A.1 tabell 10 hvor vi ser bort i fra random effects-estimering på grunnlag av diskusjon i delkapittel 4.2.2. Estimatenes er i stor grad lik grunnmodellen hvor vi brukte logaritmen til arbeidsledigheten som høyresidevariabel. Derimot er langtidskoeffisientene merkbart forskjellige. For grunnskolegruppen faller koeffisientverdien med 68 prosent til $-0,0199$. Også videregående- og høyere utdanningsgruppene får lavere verdier med et fall på henholdsvis 66 og 72 prosent til $-0,0129$ og $-0,0059$. Alle verdiene er signifikante med relativt høye t-verdier.

Resultatene for likning 4.4 er rapportert i tabell 11 i appendiks A.1. De beregnede verdiene er ikke mye forandret, men vi ser at estimatet til den inverse av ledigheten nå får positive verdier. Langtidsledighetskoeffisientene kan være vanskelig å tolke siden de fremstår med positivt fortegn, men ved hjelp av differensiering, som forklart over, vil vi oppnå de semielastiske langtidseffektene.

Det er også rapportert estimerte verdier ved bruk av gruppespesifikk ledighet for likning 4.3 og 4.4 i appendiks A.1 tabell 12 og 13. De langsiktige ledighetskoeffisientene blir generelt sett mindre i absoluttverdi sammenliknet med tabell 10 og 11, med unntak av videregående-skolegruppen som får økt koeffisient med U^{-1} som forklaringsvariabel.

For å undersøke om lønnskurvene er robuste med hensyn til funksjonsform, kan vi sammenlikne den langsiktige semielastisiteten. Vi kan da se hvilken effekt lønn har ved 1 prosents endring i arbeidsledigheten. Resultatene for *langtidsløsningen* og dens semielastisitet er rapportert i tabell 8.

Tabell 8: Sammenlikning av ulike funksjonsformer på tvers av gruppene

Funksjonsform	Gruppe								
	GRU			VGS			HØY		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
$\log U$	-0,062			-0,038			-0,021		
	(-4,19)			(-5,12)			(-3,22)		
$\log U_{Spes}$	-0,032			-0,030			-0,007		
	(-2,88)			(-5,17)			(-1,37)		
U		-0,020			-0,013			-0,006	
		(-4,34)			(-4,20)			(-2,82)	
U_{Spes}		-0,007			-0,008			-0,005	
		(-3,21)			(-6,83)			(-1,69)	
U^{-1}			0,106			0,064			0,041
			(3,74)			(4,58)			(2,74)
U_{Spes}^{-1}			0,100			0,067			0,004
			(2,10)			(3,51)			(0,70)
Semielastisiteter									
U	-0,019	-0,020	-0,010	-0,011	-0,013	-0,006	-0,006	-0,006	-0,004
U_{Spes}	-0,006	-0,007	-0,003	-0,006	-0,008	-0,003	-0,004	-0,005	-0,001
R^2	,99790	,99787	,99779	,99905	,99902	,99899	,99928	,99927	,99926
R_{Spes}^2	,99770	,99771	,99764	,99896	,99897	,99893	,99925	,99926	,99924
σ	,0072	,0073	,0074	,0049	,0049	,0051	,0051	,0051	,0051
σ_{Spes}	,0075	,0075	,0076	,0051	,0051	,0052	,0052	,0052	,0052
AIC	-9,613	-9,598	-9,559	-10,381	-10,347	-10,322	-10,313	-10,300	-10,296
AIC_{Spes}	-9,519	-9,526	-9,496	-10,289	-10,290	-10,261	-10,274	-10,276	-10,270

Merk: Estimatene er langtidsløsningen for hver gruppe. M1 er modell 1 med $\log U$ som forklaringsvariabel, M2 er modell 2 med U som forklaringsvariabel og M3 er modell 3 med U^{-1} som forklaringsvariabel. *Spes* angir gruppespesifikk arbeidsledighet. AIC er Akaikes informasjonskriterium. T-verdier er oppgitt i parantes.

I tabellen er det også inkludert langtidsløsning ved bruk av den gruppespesifikke arbeidsledigheten og er angitt med *Spes*. Vi kan da sammenlikne tre forskjeller. i) hvordan elastisiteten blir endret *innad* i en gruppe som følge av endret modell. ii) hvordan bruk av gruppespesifikk ledighet gir utslag på de beregnede elastisitetene (likning 4.5 - 4.7). Ved bruk av aggregert ledighet vil for eksempel gjennomsnittsledigheten være for høy i høyere utdanningsgruppen i forhold til hva den gruppespesifikke ledigheten viser, og derfor vil den beregnete effekten av semielastisiteten være sterkere enn hva den egentlig er. iii) hvordan modellspekifisering kan gi utslag på hvor stor elastisitetene blir, altså hvilken gruppe responderer sterkest. Først sammenlikner jeg resultatet mellom modellspekifiseringene M1, M2 og M3 med aggregert og gruppespesifikk ledighet for hver gruppe for til slutt å sammenlikne resultat på tvers av gruppene. Tallene blir omtalt som absoluttverdier.

Semielastisiteten for *grunnskolegruppen* har meget like verdier i M1 og M2 med henholdsvis verdier på -0,019 og -0,02. I M3 derimot reduseres estimatet med 50 prosent og får en verdi på -0,01. Tendensen er den samme når vi sammenlikner modellene ved bruk av gruppespesifikk arbeidsledighet. M1 og M2 gir relativt like verdier med -0,006 og -0,007, mens M3 halveres og får en verdi på -0,003. Vi kan også merke oss at semielastisiteten faller med 68 prosent fra -0,019 til -0,006 i M1, når vi benytter gruppespesifikk ledighet sammenliknet med aggregert ledighet. Samlet sett gir altså M3 halvparten så stor lønnseffekt ved endret ledighet enn M1 og M2. Valg av funksjonsform har dermed stor betydning for hvilke verdier lønnsrelasjonen får. De rapporterte langtidsløsningene i tabellen er alle signifikante på 5 prosentsnivå.

