

Sykefraværet og sammensetningen av arbeidsstyrken

av

Stian Benonisen Thormodsæter

Masteroppgave

Masteroppgaven er levert for å fullføre graden

Master i samfunnsøkonomi

Universitetet i Bergen, Institutt for økonomi

Juni 2011

UNIVERSITETET I BERGEN



Forord

Jeg vil først og fremst gi en stor takk til min veileder Kjell Vaage, for hjelp med tilrettelegging av data og gode råd underveis i oppgaveskrivingen.

Jeg vil også takke mine hyggelige medstudenter på institutt for økonomi for å ha bidratt til en fantastisk studietid.

En spesiell takk rettes også til Teis for hjelp med korrekturlesing av oppgaven.

Data som er benyttet i denne oppgaven er hentet fra Statistisk Sentralbyrå (SSB) sin database FD-Trygd. Registerdata med informasjon om inntekt og utdanning er hentet fra SSB og påkopledd datasettet. Data om arbeidsledighet er hentet fra Norsk Samfunnsvitenskapelig Datatjeneste (NSD). Ingen av de overnevnte institusjoner er ansvarlig for de analyser og tolkninger som er gjort her.

Stian B. Thormodsæter

Stian Benonisen Thormodsæter, Bergen 21. juni 2011

Sammendrag

Sykefraværet og sammensetningen av arbeidsstyrken

av

Stian Benonisen Thormodsæter, Master i samfunnsøkonomi

Universitetet i Bergen, 2011

Veileder: Kjell Vaage

Formålet med denne empiriske oppgaven er å undersøke om variasjonen i sykefraværet over konjunktursyklusene skyldes endringer i sammensetningen av arbeidsstyrken. Sykefraværet kan vises å være negativt korrelert med arbeidsledigheten. Det er vanlig å trekke frem to ulike forklaringer på denne sammenhengen. Økonomiske nedgangstider kan disiplinere arbeidstakere til å redusere sykefraværet sammenlignet med en høykonjunktur. Alternativt, kan sammenhengen reflektere endring i sammensetningen av arbeidsstyrken, hvor arbeidere med dårligere helse/funksjonsevne i større grad blir ansatt når sysselsettingen øker.

Problemstillingen bygger på en tidligere analyse av Askildsen et al. (2005), som undersøker om sykefraværet til arbeidere med stabil arbeidsmarkedstilknytning påvirkes av arbeidsledigheten. Data som er brukt i oppgaven er et paneldatasett basert på forløpsdatabasen FD-Trygd og påkoplede registeropplysninger om inntekt og utdanning. Datasettet inneholder kun informasjon om sykefravær som strekker seg over 16 dager (14 dager før 1998). I analysen er sannsynligheten for sykefravær estimert med en fast effekt logitmodell og varigheten av sykefravær estimeres med en lineær fast effekt-modell. Statistikkprogrammet Stata 11.1 brukes i beregningene. Resultatene indikerer at både sannsynligheten for sykefravær og antall sykefraværsdager påvirkes negativt av arbeidsledigheten. For arbeidere med stabil arbeidsmarkedstilknytning påvises det en tilsvarende negativ, men relativt svakere effekt.

Innholdsfortegnelse

Forord.....	ii
Sammendrag.....	iii
Innholdsfortegnelse.....	iv
Tabeller.....	vii
Figurer.....	viii
Kapittel 1: Innledning.....	1
1.1 Bakgrunn.....	1
1.2 Problemstilling.....	2
1.3 Videre oppbygning av oppgaven.....	3
Kapittel 2: Sykepengeordningen og utviklingen i sykefraværet.....	4
2.1 Sykepengeordningen.....	4
2.2 Reformer i sykepengeordningen.....	5
2.2.1 Ot. Prp 48 2003-2004.....	6
2.2.2 Intensjonsavtalen om et inkluderende arbeidsliv (IA).....	7
2.2.3 Aktiviserings- og nærværsreformen.....	7
2.2.4 Oppsummering av reformene.....	9
2.3 Sykefraværet.....	9
Kapittel 3: Sykefravær og økonomisk teori.....	13
3.1 Økonomiske modeller for sykefraværet.....	14
3.1.1 Den neoklassiske arbeidstilbudsmodellen.....	14
3.1.2 Effektivitetslønnsteori.....	15
3.1.3 Sykepengeordningen som sosialforsikring.....	15
3.2 Sykefraværet og konjunktorene.....	16
3.2.1 Sammensetningshypotesen.....	16
3.2.2 Disiplineringshypotesen.....	17

3.2.3	Stresshypotesen	18
3.3	Nyere hypoteser.....	18
3.3.1	Omorganisering og nedbemanning.....	18
3.3.2	Sosiale normer	19
3.3.3	Legenes portvaktrolle	19
3.4	Tidligere forskning	20
3.4.1	Norske studier.....	20
3.4.2	Utenlandske studier	22
Kapittel 4: Økonometrisk metode		24
4.1.	Paneldata	24
4.2.	Lineære paneldatamodeller	25
4.2.1.	MKM.....	26
4.2.2.	Tilfeldig effekt-estimatoren.....	27
4.2.3.	Fast effekt-estimatoren	27
4.3.	Ikke-lineære paneldatamodeller	28
4.3.1.	Fast effekt logitmodell.....	30
4.4.	Fast eller tilfeldig effekt?.....	30
4.4.1.	Hausmantest	31
4.5.	Estimeringsutfordringer.....	32
4.5.1.	Endogenitet.....	32
4.5.2.	Selektert utvalg.....	33
Kapittel 5: Databeskrivelse		34
5.1.	Datamaterialet og utvalgene	34
5.2.	Variablene	43
5.2.1.	Avhengig variabel: Sykefravær	43
5.2.2.	Uavhengige variabler.....	44
5.3.	Deskriptiv statistikk.....	52

Kapittel 6: Empiriske resultater.....	54
6.1. Fast eller tilfeldig effekt?.....	54
6.2. Resultater fra fast effekt logitmodellen	55
6.2.1. Tolkning av koeffisienter	58
6.3. Resultater fra den lineære fast effekt-modellen.....	64
6.3.1. Tolkning av koeffisienter	67
Kapittel 7: Oppsummering	70
Litteraturliste	72
Appendiks A: Estimering av 6.1 med lineær fast effekt-estimator	76
Appendiks B: Estimering av 6.2 med MKM.....	77
Appendiks C: Utregning av vendepunkter	78

Tabeller

Tabell 1: Fordelingen av sykefravær, 1993-2006.....	39
Tabell 2: Fordelingen av antall sykefravær per år, 1993-2006.....	40
Tabell 3 Sykefraværsdager per arbeider og antall sykefraværstilfeller per arbeider.....	42
Tabell 4: Deskriptiv statistikk.	53
Tabell 5: Resultater fra estimering av fast effekt logitmodellen. Oppgitt som logit-koeffisienter. 1994-2006.....	57
Tabell 6: Resultater fra estimering av den lineære fast effekt-modellen. 1994-2006.	66

Figurer

Figur 1: Erstattet sykefravær.	11
Figur 2: Sykefraværet i Norge sammenlignet med andre land.	12
Figur 3: Sykepengedager og nye sykepengetilfeller.....	37
Figur 4: Antall sykefraværsdager per arbeider, fordelt på kjønn.	41
Figur 5: Antall individer med sykepengetilfeller som ikke er gradert, og antall individer med graderte sykepengetilfeller.	43
Figur 6: Sykefraværsdager per arbeider og kommunal arbeidsledighet (årgjennomsnitt).....	45

Kapittel 1: Innledning

1.1 Bakgrunn

Under LO Stats Kartellkonferanse i 2009 uttrykte statsminister Jens Stoltenberg sin bekymring for det norske sykefraværet. I talen påpekte han at sykefraværet ligger på et høyere nivå enn det burde og at nivået er økende.

“Vi trenger en debatt om hvordan vi kan gjøre noe med nivå og vekst, fordi kostnadene er for høye. Det er et alvorlig problem. Både for den sykmeldte, for den enkelte arbeidsplass som mister kompetanse og for fellesskapet som mister produktiv arbeidskraft og inntekter, og som påføres nye utgifter”(Regjeringen 2009).

Høy yrkesdeltakelse, særlig blant kvinner og eldre, kan delvis forklare det høye sykefraværet i Norge sammenlignet med andre land. Stoltenberg er likevel enig med de fleste forskere innenfor dette fagfeltet om at dette ikke er en tilstrekkelig forklaring, derfor vil en meningsfull debatt kreve bedre kunnskap om faktorer som kan påvirke nivået og utviklingen i sykefraværet.

Sykefraværet i Norge er et viktig økonomisk tema fordi høyt fravær fører til kostnader både for samfunnet som helhet og for de enkelte individer som blir rammet og dermed midlertidig mister muligheten til å arbeide. Statens budsjetterte utgifter til sykepenger er anslått til 36,4 milliarder kroner for 2011. Dette er rene overføringer, altså er kostnader i forbindelse med tiltak rettet mot å redusere fraværet eller utgiftene til helsevesenet i forbindelse med sykdom hos arbeidstakere ikke medregnet. Arbeidsgivernes kostnader som følger av produksjonstap er heller ikke inkludert. Tall fra NAV viser at sykefraværet førte til over 7,1 millioner tapte dagsverk i 2010. Sykefravær har også konsekvenser for den enkelte arbeider. Selv om den norske sykepengeordningen gir 100 % lønnskompensasjon inntil et visst beløp fra første fraværsdag, indikerer nyere forskning at individer som kommer tilbake i arbeid etter sykefravær opplever en dårligere lønnsutvikling og svekkede karrieremuligheter (Markussen 2010). Langvarig sykefravær øker også sannsynligheten for å havne på mer langvarige trygdeytelser, og til slutt ende som arbeidsufør (Gjesdal & Bratberg 2003).

Utbetalinger fra Folketrygden er den desidert største utgiftsposten i Statsbudsjettet. Andelen av Folketrygden som går til sykepenger, er ca. 11 % i budsjettet for 2011. Derfor er det

knyttet stor interesse til utviklingen i sykefraværet og temaet blir ofte diskutert i nasjonale medier. En vanlig oppfatning er at Norge har et svært høyt sykefravær sammenlignet med andre land. Dette er ikke en uproblematisk påstand, og komplikasjoner knyttet til slike sammenligninger diskuteres blant annet av Bye et al. (2010). Likevel er det liten tvil om at det norske sykefraværet er høyt i forhold til land vi ønsker å sammenligne oss med. Sverige og Nederland har historisk sett hatt sykefravær på samme nivå som Norge, men har gjennomført reformer i starten av 2000-tallet som har redusert fraværet i begge land.

Debatten rundt sykefraværet kan være vanskelig på grunn av motsetningen mellom graden av sjenerøsitet i kompensasjonen og treffsikkerheten til sykepengeordningen. Treffsikkerheten vurderes ut ifra om ytelsene faktisk går til arbeidstakere som blir rammet av helseproblemer, det vil si at ordningen i minst mulig grad blir misbrukt. På den ene siden ønsker man at kompensasjonsgraden skal være så høy som mulig slik at nyttetapet som påføres arbeidstakere med uforutsette helseproblemer skal minimeres, på den andre siden ønsker man ikke utilsiktede effekter på arbeidstilbudet til disse individene. Det har blitt påvist i forskningslitteraturen at det er en positiv sammenheng mellom sjenerøsiteten i sykepengeordningen og nivået på sykefraværet¹. Dette illustrerer viktigheten av å klarlegge hvilke insentiver som ligger til grunn for enkeltindividers arbeidstilbud og på hvilken måte disse endres som følge av ulike utforminger av trygdesystemet.

1.2 Problemstilling

Sykefraværet har svingt kraftig de siste 15-20 årene. I noen land har det vist seg å være en sammenheng mellom sykefraværet og de makroøkonomiske konjunktorene. I perioder med lav arbeidsledighet har vi kunnet observere høyt sykefravær og omvendt. Denne sammenhengens var tydelig i Norge i perioden 1989-2000². Etter dette har sammenhengens vært svakere. Mest sannsynlig eksisterer det også andre faktorer som påvirker utviklingen i sykefraværet, slik som yrkesdeltakelse, sammensetningen av arbeidsstyrken, normendringer og økende grad av omstilling og effektivisering i arbeidslivet. Det er av stor interesse å finne hva som driver sykefraværet, siden resultatet vil ha betydning for hvilke typer tiltak man skal iverksette for å redusere fraværet.

¹ Se f.eks. Johanson & Palme (2005) eller Henrekson & Persson (2004).

² Se f.eks. Askildsen et al. (2005), Nossen (2009) eller Dyrstad & Ose (2002)

En teori om hvorfor sykefraværet svinger over konjunktorene tar utgangspunkt i at arbeidere med dårligere helse/funksjonsevne, såkalte “marginale arbeidere”, utgjør en større andel av arbeidsstyrken i perioder med høy sysselsetting³.

Problemstillingen i denne oppgaven bygger på en tidligere studie av Askildsen, Bratberg & Nilsen (2005), som heretter refereres til som ABN, og er formulert slik: er variasjonen i sykefraværet over konjunktursyklusene et resultat av at marginale arbeidere går inn i arbeidsmarkedet når sysselsettingen er høy, og forlater arbeidsmarkedet når sysselsettingen er lav?

For å svare på spørsmålet vil vi sammenligne to grupper arbeidere: De som ifølge vår definisjon *ikke* er marginale, i den forstand at de befinner seg i arbeidsmarkedet over hele vår observasjonsperiode, og et representativt utvalg av alle andre arbeidere (inkludert de marginale). Hvis den stabile gruppen sin tilbøyelighet for sykefravær blir påvirket negativt av lav sysselsetting/ høy arbeidsledighet, tolkes dette som et argument mot sammensetningseffekten.

1.3 Videre oppbygning av oppgaven

I kapittel 2 blir det gitt en innføring i institusjonelle forhold bak sykepengeordningen, samt de viktigste historiske reformene som har blitt gjennomført. Utviklingen i sykefraværet blir presentert i avsnitt 2.3, med fokus på det norske fraværet. Økonomisk teori for sykefravær og tidligere forskning blir presentert i kapittel 3. Den økonometriske metoden som benyttes i analysen gjennomgås i kapittel 4. En presentasjon av datamaterialet og variablene blir gitt i kapittel 5, mens de empiriske resultatene presenteres og tolkes i kapittel 6. Kapittel 7 gir en oppsummering av de viktigste funnene i oppgaven.

³ Hypotesen kalles ”sammensetningshypotesen” og blir nærmere forklart i avsnitt 3.2.1

Kapittel 2: Sykepengeordningen og utviklingen i sykefraværet

I dette kapittelet presenteres hovedinnholdet i sykepengeordningen, samt de viktigste historiske reformene. Utviklingen i det norske sykefraværet beskrives i avsnitt 2.3, og utviklingen sammenlignes med sykefraværet i andre land.

2.1 Sykepengeordningen

Sykepengeordningen er regulert av Folketrygdloven. Medlemskap i Folketrygden er, med noen unntak, automatisk og pliktig for alle som bor og arbeider i Norge. Medlemsavgiften blir betalt gjennom beskatning og Folketrygden gir økonomisk støtte ved sykdom, svangerskap og fødsel, arbeidsløshet, alderdom, uførhet, dødsfall og tap av forsørger. Ved tap av arbeidsinntekt har medlemmer av Folketrygden rett på kompensasjon avhengig av årsaken til, og varigheten på, fraværet. For selvstendig næringsdrivende gjelder egne bestemmelser. Ved midlertidig fravær fra arbeidslivet på grunn av sykdom, har arbeidstakere rett til enten sykepenger eller arbeidsavklaringspenger (AAP) så lenge målet er at arbeidstakeren skal tilbake i jobb. Ved varig uførhet skal personen over på uførepensjon. Formålet med sykepengeordningen er beskrevet i folketrygdloven § 8-1:

”Formålet med sykepenger er å gi kompensasjon for bortfall av arbeidsinntekt for yrkesaktive medlemmer som er arbeidsuføre på grunn av sykdom eller skade.”

Arbeidstaker må ha vært ansatt hos arbeidsgiveren i minst fire uker for å ha rett til sykepenger og det må foreligge tap av arbeidsinntekt. Inntektsgrunnlaget må utgjøre minst 50 % av grunnbeløpet og sykepengegrunnlaget kan ikke overstige 6 ganger grunnbeløpet i Folketrygden⁴. Store grupper i arbeidslivet har likevel inngått tariffavtaler eller arbeidsavtaler som gir full lønn under sykdom uten begrensning av 6 G-regelen, hvorav differansen utbetales av arbeidsgiver. Antall sykepengedager er begrenset til ett kalenderår i løpet av de tre siste årene. Ny rett opptjenes likevel etter 26 uker med sammenhengende arbeidsførhet. Sykepenger ytes fra første dag med 100 % kompensasjon for tapt arbeidsinntekt. De 16 første kalenderdagene dekkes av arbeidsgiver (arbeidsgiverperioden), deretter betales sykepengene

⁴ Grunnbeløpet (G) er et sentralt begrep i folketrygden og brukes som grunnlag til å beregne trygde- og pensjonsutbetalinger. Størrelsen på grunnbeløpet vedtas av Stortinget gjennom forhandlinger i det årlige trygdeoppgjøret. Fra 1.mai 2011 er grunnbeløpet på 79 216 kroner.

fra folketrygden⁵. I visse tilfeller hvor arbeidstakeren har langvarig eller kronisk sykdom, kan det søkes om at trygden dekker utgiftene også i arbeidsgiverperioden. Det kan i tillegg tegnes forsikring mot ansvar for sykepenger i arbeidsgiverperioden, dette gjelder spesielt arbeidsgivere med få ansatte.

Etter sykepengeordningens utløp kan det søkes om arbeidsavklaringspenger (AAP). AAP er en ny støtte som fra 1.mars 2010 har erstattet de tidligere ytelsene atferingspenger, rehabiliteringspenger og tidsbegrenset uførestøtte. Hensikten er å hjelpe personer tilbake i arbeid etter perioder med sykdom eller skade. Bistanden kan bestå av arbeidsrettede tiltak, medisinsk behandling eller annen oppfølging fra NAV. Dekningsgraden er i utgangspunktet 66 % av inntektsgrunnet, men flere forhold er av betydning for kompensasjonsnivået. Blant annet om personen er barneforsørger, mottar andre folketrygdytelse eller lider av yrkesrelatert skade eller sykdom. For å kunne motta AAP, må arbeidsevnen være nedsatt med minst 50 % som følge av sykdom eller skade og støtten kan vanligvis ikke mottas i mer enn fire år. Bakgrunnen for at AAP erstattet de gamle ordningene, var at Stortinget ønsket et enklere regelverk der trygdemottakere skulle slippe å bli sendt frem og tilbake mellom ulike støtteordninger.

2.2 Reformen i sykepengeordningen

I denne delen av oppgaven presenteres de viktigste reformene i sykepengeordningens historie. Utformingen av sykepengeordningen påvirker arbeidstilbudet til aktører i arbeidsmarkedet. Reformene illustrerer hvordan regjeringen kan påvirke insentivstrukturen til både arbeidere og arbeidsgivere. Innholdet i den norske sykepengeordningen illustrerer likevel at hensynet til fordeling og rettferdighet står sterkest. Først presenteres reformene, før insentivvirkningene oppsummeres i avsnitt 2.2.4.

Den første loven om sykeforsikring ble vedtatt 18. september 1909 og innførte blant annet rett til sykepenger for lavtlønnede arbeidstakere. Siden dette har ordningen blitt revidert ved flere anledninger. Sykepengeordningen ble innlemmet i folketrygden i 1971 og retten til full lønn ved sykefravær for arbeidstakere ble innført i 1978. Sistnevnte reform kan sees på som

⁵ Disse reglene gjelder for arbeidstakere. Det er utarbeidet spesielle regler for selvstendig næringsdrivende og andre grupper, som ikke vil bli videre diskutert her.

inngangen til dagens sykepengeordning der karensdagene⁶ ble fjernet og arbeidsgiverperioden på 14 dager og bruk av egenmelding ble innført. Trygdeetaten krevde fra 1988 også en særskilt begrunnet legeerklæring, "Sykemeldingsattest 2", etter 8 ukers sykemelding. Målet med denne reformen var å redusere antall langtidsfravær gjennom muligheten til å spesifisere behandlings- og attføringstiltak.

Arbeidsgiverperioden ble i 1998 utvidet fra 14 til 16 dager, og maksimaltiden for sykepenger fra trygden redusert til 248 sykepengedager etter utløpet av arbeidsgiverperioden. Fra lovteksten fremgår det at "hensikten med endringen er å motivere arbeidsgivere til å sette i verk tiltak for å forebygge sykefravær, og sikre en tidlig oppfølging av sykemeldte." Det er da underforstått at arbeidsgiverne skulle få sterkere insentiver til å redusere sykefraværet da de måtte dekke en større del av kostnadene forbundet med fravær.

2.2.1 Ot. Prp 48 2003-2004

Fra 1. januar 2004 ble opptjeningstiden for retten til sykepenger utvidet fra to til fire uker. 1. juli samme år ble det innført et sett med regelendringer i Folketrygdloven (Ot. Prp 48 2003-2004) som hadde som formål å redusere sykefraværet. Regelendringene bestod i alt av seks deler, der hovedfokuset var å forbedre legenes portvaktrolle. Det ble lagt opp til at legene sjeldnere skulle benytte *aktiv* sykemelding og oftere benytte *gradert* sykemelding. Aktiv sykemelding betyr at den ansatte kan være tilstede på arbeidsplassen og utføre alternative oppgaver, mens gradert sykemelding skal brukes i tilfeller hvor den ansatte kan utføre sine vanlige arbeidsoppgaver. Den viktigste forskjellen mellom ordningene er at NAV fullfinansierer sykepengene til arbeidstakere på aktiv sykemelding de dagene de er i arbeid, mens for gradert sykemelding må arbeidsgiver lønne arbeidet ut fra graden av arbeidsførhet. Aktiv sykemelding gir arbeidsgiveren ingen økonomiske insentiver til å legge til rette for friskmelding. Bruk av gradert sykemelding skal motivere arbeidsgiver til tettere oppfølging av den sykemeldte med sikte på raskest mulig friskmelding.

Som en del av regelendringene ble også fastlegen pålagt å sende inn en utvidet legeerklæring i de tilfeller der arbeidsrelatert aktivitet ikke er igangsatt etter 8 uker sykemelding (Aktivitetskrav). Hvis slik aktivitet ikke er mulig skal den utvidede legeerklæringen dokumentere medisinske grunner som er til hinder for slik aktivitet. Reformen åpnet også for

⁶ Karensdagene utgjorde de første dagene av et sykefravær, hvor arbeidstaker ikke fikk betalt sykepenger av arbeidsgiver.

sanksjoner mot leger og arbeidere som ikke følger bestemmelsene, der legene kan miste retten til å skrive ut sykemeldinger og arbeidere kan miste sykelønnen.

2.2.2 Intensjonsavtalen om et inkluderende arbeidsliv (IA)

Den første IA-avtalen mellom parter i arbeidslivet og regjeringen ble signert i desember 2001 og varte i fire år. Bakgrunnen for avtalen var at stadig flere arbeidstakere gikk ut av arbeidslivet og over på langvarige trygdeordninger. Avtalen har blitt videreført to ganger og den gjeldende avtalen ble signert 1.mars 2010 og skal gjelde frem til utgangen av 2013. Det overordnede målet i avtalen er å forebygge og redusere sykefraværet, styrke jobbnærværet og bedre arbeidsmiljøet, samt hindre utstøting og frafall fra arbeidslivet. Blant tiltakene som partene i avtalen har iverksatt for å nå målene, er det mest omfattende grepet opprettelsen av NAV Arbeidslivsentere i hvert fylke. Virksomheter i næringslivet kan inngå avtale med arbeidslivsenteret og får da status som IA-virksomheter. Samarbeidet fører med seg visse forpliktelser fra både myndighetene og virksomhetenes side⁷. Ansatte i IA-virksomheter kan egenmelde seg i åtte dager, mot tidligere tre dager. Et viktig delmål i avtalen er at sykefraværet skal reduseres med 20 % i forhold til nivået i andre kvartal 2001. Dette betyr i praksis at det totale sykefraværet ikke skal overstige 5,6 %.

Som en del av IA-samarbeidet foreslo Sykefraværsutvalget høsten 2006 tiltaket "Raskere tilbake" med målsetting om å hindre unødig langvarig sykefravær. Lovendringene trådte i kraft i mars 2007 og innebærer blant annet tilbud som sørger for at den sykemeldte kan få raskere avklaring, oppfølging og arbeidsrettet rehabilitering gjennom NAV. Det viktigste virkemiddelet er at trygdeetaten kjøper helsetjenester til den sykemeldte for å gi dem kortere ventetid for behandling og dermed raskere retur til arbeidslivet. Det legges også større vekt på tettere oppfølging fra arbeidsgiver gjennom dialogmøter med sykemeldt arbeidstaker.

2.2.3 Aktiviserings- og nærværsreformen

Før signeringen av den gjeldende IA-avtalen, ble det oppnevnt en faglig ekspertgruppe som skulle vurdere mulige administrative tiltak for å redusere sykefraværet. Målet om 20 % reduksjon var ikke nådd, tvert imot var sykefraværet stigende, og i tredje kvartal 2009 var det totale sykefraværet på 7,7 %. De administrative tiltakene som gruppen foreslo, bestod i alt av

⁷ Se (Regjeringen 2010) for den gjeldende IA-avtalens forpliktelser.

seks deler og ble beskrevet som en “Aktiviserings- og nærværsreform.” Det viktigste elementet er økt fokus på aktivitet der gradert sykemelding skal gjøres til hovedregel, samtidig som det skal være en tettere og hyppigere oppfølging fra arbeidsgiver gjennom møter og tilrettelegging. Ekspertgruppen legger vekt på at tiltakene som ble iverksatt i 2004 (der økt aktivitet stod sentralt) førte til et sterkt fall i sykefraværet, men at intensjonen i tiltakene bare delvis var fulgt opp i praksis (Mykletun et al. 2010). Forskning indikerer også at økte aktivitetskrav reduserer både antallet og lengden på sykemeldinger (Markussen et al. 2010). Sannsynligheten for at langtidssykemeldte må over på varige støtteordninger går også ned når aktivitetskravene økes.