Også for *videregående skolegruppen* er tendensen den samme. M1 og M2 har relativt like semielastiske verdier på -0,011 og -0,013, mens M3 halveres til -0,006. Ved bruk av gruppespesifikk ledighet gir M2 høyest estimat med -0,008. M1 har en noe lavere verdi med -0,006, en reduksjon på 25 prosent. Vi legger merke til at M3 igjen gir lavest verdi på -0,003, en reduksjon på hele 63 prosent sammenliknet med M2. Alle modellene får lavere semielastisitet ved bruk av gruppespesifikk arbeidsledighet, og M1 faller med 45 prosent fra -0,011 til -0,006 sammenliknet med aggregert ledighet. Totalt sett gir M1 og M2 relativt like semielastisiteter for både aggregert og gruppespesifikk ledighet, mens M3 skiller seg ut med cirka halvparten så store verdier. Alle langtidsløsningene gir høye t-verdier som er signifikante på 1 prosentsnivå.

For *høyere utdanningsgruppen* er semielastisiteten helt lik for M1 og M2 med en verdi på -0,006. Estimatet i M3 reduseres med 33 prosent og får en verdi på -0,004. Når vi bruker den gruppespesifikke arbeidsledigheten, gir M2 høyest verdi på -0,005. Estimatet blir noe redusert i M1 og får en verdi på -0,004, som er en nedgang på 20 prosent. Igjen er det M3 som skiller seg ut, og vi får her en nedgang på hele 80 prosent sammenliknet med M2. I M1 vil bruk av gruppespesifikk ledighet føre til et fall i semielastisiteten på 33 prosent sammenliknet med aggregert ledighet, fra -0,006 til -0,004. Totalt sett er det M3 som gir lavest semielastisitet både for aggregert og gruppespesifikk ledighet, mens M2 gir høyest estimat. Vi bør merke oss at langtidsløsningene for høyere utdanningsgruppen har generelt lavere t-verdier. Resultatene for gruppespesifikk arbeidsledighet gir insignifikante verdier, mens resultatene for aggregert ledighet gir signifikante verdier på 1 prosentsnivå.

Når vi sammenlikner gruppene seg imellom, ser vi at semielastisiteten for høyere utdanningsgruppen, i motsetning til grunnskole- og videregående skolegruppen, ikke faller i like stor grad når vi benytter gruppespesifikk ledighet sammenliknet med aggregert ledighet. Årsaken til dette kan komme av flere grunner. Blant annet vil $1/U$ i likning 4.5 gi utslag på semielastisiteten, der grunnskolen har høyere gjennomsnittlig ledighet mens høyere utdanning har lavere gjennomsnittlig ledighet enn det aggregerte ledighetsgjennomsnittet. Denne faktoren gjør at semielastisiteten faller mer hos grunnskolegruppen enn hos høyere

utdanningsgruppen.

Videre er det interessant å merke seg at forskjellen mellom gruppene er relativ liten ved bruk av gruppespesifikk ledighet uansett funksjonsform. Modell 3 gir størst forskjeller mellom gruppene hvor grunnskole og videregående skole begge får verdier på $-0,003$, mens høyere utdanning får en verdi på $-0,001$. Sammenliknet med aggregert regional ledighet ser vi altså mindre forskjeller mellom gruppene. Noe av årsaken kan igjen være at gruppespesifikk ledighet gir et korrigert estimat av semielastisitetene, fordi aggregert regional ledighet ikke korrelerer nøyaktig med de gruppespesifikke ledighetene, spesielt høyere utdanningsgruppen og forskjellen mellom gruppene minker betraktelig.

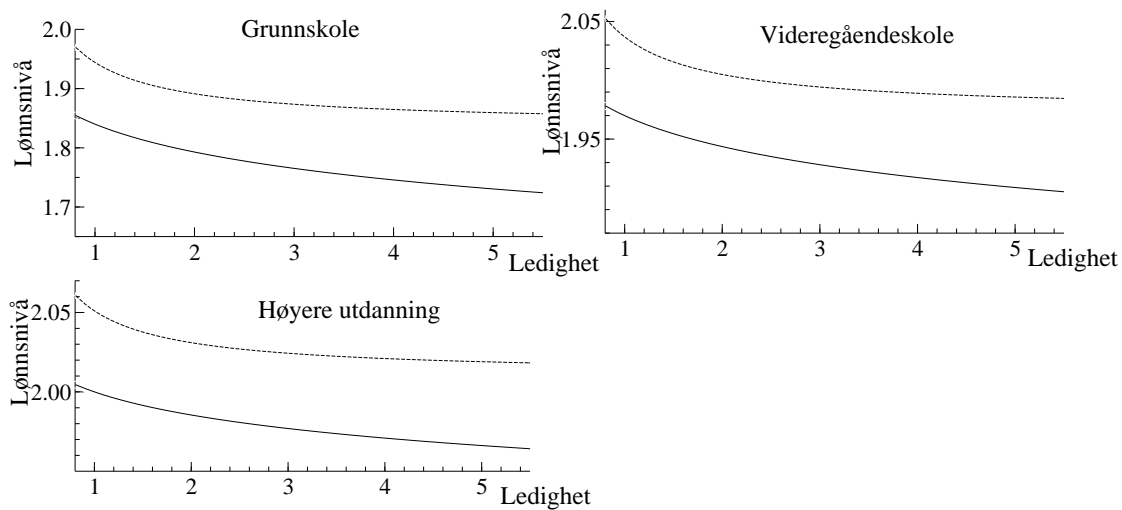
Tolkningen av semielastisitetene er lettere å forstå hvis vi multipliserer estimatene med 100. Da vil en enhets endring i ledigheten gi prosentvis endring i lønn. Ved bruk av aggregert arbeidsledighet vil det for grunnskolegruppen bety at ett prosentpoengs økning i ledigheten fører til cirka 2 prosents reduksjon i lønn ved bruk av modell 1 og 2, mens i modell 3 er nedgangen mindre og er 1 prosent. For videregående skolegruppen varierer nedgangen mellom 1,3 og 0,6 prosent, hvor modell 2 gir størst nedgang og modell 3 gir minst nedgang. Variasjonen er ikke like stor hos høyere utdanningsgruppen hvor nedgangen ligger mellom 0,6 og 0,4 prosent, igjen gir modell 3 minst nedgang. Ved bruk av gruppespesifikk ledighet vil en prosentpoengs økning i ledigheten føre til mellom 0,7 - 0,3 prosents reduksjon i lønn for grunnskolegruppen. For videregående vil økningen føre til redusert lønn med mellom 0,8 og 0,3 prosent. I høyere utdanningsgruppen faller lønnen med mellom 0,4 og 0,1 prosent. Modell 3 gir lavest lønnsrespons for alle gruppene og modell 2 gir høyest lønnsrespons for alle gruppene. Ut i fra disse resultatene ser vi at funksjonsform spiller en rolle, spesielt når vi skal undersøke hver enkelt gruppes lønnsrespons og styrken av den.

Ved valg av modell³¹ vil R^2 , AIC og standardavviket, σ , kunne fortelle hvilken spesifikkasjon vi foretrekker. Ved bruk av aggregert arbeidsledighet gir M1 høyest R^2 og lavest AIC og σ verdier for samtlige grupper med unntak av høyere utdanningsgruppen hvor σ verdiene er like. Det betyr modell 1 med log U som forklaringsvariabel har best forklaringskraft. Ved bruk av den gruppespesifikke arbeidsledigheten gir M2 høyest R^2 og lavest AIC verdier, mens σ gir lavest verdi for M1 og M2 i grunnskole- og videregående skolegruppen. Igjen er verdiene av σ like hos høyere utdanning. Vi vil derfor foretrekker modell 2 med U_{Spes} som forklaringsvariabel.

I figur 4.2 ser vi lønnskurven for M1 og M3. Ved bruk av M3 ligger kurven over M1s estimerte lønnskurve for samtlige grupper. Videre ser vi at en økning fra 1 til 2 prosents ledighet i grunnskolegruppen gir en reduksjon på 3,1 prosent i lønnen i M3, mens tilsva-

³¹En formell test av modellvalg er H_0 : M1 og M2 har samme forklaringskraft versus H_1 : M2 forklarer mer. Ratioen $\frac{WALD_{M2}}{WALD_{M1}} \sim F(5,5)$ som gir $\frac{895}{913} = 0,98$. Kritisk verdi for 5 prosentsnivå er 5,05. Testing gir at ingen av modellene har statistisk bedre forklaringskraft for noen av gruppene.

Figur 4.2: Lønnskurvene for alle grupper ved bruk av regional ledighet i M1 og M3. M1 er representert med hel linje, mens M3 har prikket linje.



rende økning i ledigheten reduserer lønnen med 2,2 prosent i M1. For ledighet over 3.5 prosent blir lønnskurven i M3 tilnærmet flat, mens kurven for M1 gir målbar nedgang i lønn. Dette gjelder for alle utdanningsgruppene.

4.6 Oppsummering

I dette delkapitlet er det prøvd ut ulike funksjonsformer av lønnskurven for å avklare kurvens robusthet. De nye funksjonsformene som ble prøvd, bruker ledighet, U , og den inverse av ledigheten, U^{-1} som høyresidevariabler.

For å kunne sammenlikne de ulike spesifikasjonene er det beregnet semielastiske verdier for hver gruppe, hvor også basismodellen med den naturlige logaritmen, $\log U$, som høyresidevariabel er inkludert. Det er også oppgitt standardavvik, R^2 og Akaikes informasjonskriterium for å kunne avgjøre hvilken modell som gir best estimat.

Totalt sett gir basismodellen med $\log U$ og modell 2 med ledighet U som forklaringsvariabler relativt like semielastiske verdier for samtlige grupper. Modell 3 med den inverse av ledighet som forklaringsvariabel, anses som et utskudd og gir markant nedgang i lønnseffekten, sammenliknet med modell 1 og modell 2. Ved bruk av føyningsmål og informasjonskriterier ble modell 1 og modell 2 foretrukket fremfor modell 3.

Ved bruk av gruppespesifikk arbeidsledighet ble forskjellen av lønnseffekten på ledighet kraftig redusert mellom gruppene, uavhengig av modell. For eksempel gir modell 2 semielastiske verdier på -0,007, -0,008 og -0,005 for henholdsvis grunnskole, videregående skole og høyere utdanning. Tilsvarende tall ved bruk av aggregert ledighet er -0,02, -0,013 og -0,006.

5 Sammenlikning med tidligere studier

Denne studien har undersøkt hvordan lønnsresponsen mellom ulike utdanningsgrupper påvirkes av aggregert regional og gruppespesifikk arbeidsledighet. Utgangspunktet for undersøkelsen er funn gjort av BO (1994). Her finner de en lønnselastisitet på -0,1 og antyder en sterkere effekt hos lavtlønnede grupper sammenliknet med høytlønnede grupper. I Norge har Johansen (1999) utført en studie hvor han bruker data for "skilled" og "unskilled" arbeidere og avdekker i tråd med andre internasjonale studier³², at lønnsresponsen er sterkere for lavt utdannede enn for høyt utdannede. Resultatene ser ut til å være konsistente for Europa og USA. Johansen påpeker at fremtidige undersøkelser bør omfatte gruppespesifikk ledighet for å avdekke eventuelle forskjeller mellom utdanningsgrupper. Min studie estimerer både for aggregert regional ledighet og gruppespesifikk regional ledighet. Resultatene antyder en sterkere respons jo lavere utdanning, mens forskjellen *mellom* gruppene avhenger av modellspesifikasjon og gjør resultatene mindre robust.

I den empiriske analysen er regresjonen estimert med både fixed og random effects (FE og RE), og det viser seg at FE-modellene ga de mest troverdige resultatene blant annet på bakgrunn av en formell Hausman-test. Grunnen er at denne metoden tillater korrelasjon mellom forklaringsvariablene og regionspesifikke effekter som ikke varierer over tid. Det ble ikke testet for potensielle endogenitetsproblemer³³ mellom ledighet og lønn, fordi vi antar at lønnen for hver gruppe ikke påvirker den regionale ledigheten som helhet og kan derfor behandle variablene som eksogene. Hvis dette ikke er tilfelle, vil estimatene være skjeve. Likevel anses ikke endogenitet å være et stort problem fordi ledighetsraten er gjennomsnittet for en region, mens lønnen representerer en definert gruppe, og en gitt gruppe dominerer ikke det lokale arbeidsmarkedet. Ved bruk av gruppespesifikk ledighet faller denne argumentasjonen, men arbeidsledighet vil uansett først og fremst påvirkes av ulike typer etterspørselssjokk.