Samtidig som den nye IA-avtalen ble signert, ble det også utarbeidet en protokoll mellom partene i arbeidslivet og myndighetene der noen av tiltakene foreslått av ekspertgruppen ble vedtatt. Dette er endringer i Arbeidsmiljøloven og Folketrygdloven som etter planen skal innføres 1.juli 2011 (Arbeidsdepartementet 2010).

- Det skal utarbeides faglig støtte/veiledning for sykemeldingsarbeidet med obligatorisk opplæring for alle som skal sykemelde.
- Oppfølgingsplanen mellom ansatt og arbeidsgiver skal nå leveres etter 4 uker mot tidligere 6. Planen beskriver hva som skal til for at den sykemeldte skal være helt eller delvis i arbeid.
- Dialogmøte 1 mellom ansatt og arbeidsgiver (og evt. bedriftshelsetjeneste eller fastlege) skal nå gjennomføres for alle sykemeldte (også graderte) senest innen 8 uker.
- Ordningen med aktiv sykemelding skal avvikles.

Ekspertgruppen foreslo også en endring i arbeidsgivernes medfinansiering av sykelønnen. Forslaget innebar en forkorting av arbeidsgiverperioden fra 16 til 10 dager, men etter 8 uker skulle arbeidsgiver finansiere 20 % av sykelønnen hvis *ikke* sykemeldingen var gradert. Dette var ment å gi arbeidsgiverne sterkere insentiver til å forebygge langtidsfravær og utstøting, samtidig som den samlede finansieringen holdes tilnærmet uendret. Partene i arbeidslivet ble enige om å diskutere en slik endring i løpet av 2011. Regler for sterkere sanksjoner ved brudd på regelverket skal også gjennomgås.

2.2.4. Oppsummering av reformene

Etter inngangen til dagens sykepengeordning har regjeringen gjennomført flere sentrale reformer. Utvidelsen av arbeidsgiverperioden fra 14 til 16 dager gir arbeidsgiver incentiver til å redusere sykefraværet blant sine ansatte. Overgangen fra aktiv til gradert sykemelding medfører sterkere incentiver for arbeidsgiver til å tilrettelegge for at den sykemeldte skal returnere i full stilling. IA-avtalen ble rettet mot å forbedre forhold på arbeidsplassen, men påvirker i liten grad de økonomiske incentivene til arbeider eller arbeidsgiver. “Raskere tilbake” går i hovedsak ut på å kjøpe behandlingsplasser i helsevesenet til sykemeldte arbeidere, mens aktiviserings- og nærværsreformen viderefører intensjonen bak 2004-reformen. Partene i arbeidslivet har også blitt enige om å diskutere om arbeidsgiver skal finansiere 20 % av sykelønnen dersom sykemeldingen ikke er gradert etter 8 uker. Når det gjelder incentivvirkningene av reformene, er arbeidsgiverne lite/moderat påvirket gjennom avviklingen av aktiv sykemelding og utvidelsen av arbeidsgiverperioden. Arbeidstakers incentiver er tilnærmet uberørt. Dette må sees i sammenheng med de sterke fagforeningene i Norge, som beskytter velferdsordningene til sine medlemmer. Ekspertgruppen bak aktiviserings- og nærværsreformen påpeker i sin rapport at *”justering av kompensasjonen ikke fremstår som aktuell politikk i Norge i dag”*, og ser derfra bort fra redusert kompensasjon som virkemiddel for å redusere fraværet (Mykletun et al. 2010).

Dersom utviklingen i sykefraværet hovedsakelig er forårsaket av marginale arbeidere med nedsatt helse/funksjonsevne som går inn i arbeidsmarkedet ved høy sysselsetting, er det ikke hensiktsmessig å bruke redusert kompensasjonsgrad for å redusere sykefraværet. Arbeidsmarkedstiltak rettet mot å forbedre forholdene til de marginale arbeiderne vil dermed være mer virkningsfulle.

2.3 Sykefraværet

Det er hensiktsmessig å presentere de ulike hovedkildene for sykefraværsstatistikk i Norge:

- *Den sentrale sykefraværstatistikken*: ble først publisert i 2001 og er i dag hovedkilden for data om sykefraværet. Statistikken publiseres av NAV og SSB hvert kvartal og omfatter både egen- og legemeldt sykefravær for alle arbeidstakere. Tall for egenmeldt fravær kartlegges av SSB gjennom en virksomhetsbasert utvalgsundersøkelse, mens det legemeldte fraværet f.o.m. dag 4 (dag 8 for IA-bedrifter) rapporteres til NAV. Målet på

sykefraværet kalles sykefraværsprosenten og måler antall tapte dagsverk på grunn av egen sykdom i prosent av avtalte dagsverk.

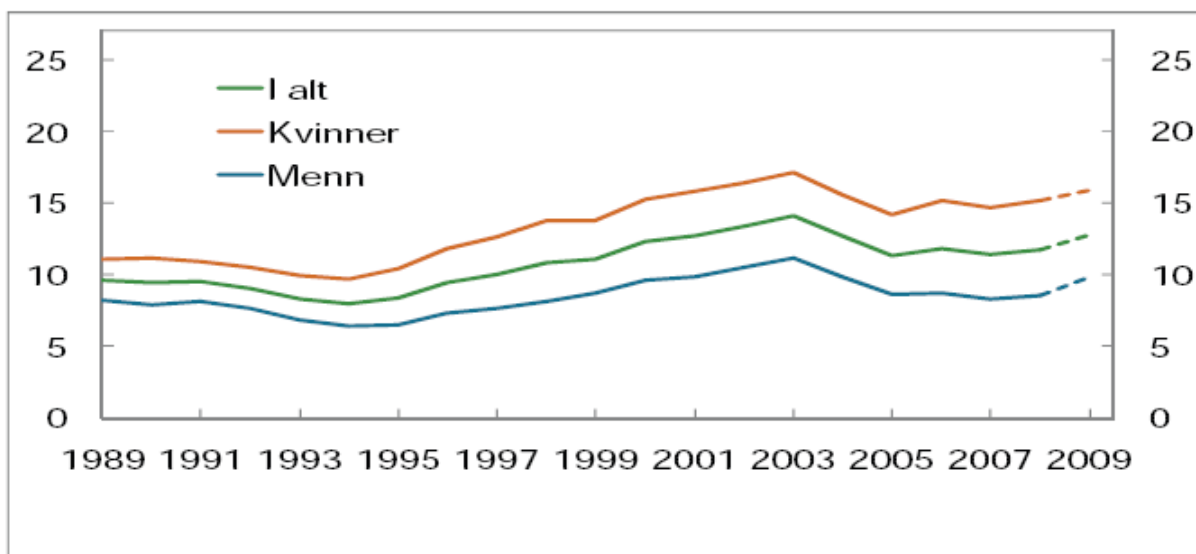
- *NHOs fraværstatistikk*: er basert på et utvalg av medlemsbedriftene og omfatter også begge typer sykefravær. Tidligere var dette den eneste datakilden for egenmeldt fravær. Statistikken ble imidlertid avsluttet i 2002 etter at den sentrale sykepengestatistikken, som omfatter alle næringer, ble innført. Fraværsmålet - sykefraværsprosenten - angir prosentvis fravær i forhold til mulige dagsverk.
- *Arbeidskraftundersøkelsen (AKU)*: er en kvartalsvis spørreundersøkelse blant et tilfeldig utvalg av befolkningen i alderen 15-74 år der intervjuene baseres på en referanseuke. Nivå tallene i AKU er forskjellig fra sykefraværstatistikken, noe som skyldes ulike definisjoner av måltallene. AKU rapporterer sykefraværsprosent beregnet som antall sysselsatte som var fraværende fra arbeid hele referanseuken pga. egen sykdom, i prosent av alle sysselsatte.
- *NAVs sykefraværstatistikk*: har registerdata over sykefravær som innebærer utbetalinger fra folketrygden (sykefravær utover 16 dager fra 1. april 1998). Blant variablene som rapporteres finner vi antall avsluttede sykepengetilfeller og antall erstattede sykepengedager.

Ifølge sykefraværstatistikken fra NAV og SSB var det totale sykefraværet på 7 % i 4. kvartal 2010, hvor 1,1 % var egenmeldt fravær. Ved å undersøke tallene for perioden fra 2001 til 2010 ser vi at det legemeldte sykefraværet varierer mellom 5,5 til 7,2 %, mens det egenmeldte fraværet ligger relativt stabilt rundt ca. 1 % når det justeres for sesongvariasjoner⁸. Annen forskning med utgangspunkt i NHOs fraværstatistikk for perioden 1975-2002 anslår at egenmeldt fravær har ligget relativt stabilt rundt 1 % også i denne perioden (Gjesdal 2005). Det er altså legemeldt fravær, og spesielt langtidsfravær, som utgjør den største delen av variasjonen. Tiltakene som blir innført for å redusere sykefraværet, bør da rettes mot å redusere langtidsfraværet.

Siden sykefraværstatistikken først ble utgitt i 2001 er det vanskelig å se på endringene over en lengre tidsperiode med utgangspunkt i dette datagrunnlaget. Nossen & Thune (2009) bruker NAVs statistikk over sykepenger betalt av folketrygden i årene 1989-2008 for å se på

⁸ Det egenmeldte fraværet er høyest de tre første månedene i året og dette har i følge statistikkbeskrivelsen fra Folkehelseinstituttet en sammenheng med influensasesongen som normalt faller på denne tiden av året.

svingninger i sykefraværet over denne perioden. De finner at det totale sykefraværet har økt fra 9,6 til 11,7 dager per lønnstaker per år, som tilsvarer en økning på 22 %. Antall fravær økte svært lite, mens lengden på hvert fravær økte med gjennomsnittlig 19 %. Den totale variasjonen innenfor denne tidsperioden kan kort oppsummeres som en nedgang frem mot 1994, etterfulgt av en kraftig økning helt til 2003 da det brått gikk ned i perioden mot 2005. Fra 2005 til 2009 har det vært en svak økning, men tall fra NAV viser at sykefraværet er på vei ned igjen i 2011. Ser vi bort fra den brå nedgangen i 2004 som etterfulgte reformen i sykepengeordningen, har det vært en svak trendmessig vekst i sykefraværet siden midten av 1990-tallet (se figur 1). Kvinner har i gjennomsnitt høyere fravær enn menn, noe som gjelder for alle yrker (Mykletun et al. 2010). I løpet av perioden 1989 til 2008 har kvinners sykefravær økt med 37 %, mens det for menn har økt med 5 %. Etter finanskrisen i 2008 økte imidlertid sykefraværet kraftig for menn, særlig langtidsfraværet, mens det var en moderat stigning i fraværet for kvinner (Nossen 2010). Dette må sees i sammenheng med at menn er overrepresentert i bygg og anleggsbransjen, som er svært konjunkturutsatt og hadde den største økningen i sykefraværet av samtlige bransjer.

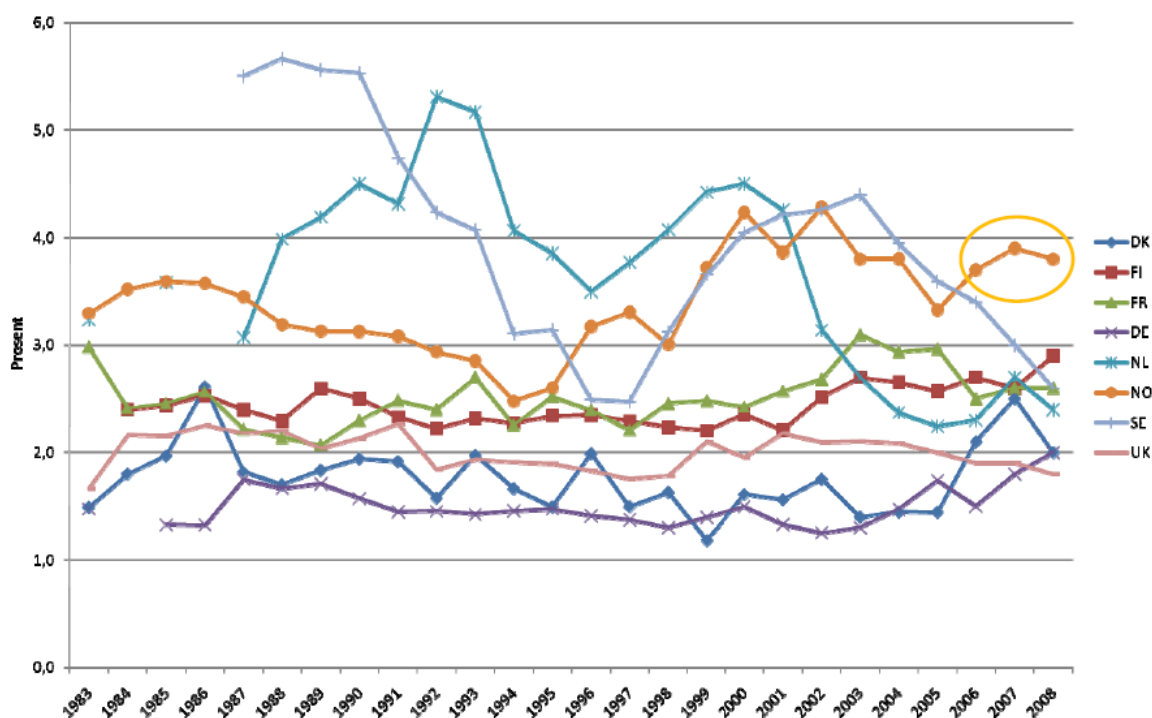


Figur 1: Erstattet sykefravær.

Sykepengedager betalt av Folketrygden pr. lønnstaker pr. år 1989-2008. Anslag for 2009 (stiplet). Korrigerert for endring i arbeidsgiverperiode fra 14 til 16 dager og inkludering av statsansatte i regnskapsstatistikken. (Kilde: Arbeids- og Velferdsdirektoratet)

Som nevnt innledningsvis er det ikke uproblematisk å sammenligne sykefraværet mellom land. Sykepengeordningene er forskjellige på flere måter, blant annet hvor lenge en ansatt kan være sykemeldt (og fremdeles klassifiseres som “ansatt”), og hvor stor del av fraværet som er finansiert av det offentlige. Basert på arbeidskraftsundersøkelsene er sykefraværet i Norge

høyt sammenlignet med andre land (se figur 2). Nederland og Sverige er de landene som historisk har hatt tilsvarende nivå som Norge i sykefraværet. Bonato og Lusinyan (2004) trekker frem høy yrkesdeltakelse og sjenerøse sykepengeordninger som forklaring på det høye fraværet i disse landene. Det synes likevel ikke å være en entydig sammenheng mellom nivået i sykefraværet og antall sysselsatte i prosent av befolkningen. Flere land har om lag like høy andel sysselsatte som Norge, samtidig som sykefraværet er lavere. Det er også kommuner i Norge med høy sysselsetting som har lavere sykefravær enn kommuner med lav sysselsetting (Mykletun et al. 2010) Det viser seg også at Norge har nesten dobbelt så stor andel av den yrkesaktive befolkningen på uføretrygd som OECD-gjennomsnittet. Det er ikke bare nivået på sykefraværet som er viktig, men også trenden i utviklingen. For Norge sin del har sykefraværet som nevnt hatt en trendmessig vekst siden starten av 90-tallet. I både Nederland og Sverige har sykefraværet avtatt de siste 10 årene, trolig på grunn av omfattende reformer i sykepengeordningen på 2000-tallet⁹.



Figur 2: Sykefraværet i Norge sammenlignet med andre land.

Basert på AKU undersøkelser i de respektive land. Kilde: EUROSTAT. Figur hentet fra Mykletun et al (2010)

⁹ Reformene er relativt ulike i de to landene. Nederland har økt finansieringsansvaret til arbeidsgiveren, mens Sverige har satset på flere områder samtidig.

Kapittel 3: Sykefravær og økonomisk teori

Individuell helsetilstand er utvilsomt den viktigste årsaken til at arbeidere går ut i sykefravær, men også andre sosiale og økonomiske forhold har en påvirkning – enten direkte eller via helsetilstanden. Blant annet vil sammensetningen av befolkningen i yrkesaktiv alder, yrkesdeltakelsen, legers sykemeldingspraksis, endringer i regelverk og oppfølgingspraksis kunne endre utviklingen i sykefraværet (Bjørnstad 2006). På tross av at individuell helsetilstand har stor påvirkning på ett enkelt sykefravær, er det lite sannsynlig at utviklingen i folkehelsen kan forklare variasjonen i det totale sykefraværet de siste årene. Undersøkelser viser at befolkningen i Norge er en av de friskeste i verden. Vi har gode arbeidsforhold, god selvrapportert helse og positiv utvikling i levealder (Mykletun et al. 2010, Vedlegg 7).

Den økonomiske teorien som omhandler sykefraværet, er ofte rettet mot å forklare hvordan individers insentiver blir påvirket av utformingen av kompensasjonsordningene. Siden sykepengeordningen erstatter tapt arbeidsinntekt, er det naturlig at den påvirker arbeidstilbudet. Det er vanlig å ta utgangspunkt i den neoklassiske modellen for arbeidstilbud for å forklare hvordan individer vil tilpasse valget mellom inntekt (konsum) og fritid gitt en rekke eksogene forhold. En kort gjennomgang av den neoklassiske modellen blir gitt i avsnitt 3.1.1.

Flere forfattere hevder at det er en positiv sammenheng mellom sykefraværet og konjunktorene¹⁰. Teorier for denne sammenhengen blir drøftet i avsnitt 3.2. Det kan se ut som at den prosykliske variasjonen mellom sykefraværet og konjunktorene har avtatt de senere årene. Nyere teorier har fulgt i kjølvannet av dette, der sykefraværet knyttes til økt konkurranse, globalisering og omstilling i arbeidsmarkedet. Endringer i sosiale normer er også en mulig forklaring på utviklingen i sykefraværet. De nye teoriene presenteres i avsnitt 3.3. Til slutt gjennomgår vi tidligere forskning, med fokus på analyser som beskriver sammenhengen mellom konjunktorene og sykefraværet.

¹⁰ Se f.eks. ABN, Nossen (2008) eller Dyrstad & Ose (2002).

3.1 Økonomiske modeller for sykefraværet

3.1.1 Den neoklassiske arbeidstilbudsmodellen¹¹

Modellen tar utgangspunkt i et representativt individ med preferanser for inntekt og fritid. Inntekten blir generert av lønnet arbeid, og lønnsatsen blir en alternativkostnad, eller pris, for fritid. Individet vil tilpasse seg slik at det på marginen er indifferent mellom å jobbe litt mer eller å ta ut litt mer fritid. Hvis individet har inntekt utenom arbeidsinntekt, er det normalt å anta at for en gitt tilpasning vil han/hun ønske å arbeide mindre hvis denne inntekten øker, med andre ord vil konsumet av fritid øke. Inntektseffekten er ifølge denne antagelsen positiv, noe som betyr at vi ser på fritid som et normalt gode. Økes lønnsatsen vil også inntekten øke, men det betyr i tillegg at alternativkostnaden ved fritid har blitt høyere. Substitusjonseffekten alene gir insentiver til å jobbe mer og konsumere mindre fritid. Ved en lønnsøkning har vi altså to motstridende effekter, men det er vanlig å anta at substitusjonseffekten dominerer slik at høyere lønn gir høyere arbeidstilbud¹². Vanligvis er det ikke mulig å velge antall arbeidstimer fritt, men arbeidsgiver tilbyr en meny av ulike kontrakter. Betingelsene i arbeidskontraktene vil påvirke det observerte fraværet (Barmby et al. 1995). Hvis individet er feiltilpasset slik at avtalte arbeidstimer ikke passer med preferansene, så kan antall arbeidstimer reduseres ved å bruke sykemelding.

I en slik enkel modell vil man ved 100 % lønnskompensasjon forvente at alle individer vil ta ut det maksimale antallet sykedager. Dette passer ikke med empirien. Modellen kan utvides ved å inkludere en straffefunksjon som gir en ulempe ved fravær som ikke er direkte knyttet til tap av arbeidsinntekt. Reduserte muligheter for forfremmelse, dårligere forhandlingskraft i fremtidige lønnsforhandlinger eller sosial stigmatisering på arbeidsplassen er noen mulige ulemper ved sykefravær. Sammenhengen mellom sykefraværet og konjunktorene kan inkluderes i modellen ved å ta med arbeidsledigheten i straffefunksjonen, der høyere arbeidsledighet øker kostnadene ved sykefravær.

¹¹ Se Allen (1981) for en detaljert fremstilling.

¹² Barmby et al. (1995) finner at en lønnsøkning påvirker fraværet negativt slik at substitusjonseffekten dominerer. Det finnes likevel teorier om en "bakoverbøyd arbeidstilbudskurve" der individer velger å arbeide mindre gitt høyere lønn når lønnen stiger over visse terskelverdier (Ellis 1981)

3.1.2 Effektivitetslønnsteori

Sykefraværet kan også forklares ved å bruke teorien om effektivitetslønn (Shapiro & Stiglitz 1984). Asymmetrisk informasjon gjør det mulig for arbeidere å yte mindre enn full innsats, og hvis man antar at det ikke er mulig for arbeidsgivere å overvåke samtlige arbeidere, så kan de i stedet bruke effektivitetslønn som et substitutt til overvåking. I en situasjon der arbeidsmarkedet klarerer og det ikke finnes arbeidsledighet, kan ikke arbeidsgivere “straffe” arbeidere som sluntrer unna. Det finnes ingen kostnad ved å få sparken hvis man kan gå rett inn i ny jobb. Hvis arbeidsgiverne i stedet setter opp lønnen slik at det blir overskuddstilbud av arbeidskraft vil det derimot være en mulighet for å ende opp som arbeidsledig hvis man blir tatt i unnasluntring eller unødvendig fravær. Denne risikoen øker jo høyere arbeidsledigheten er og derfor vil en slik modell predikere at det er et negativt forhold mellom sykefraværet og ledigheten. Effektivitetslønnsteori kan være best egnet til å forklare korttidsfraværet (ABN).

3.1.3 Sykepengeordningen som sosialforsikring

Tap av arbeidsinntekter kan få store økonomiske konsekvenser for enkelte individ. I situasjoner der et skadefall kan oppstå med en viss sannsynlighet, er det ofte effektivt å forsikre seg mot uhell gitt at tilstrekkelig mange står ovenfor en slik risiko. Eksterne effekter i markedet kan likevel føre til effektivitetsproblemer og i verste fall kan hele markedet for forsikringen bryte sammen. Bratberg og Risa (2000) diskuterer tre aktuelle problemer ved forsikring som oppstår på grunn av informasjonsasymmetrier. Obligatorisk sosialforsikring, slik som sykepengeordningen, kan løse noen, men ikke alle disse problemene.

Under perfekte markedsforhold med full informasjon, kan risikopremien settes slik at det enkelte individ kommer ut med en velferdsgevinst. Hvis premien settes lik det potensielle inntektstapet vektet med sannsynligheten for uhell, vil risikoaverse individer ønske å betale denne prisen for å unngå usikkerhet. I det stiliserte eksempelet kjenner tilbyderer av forsikringen de ulike risikoprofilene for hvert individ. I virkeligheten har de ikke denne informasjonen, derfor må markedspremie på forsikring reflektere gjennomsnittsriskoen i den forsikrede populasjonen. Asymmetrisk informasjon gjør at enkeltindivider kjenner sin egen sannsynlighet for skade bedre enn forsikringstilbyder. De med lav sannsynlighet for skade vil derfor ikke ønske å betale risikopremien som er høyere enn deres forventede tap. Forsikringsselskapet sitter tilbake med bare høyrisikoindivider og må derfor øke prisen på

forsikringen ytterligere for å dekke sine kostnader. Dette kan igjen føre til at flere individer kan vurdere forsikringen for dyr i forhold til deres risikoprofil. Denne spiralen kan føre til at markedet bryter sammen. Problemet er i forsikringslitteraturen kjent som *ugunstig utvalg*¹³. Ved å gjøre forsikringsordningen obligatorisk for alle, slik sykepengeordningen er, så forsvinner dette problemet. Da kan ikke de “gode” kundene velge vekk forsikringen slik at bare de “dårlige” sitter igjen.

Et annet problem som kan oppstå i et forsikringsmarked er *atferdsrisiko*. Problemet oppstår hvis forsikringstakeren kan påvirke sin egen risiko for uhell, enten direkte ved selv å forårsake uhellet, eller indirekte ved å opptre uaktsomt slik at ulykkesrisikoen øker. Forsikringssselskapet, eller de andre i forsikringskollektivet, kan ikke observere graden av forebygging og problemet oppstår på grunn av denne informasjonsasymmetrien. Når det gjelder sosialforsikring har likevel forsikringstakeren insentiver til å forebygge skade siden et uhell vil innebære enten sykdom eller uførhet. Mangel på insentiver til å forebygge langtidsfravær er først og fremst aktuelt for arbeidsgiver siden folketrygden finansierer hele fraværet etter dag 16. I motsetning til eksempelet med ugunstig utvalg kan ikke atferdsrisiko løses ved at forsikringsordningen gjøres obligatorisk.

De vanligste diagnosegruppene for sykefravær er muskel-, skjelett- og psykiske lidelser (Arbeidstilsynet 2007). Disse skadene kan ofte være vanskelig å observere og forsikringskollektivet må stole på den enkeltes subjektive vurdering av skadeomfanget. Problemer med *rapporteringsrisiko* kan oppstå i tilfeller hvor den forsikrede feilrapporterer ikke-observerbare skader. Det kan også være ulike oppfatninger av hva som kvalifiserer til et forsikringstilfelle, det vil si hvilke helsetilstander som er “alvorlige nok” for et sykefravær. Problemet kan motvirkes ved å innføre egenandeler (karensdager og/eller lavere kompensasjonsgrad) eller bedre kontroll ved portvaktfunksjonen.

3.2 Sykefraværet og konjunktorene

3.2.1 Sammensetningshypotesen

Sammensetningshypotesen står sentralt i denne oppgaven. Hypotesen tar utgangspunkt i at det finnes en gruppe arbeidere med dårligere helse/funksjonsevne, som har en marginal tilknytning til arbeidsmarkedet. I en høykonjunktur vil bedriftene ha større problemer med å

¹³ Se Rothschild og Stiglitz (1976) for en mer detaljert fremstilling.

skaffe seg arbeidskraft og vil ansette flere marginale arbeidere med nedsatt helse. Siden disse utgjør en større andel av arbeidsstyrken, vil sykefraværet gå opp. I lavkonjunkturer vil derimot personer med dårlig helse ha større vanskeligheter med å finne seg jobb eller holde på jobben når ledigheten er høy. Ifølge denne teorien vil altså sammensetningen av arbeidsstyrken endres over konjunktorene på grunn av helsemessig seleksjon. Seleksjon inn eller ut av arbeidsstyrken kan forekomme på ulike måter, det finnes flere studier som har påvist dette¹⁴. Spørsmålet er i hvilken grad denne seleksjonen varierer med konjunktorene. I tillegg er det usikkert om gruppen med marginale arbeidere er stor nok og har et tilstrekkelig høyt nivå av sykefravær til å kunne forklare de betydelige svingningene i observert fravær.

3.2.2 *Disiplineringshypotesen*

Hypotesen tar utgangspunkt i at økonomiske nedgangstider disiplinerer arbeidstakere til å redusere sykefraværet sammenlignet med en høykonjunktur. Arbeidstakerens forventede kostnader ved sykefravær øker når arbeidsledigheten er høy. Ved "kostnader" menes det blant annet høyere risiko for å miste jobben og lavere sannsynlighet for å finne ny jobb hvis man blir arbeidsledig. Det kan også tenkes at svekkede lønns- og karrieremuligheter eller dårlige referanser, er kostnader som øker når ledigheten er høy. Noen arbeidere kan føle et ansvar ovenfor bedriften og kollegaer og vil derfor i større grad ønske å redusere sitt fravær av hensyn til disse når økonomien er stram. Siden sykdom ofte er en subjektiv vurdering, kan det tenkes at den helsetilstanden som klassifiseres som "syk nok til ikke å jobbe" er avhengig av de forventede kostnadene ved sykefraværet. Denne typen fravær er ofte forbundet med problemet atferdsrisiko.

Disiplineringshypotesen kan også brukes til å forklare begrepet *sykenærvær*. En arbeider kan velge å gå på jobb selv om funksjonsevnen er nedsatt på grunn av sykdom. Hvis sykefraværet reduseres som følge av økt disiplinering i en lavkonjunktur, er det også rimelig å anta at sykenærværet går opp, enten som følge av at arbeidere ikke velger å sykemelde seg på tross av sykdom, eller at de kommer tidligere tilbake fra et sykefravær.

¹⁴Nossen (2009) peker på ulike deler av arbeidslivet som kan være påvirket av seleksjon, og gir en oversikt over studier som gir støtte for at slik seleksjon finner sted.

3.2.3 *Stresshypotesen*

Stresshypotesen retter fokus mot betydningen av arbeidsmiljø og arbeidsrelatert stress for helse og sykefravær. Økt aktivitetsnivå i økonomien øker arbeidspresset i bedriftene som kan videreføre høyere krav til arbeiderne. Blant annet krav om mer overtid, hyppigere kundemøter, mindre pauser og økt konkurranse både eksternt og internt i bedriften. Dette kan øke forekomsten av stressrelaterte sykdommer som nakke- og skulderplager, hjerte- og karsykdommer eller mentale lidelser som depresjon og utbrenthet (Arbeidstilsynet 2007). Stress kan også forverre andre helseplager. Arbeidsrelatert stress og betydningen av arbeidsmiljø er sentrale elementer i IA-avtalen. Det eksisterer likevel få empiriske studier som støtter stresshypotesen direkte, og hypotesen må sees i sammenheng med forskning som viser at dødeligheten og andre helsemessige risikofaktorer er positivt korrelert med konjunktorene¹⁵.

3.3 *Nyere hypoteser*

Her presenteres nyere hypoteser om hvilke effekter som påvirker sykefraværet. Hovedfokuset i disse hypotesene er ikke rettet mot den prosykliske variasjonen, men heller den generelle økningen i sykefraværet.

3.3.1 *Omorganisering og nedbemanning*

I motsetning til stresshypotesen er fokuset ikke rettet mot hvordan arbeidsmiljøet endres over konjunktorene, men heller mot den generelle økningen i graden av omorganiseringer. Økt globalisering fører med seg flytting av produksjon til utlandet, raskere teknologisk endring, flere internasjonale fusjoner osv. Bedrifter som vil konkurrere i et slikt marked må i større grad enn tidligere re-optimere for å møte endringer i markedsforholdene. Dette innebærer ofte omstruktureringer og nedbemanning. Et mer turbulent arbeidsmarked kan øke belastningen på de ansatte og i tillegg gjøre det vanskeligere for arbeidere med helseproblemer å delta i arbeidsmarkedet. Spørsmålet er om økt omstilling fører til dårligere helse og mer sykefravær, og om Norge er spesielt utsatt for slike endringer i arbeidsmarkedet. Studier indikerer at omstrukturering og nedbemanning fører til negative konsekvenser både for de som blir oppsagt og de som blir værende i virksomheten. Når det gjelder

¹⁵ Se Nossen (2009) for en oversikt.

konsekvensene for sykefraværet, gir ikke forskningen entydige resultater. Noen studier finner at nedbemanning øker forekomsten av arbeidsrelaterte helseproblemer som ender i sykefravær eller andre langvarige støtteordninger¹⁶. En stor populasjonsstudie fra Sverige indikerer at større nedbemanninger kan føre til en forverring av helsetilstanden og økning i langtidsfraværet (Westerlund et al. 2004). En annen studie indikerer at nedbemanninger reduserer sykefraværet for kvinner, mens det ikke påvirker fraværet til menn (Theorell et al. 2003). Det er imidlertid lite som tyder på at arbeidslivet i Norge alt i alt har blitt tøffere de siste årene. At utviklingen har vært særlig mye verre enn i resten av verden, er også svært tvilsomt (Mykletun et al. 2010, Vedlegg 7).

3.3.2 Sosiale normer

Sykepengeordningen i Norge er svært sjenerøs med 100 % kompensasjon og lav grad av kontroll. Likevel er det et mindretall av befolkningen som benytter seg av denne ordningen, særlig over lengre perioder. Sosiale normer kan føre til at overdreven bruk av sykemelding kan oppleves som stigmatiserende. Nyere forskning har forsøkt å avdekke om normene er i ferd med å endres, det vil si om de sosiale sanksjonene generelt er svakere enn før. Dette vil gjøre det lettere å sykemelde seg nå enn tidligere, gitt den samme helsesituasjonen. I den sammenheng er det også nærliggende å undersøke i hvilken grad arbeidere blir påvirket av arbeidsmiljøet rundt seg. Vil kollegaer påvirke hverandre, dvs. finnes det en gruppepåvirkning i sykefraværet? Studier indikerer at en slik “smitteeffekt” faktisk finner sted¹⁷. Jo flere sykemeldte som befinner seg innenfor samme arbeidsplass jo mer akseptert vil det være å bli sykemeldt. Dette omtales som “sosiale multiplikatorer” og kan føre til at utviklingen i sykefraværet forverres over tid som følge av gruppeinnflytelse.

3.3.3 Legenes portvaktrolle

De som sykemelder er de som vokter porten til sykepengeordningen. I 80 % av tilfellene er dette fastlegene. I tillegg til å være portvokter skal legene også behandle og være en støttespiller for pasienten, og dette kan være en vanskelig dobbeltrolle (Mykletun et al. 2010). Derfor er det viktig med en sykepengeordning som setter klare retningslinjer for

¹⁶ Se f.eks. Rege et al. (2009), Røed & Fevang (2005) eller Østhus (2007).

¹⁷ Se f.eks. Dale-Olsen et al. (2010), Lindbeck et al. (2009) eller Rieck og Vaage (2011).

sykemeldingspraksisen slik at fastlegers personlige egenskaper får mindre betydning, og antallet skjønnsmessige vurderinger minimeres.

Før sykemeldingsreformen i 2004 ble det foretatt en studie i Norge som avslørte at 95 % av tilfellene hvor pasientene selv tok initiativet til sykemelding faktisk endte med sykemelding (Larsen et al. 1994). Dette indikerer at pasientene har stor grad av innflytelse på fastlegen, noe som støttes av studien til Guldbrandsen et al. (2002), der over halvparten av fastlegene bekrefter at pasientens ønske ofte eller delvis går foran deres egen medisinske vurdering.

Fra avsnitt 2.2.1 vet vi at endringen i sykemeldingspraksisen i 2004 førte til klarere retningslinjer for fastleger når det gjelder bruk av gradert sykemelding, aktivitetskrav og sanksjoner mot fastleger som ikke følger regelverket. Markussen (2009) finner at reformen reduserte sykefraværsprosenten med 22,9 % de første 6 månedene etter reformen, og at langtidseffekten var på 24,5 %. En mulig forklaring på reduksjonen er at fastlegen kunne henvise til et strengere regelverk med potensielle sanksjoner og dermed oppfordre arbeidere å returnere til jobb.

Markussen (2009) påpeker at styrking av legenes forhandlingskraft, med økt fokus på medisinske vurderinger og graderte sykemeldinger, kan være en nyttig måte å redusere fraværet når man har en sjenerøs sykepengeordning. Basert på Aktiviserings- og nærversreformen foreslått av Ekspertgruppen, kan det virke som om de er av samme oppfatning.

3.4 Tidligere forskning

3.4.1 Norske studier

ABN bruker norske registerdata fra årene 1990-1995 for å undersøke om variasjoner i sykefraværet utover arbeidsgiverperioden kan forklares ved disiplinierungs- eller sammensetningshypotesen. Sannsynligheten for sykefravær estimeres betinget på et utvalg av individuelle karakteristikk og den regionale arbeidsledigheten. I tillegg estimerer de varigheten av sykefraværet gitt positive verdier av den første regresjonen. Fra hele utvalget, som inneholder individer fra 16-66 år, har de konstruert et ubalansert panel der de begrenser utvalget til individer mellom 30-55 år. Dette er fordi de ønsker å ekskludere arbeidere som beveger seg inn og ut av arbeidsstyrken grunnet livssituasjonen. De konstruerer også et begrenset utvalg som består av individer som er tilstede i arbeidsgiverens register gjennom

alle seks årene. Hvis sammensetningshypotesen gjelder, vil den antatte effekten av arbeidsledigheten på sykefraværet forsvinne eller reduseres i utvalget med “stabile arbeidere”. Resultatene viser at det er en negativ signifikant effekt av arbeidsledighet på sykefraværet. Koeffisienten er signifikant på 10 % nivå for menn og 5 % for kvinner. I det stabile utvalget, der “marginale” arbeidere er ekskludert, er effekten av arbeidsledigheten sterkere og mer signifikant. Dette indikerer at variasjoner i sannsynligheten for sykefravær ikke er drevet av endringer i sammensetningen av arbeidsstyrken. De finner ingen signifikant effekt av arbeidsledigheten på varigheten av sykefraværet for hele utvalget, men den negative effekten er tydeligere og mer signifikant for det begrensede utvalget. Resultatet indikerer at stabile arbeiderne også endrer arbeidstilbudet som følge av endringer i betingelsene på arbeidsmarkedet. Modellen forklarer likevel relativt lite av variasjonen i lengden på sykefraværet, mest sannsynlig er dette fordi helse har en sterkere påvirkning på denne variabelen. Likevel argumenterer ABN for at sammensetningshypotesen ikke kan forklare variasjonen i sykefraværet i perioden 1990-1995.

Dyrstad og Ose (2002) bruker tidsseriedata for å undersøke om disiplinerings- eller sammensetningseffekter best kan forklare variasjonen i sykefraværet blant norske arbeidere i tidsperioden 1971-1998. Ved å bruke data fra NHOs fraværstatistikk¹⁸ kan de analysere effekter på både kort- og langtidsfraværet. Resultatet viser at disiplineringshypotesen bare kan forklare langtidsfraværet. De finner ikke bevis for at noen av effektene er gjeldende for korttidsfraværet til kvinner, mens begge effektene kan være til stede for menn. Dette er overraskende, siden man forventer at bare de som ikke er alvorlig syke lar seg “disiplinere”, og som følge av dette bør denne effekten være mest synlig i korttidsfraværet. Det er likevel viktig å nevne at korttidsfravær her er definert som fravær inntil tre dager, som betyr at relativt korte fravær på mer enn tre dager vil bli klassifisert som langtidsfravær.

Biørn et al. (2010) baserer seg på norske paneldata fra perioden 1993 til 2005 for å undersøke mulige årsaker til det økende sykefraværet. Fokuset er rettet mot å skille ut effektene av bestemte karakteristikk knyttet til fødselsår, alder og tid. De finner at individuelt fravær økte mer i denne tidsperioden enn det man kan lese fra aggregerte data. Selv om de ikke klarer å identifisere hvilke mekanismer som står bak økningen i sykefraværet, har de på bakgrunn av

¹⁸ NHOs fraværstatistikk er basert på et utvalg av medlemsbedriftene. Dataene er aggregerte - ikke individbaserte - som gjør at de best egner seg til å beskrive gjennomsnittsverdier på gruppenivå. Det er vanskelig å bruke slike data til å trekke slutninger på individnivå.

resultatene utelukket noen mulige forklaringer. Hypotesen om at yngre generasjoner som deltar på arbeidsmarkedet har dårligere arbeidsmoral (normer) enn eldre generasjoner, blir avvist. Dataene støtter heller ikke sammensetningshypotesen, tvert imot viser det seg at arbeidere med høy tilbøyelighet for sykefravær blir sortert *ut* av arbeidsmarkedet.

Dale-Olsen (2009) analyserer ved hjelp av norske paneldata fra 2000-2004 hvordan endringer i skatteregler påvirker arbeideres legemeldte sykefravær og lønn. Teorien for effektivitetslønn brukes for å studere hvordan arbeideres valg av arbeidsinnsats og fravær endres som følge av endringer i marginals-katten. Resultatene viser at arbeidere som kompenseres basert på innsats (lønnsstruktur, bonuser osv.) endrer adferd som følge av skatteendringer. Økt marginals-katt fører til mindre innsats og høyere fravær, og det argumenteres for at disse effektene bør tas større høyde for i utformingen av skattesystemet.

3.4.2 *Utenlandske studier*

Arai og Thoursie (2005) tar utgangspunkt i svenske paneldata i perioden fra 1989 til 1999. De konstruerer aggregerte observasjoner fordelt på næring og region for å undersøke om det er disiplinering eller sammensetningseffekter som driver den prosykliske variasjonen i sykefraværet¹⁹. Den empiriske analysen baserer seg på forholdet mellom midlertidige arbeidskontrakter og sykefraværet, der midlertidige kontrakter er en indikator på andelen marginale arbeidere i Sverige. Ifølge sammensetningshypotesen vil det være en positiv korrelasjon mellom sykefraværet og andelen av slike kontrakter. Motsatt vil disiplinering innebære at arbeidere på midlertidige kontrakter føler mindre jobbsikkerhet og derfor vil ønske å minimere sykefraværet. Den empiriske analysen viser at det er en negativ korrelasjon mellom midlertidige kontrakter og sykefraværet. Arai og Thoursie (2005) konkluderer med at både seleksjon- og disiplineringmekanismen kan være tilstede i sykefraværet over tid, men resultatene indikerer at disiplineringmekanismen dominerer.

Henrekson og Persson (2004) baserer seg på tidsseriedata fra Sverige i perioden 1955-99 for å undersøke hvordan endringer i kompensasjonsgrad påvirker sykefraværsadferden. Det ble gjennomført flere reformer i den svenske sykepengeordningen i løpet av denne perioden, og resultatene viser at i de fleste tilfeller hadde reformene sterke effekter på sykefraværet. Når ordningen gjøres mer generøs, øker det aggregerte antallet sykefraværsdager. Omvendt fører

¹⁹ Sammenhengen mellom sykefraværet og arbeidsledigheten i Sverige er sterk og utslagene er større enn i Norge (Nossen 2008).

enn innstramning til en reduksjon i sykefraværet. Resultatene støttes av en komplementær paneldatastudie (fra 1983-91). Henrekson og Persson (2004) trekker frem disiplineringshypotesen som en plausibel forklaring på denne sammenhengen.

Johanson og Palme (2005) tar utgangspunkt i et tilfeldig utvalg svenske arbeidere (basert på en levekårsundersøkelse fra 1991) for å studere hvordan arbeidernes insentiver endres som følge av en mer restriktiv sykelønnsordning. Kompensasjonsgraden ble i 1991 redusert for fravær inntil 89 dager, mens den ble holdt uendret fra dag 90. Hvis atferdsrisiko er tilstede, vil frekvensen av sykefravær gå ned som følge av at kostnadene ved sykefravær har økt på kort sikt. I tillegg har kostnadene ved å returnere fra sykefravær etter dag 90 økt, fordi det finnes en sannsynlighet for å havne i et nytt fravær (med lavere kompensasjon enn etter 90 dager). Atferdsrisiko vil dermed føre til at lange sykefravær blir lengre. Resultatene viser at antallet korttidsfravær ble redusert og lengden på langtidsfraværene økte som følge av endringen i kompensasjonsgraden. Derfor konkluderer forfatterne med at det er et atferdsrisikoprosjekt i den svenske sykepengeordningen.

Thoursie (2007) bruker registerdata fra den svenske populasjonen i 1987 for å undersøke om det forekommer illegitim bruk av sykepengedager. Den relative verdien av fritid i forhold til arbeid er antatt å øke i tidsperioden rundt individets fødselsdag. Thoursie konstruerer en tidsserie av antall personer som hadde fravær samme uken som fødselsdagen (behandlet gruppe) og en tilsvarende tidsserie av antall personer som hadde fravær uken før og etter fødselsdagen (ubehandlet gruppe). Differansen mellom de to tidsseriene regresses mot en konstant som måler gjennomsnittsforskjellen i fraværssrapporteringen til den behandlede og den ubehandlede gruppen. Resultatene indikerer at bare unge menn har illegalt sykefravær i tidsperioden rundt fødselsdagen. Dette kan indikere at kvinner og eldre arbeidere oppfatter kostnadene ved å bruke sykefravær rundt fødselsdagen som høyere enn nytten.

Kapittel 4: Økonometrisk metode

For å undersøke hvordan sykefraværet blir påvirket av ulike faktorer, er det naturlig å skille mellom effekter på *sannsynligheten for* (insidens) og *lengden på* (varighet) sykefravær. I denne oppgaven vil vi følge ABN sin tilnærming, og analysere effektene i to ulike regresjoner.

I avsnitt 4.1 omtales paneldata og hvilke fordeler denne datatypen har i forhold til vanlige tverrsnitt- eller tidsseriedata. Vanlige metoder for paneldata omtales i avsnitt 4.2. Sannsynligheten for sykefravær i år t kan estimeres ved å bruke en lineær eller diskret sannsynlighetsmodell. Begge metodene blir diskutert nærmere i avsnitt 4.3. I avsnitt 4.4 diskuteres valget mellom to vanlige paneldatametoder – fast eller tilfeldig effekt estimatorer. Til slutt drøftes ulike estimeringsutfordringer knyttet til denne typen analyser.

Fremstillingen under bygger på Verbeek (2008).

4.1. Paneldata

Et paneldatasett inneholder gjentatte observasjoner for utvalgte enheter (individer, husholdninger, bedrifter) over en gitt tidsperiode. Denne typen data er dermed en kombinasjon av tverrsnitt- og tidsseriedata. Tilgangen på denne typen data tillater oss å estimere mer kompliserte og realistiske modeller, og gir også mer presise estimat enn ved å bruke enkle tverrsnitt- eller tidsseriedata. I tillegg er det også mulig å kontrollere for individuell uobserverbar heterogenitet. Dette er uobserverbare egenskaper ved individene som påvirker den avhengige variabelen, men som ikke er inkludert i datamaterialet. I sammenheng med sykefraværet er individuell helsetilstand et viktig eksempel på slik heterogenitet. Helsetilstand er ikke inkludert i datasettet (FD-Trygd), men har stor påvirkning på sykefraværet til enkeltindivider. Paneldata gjør det mulig å kontrollere for slike uobserverte egenskaper. Dersom vi ikke tar høyde for uobserverbar heterogenitet, vil dette føre til endogenitet i modellen, et problem som drøftes nærmere i avsnitt 4.5.

4.2. Lineære paneldatamodeller

Hvis vi lar i representere individ ($i = 1, \dots, N$) og t står for tidsperiode ($t = 1, \dots, T$), kan vi skrive den standard lineære regresjonsmodellen slik:

$$Y_{it} = \beta_0 + X'_{it}\beta + \varepsilon_{it} \quad (4.1)$$

, hvor X_{it} er en K -dimensjonal vektor av forklaringsvariabler og β er en K -dimensjonal vektor av parametere som måler X_{it} sin påvirkning på Y_{it} . I denne modellen er skjæringspunktet β_0 og helningskoeffisienten β identisk for alle individer i alle tidsperioder. Det stokastiske feilleddet ε_{it} varierer over tid og individer, og fanger opp alle uobserverte faktorer som påvirker Y_{it} . For å oppnå forventingsrette, effisiente og konsistente estimatorer ved å bruke minste kvadrats metode, kreves det at de klassiske MKM antakelsene er oppfylt.

Det er ofte problematisk å bruke MKM på paneldatamodeller. Når vi observerer de samme individ over flere tidsperioder, er det ofte urealistisk å anta at feilleddene er ukorrelerte over t . Uobserverbar heterogenitet, slik som helsetilstand, blir fanget opp av ε_{it} . Estimatorer som tar høyde for korrelasjonen i feilleddet vil være mer effektive enn MKM. Korrelasjonen kan fremstilles på denne måten:

$$\varepsilon_{it} = \alpha_i + u_{it} \quad (4.2)$$

, hvor den individspesifikke komponenten α_i er konstant over tid. Valg av paneldatametode avhenger av hvilke forutsetninger vi legger til grunn for α_i . Mer presist, om vi antar:

$$Cov(\alpha_i | X_{it}) = 0 \quad (4.3)$$

, eller
$$Cov(\alpha_i | X_{it}) \neq 0 \quad (4.4)$$

Dersom vi antar at den individspesifikke komponenten er ukorrelert med de inkluderte forklaringsvariablene (4.3) kan vi bruke MKM eller tilfeldig effekt estimator (det kommer likevel frem i neste avsnitt at MKM er ineffisient i forhold til tilfeldig effekt-estimatoren). I mange tilfeller er dette en streng antagelse, med andre ord er det (4.4) som gjelder. I det følgende presenteres tre ulike metoder for å estimere paneldatamodeller.

4.2.1. MKM

Først antar vi at α_i består av *tilfeldige* faktorer og er uavhengig og identisk fordelt over individer. Det gir oss følgende modell:

$$Y_{it} = \beta_0 + X'_{it}\beta + (\alpha_i + u_{it}), \quad \alpha_i \sim IID(0, \sigma_\alpha^2); \quad u_{it} \sim IID(0, \sigma_u^2) \quad (4.5)$$

, hvor $\varepsilon_{it} = \alpha_i + u_{it}$ blir behandlet som et feilledd som består av to komponenter: en individspesifikk komponent, som ikke varierer over tid, og et restledd, som er antatt å være ukorrelert over tid. For at MKM skal gi konsistente og forventingsrette estimatorene må (4.3) gjelde.

Problemet med å estimere paneldatamodeller med MKM kan vi finne i varians-kovariansmatrisen til ε_{it} . Variansen til feilleddet ε_{it} er gitt ved:

$$E(\varepsilon_{it}^2) = E[(\alpha_i + u_{it})^2] = E(\alpha_i^2 + 2\alpha_i u_{it} + u_{it}^2) = \sigma_\alpha^2 + \sigma_u^2 \quad (4.6)$$

Kovariansen for i'te enhet er gitt ved:

$$\begin{aligned} E(\varepsilon_{it}, \varepsilon_{is}) &= E[(\alpha_i + u_{it})(\alpha_i + u_{is})] \\ &= E[\alpha_i^2] + E[\alpha_i u_{it}] + E[\alpha_i u_{is}] + E[u_{it} u_{is}] \\ &= \sigma_\alpha^2 \end{aligned} \quad (4.7)$$

For T observasjoner for i'te enhet kan varians-kovariansmatrisen $\Omega = E[\varepsilon_i \varepsilon_i'] = V[\varepsilon_i]$ dermed skrives som:

$$\Omega = \begin{bmatrix} \sigma_\alpha^2 + \sigma_u^2 & \sigma_\alpha^2 & \dots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_u^2 & \dots & \sigma_\alpha^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \dots & \sigma_\alpha^2 + \sigma_u^2 \end{bmatrix} \quad (4.8)$$

$V[u_{it}] = \sigma_u^2 I_k$ er kovariansmatrisen til feilleddet når Gauss-Markow betingelsene er oppfylt. Vi ser at $\Omega \neq \sigma_u^2 I_k$ med mindre $\sigma_\alpha^2 = 0$, og MKM estimatoren vil derfor være ineffisient som følge av autokorrelasjonen i feilleddskomponenten α_i (Verbeek (2008), s. 364).