Når det gjelder utdanningsgruppens lønnsrespons, har vi estimert en dynamisk modell med langtidsløsninger der estimatene viser - uavhengig av metode - at responsen er avtakende i utdanningsnivået. Videre er det testet for om langtidseffektene er signifikant forskjellig mellom gruppene. Resultatene gir relativt høye t-verdier og er signifikant ved 2 prosentsnivå. Ved estimering med gruppespesifikk ledighet faller den langsiktige lønnsresponsen for alle gruppene. Mest iøynefallende er grunnskolegruppens koeffisient som halveres og blir tilnærmet lik videregående skolegruppens estimat. Selv om vi ikke har utført en formel test for signifikante gruppeforskjeller, kan vi med sikkerhet fastslå at det ikke finnes numerisk ulik langsiktig lønnsrespons for grunn- og videregående skole. Elastisitetene for gruppespesifikk regional ledighet er tilnærmet lik -0,03 for begge gruppene og

³²Blanchflower og Oswald (1994), Campbell (1997) og Turunen (1998).

³³Blanchard og Oswald (1994) tester for endogenitet og benytter instrumentvariabler som korrelerer og erstatter ledighetsvariablen, samtidig som den ikke er korrelert med restleddet og blir eksogen.

er signifikante ved 1 prosentnivå. For høyere utdanningsgruppen faller langtidselastisiteten til $-0,007$ og er ikke signifikant ved 10 prosentnivå. Forklaringen kan være at de høyt utdannede er en spesiell gruppe av lønnstakere siden det er mer sannsynlig at de migrerer hyppigere mellom regionene enn resten av befolkningen. Videre er det slik at det er større sannsynlighet for at de migrerer fra fjerntliggende til sentrale (økonomiske) regioner enn vice versa (Hynninen 2009). Det kan også tenkes at arbeidsmarkedssituasjonen for høyt utdannede er mer stabil over tid, som igjen gjør at de er mindre følsom for endringer i ledighetsraten. Ehreberg og Smith (2000) argumenterer for at høyt utdannede har bedre oversikt over potensielle arbeidsmarkeder, noe som kan bidra til å utvide det geografiske arbeidssøkeområdet. Dette gjør at gruppen flytter mer, og igjen minker innvirkningen av regional ledighet. Mine funn er i samsvar med studier fra andre land³⁴, og vi kan hen-
tyde til at det ikke finnes en signifikant lønnskurve ved bruk av gruppespesifikk regional ledighet for høyere utdanningsgruppen.

Estimatene for den utvidete modellen med aggregert regional ledighet hvor boligpris, ordinære arbeidsmarkedstiltak og klimavariabler er inkludert, ser vi en marginal nedgang i langtidselastisiteten for samtlige grupper. Testen for gruppeforskjeller viser det er signifikante ulikheter på 1 prosentnivå mellom alle gruppene.

Så langt har vi funnet eksistens av ulike lønnskurver for våre tre utdanningsgrupper, også i en utvidet modell. Resultatene overlever derimot ikke bruk av gruppespesifikk ledighet hvor grunn- og videregående skolegruppene får tilnærmet like elastisiteter og høyere utdanning får en insignifikant verdi. Dette kan implisere at høyt utdannede har en sterkere regional forhandlingsmakt enn andre grupper, fordi de blant annet er mer fleksible arbeidstakere hva gjelder migrasjon, slik at regional ledighet ikke påvirker lønnen i like stor grad.

Lønnsrelasjonens robusthet er testet for i tabell 8 ved å inkludere andre former av den aggregerte regionale arbeidsledighetsraten. I tråd med BO (1994) er det gjort regresjoner med aggregert regional ledighet og dens inverse. Ved sammenlikning med basismodellen gir modellform 2, med aggregert regional ledighet som forklaringsvariabel, relativt like semielastiske verdier innad i gruppene. Modellform 3, med den inverse av aggregert regional ledighet som forklaringsvariabel, gir derimot merkbart lavere semielastiske verdier. For grunn- og videregående skole halveres semielastisiteten, mens for høyere utdanning reduseres den med en tredel. På bakgrunn av dette vil valg av funksjonsform ha betydning for i hvor stor grad lønnsresponsen er i gruppen. Eksempelvis vil valg av modell 3 for grunnskole, modell 2 for videregående og modell 1 (grunnmodell) for høyere utdanning medføre insignifikante gruppeforskjeller hvor videregående skolegruppen har sterkest lønnsrespons. Dette er på tross av aggregert regional ledighet som ledighetsvariabel. Resultatene blir perverse ved bruk av gruppespesifikk regional ledighet. En kan manipulere

³⁴Hynninen (2009).

det slik at høyere utdanning har sterkest semielastisk lønnsrespons med $-0,005$ (modell 2) signifikant på 10 prosentnivå, mens grunn- og videregående skole får verdier på $-0,003$ (modell 3) signifikant på 5 prosentnivå. Bruk av føyningsmål og informasjonskriterier viser at basismodellen med den naturlige logaritmen av aggregert regional ledighet er best egnet for alle grupper, mens modell 2 gir best estimat for gruppespesifikk ledighet.

Når det gjelder sammenlikning med tidligere norske studier, er det så vidt meg bekjent, bare Johansen (1999) som har utført en studie hvor gruppesammenlikning mellom utdanningsnivå foretas. I hans undersøkelse ble det benyttet aggregert ledighet. Johansens resultater er en del høyere enn mine og finner en langtidselastisitet på $-0,05$ for "skilled" arbeidere, mens "unskilled" arbeidere responderer dobbelt så mye med en elastisitet på $-0,1$. Til sammenlikning finner jeg en langtidselastisitet på $-0,02$ for høyere utdanningsgruppen og $-0,06$ for grunnskolegruppen. Årsaken til ulike estimat kan være datagrunnlaget hvor jeg sammenlikner tre utdanningsgrupper med regional ledighet, mens Johansen ser på to utdanningsgrupper med aggregert ledighet. Begge studiene finner signifikante lønnskurver for alle gruppene hvor også gruppeforskjellen er signifikant.

Internasjonale studier finner blandete resultater. Turunen (1998) bruker i sin studie regional data fra USA mellom 1979-1992 og kontrollerer for ulike utdanningsgrupper. I fixed effects-estimeringen finner han en elastisitet på $-0,08$ for gruppen med mindre enn 12 års utdannelse, $-0,06$ for gruppen med 12 til 16 års utdannelse og $-0,035$ for gruppen med mer enn 16 års utdannelse. Elastisiteten for de med mer enn 16 års utdannelse er ikke signifikant ved 10 prosentnivå. Verdiene er noe høyere enn mine, men resultatene viser en avtakende lønnsrespons i utdanning. Han undersøker også hvorvidt forskjellene er signifikant mellom gruppene og fastslår at lønnsresponsen er ulik for ulike utdanningsgrupper. Turunen har hverken estimert for gruppespesifikk ledighet eller kontrollert for ulike funksjonsformer. Baltagi og Blien (1998) bruker data i perioden 1981-1990 og undersøker den tyske lønnskurven. Ved bruk av regional ledighet finner de en elastisitet på $-0,11$ i gruppen for lavere kvalifikasjoner og $-0,06$ i gruppen for høyere kvalifikasjoner etter å ha kontrollert for fixed effects og endogen ledighet. Forskjellen mellom utdanningsgruppene er signifikante. Metodevalg og datagrunnlag er sannsynligvis noe av årsaken til høyere estimat, sammenliknet med mine. Det er ikke estimert verdier ved bruk av gruppespesifikk arbeidsledighet og ingen eksperimenter av funksjonsform. Sanz-de-Galdeano og Turunen (2006) ser på euroområdet i perioden 1994-2001 med regionale ledighetsrater. De skiller mellom primært, sekundært og tertiært utdanningsnivå. Estimaten fra fixed effects-modellen gir tilnærmet like elastisiteter på henholdsvis $-0,121$, $-0,138$ og $-0,135$. De finner *ingen* signifikant forskjell mellom gruppene. Igjen er det ikke estimerte verdier ved bruk av gruppespesifikk ledighet. Campbell (1997) undersøker lønnsresponsen for USA i perioden 1975-1994. Han bruker både aggregert og gruppespesifikk ledighet i sin analyse. Den estimerte lønnslikningen har en litt annen form, men resultatene lar seg sammenlikne.

For aggregert ledighet er det signifikante forskjeller mellom høyt og lavt utdannede, hvor ledighetseffekten gir sterkest respons for sistnevnte. Ved bruk av gruppespesifikk ledighet er resultatene mer uklar. Her finner han en insignifikant forskjell mellom gruppene hvor høyt utdannede responderer mindre.

I et forsøk på en kortfattet oppsummering kan vi se en antydning til at lønnskurven varierer mellom gruppers utdanningsnivå. Hvorvidt denne forskjellen er signifikant avhenger av datagrunnlag (og mulig funksjonsform). I hvor stor grad de respektive gruppene responderer varierer, men mine estimat synes generelt sett å gi lavere verdier for samtlige grupper. Vi kan også merke oss at t-verdien for høyt utdannede inntar lave verdier, og anses ofte som insignifikant.

På grunn av manglende norske studier hvor lønnsresponsen på ledighet mellom ulike utdanningsnivå sammenliknes, er det vanskelig å si om effektene funnet i denne studien er representative. Likevel kan generelle studier være en pekepinn for mine resultater. Barth et al (2002) ser på en norsk lønnskurve og bruker regional data fra 1991-1997. De finner en signifikant elasticitet på $-0,0319$ for menn. Regresjonen er gjort ved fixed effects-estimering. Dette er meget nær mine resultater (ved bruk av gruppespesifikk ledighet) hvor både grunnskole, og videregående skolegruppen får signifikante verdier på $-0,03$. Som vi vet er ikke høyere utdanningsgruppens elasticitet signifikant ved 5 prosentnivå. Albæk et al (1999) brukte regional data fra 1989, 1991 og 1993 og finner en elasticitet på $-0,0147$. Estimaten var insignifikant ved bruk av fixed effects-estimering. Som nevnt kan årsaken være at perioden var karakterisert ved en høy grad av sentralisering og liten nominell lønnsøkning. Mine resultat er noe høyere, men anses som relativt like. Raaum og Wulfsberg (1997) innhenter data fra 1997 og estimerer lønnseffekten av regional ledighet. De finner en langtidselasticitet på $-0,011$ som er insignifikant. Også de bruker ulike funksjonsformer og deres resultater avhenger av modellform. Deres foretrukne modell gir en positiv effekt på lønn av økt ledighet. Dyrstad og Johansen (2000) finner en signifikant langtidselasticitet på $-0,02$ ved bruk av regional ledighet i ulike funksjonsformer. Dataene er fra perioden 1970-1988. Mine estimat er til sammenlikning noe høyere, men estimeringsmetoden er ulik. Wulfsberg (1997) bruker regiondata i tidsperioden 1972-1988. Han bruker ulike modeller og finner at lønnselasticiteten er mellom $-0,1$ og $-0,04$ ved bruk av fixed effects-estimering. Resultatene avhenger i stor grad av modellvalg, men Wulfsbergs funn kan sies å være generelt høy sammenliknet med andre norske studier. Brunstad og Dyrstad (1997) finner en gjennomsnittlig elasticitet på $-0,051$. De bruker gjennomsnittlig ledighet på fylkesnivå i perioden 1970-1982. Elasticiteten er høy, men modell og metode gjør resultatene vanskelig å sammenlikne direkte med mine.

Kort oppsummert ser vi konturene av den norske lønnskurven ved regional ledighet varierer i modell- og metodevalg. Tidligere studier gir en lønnselasticitet mellom $-0,01$ og $-0,1$. De laveste estimatene er til gjengjeld insignifikant. Studier som har brukt samme

modell og metode, finner en elastisitet på rundt $-0,03$, som er nær mine estimat hvor jeg bruker gruppespesifikk ledighet. Slik sett virker mine resultat å være på linje med tidligere studier.