4.2.2. Tilfeldig effekt-estimatoren

Tilfeldig effekt-estimatoren (forkortet som RE fra engelsk “random effect”) tar høyde for autokorrelasjonen ved å transformere feilleddet.

Ved å transformere feilleddet og bruke generalisert MKM (GMKM) på det transformerte uttrykket blir tilfeldig effekt-estimatoren effisient. GMKM-estimatoren får vi ved å estimere følgende transformerte modell:

$$(Y_{it} - \vartheta \bar{Y}_i) = \beta_0(1 - \vartheta) + \beta(X_{it} - \vartheta \bar{X}_i) + (\varepsilon_{it} - \vartheta \bar{\varepsilon}_i) \quad (4.9)$$

, hvor en fast andel ϑ av de individuelle gjennomsnittene trekkes fra observasjonene for å oppnå feilleddet $(\varepsilon_{it} - \vartheta \bar{\varepsilon}_i)$, som er uavhengig og identisk fordelt over i og t . Den faste andelen er gitt ved:

$$\vartheta = 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\alpha^2}} \quad (4.10)$$

, som gir et mål på de relative størrelsene til feilleddsvariansene σ_α^2 og σ_u^2 . Transformasjonen sørger for at $\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$. Begge variansene er ukjente størrelser som må estimeres.

4.2.3. Fast effekt-estimatoren

Fast effekt-estimatoren (FE) løser problemet med uobserverbar heterogenitet ved å transformere bort all ren tverrsnittvariasjon, slik at også den individspesifikke restleddskomponenten α_i forsvinner fra uttrykket gjennom transformasjonen. Vi trenger dermed ikke pålegge modellen den strenge antagelsen (4.3), fordi modellen også er robust under (4.4). Ved å inkludere et individ-spesifikt skjæringspunkt i modellen, kan vi skrive (4.1) på denne måten:

$$Y_{it} = \alpha_i + X'_{it}\beta + u_{it}, \quad u_{it} \sim IID(0, \sigma_u^2) \quad (4.11)$$

, der α_i er konstant over tid for individ i . Feilleddet u_{it} er antatt å være uavhengig og identisk fordelt (IID) over individer og tidsperioder, med forventning lik null og varians gitt ved σ_u^2 . En fast effekt transformasjon innebærer å trekke de individspesifikke gjennomsnittene fra de opprinnelige variablene i (4.11) slik at vi får:

$$(Y_{it} - \bar{Y}_i) = (\alpha_i - \alpha_i) + (X_{it} - \bar{X}_i)' \beta + (u_{it} - \bar{u}_i) \quad (4.12)$$

, hvor $\bar{Y}_i = T^{-1} \sum_t Y_{it}$ og tilsvarende gjelder for \bar{X}_i og \bar{u}_i . Den faste effekten forsvinner fra modellen siden α_i er konstant over tid. Modellen i (4.12) kan estimeres med MKM og estimatoren kalles for *fast effekt-estimatoren* ($\hat{\beta}_{FE}$)²⁰.

Fast effekt-modellen utnytter variasjonen *innenfor* hvert individ. Modellen forklarer i hvilken grad Y_{it} er forskjellig fra \bar{Y}_i , men ikke hvorfor \bar{Y}_i er forskjellig fra \bar{Y}_j . En ulempe med fast effekt-estimatoren er at alle tidskonstante variabler faller bort med transformasjonen i (4.12). Variabler som varierer lite, vil få estimatører med høye standardavvik som følge av FE-transformasjonen.

Tilfeldig effekt-modellen utnytter variasjonen både *innenfor* og *mellom* individ og gir derfor en mer effisient estimator enn fast effekt-modellen, gitt at den er konsistent.

4.3. Ikke-lineære paneldatamodeller

Sannsynligheten for sykefravær i år t kan estimeres ved å bruke en lineær eller diskret sannsynlighetsmodell. Sannsynlighetsmodeller kan baseres på en underliggende latent modell, som representerer antagelser om individ i 's nyttemaksimering²¹. La Y_{it}^* representere individ i 's differanse i nytte mellom sykefravær og ikke sykefravær i år t . Denne differansen kan ikke observeres, derfor definerer vi en indikatorvariabel som måler om individ i har fravær i år t eller ikke. Dette kan vi skrive som:

$$Y_{it} \begin{cases} 1 \Leftrightarrow Y_{it}^* > 0 \\ 0 \Leftrightarrow Y_{it}^* \leq 0 \end{cases} \quad (4.13)$$

Hvis vi tar utgangspunkt i (4.1) og måler $Y_{it} = 1$, kan dette tolkes som at differansen i nytte er positiv, og sykefravær gir dermed positiv nytte. Sannsynligheten for et positivt utfall er lik forventningen til den avhengige variabelen betinget på de uavhengige variablene. Dette kan fremstilles som en lineær funksjon av forklaringsvariablene:

²⁰ Estimatoren kalles også også LSDV-estimatoren (Least Squares Dummy Variable) siden vi kan oppnå samme estimator ved å estimere modellen med en dummyvariabel for hver observasjonseenhet. Denne metoden er imidlertid svært upraktisk hvis N er stor.

²¹ I studier av sykefraværet brukes ofte den neoklassiske arbeidstilbudsmodellen fra Allen (1981).

$$E\{Y_{it}|X_{it}\} = P\{Y_{it} = 1|X_{it}\} = \beta X'_{it} \quad (4.14)$$

Den lineære modellen impliserer at $X'_{it}\beta$ er en sannsynlighet og derfor bør ligge mellom 0 og 1. Det gir derfor ikke alltid mening å tolke resultater fra lineære sannsynlighetsmodeller, siden predikerte verdier av $X'_{it}\beta$ kan ligge utenfor dette intervallet. Et annet problem med lineære sannsynlighetsmodeller er at feilleddet er heteroskedastisk, som betyr at $Var(\varepsilon_{it})$ ikke er konstant. Dette skaper problemer for hypotesetestingen, men problemet kan reduseres ved hjelp av en transformasjon som gjør standardfeilene robuste ovenfor heteroskedastisitet (White 1980).

En diskret sannsynlighetsmodell kan på samme måte som i det lineære tilfelle utledes fra en latent underliggende modell. Generelt kan vi skrive:

$$P\{Y_{it} = 1|X_{it}\} = F(X'_{it}\beta) \quad (4.15)$$

, hvor $F(\cdot)$ er en fordelingsfunksjon for ε_{it} som bare kan ta verdier i intervallet $[0,1]$. Modellen måler sannsynligheten for at $Y_{it} = 1$ gitt vektoren av forklaringsvariabler X_{it} . Valg av fordelingsfunksjon for feilleddet bestemmer hvilken type modell vi bruker.

De to vanligste fordelingsfunksjonene er standard normalfordelingsfunksjonen og den logistiske funksjonen. Disse gir oss henholdsvis probit- og logit-modellene. Begge modellene gir en S-formet kumulativ fordelingsfunksjon, noe som sørger for at den estimerte sannsynligheten holder seg innenfor $[0,1]$ intervallet.

Ikke-lineære modeller estimeres ved å bruke sannsynlighetsmaksimeringsprinsippet (SM). I prinsippet bestemmes parameterne slik at sannsynligheten for å tilfeldig trekke de observasjonene vi har fra populasjonen, blir maksimert. Sannsynligheten avhenger av fordelingen til feilleddet. I paneldataanalyse oppstår det visse problemer når man skal estimere diskrete sannsynlighetsmodeller ved å bruke SM. Det kan vises at den vanlige SM-estimatoren er inkonsistent for fast T og $N \rightarrow \infty$. Problemet oppstår når vi skal estimere α_i , hvor inkonsistensen i α_i generelt vil overføres til estimatorene i vektoren β .

Vi kan bruke *betinget* sannsynlighetsmaksimering til å eliminere α_i . Dette gjøres ved å maksimere sannsynlighetsfunksjonen betinget på et tilstrekkelig instrument for α_i . For fast effekt logitmodellen kan vi bruke $\sum_{t=1}^T Y_{it}$, det vil si antall utfall som er lik 1 for individ i over tidsperioden T (Chamberlain 1980). For probitmodellen finnes det ingen tilstrekkelige

instrument for α_i . Derfor kan vi ikke estimere en konsistent fast effekt probitmodell for paneldata med en endelig tidsperiode T .

Det er også lite fordelaktig å estimere tilfeldig effekt logitmodeller. Når vi antar at feilleddet $\varepsilon_{it} = \alpha_i + u_{it}$ har en logistisk fordeling, så er alle korrelasjoner begrenset til å være $1/2$ (Maddala 1987).

4.3.1. Fast effekt logitmodell

Tar vi utgangspunkt i (4.15), setter inn den logistiske funksjonen for $F(\cdot)$ og i tillegg inkluderer et individspesifikt skjæringspunkt slik som i (4.11) får vi:

$$P\{Y_{it} = 1|X_{it}\} = \frac{\exp(\beta X'_{it} + \alpha_i)}{1 + \exp(\beta X'_{it} + \alpha_i)} \equiv \Lambda(\beta X'_{it} + \alpha_i) \quad (4.16)$$

, hvor $\Lambda(\cdot)$ betegner en kumulativ logistisk fordeling. Modellen estimeres med betinget SM, der vi bruker $\sum_{t=1}^T Y_{it}$ som instrument for α_i . For at individer skal bidra til sannsynlighetsfordelingen, krever vi at de skifter status minst en gang i løpet av hele tidsperioden. Det betyr at individer som enten har $Y_{it} = 1$ eller $Y_{it} = 0$ for alle t , vil utelates fra modellen. Vi får dermed et mindre utvalg enn ved å bruke tilfeldig effekt-modeller eller MKM. Individer som ikke endrer status, blir droppet fordi de ikke gir informasjon som bidrar til estimeringen av parameteren β . Resultatene i logistiske modeller oppgis logit-koeffisienter, som er vanskelig å tolke direkte. Det er vanlig å fokusere på fortegn og signifikansnivå. Alternativt kan man regne ut marginaleffekter, men dette er problematisk i fast effekt-modeller grunnet tilstedeværelsen av α_i (Cameron & Trivedi (2009), s.616).

4.4. Fast eller tilfeldig effekt?

En viktig forskjell på paneldatametodene vi har gjennomgått er hvorvidt den individspesifikke effekten α_i er antatt å være en fast estimerbar effekt eller en individuell stokastisk feilleddskomponent. Generelt er det ønskelig å bruke en fast effekt-estimator hvis individene i utvalget er "unike" individer med individuelle karakteristikk. Ved å bruke fast effekt-estimator kan effekten av de individuelle karakteristikkene estimeres. I motsetning er tilfeldig effekt-estimatoren av størst interesse hvis individene ikke har spesielle karakteristikk som kan påvirke de andre forklaringsvariablene. Tilfeldig effekt-modeller gir ofte mer statistisk

signifikante estimater enn fast effekt-modeller, siden tilfeldig effekt-modellen utnytter mer av variasjonen i variablene.

Som vi allerede har sett, er en viktig betingelse for konsistente estimatorer i tilfeldig effekt-modellen at $Cov(\alpha_i|X_{it}) = 0$. Når man analyserer sykefraværet, er det imidlertid naturlig å anta at den uobserverte individspesifikke karakteristikken “helse” vil påvirke variabler som blant annet inntekt, utdanning og alder. Det finnes metoder for å teste om det er korrelasjon mellom forklaringsvariablene og den individspesifikke effekten. Dette kan f. eks gjøres ved å bruke en Hausmantest.

4.4.1. Hausmantest

Vi ønsker å sammenligne fast og tilfeldig effekt-estimatorene. Hausman (1978) foreslo en test som utnytter at fast effekt-estimatoren er konsistent under både null- og alternativhypotesen, mens tilfeldig effekt-estimatoren er konsistent (og typisk effesient) bare under nullhypotesen. Null- og alternativhypotesene er definert slik:

$$H_0: Cov(\alpha_i|X_{it}) = 0 \quad H_A: Cov(\alpha_i|X_{it}) \neq 0 \quad (4.17)$$

Under H_0 gir begge metodene rimelig like estimatorverdier, som innebærer at $plim(\hat{\beta}^{FE} - \hat{\beta}^{RE}) = 0$. Nullhypotesen forkastes dersom estimatorene er signifikant forskjellige, som betyr at testen indikerer at det er korrelasjon mellom den individspesifikke effekten og forklaringsvariablene. Da er tilfeldig effekt-estimatoren inkonsistent og vi bør velge fast effekt-estimatoren.

Testobservatoren defineres med utgangspunkt i kovariansmatrisen til differansevektoren $(\hat{\beta}^{FE} - \hat{\beta}^{RE})$. Hausman sitt hovedresultat er at kovariansen til en effisient estimator med sin differanse i forhold til en ineffisient estimator, er null (Greene (2003), s.301). Ved å bruke dette resultatet kan det vises at kovariansmatrisen er gitt ved:

$$Var(\hat{\beta}^{FE} - \hat{\beta}^{RE}) = Var(\hat{\beta}^{FE}) - Var(\hat{\beta}^{RE}) = \varphi \quad (4.18)$$

Testobservatoren er χ^2 -fordelt med K-frihetsgrader og er gitt ved:

$$W = (\hat{\beta}^{FE} - \hat{\beta}^{RE})' \hat{\varphi}^{-1} (\hat{\beta}^{FE} - \hat{\beta}^{RE}) \quad (4.19)$$

, hvor den estimerte kovariansmatrisen er gitt ved $\hat{\phi}$. Den estimerte kovariansmatrisen er basert på $\hat{\beta}^{FE}$, som er helningsvektoren fra FE-estimeringen, og $\hat{\beta}^{RE}$, som er helningsvektoren fra RE-estimeringen. Hvis testobservatoren W er større enn den kritiske verdien, forkaster vi nullhypotesen om at den individspesifikke effekten er ukorrelert med forklaringsvariablene.

4.5. Estimeringsutfordringer

I dette avsnittet redegjøres det for utfordringer som kan oppstå i forbindelse med analysen i kapittel 6.

4.5.1. Endogenitet

Problemet med endogenitet oppstår dersom feilledet i regresjonsmodellen er korrelert med en eller flere av forklaringsvariablene i modellen. Dette fører til brudd på forutsetningen om at ingen av forklaringsvariablene skal være korrelert med feilledet, og estimatorene som er gjennomgått i dette kapitlet vil gi forventningsskeive og inkonsistente estimat.

Blant årsakene som kan føre til endogenitet er det verdt å nevne utelatte variabler, målefeil og simultane ligninger. I denne analysen er problemet med utelatt variabel mest relevant. Som tidligere nevnt påvirker variabelen “helse” sykefraværet, men er ikke inkludert i datasettet. Individuell helsetilstand vil være sterkt korrelert med forklaringsvariabler som alder, inntekt og utdanning.

Bruk av instrumentvariabel er en vanlig metode for å korrigere for endogenitet. Kravet til et godt instrument er at det skal være korrelert med den endogene forklaringsvariabelen, men ikke korrelert med feilledet. I stedet for å løse problemet med mulig endogen inntekt ved å bruke instrumentvariabelmetoden, har vi valgt å bruke lagget inntektsvariabel. Vi bruker inntekten i år $t-1$ som en proxyvariabel for inntekten i år t . Helsetilstanden i år t vil være korrelert med inntekten i år t , men mindre korrelert med inntekten i år $t-1$. I tillegg vil FE-estimatoren fjerne alle helsefaktorer som er tidskonstant gjennom FE-transformasjonen. Lagget inntektsvariabel og FE-estimator bør derfor gjøre modellen mer robust ovenfor mulig endogenitet forårsaket av den utelatte forklaringsvariabelen “helse”. ABN lagger også inntektsvariabelen en periode for å motvirke endogenitetsproblemet.

4.5.2. Selektert utvalg

I empiriske analyser med paneldata er det vanlig å ha problemer med manglende variabler. Dette kan skyldes problemer med innsamling av data eller at individer forlater utvalget. Hvis individer forlater utvalget på en ikke-tilfeldig måte, kan dette ha uheldige konsekvenser for modellen i form av forventningsskeive estimatorer og ugyldige tester. Problemer med frafall er imidlertid små ved bruk av registerdata. Metoden som brukes for å begrense utvalgene i denne oppgaven kan imidlertid føre til brudd på forutsetningen om tilfeldig utvalg.

I kapittel 5 presenteres utvalgene som brukes i analysedelen. Som det fremgår av problemstillingen i oppgaven, ønsker vi spesielt å undersøke om effekten av arbeidsledigheten er negativ også for individer som faller utenfor den “marginale” arbeiderdefinisjonen. Først betinger vi at arbeiderne i utvalget skal være mellom 30-55 år. Dermed konstruerer vi et sekundært utvalg av stabile arbeidere, som tjener pensjonsgivende inntekt alle årene de er tilstede i utvalget. Dersom det ikke er tilfeldig hvem som går ut av arbeidsstyrken, kan vi risikere at utvalget er selektert og vi har dermed brudd på forutsetningen om tilfeldig utvalg. Eksempelvis kan individer med dårlig helse selekteres ut av arbeidsstyrken.

Kapittel 5: Databeskrivelse

Dette kapittelet inneholder en beskrivelse av datamaterialet som ligger til grunn for analysen i kapittel 6. I neste avsnitt presenteres datagrunnlaget som utvalgene er basert på, samt utviklingen i sykefraværet til individene som inngår i utvalgene. I avsnitt 5.2 diskuteres variablene som brukes i analysen og i 5.3 presenteres deskriptiv statistikk for disse variablene.

5.1. Datamaterialet og utvalgene

Datamaterialet som benyttes i denne oppgaven, er hovedsakelig hentet fra forløpsdatabasen FD-Trygd, som inneholder data om hele den norske befolkningen fra og med 1992 til 2008. De viktigste datatypene som inngår i databasen er tryggedata, demografidata og sysselsettingsdata. FD-Trygd er etablert og drives av Statistisk sentralbyrå (SSB). Opplysningene er hentet fra interne registre hos SSB samt registre fra NAV (tidligere Rikstrygdeverket og Aetat). Dataene er koplet sammen med tall fra inntekts- og formuesregisteret og utdanningsdata fra den nasjonale utdanningsdatabasen (NUDB).

Forløpsdata registrerer endringer i enkeltindividers livsløp. Endringer i individets "tilstand" markeres med en start og sluttdato, som for eksempel gjør det mulig å beregne hvor lenge individet er mottaker av en offentlig støtteordning. Inntekt- og formuesdataene inneholder bare årsopplysninger. En svakhet ved FD-Trygd i denne sammenhengen er at kun sykefraværsdager betalt av Folketrygden inngår i datamaterialet. Vi har ikke informasjon om korttidsfraværet, det vil si fravær som ikke overstiger arbeidsgiverperioden²².

Basert på dette datagrunnlaget har vi trukket et 30 % randomisert utvalg²³ som vi observerer i årene 1993 til 2006. På grunn av manglende variabler utelukker vi observasjoner fra årene 1992, 2007 og 2008. I samsvar med ABN inkluderer vi bare individer i alderen 30-55 år. Grunnen til at vi ekskluderer individer utenfor denne aldersgruppen, er at vi ønsker å undersøke om det er marginale arbeidere som skaper variasjonen i sykefraværet over konjunktursyklusene (jf. sammensetningshypotesen). Derfor ønsker vi ikke å inkludere individer som beveger seg inn og ut av arbeidsstyrken grunnet livssyklusfenomener som f.eks. utdanning, førstegangstjeneste, tidligpensjonering og helseproblemer knyttet til høy alder. En

²² Som forklart i avsnitt 2.2 ble arbeidsgiverperioden utvidet fra 14 til 16 dager i 1998.

²³ Utvalgsstørrelsen ble valgt pga. kapasitetsbegrensninger ved dataprogrammene brukt i analysen.

naturlig konsekvens av at vi fokuserer på den spesifikke aldersgruppen, er at individene som er 55 år i 1993 forsvinner ut av utvalget i 1994, og individer som er 55 år i 1994 forsvinner ut i 1995 osv. For at ikke utvalget skal inneholde færre og færre alderskohorter etter hvert som 56 åringene forsvinner ut for hvert år, supplerer vi kontinuerlig med 30 % av de nye 30 åringene. Mer presist trekker vi et 30 % tilfeldig utvalg av 30 åringene i 1994 som knyttes på panelet for å kompensere for 56 åringene som forsvinner ut. Denne prosessen gjentas for alle de påfølgende årene.

Selvstendig næringsdrivende er ikke tatt med i utvalget siden de er omfattet av andre regler enn vanlige arbeidstakere når det gjelder sykepengeordningen²⁴. For eksempel er dekningsgraden for selvstendig næringsdrivende 65 % fra og med 15 sykefraværsgener, mot 100 % fra første dag for arbeidstakere. Incentivene for arbeid og sykemelding vil være ulike mellom selvstendig næringsdrivende og arbeidstakere (Allen 1981). Prosessen med å fjerne de selvstendig næringsdrivende er likevel ikke uproblematisk. Variabelen som brukes for å identifisere de selvstendig næringsdrivende, mangler for årene 2004-2006. Siden utvalget vårt er basert på 30 % av vanlige arbeidstakere i 1993, som suppleres med nye 30 åringene hvert år, har bortfallet av den aktuelle variabelen konsekvenser som er verdt å nevne. Alle individer i utvalget er bekreftet å være vanlige arbeidstakere før 2004. Likevel kan vi ikke kontrollere for om individene har skiftet status til selvstendig næringsdrivende i løpet av årene 2004-2006. Dette gjelder for individene som er tilstede i utvalget over hele analyseperioden og for de nye 30 åringene som blir lagt til hvert år. For eksempel vil en 30 åring som inkluderes i 2006 ha status som vanlig arbeidstaker i en alder av 27 år i 2003, men kan ha skiftet status i løpet av den påfølgende treårsperioden. Effektene av mulige statusendringer antas likevel å være små. Noen studier indikerer at sykefraværet hos lønntakere og selvstendig næringsdrivende varierer på samme måte som følge av konjunktursvingninger (Nossen & Thune 2010).

Basert på dette utvalget har vi konstruert et ubalansert panel for årene 1993-2006. Panelet består av 728058 individer, der 228703 (ca. 31 %) av disse er tilstede gjennom hele perioden. Individene i utvalget er født mellom 1938 og 1976. Vi har i tillegg konstruert et begrenset utvalg av "stabile arbeidere" som er registrert med pensjonsgivende inntekt alle årene de inngår i utvalget. Det stabile utvalget består av 495938 individer, hvor 151691 (ca. 31 %) av

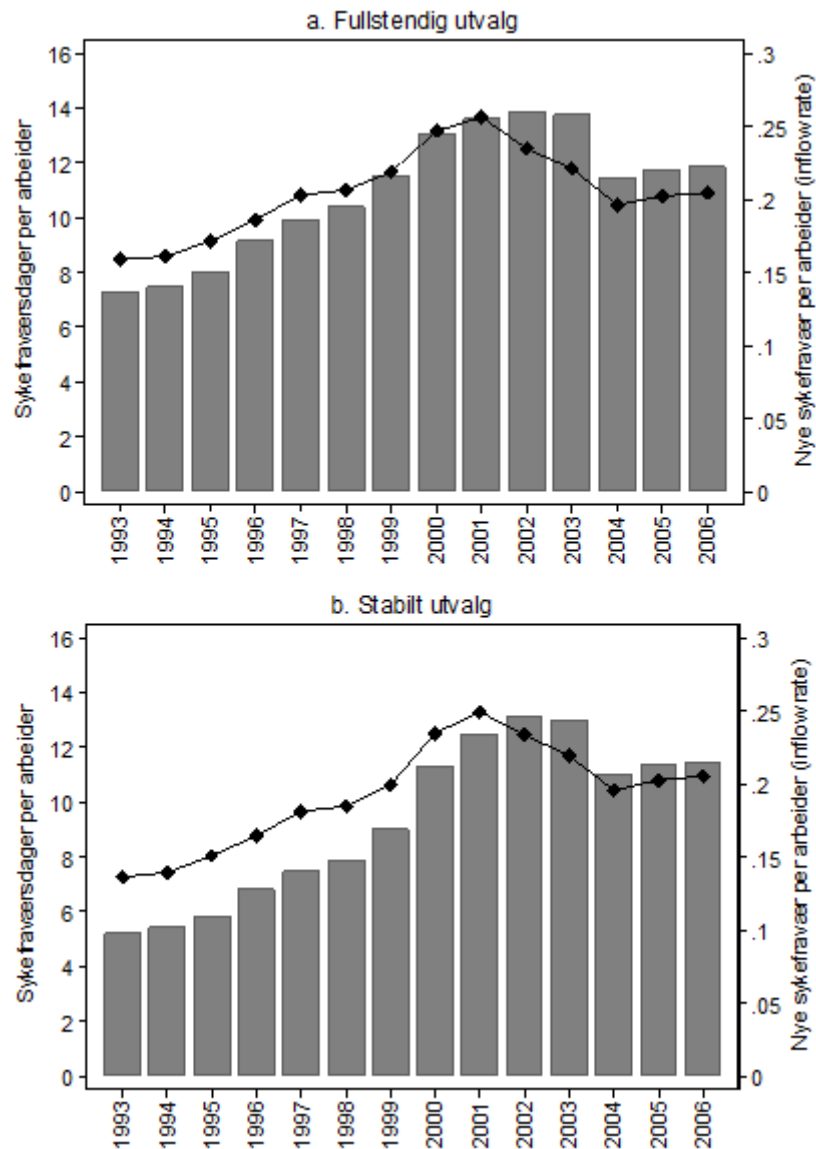
²⁴ Individer med registrerte arbeidskommuner innenfor norske områder som Svalbard, Jan Mayen og Kontinentalsokkelen er også fjernet fra utvalget. Grunnen er at vi ikke har tall for arbeidsledigheten i disse kommunene.

disse har observasjoner alle 14 årene. ABN betinger sitt stabile utvalg på at arbeiderne skal stå registrert i arbeidsgivernes register gjennom *alle* årene mellom 1990-1995. Dersom vi bruker en tilsvarende streng betingelse, dvs. at arbeiderne må tjene pensjonsgivende inntekt alle årene mellom 1993-2006, vil utvalget bare bestå av alderskohorten født mellom 1951 og 1963. Likevel er vår definisjon av det stabile utvalget strengere enn i ABN, fordi arbeiderne i gjennomsnitt har vært i arbeidsstyrken lengre. Et tilfeldig individ i det stabile utvalget har i gjennomsnitt hatt et fast arbeidsforhold i 11,3 påfølgende år. Tilsvarende gjennomsnitt i ABN kan maksimalt ta verdien 6. En mulig konsekvens av at arbeiderne i vårt utvalg er “mer stabile”, er at de har andre egenskaper enn utvalget i ABN.

En alternativ tilnærming til definisjonen av utvalgene, er å definere et stabilt og et marginalt utvalg, for deretter å sammenligne effekten av arbeidsledigheten på sykefraværet. Tilnærmingen i denne oppgaven følger ABN slik at resultatene er enklere å sammenligne. I tillegg kan det være problematisk å identifisere de marginale arbeiderne. Ose (2010) drøfter noen av disse problemene, f.eks. om det er det mulig at marginale arbeidstakere har ulike egenskaper avhengig av sektortilhørighet. Et eksempel på en indikator på andelen marginale arbeidere, er antall midlertidige arbeidskontrakter²⁵.

Hvis effekten av arbeidsledigheten på sykefraværet hovedsakelig er forårsaket av at marginale arbeidere går inn og ut av arbeidsstyrken, så bør denne effekten forsvinne, eller i det minste reduseres kraftig i utvalget med “stabile arbeidere”. Den negative sammenhengen mellom sykefraværet og arbeidsledigheten er tydelig gjennom hele analyseperioden til ABN (1990-1995). I deler av vår analyseperiode, nærmere bestemt 1999-2004, er derimot sammenhengen svak. Dette kan medføre at vi estimerer en svakere sammenheng mellom arbeidsledigheten og sykefraværet i forhold til resultatene i ABN.

²⁵ Se f.eks. Arai og Thoursie (2005).



Figur 3: Sykepengedager og nye sykepengetilfeller.

Antall sykepengedager betalt av folketrygden per arbeider (søyler) og antall nye sykefravær per arbeider (kurve). Tallene er ikke korrigert for at arbeidsgiverperioden ble utvidet fra 14 til 16 dager i 1998.

Figurene 3a og 3b viser utviklingen i antall sykefraværsdager (søyler) og antall nye sykefravær per arbeider, eller “inflow rates” (kurve). Hensikten med å inkludere begge ratene er for å undersøke om utviklingen i varigheten er ulik utviklingen i frekvensen av sykefravær. Dersom økningen i det totale sykefraværet skyldes at fraværene har blitt lengre, vil dette innebære andre politikimplikasjoner enn dersom årsaken er økning i antall tilfeller. Lengre fravær kan f. eks. reduseres ved økt bruk av graderte sykemeldinger, mens for å redusere

antall korte sykefravær er det mer effektivt å endre de økonomiske insentivene (gitt at det eksisterer disiplineringseffekter).

Antall sykefraværsdager økte frem mot 2004, hvor fraværet sank betraktelig fra 2003 til 2004. Fallet i sykefraværet må sees i sammenheng med regelendringene fra 1. juli 2004 som hadde en betydelig negativ effekt på sykefraværet (Markussen 2009). Antall sykefravær per arbeider-raten er konstruert ved å dividere det totale antall nye sykefravær hvert år på antall individer med pensjonsgivende inntekt samme året. Kurven øker frem mot 2001, for deretter å synke markant frem mot 2004. Figuren illustrerer at utviklingen i antall sykefraværsdager og antall tilfeller er relativt lik frem til 2001. Årsaken til den ulike utviklingen f.o.m. 2002 blir nærmere forklart under.

Når det gjelder forskjellen mellom utvalgene, kan vi legge merke til følgende: antall sykefraværsdager per arbeider er høyere i det fullstendige utvalget, mens forskjellen i antall nye sykefravær per arbeider (inflow rate) ikke er like tydelig. Utregning viser imidlertid at gjennomsnittlig inflow rate er 0,21 i det fullstendige utvalget mot 0,19 i det stabile. Dette indikerer at arbeidere med en stabil arbeidsmarkedstilknytning også har færre sykefravær enn referansegruppen (det fullstendige utvalget). Økningen i antall sykefraværsdager per arbeider har likevel vært størst i det stabile utvalget. Dette indikerer at økningen i sykefraværet i tidsperioden 1993-2006 ikke kan forklares med en økning i marginale arbeidere med dårligere helse.

Hovedårsaken til at utviklingen i de to grafene (søyle og kurve) er ulik f.o.m. 2002 er at registreringsrutinene til Rikstrygdeverket ble endret dette året (Rikstrygdeverket (2003), s.14). Før endringen ble forlengelser registrert som nye fravær, som betyr at en overgang til aktiv sykemelding vil registreres som et nytt sykefraværstilfelle. Derfor vil grafen for antall sykefraværstilfeller synke etter 2001 på grunn av minst to ulike årsaker:

- Rutineendringene til Rikstrygdeverket fører til en *ikke-reell* reduksjon i antall sykefraværstilfeller. Grafen synker til et vedvarende lavere nivå i 2002.
- Regelendringene for sykemeldingspraksis i 2004 fører i stor grad til en *reell* reduksjon i sykefraværet.

Rikstrygdeverkets statistikk har overestimert antall episoder før 2002, mens etter endringen vil et sykdomstilfelle omfatte hele den sammenhengende sykefraværsperioden. Antall

sykefraværsdager betalt av Folketrygden er mindre berørt av endringen. Før regelendringen ble et sykdomstilfelle som startet i år t og endret til aktiv/gradert sykemelding i år $t+1$, registrert som to tilfeller, ett i hvert år. Etter regelendringen vil det samme fraværet kun registreres som ett fravær i år t . Den binære avhengige variabelen i våre modeller indikerer hvorvidt individ i har hatt minst ett sykefravær, og tar verdien én dersom antall sykefraværsdager er større enn null i år t . Regelendringen kan dermed føre til at gjennomsnittsverdien for den avhengige variabelen blir relativt lavere etter 2002 når vi kontrollerer for andre faktorer.

Tabell 1: Fordelingen av sykefravær, 1993-2006.

Antall sykefravær fordelt på individer									
Sykefravær	Fullstendig utvalg				Stabilt utvalg				
	Menn		Kvinner		Menn		Kvinner		
	Frekvens	Prosent	Frekvens	Prosent	Frekvens	Prosent	Frekvens	Prosent	
0	208 229	57.16	155 801	42.86	139 710	53.19	79 408	34.04	
1	61 756	16.95	64 693	17.78	50 117	19.08	47 393	20.32	
2	33 527	9.20	43 425	11.94	26 291	10.01	31 844	13.65	
3	20 226	5.55	29 574	8.13	15 522	5.91	21 654	9.28	
4	12 940	3.56	20 372	5.60	9 740	3.71	14 889	6.38	
5	8 593	2.38	14 613	4.02	6 537	2.49	10 837	4.65	
>5	18 999	5.20	35 310	9.67	14 766	5.61	27 230	11.68	
Totalt	364 270	100	363 788	100	262 683	100	233 255	100	

I tabell 1 illustreres fordelingen av antall sykefravær over individer. Ved å summere antall sykefravær per individ over hele perioden 1993-2006, viser det seg at 57 % av menn i det fullstendige utvalget har ingen sykefravær over 16 dager (14 dager før 1998), mens ca. 43 % av kvinnene unngår sykefravær i samme periode. I det stabile utvalget reduseres andelen til 53 % for menn, og 34 % for kvinner. Dette betyr ikke at stabile arbeidere er sykere, figur 3 illustrerer det motsatte. Derimot har flere i det stabile utvalget tilgang på sykepenger siden de har en stabil arbeidsmarkedstilknytning. Forholdet mellom gjennomsnittlig sykefravær i motsetning til gjennomsnittlig sykefravær per arbeider drøftes nærmere i avsnittet med deskriptiv statistikk (5.3).

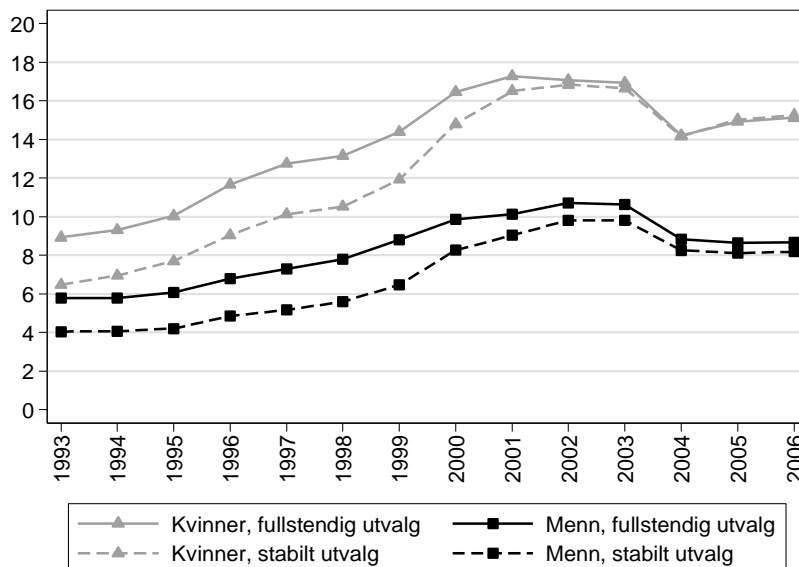
Tabell 2 illustrerer den årlige fordelingen av sykefravær for individene i begge utvalgene. Hovedtrekkene er de samme som for tabell 1, hvor menn har mindre fravær enn kvinner, og stabile arbeidere med større tilgang på sykepenger har flere erstattede sykefravær utover arbeidsgiverperioden.

Tabell 2: Fordelingen av antall sykefravær per år, 1993-2006

Fordeling av årlige sykefravær									
Sykefravær	Fullstendig utvalg				Stabilt utvalg				
	Menn		Kvinner		Menn		Kvinner		
	Frekvens	Prosent	Frekvens	Prosent	Frekvens	Prosent	Frekvens	Prosent	
0	3 056 617	89.68	2 886 220	84.26	2 125 823	88.44	1 689 056	80.48	
1	281 628	8.26	421 592	12.31	225 617	9.39	322 703	15.37	
2	54 814	1.61	92 554	2.70	41 401	1.72	68 930	3.28	
3	11 657	0.34	19 495	0.57	8 372	0.35	14 124	0.68	
4	2 444	0.07	4 023	0.12	1 716	0.07	2 767	0.14	
5	653	0.03	1 066	0.03	457	0.02	744	0.04	
6	203	0.01	324	0.01	155	0.01	243	0.01	
Totalt	3 408 016	100	3 425 274	100	2 403 541	100	2 098 567	100	

Figur 4 på neste side, illustrerer antall sykefraværsdager per arbeider fordelt på kjønn for begge utvalgene. Figuren viser at kvinner har høyere sykefravær enn menn. Samtidig kan vi se at kvinner også har hatt en kraftigere økning i antall sykefraværsdager i løpet av den aktuelle tidsperioden. Forskjellen mellom kjønnene er den samme i begge utvalgene, men den generelle økningen i sykefraværsdager har vært størst i det stabile utvalget (jf. figur 3). Figur 4 indikerer at forskjellen i sykefraværet mellom de to utvalgene var størst i starten av perioden, mens den gradvis har minsket fram mot 2006. I 2006 er det kun en marginal forskjell mellom de to utvalgene.

Kostøl & Telle (2011) viser at sykefraværet til menn har vært tilnærmet uendret mellom 1979 og 2009, mens det for kvinner har økt med nesten 70 prosent i samme periode. En av årsakene til at deres funn indikerer fallende fravær blant menn, kan være at de fokuserer på en lengre tidshorisont enn vi har tilgjengelig i våre data. Figurene deres indikerer at menns fravær varierer mest etter 1991. En alternativ forklaring er at Kostøl & Telle (2011) bruker data fra AKU-undersøkelsen, mens dataene i FD-Trygd er basert på NAV sine register for fødsels- og sykepenger.



Figur 4: Antall sykefravær dager per arbeider, fordelt på kjønn.

Tallene er ikke korrigert for at arbeidsgiverperioden ble utvidet fra 14 til 16 dager i 1998.

Det finnes flere ulike teorier for hvorfor kvinner har høyere sykefravær²⁶. Teorien om “den doble byrde” fokuserer på at kvinner med barn bruker nesten dobbelt så mye tid på hjemmearbeid som menn med barn. Derfor er belastningen ved å balansere små barn med en karriere større for kvinner enn for menn. For å unngå dobbeltarbeid, vil noen kvinner i større grad redusere sitt arbeidstilbud for å fokusere på arbeid i hjemmet, mens andre velger motsatt. Den deskriptive statistikken i avsnitt 5.3 viser at kvinner med stabil arbeidsmarkedstilknytning i gjennomsnitt har færre barn enn ellers. Hvis teorien om den doble byrde er sann, er det sannsynlig at belastningen øker med antall barn. I modellen kontrollerer vi for antall barn under 11 år.

En alternativ forklaring på kvinners høye sykefravær er at kvinner er mindre arbeidsføre enn menn, et argument som i stor grad knytter seg til svangerskapsrelaterte sykdommer. Ifølge tall fra Statistisk Sentralbyrå (SSB 2011) har antall fødsler vært relativt stabilt siden begynnelsen av 1990-tallet. Svangerskapsrelaterte sykdommer kan neppe forklare den observerte økningen i kvinners fravær, med mindre gravide kvinner er syke i mye større grad enn tidligere. Forskning indikerer at svangerskapsrelatert fravær ikke kan forklare økningen i sykefraværet til kvinner (Kostøl & Telle 2011). Kvinner er overrepresentert i yrker eller sektorer med høyt sykefravær, slik som helse- og omsorgssektoren. De fysiske og mentale kravene er høye i

²⁶ Kostøl og Telle (2011) gir en detaljert oversikt over disse teoriene.

disse yrkene, hvor 90 % rapporterer at de har hatt plager knyttet til muskel- og skjelettlidelser de siste 14 dagene (Eriksen 2003). Kombinert med høyt krav til tidspress kan dette føre til at arbeidsbelastningen er helseskadelig.

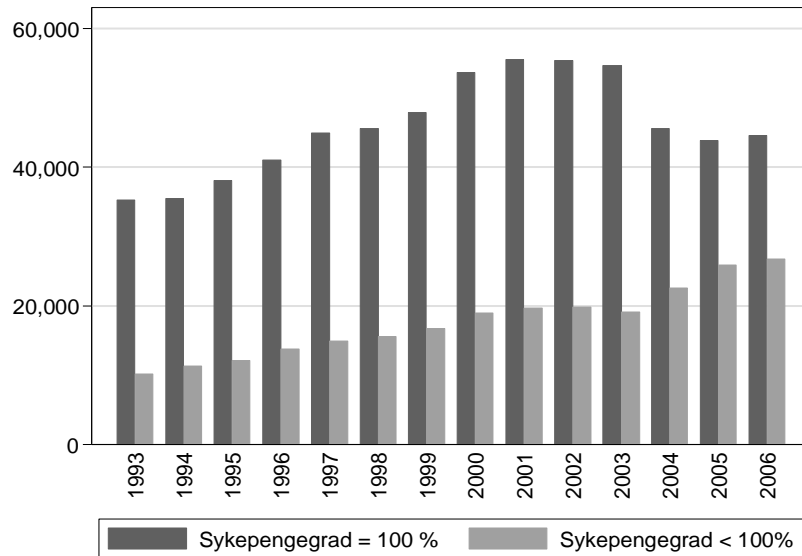
Tabell 3 Sykefraværsdager per arbeider og antall sykefraværstilfeller per arbeider.
Gjennomsnittsverdier 1993-2006.

	Hele utvalget		
	Observasjoner	Sykefraværsdager per arb.	Antall tilfeller per arb.
Offentlig sektor	1796867	12.28	0.25
Privat sektor	3003855	10.21	0.19
	Stabile arbeidere		
Offentlig sektor	1489507	10.82	0.23
Privat sektor	2546686	8.54	0.18

Tabell 3 illustrerer forskjellen i sykefraværet mellom privat og offentlig sektor²⁷, fordelt på begge utvalgene. Ansatte i offentlig sektor har flere sykefraværstilfeller enn ansatte i privat sektor, og sykefraværene er i gjennomsnitt av lengre varighet. Tabellen illustrerer også at det gjennomsnittlige sykefraværet er lavere i utvalget med stabile arbeidere enn i det fullstendige utvalget.

Andelen graderte sykemeldinger har økt i løpet av analyseperioden. Figur 5 på neste side, illustrerer denne utviklingen, og vi ser at økningen har vært særlig kraftig etter regelendringene i Ot. Prp 48 2003-2004 ble innført (jf. avsnitt 2.2.1).

²⁷ Offentlig sektor kan defineres som “den sektoren som produserer fellestjenester som i hovedsak er finansiert via skatter og avgifter”. Hvilke næringer vi har definert som offentlig sektor i denne oppgaven presenteres i avsnitt 5.2.2.



Figur 5: Antall individer med sykepengetilfeller som ikke er gradert, og antall individer med graderte sykepengetilfeller.

5.2. Variablene

I denne delen av oppgaven vil vi presentere de ulike variablene som blir brukt i regresjonsanalysen i neste kapittel. Vi antar på forhånd at de uavhengige variablene har en signifikant påvirkning på sannsynligheten for sykefravær og lengden på fraværene. På bakgrunn av tidligere analyser og økonomisk teori vil vi drøfte forventede verdier og fortegn på de ulike forklaringsvariablene. Selve analysen vil bestå av to modeller for sykefravær, en som estimerer insidens og en som estimerer varighet.

5.2.1. Avhengig variabel: Sykefravær

Den første modellen skal estimere sannsynligheten for sykefravær gitt den regionale arbeidsledigheten og en rekke personlige karakteristika. I denne modellen er sannsynligheten for sykefravær definert som avhengig variabel. Variabelen tar verdien 1 dersom individ i starter minst ett sykefravær utover arbeidsgiverperioden i løpet av år t . Motsatt, dersom individet ikke starter minst ett sykefravær i år t , vil variabelen ta verdien 0. Denne definisjonen er konsistent med analysen til ABN og passer best med rådata. Årsaken er at i FD-Trygd er antall sykefraværsdager registrert i år t , selv om fraværet strekker seg til år $t+1$. Alternativt kunne vi definert den diskrete variabelen til å indikere hvorvidt individ i var

fraværende minst en dag i løpet av år t . Begge metodene gir samme resultat basert på rådata; ett registrert fravær i år t , men den sistnevnte definisjonen er ikke konsistent med det vi faktisk måler.

Den andre modellen skal estimere varigheten (antall dager) på sykefravær som starter i år t . Variabelen er definert slik at den summerer antall dager for alle fravær som starter i år t . Selv om et fravær strekker seg til år $t+1$ vil alle dagene registreres i det året fraværet starter. Dette vil ikke gi systematiske skjevheter i vårt utvalg slik vi har definert den avhengige variabelen. Det er imidlertid viktig å påpeke at variabelen “antall dager” kan være summen av flere sykefravær utover arbeidsgiverperioden. Derfor kan ikke resultatene fra denne modellen direkte tolkes som varighetsmodeller (ABN).

5.2.2. Uavhengige variabler

Arbeidsledighet

Arbeidsledige blir definert som arbeidsføre personer som søker inntektsgivende arbeid ved arbeidskontorene²⁸. I tillegg må vedkommende ha vært uten inntektsgivende arbeid de siste to ukene. Variabelen måler arbeidsledigheten som prosent av den registrerte arbeidsstyrken for hver kommune²⁹. Tallene er hentet fra Norsk Samfunnsvitenskapelig Datatjeneste (NSD). ABN bruker fylkesvis i stedet for kommunal ledighet. De begrunner valget med at fylke, i større grad enn kommuner, konstituerer et geografisk arbeidsmarked. Vi velger å bruke kommunal ledighet fordi den gir mer detaljert informasjon. Det kan tenkes situasjoner der arbeidsledigheten på fylkesnivå er lav, mens noen kommuner har høy arbeidsledighet, og omvendt. Som kontroll har vi utført regresjonene med fylkesvis arbeidsledighet, uten at resultatene endrer seg nevneverdig.

Figur 6 illustrerer arbeidsledigheten og antall sykefraværsdager per arbeider i perioden 1993-2006. Arbeidsmarkedet er påvirket av konjunktursvingninger i økonomien og det fremgår av figuren at det i perioden fra 1993 til 1998 var en sterk sysselsettingsvekst som førte til at

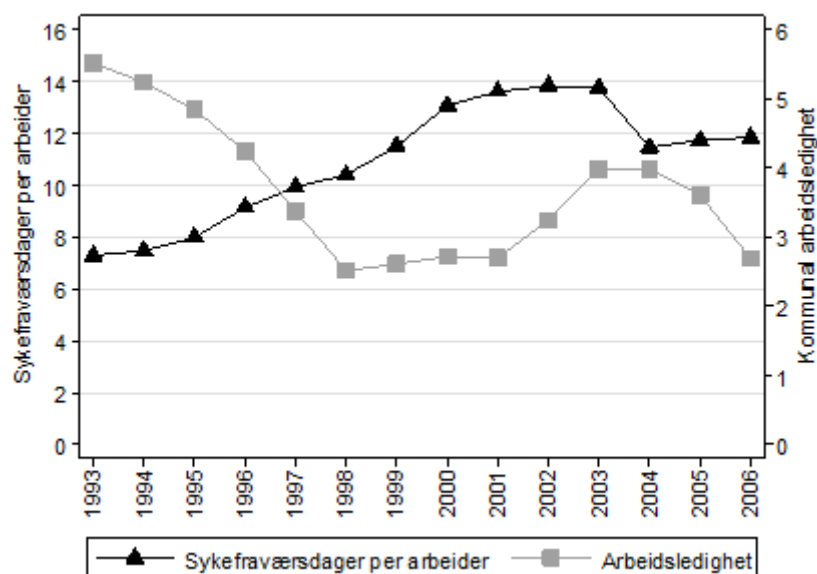
²⁸ Et annet mål på arbeidsledighet er SSB sine tall som er basert på Arbeidskraftundersøkelsen (AKU). AKU-tallene inkluderer også arbeidsledige som ikke registrerer seg ved arbeidskontorene, og en del av de som går på arbeidsmarkedstiltak.

²⁹ NSD oppgir kvartalsvis ledighet. Vi har regnet ut gjennomsnittlig ledighet for hvert år (ikke sesongjustert). Det mangler observasjoner for 1. kvartal 1993, derfor er ledigheten dette året beregnet som gjennomsnittet av ledigheten i 2., 3. og 4. kvartal.

arbeidsledigheten sank kraftig. Ledigheten var deretter stabil rundt 2,5 %, frem til den steg moderat i perioden fra 2002-2004. Økonomien gikk deretter inn i en høykonjunktur med lav ledighet frem til finanskrisen i 2008. Figuren viser at det er en negativ sammenheng mellom arbeidsledigheten og sykefraværet i perioden fra 1993-1998. Den negative sammenhengen er imidlertid ikke like fremtredende etter denne perioden. Fra og med 1999 stiger både arbeidsledigheten og sykefraværet frem mot 2003. Etter 2004 synker arbeidsledigheten kraftig, mens sykefraværet stiger moderat.

En prosyklisk sammenheng mellom arbeidsledigheten har blitt påvist av blant annet ABN, Dyrstad & Ose (2002) og Nossen (2009).

Korrelasjonskoeffisienten mellom antall sykefraværsdager per arbeider og gjennomsnittlig kommunal arbeidsledighet er -0.73 i perioden 1993-2006. Vi forventer dermed at arbeidsledigheten skal ha en negativ påvirkning på sykefraværet.



Figur 6: Sykefraværsdager per arbeider og kommunal arbeidsledighet (årgjennomsnitt).

Tall for arbeidsledigheten er hentet fra NSDs kommunedatabase. Tallene for sykefravær er ikke korrigert for at arbeidsgiverperioden ble utvidet fra 14 til 16 dager i 1998.

Offentlig sektor

Denne variabelen er en dummyvariabel som indikerer om individ i er ansatt i offentlig sektor i år t . Variabelen defineres med utgangspunkt i hvilken arbeidsnæring individet er ansatt i³⁰. Vi definerer i denne oppgaven offentlig sektor som følgende næringer:

- Offentlig administrasjon, forsvar, politi og rettsvesen
- Undervisning, helse- og andre sosialtjenester
- Drift av bibliotek og museer
- Vannforsyning
- Annen offentlig tjenesteyting

Privat sektor defineres som følgende næringer:

- Jordbruk, skogbruk, fiske og fangst
- Oljeutvinning og bergverksdrift
- Industri
- Kraftforsyning
- Bygge- og anleggsvirksomhet
- Varehandel, hotell- og restaurantvirksomhet
- Bank- og finansieringsvirksomhet, forsikring, eiendomsdrift m.m.
- Annen privat tjenesteyting

Studier indikerer at sykefraværet i offentlig sektor er høyere enn i privat sektor. Lindbeck et al. (2009) betegner dette som en “empirisk regularitet”, og forklarer forskjellen med at arbeidsgivere i privat sektor har sterkere insentiver til å forhindre sykefravær. Det kan også være slik at arbeidere med høyt fravær foretrekker stillinger med høyere jobbsikkerhet, og at seleksjon sørger for at disse blir ansatt i offentlig sektor.