Tabell 9: Utvalgte resultat for Norge

Forfatter	Tidsperiode	Elastisitet	FE-estimering	Ledighetsvariabel
Blanchflower og Oswald (1994)	1989-1991	-0,08	Delvis	Regional
Barth et al (2002)	1991-1997	Menn -0,0319	JA	Regional
		Kvinner -0,0194		
Albæk et al (1999)	1989, 91, 93	-0,0147	JA	Regional
Raaum og Wulfsberg (1997)	1997	-0,011	NEI	Regional
Johansen (1997)	1989	-0,06	NEI	Regional
Johansen (1999)	1966-1978	Skilled -0,05	NEI	Aggregert
		Unskilled -0,1		
Dyrstad og Johansen (2000)	1970-1988	-0,02	JA	Regional
Wulfsberg (1997)	1972-1988	-0,1 til -0,04	JA	Regional

6 Konklusjon

Gjennom denne studien er det forsøkt å avdekke eventuelle gruppeforskjeller av lønnsrespons på regional ledighet ved bruk av regional og gruppespesifikk arbeidsledighet. Videre er det utprøvd ulike funksjonsformer som en robusthetstest. Jeg kan på bakgrunn av dette trekke følgende konklusjon:

1. Det eksisterer ulike lønnskurver for hver utdanningsgruppe hvor lavt utdannede responderer sterkere enn høyt utdannede og hver enkelt kurve er signifikant. Det er også signifikante forskjeller mellom utdanningsgruppene. Disse resultatene er oppnådd ved bruk av aggregert regional ledighet.
2. Det eksisterer ikke signifikante lønnskurver for alle utdanningsgruppene. Høyt utdannede responderer minst, men verdiene er insignifikante. Det er heller ikke signifikante forskjeller mellom utdanningsgruppene. Disse resultatene er oppnådd ved bruk av gruppespesifikk ledighet.
3. Lønnskurvene overlever ikke alle modellspesifikasjonene. Resultatene er avhengig av funksjonsform, og perverse resultater kan fremkomme.

Til slutt kan vi merke oss at disse konklusjonene stemmer overens med nasjonal og internasjonal forskning. Funnene tyder på at det ikke eksisterer en lønnskurve for høyt utdannede og forskjeller mellom utdanningsgruppene nesten forsvinner ved bruk av gruppespesifikk ledighet. En forståelse av disse resultatene er at lønnsresponsen for grunn- og videregående skole er like og signifikant, men signifikant ulik høyere utdanningsgruppen. Som Johansen (1999) konkluderer, kan dette bety at høyt utdannede står overfor liten risiko for å bli permittert når permitteringer er gjort, og dermed blir lønnen uforandret. Lønnen til lavere utdannede er mer følsom for variasjon i arbeidsledigheten siden forventet kostnad av å være permittert er tett knyttet til tilstanden i arbeidsmarkedet. En av grunnene til at høyere utdannede ikke blir permittert, kan være at gruppen generelt sett er mer fleksibel mellom regionale arbeidsmarkeder og dermed får større forhandlingsmakt.

Refferanseliste

Albæk, K., Asplund, R., Blomskog, S., Barth, E., Gudmundsson, B., Karlsson, V. og Madsen, E. (1999), "*The wage curve in the Nordic Countries*", TemaNord serien 1999:597

Albæk, K., Asplund, R., Blomskog, S., Barth, E., Gudmundsson, B., Karlsson, V. og Madsen, E. (2000), "*Dimension of the Wage-Unemployment Relationship in the Nordic Countries: Wage Flexibility without Wage Curves*", Research in Labor Economics, Emerald Group Publishing Limited, 19:345-381

Azariades, C. (1975), "*Implicit Contracts and Underemployment Equilibria*", Journal of Public Economy, 83:1183-1202.

Baltagi, B. og Blien, U. (1998), "*The German wage curve: evidence from the IAB employment sample*", Economics Letters 61:135-142.

Barth, E., Bratsberg, B., Naylor, R. og Raaum, O. (2002), "*Explaining variations in wage curves: theory and evidence*", Working paper 03/2002, Department of Economics, University of Oslo.

Bell, B., Nickell, S. og Quintini, G. (2002), "*Wage equations, wage curves and all that*", Labour Economics 9:341-360.

Blanchflower, D., og Oswald, A. (1994), "*The wage curve*", MIT Press, Cambridge.

Blanchflower, D., og Oswald, A. (2005), "*The wage curve reloaded*", Working paper 11338, National Bureau of Economic Research, Cambridge.

Blomquist, G., Berger, M. og Hoehn, J. (1988), "*New estimates of Quality of Life in Urban Areas*", American Economic Review, 78:89-107

Bratsberg, B., og Turunen, J. (1996), "*Wage curve evidence from panel data*", Economics Letters, 51:345-353.

Brunstad, R. og Dyrstad, J. (1997), "*Boming Sector and Wage Effects: An Empirical Analysis on Norwegian Data*", Oxford Economic Papers, 49:89-103.

Calmfors, L., (1995), "*Labour Market Policy and Unemployment*", European Economic Review, 39:583-592.

Calmfors, L. og Driffill, J. (1988), "*Centralisation of wage bargaining*", Economic Policy, 6:14-61.

Calmfors, L. og Nymoene, R., (1990), "*Real wage adjustment and employment policies in the Nordic countries*", Economic Policy, 11:397-448.

- Campbell, C. (1997), *"The variation in wage rigidity by occupation and union status in the us"*, Oxford Bulletin of Economics and Statistics, 59:134-147.
- Carlsen, F. og Johansen, K., (2004), *"Subjective Measure of Employment Opportunities and Interregional Migration"*, Labour, 18:563-589.
- Card, D. (1995), *"The wage curve: A review"*, Journal of Economic Literature, 33:785-799.
- Dystad, J., og Johansen, K. (2000), *"Regional wage responses to unemployment and profitability: Empirical evidence from Norwegian manufacturing industries"*, Oxford Bulletin of Economics and Statistics, 62:101-117.
- Ehreberg, R. og Smith, R. (2000), *"Modern Labour Economics"*, Addison Wesley Longman, Inc.
- Friedman, M. (1968), *"The role of monetary policy"*, American Economic Review, 58:1-17.
- Groot, W., Mekkelholt, E. og Oosterbeek, H. (1992), *"Further evidence on the wage curve"*, Economics Letters, 38:355-359.
- Harris, J. og Todaro, M. (1970), *"Migration, Unemployment and Development: A Two-Sector Analysis"*, American Economic Review, 60:126-142.
- Haavelmo, T. (1992), *"Arbeidsmarkedstiltak og ledighet i Norge etter fylke, 1980 - 1994"* I Bakke, H., *"Arbeidsledighet"*, Universitetsforlaget.
- Hoel, M. og Nymoene, R. (1988), *"Wage formation in Norwegian manufacturing. An empirical application of a theoretical bargaining model"*, European Economic Review 32:977-998.
- Hynninen, S. (2009), *"Is there a wage curve for the highly educated?"*, ISER Working Paper Series, no.2009/17, University of Jyväskylä, Finland
- Hægeland, T., og Kirkebøen, L., (2007), *"Lønnsforskjeller mellom utdanningsgrupper"*, Rapport 2007/36, Statistisk sentralbyrå.
- Jansens, S. og Konings, J. (1998), *"One more wage curve: the case of Belgium"*, Economics Letters, 60:223-227.
- Johansen, K. (1997), *"The wage curve: convexity, kinks, and composition effects"*, Applied Economics, 29:71-78.
- Johansen, K. (1999), *"Wage flexibility for skilled and unskilled workers: new evidence on the Norwegian wage curve"*, Labour 13:413-432.
- Johansen, K. (2002), *"Regional wage curves empirical evidence from Norway"*, Working paper 03/2002, Department of Economics, Norwegian University of Science and Technology, Trondheim.

- Johansen, K. (2003), *"Regional wages, unemployment and labour market programmes"*, Mimeo, Department of Economics, Norwegian University of Science and Technology, Trondheim.
- Layard, L., Nickell, S. og Jackman, R. (1991), *"Unemployment, macroeconomics performance and the labour market"*, Oxford University Press, Oxford.
- Lindbeck, A. og Snower, D. (1988), *"The insider-outsider theory of employment and unemployment"*, MIT Press, Cambridge, MA.
- Lucas, R. og Rapping, L., (1969), *"Real wages, Employment, and Inflation"*, Journal of Political Economy, 77:721-754.
- Manning, A., (1993), *"Wage Bargaining and the Phillips Curve: The Identification and Specification of Aggregate Wage Equations"*, Economic Journal 103:98-118.
- Moene, K., Wallerstein M. og Hoel M., (1993) *"Bargaining Structure and Economic Performance"* i Flanagan R., Moene K. og Wallerstein M. (red.) *"Trade Union Behaviour, Pay Bargaining and Economic Performance"*, Oxford, Clarendon Press.
- Nicolaisen, S. og Tranæs, T. (1996), *"Wage curves for Denmark"*, Nationaløkonomisk Tidsskrift 134:223-237.
- Nijkamp, P. og Poot, J., (2005), *"The last word on the wage curve? A meta-analytic assessment"*, Journal of Economics Surveys, 19:421-450.
- Phillips, J. (1958), *"The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861-1957"*, Economica 25:283-299.
- Raaum, O. og Wulfsberg, F. (1997), *"Unemployment, Labour Market Programmes and Wages in Norway"*, Mimeo, Central Bank of Norway, Oslo.
- Sanz-de-Galdeano, A. og Turunen, J. (2006), *"The euro area wage curve"*, Economics Letters 92:93-98.
- Shapiro, C., og Stiglitz, J. (1984), *"Equilibrium unemployment as a worker discipline device"*, The American Economic Review, 74:433-444.
- Stiglitz, J. (1986), *"Theories of wage rigidity"*. In *"Butkiewicz, J., Koford, K., og Miller, J., Keynes' economic legacy"*, Praeger, New York.
- Townsend, R. (2005), *"Choosing Among Competing Theories of the Wage Curve"*, Department of Economics, Stanford University.
- Turunen, J. (1998), *"Disaggregated wage curves in the United States: evidence from panel data of young workers"*, Applied Economics, 30:1665-1677

Vermeulen, W. og Ommeren, J. (2009), "*Compensation of Regional Unemployment in Housing Market*", *Economica*, 76:71-88.

Wagner, J. (1994), "*German wage curves, 1979-1990*", *Economics Letters*, 44:307-311.

Winter-Ebmer, R. (1996), "*The wage curve, unemployment duration and compensating differentials*", *Labour Economics* 3:425-434.

Wooldridge, J. (2006), "*Introductory Econometrics - A Modern Approach*", Thomson South-Western, Ohio.

Wulfsberg, F. (1997), "*An Application of Wage Bargaining Models to Norwegian Panel Data*", *Oxford Economic Papers*, 49:419-440.

A Appendiks

A.1 Andre funksjonsformer

Tabell 10: Effekt av regional arbeidsledighet og vakanseratestrøm på lønn for ulike utdanningsgrupper. Avhengig variabel $\log\text{l\o}nn[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
$\log(W_{it-1})$	0,778 (29,2)	0,757 (32)	0,75 (31,4)
U_{it}	-0,008 (-8,62)	-0,006 (-8,47)	-0,003 (-4,20)
U_{it-1}	0,003 (3,44)	0,003 (4,00)	0,002 (2,17)
$\log(VakS_{it})$	-0,0004 (-0,194)	0,0007 (0,428)	0,003 (2,31)
$\log(VakS_{it-1})$	0,002 (0,76)	-0,002 (-1,23)	-0,002 (-1,42)
Oservasjoner	810	810	810
R^2	0,998	0,999	0,999
σ	0,007	0,005	0,005
Wald (joint)	913,4**	1579**	1021**
AR(1)-test	-0,234	-1,096	0,184
AR(2)-test	-1,682*	-3,06**	-2,721**
	Langtidsledighetskoeffisient		
	-0,0199 (-4,34)	-0,0129 (-5,48)	-0,0059 (-2,82)

Tabell 11: Effekt av den inverse regionale arbeidsledigheten og vakanseratestrøm på lønn for ulike utdanningsgrupper. Avhengig variabel $\log\text{l\o}nn[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
$\log(W_{it-1})$	0,764 (27,8)	0,743 (28,8)	0,756 (30,4)
U_{it}^{-1}	0,042 (5,60)	0,034 (6,72)	0,018 (3,71)
U_{it-1}^{-1}	-0,017 (-1,68)	-0,017 (-3,22)	-0,008 (-1,77)
$\log(VakS_{it})$	-0,0005 (-0,300)	0,0004 (0,273)	0,003 (2,22)
$\log(VakS_{it-1})$	0,002 (0,91)	-0,001 (-0,992)	-0,002 (-1,38)
Oservasjoner	810	810	810
R^2	0,998	0,999	0,999
σ	0,007	0,005	0,005
Wald (joint)	835,8**	1182**	926**
AR(1)-test	0,197	-0,705	0,266
AR(2)-test	-1,715	-2,833**	-2,718**
	Langtidsledighetskoeffisient		
	0,106 (3,74)	0,064 (4,58)	0,041 (2,74)