Offentlig sektor er antatt å være mindre påvirket av konjunktursvingninger enn privat sektor. En av årsakene til dette er at varene og tjenestene som produseres i de to sektorene er forskjellige. Etterspørsel etter varer og tjenester fra privat sektor påvirkes i større grad av svingninger i den internasjonale økonomien gjennom eksporten. Offentlige goder blir sjelden

³⁰ For årene 1993-1995 er det næringskoder fra SN82 som ligger til grunn, mens for 1996-2006 bygger næringskodene på SN96 (med en mindre oppdatering av standarden i 2002).

eksportert og etterspørselen etter f. eks. helse- og sosialtjenester påvirkes i større grad av underliggende strukturelle faktorer i økonomien. En annen årsak er at myndighetene ofte bruker finanspolitikk som virkemiddel for å motvirke konjunktursvingninger. I en nedgangskonjunktur vil dette innebære en økning i utgifter for å stimulere produksjonen i offentlig sektor. Vi antar derfor at effekten av arbeidsledigheten er mindre i offentlig enn privat sektor, og vil derfor kontrollere for dette i regresjonene ved å inkludere et interaksjonsledd.

Alder

Individene i begge utvalgene er mellom 30-55 for hele perioden 1993-2006. Det er rimelig å anta at sannsynligheten for sykefravær øker med alderen. Hovedsakelig er dette grunnet en generell forverring av helsetilstanden. Vi forventer at risikoen for visse sykdommer som for eksempel hjerte- og karsykdommer er høyere for eldre mennesker. Arbeidere som er ansatt i sektorer med høy arbeidsbelastning vil også oppleve en forverring av helsen jo lenger de har vært ansatt.

Det kan også eksistere effekter som fører til en negativ sammenheng mellom alder og sykefravær. Betydningen av å akkumulere pensjonspoeng i slutten av karrieren kan gi økte insentiver til arbeid. Ifølge hypotesen om forverringen av sosiale normer (jf. 3.3.2) har eldre arbeidere i tillegg høyere arbeidsmoral enn yngre arbeidere.

Det er viktig å kontrollere for alderseffekten i modellen, men basert på argumentasjonen ovenfor er fortegnet til koeffisienten vanskelig å predikere. Siden vi har ekskludert individer i begge ender av aldersskalaen, vil effekten av alder sannsynligvis underestimeres i forhold til den arbeidende populasjonen. Fast effekt-modeller bruker ikke variasjonen mellom individ. Effekten av alder på sykefravær måles derfor kun som effekten av å bli ett år eldre.

Studier indikerer at effekten av alder er tvetydig. Biørn et al. (2010) finner at alder øker sannsynligheten for sykefravær, mens ABN estimerer en negativ, avtakende effekt. Sistnevnte studie forklarer den negative effekten med at aldersgruppen over 55 år er ekskludert i utvalget. Forskning indikerer at helsen, målt ved en indeks, øker de første 25 arbeidsårene (Lindeboom and Kerkhofs (2002), referert til i ABN). Det er forventet at individer over 55 år har dårligst helse, dermed ekskluderes de fleste individer som ville bidratt til en positiv effekt

av alder på sannsynligheten for sykefravær. Siden vårt utvalg også er begrenset til individer mellom 30-55 år, forventes lignende effekter.

Variabelen alder vil mest sannsynlig være sterkt korrelert med andre variabler som f. eks. arbeidserfaring og inntekt. Vi har tilgang på variabelen arbeidserfaring i datamaterialet, men har droppet denne i modellene. I det stabile utvalget, slik vi har definert det, vil alder og arbeidserfaring lide under perfekt multikollinearitet. Ved å utelukke variabelen arbeidserfaring, kan det tenkes at aldersvariabelen vil fange opp noe av denne effekten siden disse er sterkt korrelert. Avhengig av fortegnet til koeffisienten for arbeidserfaring, vil koeffisienten til alder bli forventningsskjev oppover eller nedover. Som en test, har vi estimert begge modellene inkludert variabelen arbeidserfaring for det fullstendige utvalget, og funnet at koeffisienten er positiv. Dette vil medføre at variabelen alder er positivt forventningsskjev.

For å kontrollere for eventuelle ikke-lineære effekter av alder, inkluderer vi et annengradsledd ($alder^2$) i regresjonene.

Utdanning

Variabelen måler normert antall år for høyeste oppnådde grad eller utdanning (etter NUS2000 standarden). Lavt utdannede yrker kan være mer fysisk krevende enn høyt utdannede yrker. Dette øker sannsynligheten for muskel- og skjelettlidelser som utgjør en stor del av sykefraværet i Norge. Høyt utdannede yrker innebærer ofte mer varierte arbeidsoppgaver og større grad av selvstendighet, noe som øker motivasjonen for arbeid.

Utdanning er også en indikator på sosiale ulikheter. Studier viser at grupper med lav sosial status og lav utdanning har et dårligere kosthold og generelt dårligere helse enn grupper med høy sosial status (Holmboe-Ottesen et al. 2004).

Lav utdanning kan medføre mindre jobbsikkerhet. Dette kan gi lavt utdannede arbeidere insentiver til å redusere sitt sykefravær (jf. disiplineringshypotesen). Selv om disiplineringshypotesen gjelder, er det lite sannsynlig at denne effekten vil være sterkere enn de to førstnevnte. Vi predikerer derfor at utdanning har en negativ effekt på sykefraværet.

I fast effekt modeller kan det oppstå problemer med variabler som har lite innenfor-individ variasjon (jf. avsnitt 4.2.3). Siden utvalget er begrenset til individer mellom 30-55 år er dette særlig problematisk for variabelen utdanning. Få individer i denne aldersgruppen tar flere år

med utdanning. ABN har droppet variabelen i regresjonene, mest sannsynlig på grunn av lite variasjon. Bratberg et al. (2002) finner at kvinner med høy utdanning har lavere sykefravær enn kvinner med lav utdanning.

Inntekt

Inntektsvariabelen måler brutto lønnsinntekt inkludert sykepenger. Tallene er oppgitt per 10 000 og omregnet til 2006 kroner. Inntekt kan påvirke sykefraværet på ulike måter, men sykefraværet vil også påvirke inntekten (jf. Markussen (2010)).

På den ene siden er lavtlønnede yrker mer fysisk krevende enn høytlønnede, og arbeidsoppgavene er ofte mindre varierte. Dette kan resultere i lavere motivasjon for arbeid og mer sykefravær. På samme måte som utdanning, er inntekt en indikator på sosial status og mange av argumentene for en negativ effekt av utdanning gjelder også for inntekt. Inntekt og utdanning vil også være sterkt korrelert, et problem som beskrives nærmere i neste kapittel.

På den andre siden, vil høyt sykefravær kunne føre til reduserte karrieremuligheter og lavere lønnsvekst. I tillegg eksisterer det sannsynligvis en seleksjon av arbeidere med nedsatt helse/funksjonsevne til lavtlønnede yrker. Denne to-veis kausaliteten gjør koeffisientene for inntekt vanskelige å tolke presist.

Inntekt er korrelert med uobserverbare forhold, som f.eks. helse. Som diskutert i 4.5.1 lagges inntektsvariabelen ett år for å kontrollere for endogenitet, som betyr at vi ikke har observasjoner for inntekt i 1993. Det er viktig å merke seg at regresjonen derfor bare blir utført for årene 1994-2006. Vi kontrollerer for eventuelle ikke-lineære effekter av inntekt ved å inkludere et annengradsledd, $inntekt^2$, i regresjonene³¹.

Deltid

”Deltid” er en dummyvariabel som tar verdien 1 dersom individet arbeider mindre enn 20 timer i uken. I tidsperioden før 2003 har vi tatt utgangspunkt i en variabel som viser forventet arbeidstid, mens fra og med 2003 bruker vi en variabel for faktisk arbeidstid. Ifølge arbeidstilbudsmodellen kan individer ønske å bruke sykefravær for å korrigere en mulig ”mismatch” mellom arbeidskontrakten og preferansene for fritid og arbeid. Deltidsansatte vil i

³¹ Regresjonene er også utført med inntektsvariabelen fra år t , uten at koeffisientene endrer seg nevneverdig.

utgangspunktet ha mer fritid til disposisjon, som kan medføre at de har lavere sykefravær enn heltidsansatte. Motsatt kan sammenhengen mellom deltidsarbeid og sykefravær være positiv som følge av seleksjon av arbeidere med nedsatt helse/funksjonsevne inn til deltidsarbeid. Studier indikerer likevel at sykefraværet er lavere i bedrifter med mange deltidsansatte³². Vi forventer derfor at koeffisienten for deltidsarbeid er negativ.

Gift

Gift er en dummyvariabel som tar verdien 1 hvis personen er gift, og 0 ellers. Flere studier som undersøker sammenhengen mellom sivil status og sykefravær indikerer at gifte har lavere sykefravær enn andre grupper³³. Det kan tenkes at det foregår en seleksjon til ekteskapelig status ved at de med dårlig helse har mindre sannsynlighet for å bli gift enn de med god helse, og at helseproblemer øker sannsynligheten for at ekteskap skal gå i oppløsning. Dette vil føre til en negativ effekt av ekteskapelig status. Undersøkelser viser at sannsynligheten for å bli uførepensjonert også er påvirket av sivil status.

Bratberg et al. (2002) undersøker om dobbeltarbeidshypotesen kan forklare kvinners relativt høye sykefravær. De finner at en livsstil som innebærer barn og ekteskap, kombinert med arbeidsmarkedstilknytning, medfører høyere sykefravær for kvinner. Dersom dobbeltarbeidshypotesen er sann vil den totale arbeidsbyrden for gifte kvinner øke relativt til menn og enslige kvinner uten barn. Dette kan medføre en positiv effekt av ekteskapelig status for kvinner.

Antall barn under 11 år (barnu11)

Variabelen gir informasjon om hvor mange barn individet har som er under 11 år. Folketrygden gir rett til omsorgspenger i 10 dager for hver av foreldrene i forbindelse med sykdom hos barn (under 12 år) eller barnepasser. Foreldres fravær ved barns sykdom registreres ikke som sykefravær i statistikken. Det er likevel sannsynlig at antall små barn i familien påvirker sykefraværet til foreldrene. Hvis hypotesen om “den doble byrde” er sann, vil dette spesielt gjelde kvinners sykefravær.

³² Se f.eks. Moland (2007) eller Chaudhury og Ng (1992).

³³ Se Hansen (1999) for en oversikt.

Studier viser ofte at kvinner med flere barn har lavere fravær enn kvinner med få barn. Bratberg et al. (2002) argumenterer for at disse resultatene kan skyldes seleksjon, og dermed ikke representerer en kausaleffekt. Ressurssterke individer med god helse kan tenkes å kombinere karriere med flere barn. Hvis de uobserverte ressursene reduserer fraværet direkte, er det ikke effekten av antall barn som reduserer fraværet. Når det kontrolleres for slik seleksjon, finner Bratberg et al. (2002) at antall barn under 11 øker sykefraværet for kvinner.

Vi inkluderer også en interaksjonsdummy for å undersøke om effekten av antall barn under 11 på sykefraværet er forskjellig for gifte og ugifte. Variabelen tar verdien 1 hvis individet både er gift og har minst ett barn under 11 år. Ser vi bort fra samboerpar med barn, antar vi at belastningen ved å ha barn er større for ugifte enn for gifte. Belastningen kan resultere i økt sykefravær³⁴.

Graderte sykemeldinger

Variabelen *gradert* er en dummyvariabel som tar verdien 1 hvis sykepengegraden er mindre enn 100 %. Denne variabelen brukes kun i modellen som estimerer antall sykefraværsdager. *Gradert* er ikke inkludert i sannsynlighetsmodellen siden det ikke er en ex ante faktor, dvs. et sykefravær blir klassifisert som gradert etter at selve sykefraværet har inntruffet og vil således ikke påvirke ex ante sannsynligheten for sykefravær i en statisk paneldatamodell.

Hvordan sykepengegraden påvirker lengden på sykefraværet er et viktig tema, siden de nyeste reformene i sykepengeordningen fokuserer på økt bruk av graderte sykemeldinger. Det eksisterer likevel et mulig seleksjonsproblem ved å bruke denne variabelen i modellen. Graderte sykemeldinger vil ofte brukes i tilfeller hvor man forventer at sykefraværet skal vare lengre enn normalt. Arbeidere med graderte sykemeldinger kan være generelt sykere enn arbeidere med vanlige sykemeldinger. Koeffisienten kan dermed være positiv som følge av seleksjon.

Bruk av gradert sykemelding skal opprettholde kontakten mellom den sykemeldte og arbeidsplassen. Dette kan redusere insentivene til overforbruk av sykepengeordningen siden man må møte på jobb uansett. Studier indikerer at det å jobbe kan gi en positiv effekt på rygg-

³⁴Mastekaasa (2000) bruker "Antall barn under 16" som proxyvariabel for den doble byrde. Resultatet av analysen indikerer at det er lite sammenheng mellom antall barn og sykefravær for gifte kvinner, mens koeffisienten er mer signifikant for enslige mødre.

og skjelettlidelser og lettere psykiske problemer³⁵. Vi forventer derfor at graderte sykemeldinger vil redusere langtidsfraværet, men seleksjonsproblemet kan føre til at koeffisienten får positivt fortegn.

5.3. Deskriptiv statistikk

Tabell 4 inneholder deskriptiv statistikk for de uavhengige variablene. Vi inkluderer i tillegg tre variabler for sykefravær. Den første illustrerer gjennomsnittet av alle sykefraværsdager, den andre illustrerer gjennomsnittet for alle sykefraværsdager betinget på at antall dager er større enn null, mens den siste illustrerer antall sykepengetilfeller. Variabelen *arbeidserfaring* er inkludert i tabell 4 for å gi et mer detaljert bilde på forskjellen mellom utvalgene, men er likevel ekskludert i regresjonene grunnet perfekt multikollinearitet.

Vi ser at kvinner i det stabile utvalget har flere gjennomsnittlige sykefraværsdager enn i det fullstendige utvalget, mens gjennomsnittet er uendret for menn. Dette betyr ikke at kvinnelige arbeidere i det stabile utvalget har høyere fravær enn kvinnelige arbeidere i det fullstendige utvalget. Blant de ikke-stabile arbeiderne står flere individer utenfor arbeidsmarkedet, og disse har mindre tilgang på sykepenge. Regner vi ut fraværstallene per arbeider, er disse lavere i det stabile utvalget (11,5 mot 13,7), som betyr at arbeidere i det fullstendige utvalget har høyest gjennomsnittlig fravær. Gjennomsnittet for alle sykefraværsdager betinget større enn null, illustrerer at lengden på sykefraværene også er kortere i det begrensede utvalget, både for kvinner og menn. Antall sykefraværstilfeller er også lavere for stabile arbeidere.

Individene i det stabile utvalget har i gjennomsnitt flere år med utdanning, lengre arbeidserfaring og høyere inntekt enn individer i det fullstendige utvalget. Gjennomsnittsalderen er litt høyere blant de stabile arbeiderne, og andelen individer med ekteskapeelig status er også høyere.

Kvinner har i gjennomsnitt flere sykefraværsdager enn menn. Kjønnforskjellen er tilstede i alle sykefraværsvariablene i tabell 4, dette indikerer at kvinner både har flere sykefravær enn menn og at fraværene har lengre varighet.

I det fullstendige utvalget er antall individer jevnt fordelt på kvinner og menn, mens i det stabile utvalget er ca. 30 000 flere menn enn kvinner. Forskjellen indikerer at kvinner har en

³⁵ Se f.eks. Waddel (2004) og Waddel og Burton (2006), sitert i Markussen et al. (2010).

løsere arbeidsmarkedstilknytning enn menn. Dette må sees i sammenheng med at kvinner har lavere gjennomsnittlig utdannelse, færre år med arbeidserfaring, lavere inntekt og jobber mer deltid enn menn. Andelen kvinner ansatt i offentlig sektor er i gjennomsnitt over dobbelt så høy som andelen menn ansatt i offentlig sektor. I det fullstendige utvalget har kvinner i gjennomsnitt flere barn under 11 år enn menn, mens i det begrensede utvalget er det motsatt.

Tabell 4: Deskriptiv statistikk.

Utvalgsgjennomsnitt og standardavvik for individer i alderen 30-55 for årene 1993-2006.

Variabel	Hele utvalget				Begrenset utvalg			
	Menn		Kvinner		Menn		Kvinner	
	Gj.snitt	St. avvik	Gj.snitt	St. avvik	Gj.snitt	St. avvik	Gj.snitt	St. avvik
Sykefraværsdager	6.95	32.71	10.91	40.12	6.94	31.60	12.39	41.20
Syk.fra.dager > 0	67.52	79.41	69.29	78.56	60.02	73.82	63.42	73.84
Antall tilfeller	1.26	0.62	1.29	0.67	1.24	0.60	1.28	0.67
Alder	41.53	7.44	41.75	7.46	41.72	7.55	42.05	7.60
Utdanning	12.60	3.63	12.35	3.51	13.02	3.27	12.87	3.22
Arbeidserfaring ³⁶	19.90	8.49	16.07	8.02	21.39	7.62	18.70	6.95
Inntekt	29.05	24.88	18.21	14.45	36.88	21.24	25.05	11.87
Deltid	0.03	0.18	0.17	0.38	0.03	0.16	0.14	0.35
Gift	0.52	0.50	0.57	0.50	0.62	0.49	0.62	0.49
Barn under 11	0.60	0.91	0.65	0.92	0.64	0.92	0.60	0.87
Offentlig sektor	0.21	0.40	0.54	0.49	0.21	0.41	0.55	0.51
Gradert	0.02	0.15	0.05	0.22	0.03	0.16	0.07	0.25
Arbeidsledighet	3.63	1.39	3.61	1.40	3.62	1.38	3.63	1.41
Observasjoner	3409321		3426885		2403503		2098927	
Individer	364270		363788		262683		233255	

³⁶ Ikke inkludert i regresjonene grunnet perfekt multikollinearitet.

Kapittel 6: Empiriske resultater

I dette kapittelet presenteres resultatene fra den empiriske analysen. Som nevnt skilles det mellom to effekter: effekter på *sannsynligheten for* (insidens) og *lengden på* (varighet) sykefravær. Sannsynligheten for sykefravær estimeres ved å bruke en ikke-lineær diskret sannsynlighetsmodell³⁷ og lengden på sykefraværet estimeres ved å bruke en lineær modell.

Begge modellene estimeres for både det fullstendige og det begrensede utvalget. Hvis sammensetningshypotesen er sann, forventes det at effekten av arbeidsledighet forsvinner eller reduseres i utvalget med stabile arbeidere. Da kan det ikke utelukkes at variasjonen i sykefraværet over konjunktorene skyldes at arbeidere med dårligere helse/funksjonsevne utgjør en større andel av arbeidsstyrken i perioder med høy sysselsetting. Den deskriptive statistikken i kapittel 5 avslørte signifikante forskjeller mellom menn og kvinners sykefravær. På bakgrunn av dette er regresjonene utført separat for kvinner og menn.

I kapittel 6.2 presenteres resultater fra estimeringen av den diskrete sannsynlighetsmodellen, mens resultater fra den lineære modellen presenteres i kapittel 6.3. Først diskuteres valget mellom fast og tilfeldig effekt (jf. diskusjonen i avsnitt 4.4).

6.1. Fast eller tilfeldig effekt?

Avhengig av antagelsene om den individspesifikke uobserverte effekten, kan vi velge å bruke fast eller tilfeldig effekt-estimatorer for paneldatamodellene. Hvis det eksisterer korrelasjon mellom forklaringsvariablene og den uobserverte effekten, vil RE-estimatoren være inkonsistent, mens FE-estimatoren er robust ovenfor denne typen korrelasjon. Hvis det ikke eksisterer korrelasjon, er RE-estimatoren mer effisient enn FE-estimatoren siden den utnytter mer av variasjonen i datamaterialet.

Basert på spesifikasjonen av modellene er det ikke rimelig å forvente at den uobserverte effekten er ukorrelert med forklaringsvariablene. Studier indikerer at individuell helsetilstand er korrelert med sosioøkonomiske faktorer som inntekt og utdanning³⁸. Til tross for at vi har

³⁷ For å sammenligne har vi også estimert en lineær fast effekt sannsynlighetsmodell. Modellen gir tilsvarende resultater når det gjelder signifikansnivå og fortegn for variabelen arbeidsledighet. Resultatene for resten av variablene er også relativt like som i det ikke-lineære tilfellet (se appendiks A).

³⁸ Se f.eks. (Wagstaff & Doorslaer 2000).

et velutbygd helsesystem som er tilgjengelig for alle, er de sosiale ulikhetene i helse dokumentert også i Norge (Folkehelseinstituttet 2007). Hausmantesten kan brukes for å avgjøre hvilken estimator som egner seg best basert på kriteriene ovenfor. Testen kontrollerer om det eksisterer korrelasjon mellom forklaringsvariablene og de uobserverte egenskapene.

Ved å utføre en Hausmantest på nullhypotesen $H_0: \text{Corr}(\eta_i | X_{it}) = 0$ for den lineære modellen (se modell 6.2), forkastes H_0 med en testobservator 12442.98 og p-verdi 0,0000³⁹. Testen indikerer systematiske forskjeller mellom FE- og RE-estimatene, og vi forkaster dermed hypotesen om at det ikke eksisterer korrelasjon mellom forklaringsvariablene og uobserverte individsesifikke egenskaper. Tilfeldig effekt-estimatoren er inkonsistent for den lineære modellen⁴⁰. Det er en rimelig antagelse at resultatet også gjelder for fast effekt logitmodellen, siden forklaringsvariablene er de samme (unntatt *gradert*).

6.2. Resultater fra fast effekt logitmodellen

For å isolere effekten av arbeidsledighet på sykefraværet, kontrolleres det for individsesifikke karakteristikk. Siden vi estimerer en fast-effekt modell, kontrolleres det også for uobserverbare individsesifikke egenskaper som er konstante over tid. Den første regresjonsmodellen estimerer sannsynligheten for sykefravær:

$$\text{Sykefravær}_{it}^* = \beta X'_{it} + \delta \text{arbeidsledighet}_{it} + \alpha_i + u_{it} \quad (6.1)$$

Sannsynligheten for sykefravær kan ikke observeres. Derfor defineres en indikatorvariabel som indikerer hvorvidt individ i har hatt minst ett sykefravær som starter i år t , $\text{Sykefravær}_{it} = 1(\text{Sykefravær}_{it}^* > 0)$. X'_{it} er en vektor av de individsesifikke forklaringsvariablene beskrevet i kapittel 5 (unntatt *gradert*), og β er en K -dimensjonal vektor av parametere. Parameteren δ måler effekten av arbeidsledigheten på sannsynligheten for sykefravær, α_i er en individsesifikk komponent som er konstant over tid og u_{it} er hvit støy.

³⁹ Kritisk verdi for Hausmantesten er $\chi^2_{0.01}(13) = 27,69$.

⁴⁰ Årsaken til at vi ikke utfører Hausmantesten på sannsynlighetsmodellen (6.1), er at sannsynlighetsmaksimering av en tilfeldig effekt probitmodell ikke konvergerer i vårt datasett. Vi får dermed ingen resultater fra denne modellen som vi kan sammenligne med fast effekt logitmodellen.

Modellen estimeres med betinget sannsynlighetsmaksimering (betinget på $\sum_{t=1}^T Y_{it}$), som betyr at bare individer som endrer status minst en gang brukes i estimeringen.

Resultatene fra fast effekt logitmodellen er presentert i tabell 5. Koeffisientene er på samme måte som i ABN, oppgitt som logit-koeffisienter og er derfor vanskelige å tolke direkte. Diskusjonen baseres derfor på fortegn og signifikansnivå. Generelt er estimatorene i det stabile utvalget (nedre panel) mindre presise enn i det fullstendige utvalget (øvre panel). Større utvalg gir mer variasjon i X_{it} som igjen gir koeffisientene i β lavere varians og påfølgende standardfeil (Wooldridge (2009) s.56). Høye standardfeil gir lavere t-verdier og mindre signifikante estimatorene. Dette er særlig problematisk for variabler som har lite innenfor-individ variasjon, som for eksempel sivilstatus og antall år med utdanning.

Tabell 5: Resultater fra estimering av fast effekt logitmodellen. Oppgitt som logit-koeffisienter. 1994-2006.