Tabell 12: Effekt av gruppebestemt regional arbeidsledighet og vakanseratestrøm på lønn for ulike utdanningsgrupper. Avhengig variabel $\log\text{lønn}[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
$\log(W_{it-1})$	0,760 (28,9)	0,727 (25,8)	0,757 (31,7)
U_{it}	-0,002 (-4,87)	-0,002 (-6,68)	-0,002 (-3,47)
U_{it-1}	0,0004 (0,93)	0,0003 (0,689)	0,0002 (0,494)
$\log(\text{Vak}S_{it})$	0,0005 (0,268)	0,001 (0,874)	0,004 (2,64)
$\log(\text{Vak}S_{it-1})$	0,002 (0,72)	-0,002 (-1,1)	-0,002 (-1,5)
Oservasjoner	810	810	810
R^2	0,998	0,999	0,999
σ	0,008	0,005	0,005
Wald (joint)	951**	1394**	1051**
AR(1)-test	0,572	-0,986	0,732
AR(2)-test	-2,349*	-3,257**	-2,788**
	Langtidsledighetskoeffisient		
	-0,007 (-3,21)	-0,008 (-6,83)	-0,005 (-1,69)

Tabell 13: Effekt av den inverse gruppebestemte regionale arbeidsledigheten og vakanseratestrøm på lønn for ulike utdanningsgrupper. Avhengig variabel $\log\text{lønn}[X]_{it}$.

Variabler	Gruppe		
	GRU	VGS	HØY
$\log(W_{it-1})$	0,762 (25,8)	0,732 (27,6)	0,752 (32,5)
U_{it}^{-1}	0,024 (2,44)	0,03 (5,46)	0,002 (2,07)
U_{it-1}^{-1}	-0,001 (-0,113)	-0,012 (-2,15)	-0,001 (-1,14)
$\log(\text{Vak}S_{it})$	0,001 (0,775)	0,002 (1,08)	0,004 (2,69)
$\log(\text{Vak}S_{it-1})$	0,002 (0,938)	-0,001 (-0,954)	-0,002 (-1,40)
Oservasjoner	810	810	810
R^2	0,998	0,999	0,999
σ	0,008	0,005	0,005
Wald (joint)	794,6**	1120**	1075**
AR(1)-test	1,104	-0,518	1,027
AR(2)-test	-2,045*	-3,038**	-2,813**
	Langtidsledighetskoeffisient		
	0,100 (2,10)	0,067 (3,51)	0,004 (0,70)

A.2 Korrelasjonsmatrise.

Tabell 14: Korrelasjonsmatrise

	IWG	IWV	IWH	IU	IUG	IUV	IUH	LVakS	IBPNY	OT	Sommertemp	Vintertemp
IWG	1.0000	0.98753	0.97502	-0.16702	-0.18669	-0.16915	0.11998	-0.27296	0.72072	-0.65637	0.32542	0.051128
IWV	0.98753	1.0000	0.98472	-0.17609	-0.16753	-0.16453	0.12685	-0.26175	0.76526	-0.64980	0.36423	0.041580
IWH	0.97502	0.98472	1.0000	-0.15038	-0.14355	-0.13556	0.15538	-0.28929	0.78837	-0.62359	0.36444	-0.021430
IU	-0.16702	-0.17609	-0.15038	1.0000	0.84817	0.95134	0.70363	-0.28311	-0.19809	0.65171	-0.070605	-0.034072
IUG	-0.18669	-0.16753	-0.14355	0.84817	1.0000	0.86621	0.59110	-0.094051	-0.14945	0.51973	0.0098975	-0.063277
IUV	-0.16915	-0.16453	-0.13556	0.95134	0.86621	1.0000	0.69704	-0.17817	-0.13212	0.59109	-0.020714	-0.039965
IUH	0.11998	0.12685	0.15538	0.70363	0.59110	0.69704	1.0000	-0.29974	0.17824	0.33495	0.14003	0.0029372
LVakS	-0.27296	-0.26175	-0.28929	-0.28311	-0.094051	-0.17817	-0.29974	1.0000	-0.21604	-0.17152	-0.16855	0.18719
IBPNY	0.72072	0.76526	0.78837	-0.19809	-0.14945	-0.13212	0.17824	-0.21604	1.0000	-0.47558	0.43401	0.033050
OT	-0.65637	-0.64980	-0.62359	0.65171	0.51973	0.59109	0.33495	-0.17152	-0.47558	1.0000	-0.19095	-0.063977
Sommertemp	0.32542	0.36423	0.36444	-0.070605	0.0098975	-0.020714	0.14003	-0.16855	0.43401	-0.19095	1.0000	0.022894
Vintertemp	0.051128	0.041580	-0.021430	-0.034072	-0.063277	-0.039965	0.0029372	0.18719	0.033050	-0.063977	0.022894	1.0000

A.3 Hausman test

Tabell 15: Hausman test

	Koeffisienter		(FE - RE)	Standardavvik
	FE	RE		
$\log(W_{it-1})$.7728742	.9890659	-.2161917	.0253922
$\log(U)$	-.0259797	-.0236483	-.0023314	.0008024
$\log(U_{it-1})$.0119904	.0230236	-.0110332	.0011545
$\log(VakS)$	-.0012128	-.0010798	-.000133	.0008201
$\log(VakS_{it-1})$.001781	.0025656	-.0007846	.0005772
Årsdummy				
1997	.017262	.0079622	.0092998	.0013394
1998	.0410822	.0213564	.0197259	.0029752
1999	.0431426	.0095846	.033558	.0047465
2000	.0548289	.0081575	.0466714	.0062193
2001	.0638402	.0047809	.0590592	.007546
2002	.0882805	.0179121	.0703684	.0087548
2003	.0871229	.0016251	.0854978	.0102945
2004	.0937424	-.0029254	.0966678	.0114944

Merk: Hausman test benytter en kjikvadratfordelt test der H_0 : «begge metodene er riktige og gir like koeffisienter». Hypotesten forkastes med en høy og signifikant verdi: $\chi^2 = 95.20$.