Fullstendig utvalg									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0380	0.0021	0.0000	-0.0215	0.0028	0.0000	-0.0569	0.0031	0.0000
Offentlig sektor	-0.1014	0.0142	0.0000	-0.1029	0.0269	0.0001	-0.1066	0.0180	0.0000
Off.sekt*ledighet	-0.0017	0.0031	0.5965	-0.0324	0.0061	0.0000	0.0160	0.0041	0.0001
Alder	-0.1128	0.0039	0.0000	-0.0490	0.0061	0.0000	-0.1673	0.0051	0.0000
Alder ²	0.0017	0.0000	0.0000	0.0012	0.0001	0.0000	0.0021	0.0001	0.0000
Utdanning	-0.0093	0.0029	0.0014	-0.0215	0.0052	0.0000	-0.0058	0.0036	0.1048
Inntekt	0.0377	0.0004	0.0000	0.0221	0.0005	0.0000	0.0585	0.0007	0.0000
Inntekt ²	-0.0002	0.0000	0.0000	-0.0001	0.0000	0.0000	-0.0004	0.0000	0.0000
Deltid	-0.3492	0.0073	0.0000	-0.4684	0.0183	0.0000	-0.2852	0.0080	0.0000
Gift	0.0840	0.0080	0.0000	0.0349	0.0119	0.0035	0.1137	0.0108	0.0000
Barn under 11	-0.0992	0.0054	0.0000	0.0172	0.0083	0.0391	-0.1840	0.0072	0.0000
Gift*barnu11	0.0026	0.0061	0.6703	0.0157	0.0095	0.0990	-0.0137	0.0081	0.0903
Log likelihood	-917779			-377560			-538422		
Observasjoner	2507205			1084227			1422978		
Individer	291374			122737			168637		
Stabilt utvalg (pensjonsgivende inntekt alle år de inngår i utvalget)									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0336	0.0023	0.0000	-0.0170	0.0030	0.0000	-0.0532	0.0035	0.0000
Offentlig sektor	-0.0923	0.0161	0.0000	-0.0711	0.0293	0.0153	-0.1074	0.0207	0.0000
Off.sekt*ledighet	-0.0052	0.0035	0.1333	-0.0359	0.0066	0.0000	0.0145	0.0045	0.0015
Alder	-0.1029	0.0042	0.0000	-0.0396	0.0064	0.0000	-0.1625	0.0056	0.0000
Alder ²	0.0016	0.0000	0.0000	0.0011	0.0001	0.0000	0.0021	0.0001	0.0000
Utdanning	-0.0031	0.0035	0.3836	-0.0125	0.0060	0.0381	-0.0006	0.0044	0.8884
Inntekt	0.0288	0.0005	0.0000	0.0164	0.0006	0.0000	0.0473	0.0009	0.0000
Inntekt ²	-0.0001	0.0000	0.0000	-0.0001	0.0000	0.0000	-0.0003	0.0000	0.0000
Deltid	-0.3318	0.0086	0.0000	-0.4234	0.0210	0.0000	-0.2753	0.0095	0.0000
Gift	0.0771	0.0088	0.0000	0.0298	0.0129	0.0211	0.1094	0.0121	0.0000
Barn under 11	-0.1017	0.0060	0.0000	0.0158	0.0089	0.0774	-0.2016	0.0082	0.0000
Gift*barnu11	0.0122	0.0068	0.0732	0.0188	0.0102	0.0660	-0.0003	0.0092	0.9734
Log likelihood	-771563			-330922			-439420		
Observasjoner	2118055			959587			1158468		
Individer	232419			102828			129591		

6.2.1. Tolkning av koeffisienter

Arbeidsledighet

Øverste del av tabell 5 viser at når det kontrolleres for individuelle karakteristikk og uobserverte tidskonstante effekter, så finner vi en negativ effekt av kommunal arbeidsledighet⁴¹. Den negative koeffisienten er signifikant i alle tre regresjonene. I nederste del av tabellen rapporteres regresjonskoeffisientene for utvalget med stabile arbeidere. Den negative effekten av arbeidsledigheten eksisterer fortsatt, men er litt svakere enn i det fullstendige utvalget. Også her er alle koeffisientene signifikante.

Resultatene kan knyttes til problemstillingen som ble formulert innledningsvis i oppgaven. En negativ, signifikant effekt av arbeidsledighet på sannsynligheten for sykefravær blant arbeidere som har en stabil arbeidsmarkedstilknytning, indikerer at variasjoner i sykefraværet ikke er en ren sammensetningseffekt. Basert på våre resultater kan vi likevel ikke utelukke sammensetningshypotesen i samme grad som ABN. De finner en *sterkere* og mer signifikant effekt av arbeidsledigheten på sannsynligheten for sykefravær blant stabile arbeidere. Motsatt finner vi en *svakere* effekt av arbeidsledigheten blant de stabile arbeiderne. Sammenligner vi koeffisienten med resultatene i ABN, finner vi at koeffisienten er tilnærmet identisk for de fullstendige utvalgene (-0,036 i ABN og -0,038 i tabell 5). I det stabile utvalget derimot, estimerer de en koeffisient på -0,048, mens vår modell gir -0,034. Effekten av arbeidsledigheten på sykefraværet til stabile arbeidere er dermed svakere i tidsperioden 1994-2006 enn den var i 1990-1995. Disiplinerings effekter eller sykefravær forårsaket av “stress” har tilsynelatende blitt redusert.

Det er mulig å teste hvorvidt koeffisientene er signifikant forskjellige mellom utvalgene⁴². En slik test avslører at koeffisientene er ulike med t-verdier lik 3.23 for alle og 3.55 for menn. For kvinner derimot, er den tilhørende t-verdien lik 1.49. Det betyr at vi ikke kan forkaste nullhypotesen om at effekten av arbeidsledighet på kvinners sykefravær er den samme i det

⁴¹ Vi oppnår lignende koeffisienter og signifikansnivå ved å bruke fylkesvis arbeidsledighet. Tall for fylkesledigheten (årgjennomsnitt) er hentet fra NAV og illustrerer antall individer som er helt arbeidsledige i hvert fylke.

⁴² Testen gjennomføres ved å lage en dummyvariabel for individer i det stabile utvalget og interagere denne med arbeidsledighetsvariabelen. Den tilhørende t-verdien indikerer om de to utvalgene responderer signifikant forskjellig på endringer i arbeidsledigheten.

stabile og fullstendige utvalget. Våre resultater gir dermed ikke en ubetinget støtte til sammensetningshypotesen. Dersom den prosykliske variasjonen i sykefraværet er forårsaket av marginale arbeidere, forventer vi at den negative effekten av arbeidsledigheten forsvinner eller reduseres i et utvalg av stabile arbeidere. For kvinner finner vi ingen signifikant ulik effekt i det stabile og det fullstendige utvalget. Dette resultatet indikerer at arbeidere med stabil arbeidsmarkedstilknytning bidrar på lik linje med resten av befolkningen når det gjelder den prosykliske variasjonen i sykefraværet til kvinner.

Vi finner indikasjoner på at arbeidere mellom 30 og 55 år, som deltar i arbeidsmarkedet alle årene de inngår i utvalget i perioden fra 1994-2006, endrer sitt arbeidstilbud som følge av variasjoner i arbeidsledigheten⁴³. Dersom vi ser bort fra sammensetningshypotesen, kan den prosykliske variasjonen i sykefraværet best forklares som stress- og/eller disiplineringseffekter. Betingelsene på arbeidsmarkedet kan dermed disiplinere arbeidere, som endrer sin tilbøyelighet til sykefravær. Dersom disiplineringshypotesen er sann, forventer vi også at sykenærværet øker i perioder med høy arbeidsledighet.

Effekten av arbeidsledigheten er i begge utvalgene relativt sterkere for kvinner enn for menn. Forskjellen mellom kjønnene går i samme retning som i resultatene til ABN, men i deres analyse er koeffisientene mindre signifikante for menn. Rikstrygdeverket (1998) har estimert sannsynligheten for sykefravær for stabile og nyansatte arbeidstakere, og finner at for kvinner er sannsynligheten høyere for stabile arbeidstakere, mens for menn er sannsynligheten høyere for nyansatte enn for stabile. Resultatene i vår analyse trekker i samme retning.

Offentlig sektor

Når vi kontrollerer for andre individspesifikke komponenter viser resultatene at sannsynligheten for sykefravær reduseres hvis individet er ansatt i offentlig sektor. Koeffisienten er kun signifikant på 5 % nivå for menn⁴⁴. Basert på den deskriptive statistikken

⁴³ Regresjonen er også utført på et utvalg bestående av arbeidere som er tilstede i arbeidsstyrken alle årene mellom 1993-2006. Resultatet gir tilsvarende (negative) koeffisienter og signifikansnivå for variabelen arbeidsledighet, men er bare gyldig for alderskohorten født mellom 1951-1963 (jf. diskusjonen i 5.1)

⁴⁴ I store utvalg ($n > 1000$) kan det argumenteres for å bruke lavere signifikansnivå enn i mindre utvalg. T-testen dividerer verdien til den estimerte koeffisienten på dens standardavvik, og som nevnt øker variasjonen i X_{it} proporsjonalt med n . Dette gir lavere standardavvik og høyere t-verdier, gitt samme koeffisientestimatet (Wooldridge (2009), s.136).

og tidligere studier er resultatet uventet. Vi kan imidlertid ikke utelukke at vår definisjon av hvilke næringer som faller under offentlig sektor påvirker validiteten til resultatene.

Ved å interagere dummyvariabelen *offentlig sektor* med variabelen for arbeidsledighet, får vi negative koeffisienter for menn og positive koeffisienter for kvinner. Resultatet indikerer at effekten av arbeidsledigheten er ulik i offentlig og privat sektor. For menn ansatt i offentlig sektor er effekten av arbeidsledigheten mer negativ enn menn i privat sektor. Motsatt responderer kvinner i offentlig sektor mindre på endring i arbeidsledigheten enn kvinner i privat sektor. Likevel er nettoeffekten fremdeles negativ.

Siden privat sektor er mer konjunkturutsatt, forventer vi at ansatte i privat sektor er mer påvirket av arbeidsledigheten enn ansatte i offentlig sektor. Ifølge våre resultater gjelder dette bare for kvinner, men som nevnt er det usikkerhet knyttet til vår definisjon av næringene som klassifiseres som offentlig sektor.

Alder

Effekten av alder er negativ både for kvinner og menn. Dette indikerer at sannsynligheten for et nytt sykefravær avtar når individene blir eldre. Vi har inkludert et annengradsledd for å fange opp eventuelle ikke-lineære effekter på sannsynligheten for sykefravær. Fra tabell 5 ser vi at $alder^2$ er positiv for alle modellene. En negativ koeffisient for *alder* og et positivt annengradsledd, indikerer at effekten av alder på sannsynligheten for sykefravær er negativ, men at den negative effekten avtar med alderen. Vi kan finne vendepunktet for når effekten av alder går fra å være negativ til positiv (se appendiks C). For kvinner blir effekten av alder på sannsynligheten for sykefravær positiv ved 40 år i det fullstendige utvalget og 39 år i det stabile.

Basert på diskusjonen i 5.2.2, er det muligheter for at koeffisienten for alder er positivt forventningsskeiv, som indikerer at den sanne effekten er mer negativ enn koeffisienten i tabell 5 viser. Årsaken til den negative effekten kan være at vi ekskluderer de fleste individene som ville bidratt til en positiv effekt av alder på sykefravær, nemlig individer over 55 år. Hvis det er slik at helsen forbedres de første 25 arbeidsårene (Lindeboom & Kerkhoffs (2002) referert til i ABN), men forverres etter en viss alder, vil dette samsvare med den negativt avtagende effekten i våre resultater. Hauge og Kann (2007) viser at sykefraværsprosenten er relativt stabil for menn mellom 30 og 52 år, mens den er synkende for kvinner i samme alder (målt i 2002 og 2005). Etter denne alderen øker sykefraværet

kraftig, særlig for kvinner. Utviklingen samsvarer med våre resultater. ABN finner en tilsvarende negativ, men avtagende effekt av alder på sannsynligheten for sykefravær.

Utdanning

Koeffisienten for utdanning er negativ for menn, men ikke signifikant for kvinner. Estimatet for menn er bare signifikant på 5 % nivå i det stabile utvalget. Som nevnt, kan lave signifikansnivå sees i sammenheng med at fast effekt-modeller krever variasjon i X_{it} innenfor hvert individ for å oppnå presise estimatorer. Bare et mindretall av individer mellom 30 og 55 år tar flere år med utdanning. ABN har droppet utdanningsvariabelen i regresjonene, sannsynligvis på grunn av for lite innenfor-individ variasjon.

Antall år utdanning kan redusere sannsynligheten for sykefravær som dersom lavt utdannede yrker er mer fysisk og psykisk krevende enn høyt utdannede yrker. Mer spennende arbeidsoppgaver og mer ansvar i stillinger som krever høy utdanning kan også øke motivasjonen for arbeid.

Inntekt

Koeffisientene for inntektsvariabelen er positive og signifikante for både kvinner og menn i begge utvalgene. Resultatene indikerer at sannsynligheten for sykefravær øker når inntekten stiger. Hvis variasjonen i inntekt hovedsakelig stammer fra variasjon i lønnen, er dette resultatet motsatt av hva effektivitetslønnsteorien predikerer: arbeideres sykefravær reduseres når kostnaden ved en potensiell avskjedigelse er høyere (ABN).

På samme måte som for variabelen *alder*, har vi har inkludert et annengradsledd for å kontrollere for potensielle ikke-lineære effekter av inntekt. Koeffisientene for *inntekt*² er negative i alle regresjonene. Positiv koeffisient og negativt annengradsledd indikerer at en inntektsøkning gir høyere sannsynlighet for sykefravær, men at effekten er avtagende. Vendepunktet for kvinner inntreffer ved en inntekt på 731 250 kr. i det fullstendige utvalget, og 788 333 kr. i det begrensede. Sykepengeordningen gir 100 % lønnskompensasjon for inntektstap inntil 6G. Det eksisterer en alternativkostnad ved sykefravær for individer med inntekter over dette beløpet (377 352 kr. i 2006) i form av tapt arbeidsinntekt. Man skulle da forvente at vendepunktet befinner seg i et intervall rundt denne inntekten.

En positiv, avtagende effekt av inntekt på sannsynligheten for sykefravær er påvist i lignende studier⁴⁵. I ABN er vendepunktet lik 295 000 for menn, og forfatterne påpeker at den positive sammenhengen kan forklares som en inntektseffekt hvis fritid klassifiseres som et normalt gode (jf. arbeidstilbudsmodellen). En alternativ- eller supplerende forklaring kan være at høy inntekt er et resultat av lange arbeidsdager, som kan føre til overbelastning og økt fravær.

Som nevnt i 5.2.2 har vi forsøkt å løse endogenitetsproblemet ved å bruke en fast effekt modell samtidig som vi lagger inntektsvariabelen med én periode. Dette medfører at observasjonene fra 1993 ikke er inkludert i regresjonene. Problemet med toveis kausalitet er imidlertid vanskelig å unngå når man bruker inntekt som forklaringsvariabel i en modell for sykefravær. Likevel, dersom det eksisterer et problem med toveis kausalitet, forventer vi at dette medfører en negativ effekt av inntekt på sykefravær. Forklaringen er at kausalitetsproblemet bygger på argumentet om at lavtlønnede yrker har høyere fravær, og høyt sykefravær gir lavere (fremtidig) inntekt (Markussen 2010). Vi kan ikke utelukke at toveis kausalitet påvirker resultatene i denne analysen, men kausalitetsproblemet synes ikke å være årsaken til at vi finner en positiv effekt av inntekt på sannsynligheten for sykefravær.

Deltid

Dummyvariabelen *deltid* gir negative og signifikante koeffisienter for både menn og kvinner i begge utvalgene. Effekten er sterkere for menn enn for kvinner. Deltidsansatte har i utgangspunktet mer fritid enn heltidsansatte. Nytteøkningen av en ekstra enhet fritid vil derfor være lavere for deltidsansatte, gitt at nytten av fritid er avtakende i konsum. Hvis det er slik at arbeidere bruker sykefravær for å regulere “mismatch” mellom avtalt arbeidstid og preferansene for fritid, vil denne effekten være lavere for deltidsansatte (jf. arbeidstilbudsmodellen). Resultatene støtter denne hypotesen.

Gift

Når det gjelder effekten av sivil status på sannsynligheten for sykefravær, viser resultatene at sannsynligheten øker for gifte menn og kvinner. I det stabile utvalget er koeffisienten for menn bare signifikant på 5 % nivå. Resultatene må sees i sammenheng med at innenfor-individ variasjonen i sivil status er liten.

⁴⁵ Se f.eks. Bratberg et al. (2002) og ABN.

Kvinner sitt sykefravær er sterkest påvirket av ekteskapeleg status. Resultatet kan knyttes til teorien om den doble byrde. Hvis det er slik at husholdningsarbeid er ulikt fordelt mellom kjønnene, så vil gifte kvinner oppleve et større samlet arbeidspress i hjemmet og på jobben, enn enslige kvinner. Overanstrengelse og stress kan dermed øke sannsynligheten for sykefravær. Det er rimelig å anta at denne effekten øker hvis husholdningen også består av små barn.

Alternativt kan den positive effekten av ekteskapeleg status for kvinner være et uttrykk for mindre (økonomisk) forsørgeransvar. Gifte kvinner deler forsørgeransvaret med sin mann, og kan ifølge disiplineringshypotesen "tillate" seg mer sykefravær uten at det får alvorlige økonomiske konsekvenser. I tillegg kan det være slik at gifte kvinner har en løsere arbeidsmarkedstilknytning, siden mange vil redusere sitt arbeidstilbud enten som følge av mindre forsørgeransvar, eller for å unngå "dobbeltarbeid" (jf. Bratberg et al. 2002).

Ved å interagere dummyvariabelen for gift med variabelen for antall barn under 11, kan vi undersøke dette nærmere. Interaksjonsleddet gir ingen signifikante resultater. Dette gir støtte til forklaringen om forsørgeransvar/arbeidsmarkedstilknytning, siden denne effekten er mindre påvirket av antall barn. Som nevnt kan høye standardfeil og upresise estimat være et resultat av multikollinearitet. I modeller med interaksjonsledd er dette et særlig stort problem (Wooldridge (2009), s.242).

Antall barn under 11 (barnu11)

Effekten av antall barn under 11 år på sannsynligheten for sykefravær er negativ for kvinner. For menn er effekten positiv og signifikant på 5 % nivå i det fullstendige utvalget, men ikke signifikant i det stabile utvalget. Vi forventer i utgangspunktet at antall barn i husholdningen vil øke sannsynligheten for sykefravær. Som tidligere nevnt, er det mulig at individer som kombinerer en karriere med flere barn er en selektert gruppe. Dersom de innehar egenskaper, f.eks. bedre helse, som gjør dem bedre utrustet til å håndtere den ekstra arbeidsbyrden, vil dette kunne redusere sykefraværet direkte. Dersom dobbeltarbeidshypotesen er sann, vil problemet være større for kvinner enn for menn. Menn trenger ikke være ekstra ressurssterke for å kombinere karriere med små barn hvis husholdningsarbeidet er skeivt fordelt. Dette kan være årsaken til at vi finner en negativ effekt av antall barn for kvinner, og en mindre signifikant positiv effekt for menn.

Resultatene støttes av funnene til Bratberg et al. (2002), som finner en negativ effekt for kvinner før de kontrollerer for seleksjon. Ved å ta høyde for at kvinner med mange barn i større grad forlater arbeidsmarkedet (for å unngå “dobbeltarbeid”), viser resultatene av modellen at flere barn øker sannsynligheten for sykefravær for kvinner. ABN finner ingen signifikant effekt av antall barn under 11 år på sannsynligheten for sykefravær.

6.3. Resultater fra den lineære fast effekt-modellen

Den andre regresjonsmodellen estimerer (logaritmen av⁴⁶) lengden på sykefravær som starter i år t , og er gitt ved:

$$\log(\text{antall dager}_{it}) = \gamma X'_{it} + \theta \text{arbeidsledighet}_{it} + \eta_i + v_{it} \quad (6.2)$$

Ser vi bort fra at X'_{it} også inkluderer variabelen *gradert*, er forklaringsvariablene de samme som for (6.1). Modellen estimeres betinget på positivt resultat i logitmodellen ($Y_{it} = 1$), og utnytter dermed færre person-år observasjoner. I logitmodellen kreves det at individet skifter status (fra $Y_{it} = 1$ til $Y_{it+1} = 0$, eller omvendt) minst en gang i løpet av analyseperioden for at individet skal bidra til den betingede sannsynlighetsmaksimeringen. Dette er ikke en nødvendig betingelse i den lineære modellen, og et individ som er tilstede i utvalget i f. eks. to år og har fravær begge årene, vil inkluderes i regresjonene (ABN). Antall individer er derfor høyere enn i logitmodellen.

Ved å betinge på at $\log(\text{antall dager}_{it}) > 0$, estimerer vi den rene effekten på sykefraværets varighet. Den totale effekten på sykefraværet kan fremstilles som

$E(Y_{it}) = pr(Y_{it} > 0) \cdot E(Y_{it} | Y_{it} > 0)$, der vi dekomponerer effekten i to deler. Det første leddet representerer sannsynligheten for sykefravær, som estimeres i logitmodellen, mens den andre delen representerer varigheten av det erstattede sykefraværet gitt at antall sykefraværsdager er større enn null. Siden det er effekten på varigheten vi ønsker å estimere i denne modellen, betinges regresjonen (6.2) på at antall dager er større enn null. Et annet argument for å eliminere nullene er at forskning indikerer at sykefraværet er ujevnt fordelt i befolkningen (Tidsskriftet Arbeidsmiljø 2010). Påstanden støttes av tabell 1 i kapittel 5, som illustrerer at ca. halvparten av utvalget ikke har sykefravær betalt av Folketrygden i løpet av

⁴⁶ Årsaken til at den avhengige variabelen uttrykkes på logaritmisk form, er at dette forenkler tolkningen av koeffisienter og gir den avhengige variabelen flere ønskede egenskaper (se Wooldridge (2009), s.191).

perioden 1993-2006. Som nevnt, utnytter ikke fast effekt-modellen mellom-individ variasjon. En slik modell forklarer i hvilken grad Y_{it} er forskjellig fra \bar{Y}_i , men ikke hvorfor \bar{Y}_i er forskjellig fra \bar{Y}_j . Dersom individ i har minst ett sykefravær, men individ j har null, vil denne variasjonen ikke tas høyde for i fast effekt-modellen. Derfor mister vi mindre informasjon ved å utelukke individer uten sykefravær enn dersom vi hadde gjort tilsvarende i en tilfeldig effekt- eller MKM-modell.

Tabell 6 viser resultatene fra den lineære fast effekt modellen for sykefraværsdager. Modellen er estimert med panel-robuste standardfeil for å kontrollere for arbitrær autokorrelasjon og generelle former for heteroskedastisitet (Verbeek (2008), s.372). Også i den lineære modellen er resultatene fra det stabile utvalget mindre signifikante. Tabell 6 viser at regresjonene gir høye konstantledd, som indikerer at modellene forklarer relativt lite av variasjonen i antall sykefraværsdager utover arbeidsgiverperioden. En naturlig årsak er at helsetilstand har en sterkere effekt på lengden av sykefravær (ABN).

Vi har også estimert modell (6.2) ved å bruke pooled-MKM. Estimering av paneldatamodeller med MKM gir mindre effisiente estimater, siden MKM ikke tar høyde for uobserverbar heterogenitet. Hvis uobserverbare individuelle egenskaper i tillegg er korrelert med forklaringsvariablene, medfører dette endogenitet. Resultatene fra MKM estimeringen er oppgitt i appendiks B, og resultatene varierer i forhold til fast effekt modellen. Koeffisientene for variabler som inntekt og deltid har motsatt fortegn i de to modellene. Likevel er effekten av arbeidsledigheten den samme. En årsak til at det er interessant å sammenligne med MKM-resultatene, er at vi får et bilde av hvordan den uobserverbare heterogeniteten påvirker resultatene. Når vi ikke kontrollerer for individspesifikke karakteristikk som helsetilstand, evner og motivasjon, har inntekt en negativ påvirkning på sykefraværet. Det kan likevel ikke utelukkes at de forskjellige resultatene skyldes at MKM også utnytter mellom-individ variasjonen i variablene.

Tabell 6: Resultater fra estimering av den lineære fast effekt-modellen. 1994-2006.

Fullstendig utvalg									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0327	0.0028	0.0000	-0.0225	0.0039	0.0000	-0.0388	0.0042	0.0000
Offentlig sektor	-0.1798	0.0197	0.0000	-0.2843	0.0415	0.0000	-0.1366	0.0239	0.0000
Off.sekt*ledighet	-0.0033	0.0042	0.4318	0.0044	0.0094	0.6419	-0.0051	0.0053	0.3426
Alder	-0.0483	0.0056	0.0000	0.0044	0.0094	0.6416	-0.0806	0.0070	0.0000
Alder ²	0.0012	0.0001	0.0000	0.0007	0.0001	0.0000	0.0014	0.0001	0.0000
Utdanning	-0.0319	0.0043	0.0000	-0.0277	0.0083	0.0008	-0.0325	0.0050	0.0000
Inntekt	0.0117	0.0006	0.0000	0.0061	0.0007	0.0000	0.0175	0.0018	0.0000
Inntekt ²	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	-0.0001	0.0000	0.0172
Deltid	-0.0629	0.0100	0.0000	-0.0843	0.0272	0.0019	-0.0483	0.0109	0.0000
Gift	0.0232	0.0107	0.0303	-0.0451	0.0170	0.0079	0.0642	0.0138	0.0000
Barn under 11	-0.0263	0.0073	0.0003	0.0012	0.0120	0.9190	-0.0382	0.0092	0.0000
Gift*barnu11	0.0051	0.0083	0.5404	0.0170	0.0139	0.2211	-0.0035	0.0104	0.7344
Gradert	0.9071	0.0049	0.0000	0.9037	0.0087	0.0000	0.9071	0.0059	0.0000
Konstant	3.3113	0.1273	0.0000	1.8377	0.2175	0.0000	4.1516	0.1580	0.0000
R ² -within	0.107			0.100			0.112		
F-test	3730			1274			2489		
Individer	304272			127759			176513		
Observasjoner	691291			266734			424557		
Stabilt utvalg (pensjonsgivende inntekt alle år de inngår i utvalget)									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0121	0.0031	0.0001	-0.0006	0.0042	0.8963	-0.0209	0.0046	0.0000
Offentlig sektor	-0.1241	0.0220	0.0000	-0.2239	0.0457	0.0000	-0.0918	0.0270	0.0007
Off.sekt*ledighet	-0.0040	0.0046	0.3898	0.0045	0.0102	0.6589	-0.0030	0.0059	0.6047
Alder	-0.0412	0.0060	0.0000	0.0105	0.0099	0.2908	-0.0751	0.0075	0.0000
Alder ²	0.0011	0.0001	0.0000	0.0006	0.0001	0.0000	0.0013	0.0001	0.0000
Utdanning	-0.0104	0.0049	0.0333	-0.0051	0.0091	0.5793	-0.0124	0.0058	0.0320
Inntekt	0.0102	0.0008	0.0000	0.0047	0.0009	0.0000	0.0154	0.0017	0.0000
Inntekt ²	0.0000	0.0000	0.0000	0.0000	0.0000	0.0149	-0.0001	0.0000	0.0197
Deltid	-0.0626	0.0114	0.0000	-0.0897	0.0307	0.0035	-0.0475	0.0125	0.0001
Gift	0.0222	0.0117	0.0577	-0.0424	0.0183	0.0207	0.0626	0.0151	0.0000
Barn under 11	-0.0337	0.0080	0.0000	0.0047	0.0129	0.7131	-0.0552	0.0103	0.0000
Gift*barnu11	0.0041	0.0091	0.6501	0.0103	0.0148	0.4869	-0.0011	0.0116	0.9229
Gradert	0.9584	0.0053	0.0000	0.9505	0.0093	0.0000	0.9605	0.0065	0.0000
Konstant	2.7603	0.1378	0.0000	1.3401	0.2302	0.0000	3.6490	0.1726	0.0000
R ² -within	0.117			0.107			0.125		
F-test	3374			1160			2242		
Individer	240188			105661			134527		
Observasjoner	557320			223080			334240		

6.3.1. Tolkning av koeffisienter

Resultatene i den lineære fast effekt-modellen er relativt like resultatene fra fast effekt logitmodellen. Derfor vil vi ikke drøfte alle koeffisientene som er oppgitt i tabell 6 like inngående som i 6.2.1, men vi fokuserer på koeffisientene for arbeidsledighet.

Effekten av kommunal arbeidsledighet på antall dager med sykefravær betalt av Folketrygden er negativ for både kvinner og menn. I det stabile utvalget er koeffisienten også negativ, men ikke signifikant for menn. Totalt vil en økning i arbeidsledigheten på ett prosentpoeng føre til en 3,3 % reduksjon i antall fraværsdager for de som er sykemeldt i hele utvalget, og 1,2 % reduksjon i utvalget med stabile arbeidere.

Ser vi bort fra at koeffisienten for menn ikke er signifikant, ligner resultatene fra den lineære modellen på resultatene fra sannsynlighetsmodellen. Når vi holder andre faktorer konstante, vil også arbeidere som befinner seg utenfor den marginale arbeiderdefinisjonen endre atferd som følge av at betingelsene på arbeidsmarkedet endrer seg. Som i logitmodellen, er kvinners sykefravær generelt sterkere påvirket av arbeidsledigheten enn menn. Resultatene våre indikerer at lengden på sykefraværet til menn med stabil arbeidsmarkedstilknytning ikke påvirkes av arbeidsledigheten.

Effekten av arbeidsledigheten er mindre i utvalget med stabile arbeidere enn i det fullstendige, og koeffisientene er signifikant forskjellige mellom utvalgene⁴⁷. Dermed kan vi ikke utelukke sammensetningseffekter i variasjonen til sykefraværenes varighet. Den prosykliske variasjonen i lengden på sykefravær utover arbeidsgiverperioden er likevel ikke en ren sammensetningseffekt, siden effekten av arbeidsledigheten på sykefraværet negativ også i det stabile utvalget. Sykefraværet til stabile arbeidere kan per definisjon ikke være forårsaket av sammensetning.

Sammenlignes resultatene med tidligere forskning (jf. 3.4), kan funnene knyttes til resultatene i Arai og Thoursie (2005). De finner at både seleksjon- og disiplineringsmekanismen kan være tilstede i sykefraværet over tid, men resultatene i deres analyse indikerer at disiplineringsmekanismen dominerer.

⁴⁷ Med tilhørende t-verdier: 20.80 for alle, 14.88 for menn og 14.90 for kvinner

Resultatene fra fast effekt modellen skiller seg fra tilsvarende modell i ABN. De finner ingen signifikant effekt av arbeidsledigheten på antall sykefraværsdager utover arbeidsgiverperioden for det fullstendige utvalget, men effekten i det stabile utvalget er sterkere og mer signifikant. Dermed konkluderer de med at betingelsene på arbeidsmarkedet også påvirker lengden på fraværene til stabile arbeidstakere som er sykemeldt. En mulig årsak til at resultatene ikke samsvarer med våre, kunne være at modellspesifikasjonen til ABN ikke inkluderer et interaksjonsledd mellom offentlig sektor og arbeidsledighet. Generelt blir koeffisientene mindre presise som følge av denne spesifikasjonen, siden *arbeidsledighet* og *arbeidsledighet*offentlig sektor* er høyt korrelert (Wooldridge (2009), s.242). Ved å estimere modellen uten interaksjonsledd, blir koeffisienten for menn fremdeles ikke signifikant i det stabile utvalget. En annen mulig årsak til at resultatene ikke samsvarer med ABN, er at vår definisjon av utvalgene medfører at de stabile arbeiderne har en lengre arbeidsmarkedstilknytning enn i deres analyse. I tillegg illustrerer figur 4 i kapittel 5 at antall sykefraværsdager har økt klart mest for stabile arbeidstakere gjennom analyseperioden, og i 2006 er nivået bare marginalt forskjellig fra det fullstendige utvalget.

Når det gjelder forskjeller mellom resultatene for de andre variablene i fast effekt logit og den lineære fast effekt-modellen, finner vi noen forhold som bør diskuteres. Vi har inkludert en ny variabel i den lineære fast effekt-modellen. Dummyvariabelen *gradert* indikerer om sykepengegraden for det aktuelle fraværet er mindre enn 100 %. Tabell 6 viser at koeffisientene for *gradert* er sterkt positive både for kvinner og menn. Dersom sykepengetilfellet er *gradert*, gir dette en kraftig økning i antall sykepengedager. Sannsynligheten for at den positive sammenhengen representerer et kausalt forhold, er derimot liten. *Graderte* og *ikke-graderte* sykemeldinger blir ikke tildelt arbeidere på en tilfeldig måte. Legen avgjør om fraværet skal være *gradert* eller ikke, basert på en forventning om rehabiliteringsprosessen. Hovedsakelig er det sykefravær som forventes å være langvarige som blir *gradert*. Vi har ikke kontrollert for dette seleksjonsproblemet i vår analyse.

Markussen et al. (2010) estimerer effekten av *graderte* sykemeldinger på fraværets varighet. De bruker legens tilbøyelighet til å *gradere andre pasienters* sykefravær som instrument på legens forventning om rehabiliteringsprosessen. Resultatene indikerer at antall erstattede sykefraværsdager halveres dersom man substituerer en *ikke-gradert* mot en *gradert* sykemelding. Samtidig øker sannsynligheten for fremtidig sysselsetting med 12-14 %.

Andre forskjeller mellom logitmodellen og den lineære, er at den negative effekten av utdanning er signifikant for kvinner, men ikke for menn. I det stabile utvalget er den negative koeffisienten for kvinner kun signifikant på 5 % nivå. Bratberg et al. (2002) finner også at utdanning reduserer lengden på kvinners sykefravær.

Den estimerte sammenhengen mellom inntekt og sannsynligheten for sykefravær ble funnet å være positiv i logitmodellen. Ifølge den lineære modellen har inntekt også en positiv effekt på antall sykefraværsdager. En inntektsøkning på 10 000 kr. vil øke sykefraværet med 1,2 % i det fullstendige utvalget, og 1 % i det stabile. Både ABN og Bratberg et al. (2002) finner at inntekten øker lengden på sykefravær. Dersom vi sammenligner resultatene i fast effektmodellen med resultatene i MKM-modellen (appendiks B), er det spesielt noen forhold som er av interesse. Fortegnet til koeffisienten for inntekt går fra negativt til positivt når vi kontrollerer for individspesifikke egenskaper som helse, motivasjon og evner. Samtidig bruker MKM både variasjonen mellom- og innenfor individ. MKM tar dermed høyde for at individer med høy inntekt kan ha lavere sykefravær enn individer med lav inntekt. Årsaken til denne forskjellen kan likevel være at individer med høy inntekt har bedre helse, motivasjon, evner osv. Basert på diskusjonen i 5.2.2 forventer vi at inntekt påvirker sykefraværet negativt, og at to-veis kausalitet vil forsterke den negative effekten. Ved å holde helse, evner og motivasjon konstant (faktorer som påvirker inntekt i stor grad), indikerer resultatene at inntekt påvirker sykefraværet positivt. Dersom fritid er et normalt gode, vil man ønske å konsumere mer av godet når inntekten øker. Ved 100 % kompensasjon eksisterer det ingen direkte substitusjonseffekt opp til 6 G (jf. diskusjonen i 6.2.1). Dette kan bidra til å forklare den uventede positive effekten av inntekt på både sannsynligheten og lengden på sykefravær. Dersom faktorer som påvirker inntekt holdes konstant, kan den rene effekten av inntekt på sykefravær være positiv.

Kapittel 7: Oppsummering

Formålet med denne oppgaven har vært å undersøke om variasjonen i sykefraværet over konjunktursyklusene skyldes at andelen marginale arbeidere med nedsatt helse/funksjonsevne øker ved høy sysselsetting, og reduseres ved lav sysselsetting. Datamaterialet som brukes i analysen er hentet fra forløpsdatabasen FD-Trygd og koplet sammen med variabler for inntekt og utdanning. Resultatene fra analysen gjelder for sykefravær utover 16 dager (utvidet fra 14 dager i 1998). Samme problemstilling analyseres i Askildsen et al. (2005), som i teksten refereres til som ABN, og det er naturlig å sammenligne resultatene med denne analysen.

Sykefraværet i Norge har variert prosyklisk med arbeidsledigheten, men sammenhengen har vært tydeligst før år 2000. ABN baserer seg på data fra perioden 1990-1995, og finner at sammensetningshypotesen ikke kan forklare variasjonen i sykefraværet. Det sentrale målet i denne oppgaven har vært å undersøke om deres resultater også gjelder dersom vi utfører en tilsvarende analyse på nyere data som strekker seg over en lengre tidsperiode (1994-2006).

Vi har brukt to modeller, en som estimerer sannsynligheten for sykefravær og en som estimerer varigheten til sykefravær gitt at individet har rapportert minst ett fravær i år t . Modellene estimeres først på et tilfeldig utvalg, deretter på et begrenset utvalg bestående av arbeidere som befinner seg utenfor den marginale arbeiderdefinisjonen. Ifølge sammensetningshypotesen, vil en eventuell negativ effekt av arbeidsledigheten på sykefraværet reduseres kraftig eller forsvinne når vi estimerer modellene på det stabile utvalget.

Resultatene indikerer at også arbeidere med stabil arbeidsmarkedstilknytning har flere og lengre sykefravær når den kommunale arbeidsledigheten reduseres. Endringen i arbeidstilbudet til denne gruppen kan per definisjon ikke være forårsaket av sammensetningen. Den kan derimot være et resultat av disiplinering, eller være forårsaket av arbeidsrelatert stress som varierer over konjunktorene. Disiplinerings effekter innebærer at arbeidere reduserer arbeidstilbudet når betingelsene på arbeidsmarkedet er gode, da kostnadene forbundet med å miste jobben er lavere. Den norske sykepengeordningen, med 100 % kompensasjon, øker mulighetene for denne typen atferdsrisiko. Sammenlignet med hele utvalget, finner vi at effekten av arbeidsledigheten er svakere for utvalget med stabile arbeidere, og kan dermed ikke utelukke at den prosykliske variasjonen i sykefraværet til hele utvalget også påvirkes av sammensetningseffekter. Modellene er også estimert for kvinner og

menn separat. Generelt er kvinners sykefravær sterkere påvirket av arbeidsledigheten. Resultatene fra sannsynlighetsmodellen indikerer til og med at den prosykliske variasjonen i kvinners tilbøyelighet for sykefravær er lik for begge utvalgene. Dette er per definisjon ikke en sammensetningseffekt, og resultatet kan best forklares med alternative hypoteser som f.eks. stress- eller disiplineringshypotesen.

I denne oppgaven har vi fokusert på disiplinering som det mest aktuelle alternativet til sammensetning når det gjelder å forklare konjunkturvariasjonen i sykefraværet. Våre resultater indikerer at effekten av arbeidsledighet i utvalget av *stabile* arbeidere er svakere enn den tilsvarende effekten i ABN. Siden atferden til de stabile arbeiderne ikke kan forklares med endring i sammensetningen, må disiplineringseffekten være mindre. Ser vi på effekten av arbeidsledighet på *hele* utvalget, inkludert de stabile, er denne på nivå med resultatene i ABN. Når effekten av disiplinering har avtatt og den totale effekten (sammensetning pluss disiplinering) er uendret, skulle dette indikere at sammensetningseffekten er blitt sterkere i befolkningen som sådan.

Sykefraværet i Norge er et viktig politisk og økonomisk tema. Reformen i sykepengeordningen som er rettet mot å redusere sykefraværet må ta utgangspunkt i hvilke mekanismer som påvirker variasjonen i sykefraværet. Dersom hovedårsaken er sammensetningseffekter, vil økning i sykefraværet være uunngåelig når etterspørselen etter arbeid øker, og tiltak bør rettes mot de marginale arbeiderne (ABN). Resultatene til ABN kan i stor grad tolkes som bevis mot sammensetningshypotesen og de argumenterer for at disiplineringseffekter bedre kan forklare variasjoner i sykefraværet. Våre resultater indikerer snarere at variasjonen ikke er en *ren* sammensetningseffekt, dermed bør alternative eller supplementære forklaringer legges til grunn for tiltakene. Styrking av legenes forhandlingskraft, med økt fokus på medisinske vurderinger og graderte sykemeldinger, kan være en nyttig måte å redusere fraværet dersom man mistenker atferdsrisiko som følge av sjenerøs sykeordning. Disiplinerings-effekter kan også reduseres ved å innføre egenandeler for arbeidere ved sykefravær, men ekspertutvalget som står bak den siste reformen i sykepengeordningen konstaterer at *"justering av kompensasjonen ikke fremstår som aktuell politikk i Norge i dag"* (Mykletun et al. 2010).

Litteraturliste

- Allen, S. G. (1981). An Empirical Model of Work Attendance. *The Review of Economics and Statistics*, 63 (No. 1): 77-87.
- Arai, M. & Thoursie, P. S. (2005). Incentives and selection in cyclical absenteeism. *Labour Economics*, 12 (2): 269-280.
- Arbeidsdepartementet. (2010). *Protokoll mellom partene i arbeidslivet og myndighetene om felles innsats for å forebygge og redusere sykefraværet og styrke inkluderingen*. Tilgjengelig fra: <http://www.regjeringen.no/upload/SMK/Vedlegg/2010/IA%20protokoll%2024022010.pdf> (lest 21.06.11).
- Arbeidstilsynet. (2007). *Langtidssykefravær og uførepensjonering*. Tilgjengelig fra: <http://www.arbeidstilsynet.no/artikkel.html?tid=97039> (lest 21.06.11).
- Askildsen, J. E., Bratberg, E. & Nilsen, Ø. A. (2005). Unemployment, labor force composition and sickness absence: A panel data study. *Health Econ.*, 14: 1087–1101.
- Barmby, T., Orme, C. & Treble, J. (1995). Worker absence histories: a panel data study. *Labour Economics*, 2 (1995): 53-65.
- Biørn, E., Gaure, S., Markussen, S. & Røed, K. (2010). The Rise in Absenteeism: Disentangling the Impact of Cohort, Age and Time. *IZA DP No. 5091*.
- Bjørnstad, R. (2006). Er det økte sykefraværet tegn på et mer inkluderende eller ekskluderende arbeidsliv? *Økonomiske Analyser* 6(2006).
- Bonato, L. & Lusinyan, L. (2004). Work absence in Europe. *IMF Working paper No. 04/193*.
- Bratberg, E. & Risa, A. E. (2000). Incentivvirkninger i helserelaterte stønadsordninger. *Rapport skrevet på oppdrag av Sosial- og Helsedepartementet: Institutt for økonomi, Universitetet i Bergen*.
- Bratberg, E., Dahl, S. A. & Risa, A. E. (2002). The double burden – Do combinations of career and family obligations increase sickness absence among women? *European Sociological Review*, 18: 233-249.
- Bye, T., Berge, C. & Næsheim, H. (2010). Det komplekse sykefraværet. *Samfunnsøkonomen*, Nr. 3.
- Cameron, A. C. & Trivedi, P. K. (2009). *Microeconometrics using Stata*. A Stata Press Publication: Statacorp LP.
- Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *The Review of Economic Studies*, 47 (1): 225-238.
- Chaudhury, M. & Ng, I. (1992). Absenteeism Predictors: Least Squares, Rank Regression, and Model Selection Results. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 25 (3): 615-635.
- Dale-Olsen, H. (2009). Absenteeism, efficiency wages and economic incentives: Institute for Social Research.
- Dale-Olsen, H., Nilsen, K. M. & P Schøne Memo, I. f. S. R., Oslo. (2010). Imitation, Contagion, and Exertion – Do Colleagues' Sickness Absences Influence your Absence Behaviour? Oslo: Institute for Social Research.

- Dyrstad, J. M. & Ose, S. O. (2002). Non-linear Unemployment Effects in Sickness Absence: Discipline or Composition Effects? *Working Paper Series* (Nr 25/2002).
- Ellis, G. (1981). The Backward-Bending Supply Curve of Labor in Africa: Models, Evidence, and Interpretation: And Why It Makes a Difference. *The Journal of Developing Areas*, 15: 251-274.
- Eriksen, W. (2003). The prevalence of musculoskeletal pain in Norwegian nurses' aides. *International Archives of Occupational and Environmental Health*, 76 (8): 625-630.
- Folkehelseinstituttet. (2007). Sosial ulikhet i helse. En faktarapport Nasjonalt folkehelseinstitutt.
- Gjesdal, S. & Bratberg, E. (2003). *Diagnosis and duration of sickness absence as predictors for disability pension: results from a three-year, multi-register based and prospective study*. Reprint series (Universitetet i Bergen. Institutt for økonomi). Bergen: Instituttet. S. [246]-254 s.
- Gjesdal, S. (2005). Sykefraværets utvikling i Norge 1975-2002. *Tidsskrift for Den norske legeforening* (125): 742-5.
- Greene, W. H. (2003). *Econometric Analysis*. Fifth Edition utg.: Prentice Hall.
- Guldbrandsen, P., Førde, R. & Aasland, O. A. (2002). Hvordan har legen det som portvakt? *Tidsskrift for Den norske legeforening* (19): 1874 – 1879.
- Hansen, H.-T. (1999). Kunnskapsstatus for den nordiske trygdeforskningen på 1990-tallet. *SNF-rapport*.
- Hauge, L. & Kann, I. C. (2007). Hva betyr alders- og kjønns sammensetningen for sykefraværet? *Arbeid og velferd nr. 1-2007*.
- Hausman, J. A. (1978). Specification Tests in Econometrics. *Econometrica*, 46 (6): 1251-71.
- Henrekson, M. & Persson, M. (2004). The Effects on Sick Leave of Changes in the Sickness Insurance System. *Journal of Labor Economics*, Vol. 22: 87-113.
- Holmboe-Ottesen, G., Wandel, M. & Mosdøl, A. (2004). Sosiale ulikheter og kosthold. *Tidsskrift for den Norske Lægeforening* 124: 1526-8.
- Johanson, P. & Palme, M. (2005). Moral hazard and sickness insurance. *Journal of Public Economics*, Vol 89: 1879-1890.
- Kostøl, A. R. & Telle, K. (2011). Sykefraværet i Norge de siste tiårene – Det handler om kvinnene. *Samfunnsøkonomen*, 1/2011.
- Larsen, B., Førde, O. & Tellnes, G. (1994). Legens rolle i sykemelding. *Tidsskrift for Den norske legeforening*, 114 (12): 1442 - 1444.
- Lindbeck, A., Palme, M. & Persson, M. (2009). Social Interaction and Sickness Absence. *CESIFO Working Paper*.
- Lindeboom, M. & Kerkhofs, M. (2002). Health and Work of the Elderly Subjective Health Measures: Reporting Errors and the Endogenous Relationship between Health and Work. *IZA Discussion Paper No. 457*.
- Maddala, G. S. (1987). Limited Dependent Variable Models Using Panel Data. *The Journal of Human Resources*, 22 (3): 307-338.
- Markussen, S. (2009). How physicians can reduce sick leave - evidence from a natural experiment: Ragnar Frisch Centre for Economic Research.

- Markussen, S. (2010). The Individual Cost of Sick Leave: Ragnar Frisch Centre for Economic Research.
- Markussen, S., Mykletun, A. & Røed, K. (2010). The Case for Presenteism. *IZA DP No. 5343*.
- Mastekaasa, A. (2000). Parenthood, gender and sickness absence. *Social Science & Medicine*, 50 (12): 1827-42.
- Moland, L. E. (2007). Deltidsarbeid og sykefravær i Oslo kommune. *Fafo-notat 2007:01*.
- Mykletun, A., Eriksen, H. R., Røed, K., Schmidt, G., Fosse, A., Damberg, G., Christiansen, E. & Guldvog, B. (2010). Tiltak for reduksjon i sykefravær: Aktiviserings- og nærværsreform: Ekspertgrupperapprt til Arbeidsdepartementet.
- Nossen, J. P. (2008). Sykefraværet og konjunktorene – hva vet vi om sammenhengen? *Arbeid og Velferd*, 4 (2008): 25-36.
- Nossen, J. P. (2009). Sykefravær og konjunkturer – en oversikt. *Norsk Epidemiologi*, 19 (2): 115-125.
- Nossen, J. P. & Thune, O. (2009). Utviklingen i sykefraværet de siste 20 år. *Arbeid og velferd*, Nr. 3.
- Nossen, J. P. (2010). Økt legemeldt sykefravær etter finanskrisen: Flere langvarige sykefravær for menn. *Arbeid og Velferd*, 1/2010.
- Nossen, J. P. & Thune, O. (2010). Sykefravær blant selvstendig næringsdrivende – er nivå og utvikling forskjellig fra lønnstakerne? *Arbeid og Velferd*, 2/2010: 60-67.
- Ose, S. O. (2010). Kunnskap om sykefravær: nye norske bidrag. *SINTEF Helse*: Arbeidsdepartementet.
- Rege, M., Telle, K. & Votruba, M. (2009). The Effect of Plant Downsizing on Disability Pension Utilization. *Journal of the European Economic Association*, 7 (4): 754-785.
- Regjeringen. (2009). *Utdrag fra statsministerens tale på LO Stats kartellkonferanse*. Tilgjengelig fra: http://www.regjeringen.no/mobil/nn/dep/smk/aktuelt/taler_og_artikler/statsministeren/statsminister_jens_stoltenberg/2009/utdrag-fra-statsministerens-tale-pa-lo-s.html?id=586591 (lest 21.06.11).
- Regjeringen. (2010). *Intensjonsavtale om et mer inkluderende arbeidsliv 1. mars 2010 – 31. desember 2013*. Tilgjengelig fra: http://www.regjeringen.no/upload/AD/publikasjoner/web-publikasjoner/2010/IA-avtale_24022010.pdf (lest 21.06.11).
- Rieck, K. M. E. & Vaage, K. (2011). *Social Interactions at the Workplace: Exploring Sickness Absence Behaviour*: Institutt for Økonomi, Universitetet i Bergen. Upublisert manuskript.
- Rikstrygdeverket. (1998). Basisrapport 1997 - sammendrag.
- Rikstrygdeverket. (2003). Trygdestatistisk Årbok.
- Rothschild, M. & Stiglitz, J. E. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, 90: 630-649.

- Røed, K. & Fevang, E. (2005). Organizational Change, Absenteeism, and Welfare Dependency. *Journal of Human Resources*, 42 (1): 156-193.
- Shapiro, C. & Stiglitz, J. E. (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review*, 74 (3): 433-444.
- SSB. (2011). *Levendefødte og dødfødte. 1951-2010* Tilgjengelig fra: <http://www.ssb.no/fodte/tab-2011-04-07-01.html> (lest 21.06.11).
- Theorell, T., Oxenstierna, G., Westerlund, H., Ferrie, J., Hagberg, J. & Alfredsson, L. (2003). Downsizing of staff is associated with lowered medically certified sick leave in female employees. *Journal of occupational and environmental medicine*, 60:E9.
- Thoursie, P. S. (2007). Happy Birthday! You're insured! Gender Differences in work ethics. *Economic Letters*, 94 (1): 141-145.
- Tidsskriftet Arbeidsmiljø. (2010). I: Arbeidsmiljøsentret (red.). Tilgjengelig fra: http://www.arbeidsmiljo.no/xp/pub/hoved/tidsskrift/temaer/sykefravar_og_uforhet/478990 (lest 21.06.11).
- Verbeek, M. (2008). *A Guide to Modern Econometrics*. 3rd Edition utg.: John Wiley & Sons Ltd.
- Waddell, G. (2004). *The Back Pain Revolution*. Elsevier Health Sciences.
- Waddell, G. & Burton, A. K. (2006). Is work good for your health and well-being? I: Office), T. T. S. (red.). *Report for Department for Work and Pensions*. London.
- Wagstaff, A. & Doorslaer, E. (2000). Equity in health care finance and delivery. I: Culyer, A. J. & Newhouse, J. P. (red.) *Handbook of Health Economics*, s. 1803-1862: Elsevier.
- Westerlund, H., Ferrie, J., Hagberg, J., Jeding, K., Oxenstierna, G. & Theorell, T. (2004). Workplace expansion, long-term sickness absence, and hospital admission. *The Lancet* 363 (9416): 1193-1197.
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48 (4): 817-38.
- Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach*, b. Fourth Edition.
- Østhus, S. (2007). For better or worse? Workplace changes and the health and well-being of Norwegian workers. *Work Employment and Society*, 21 (4): 731-750

Appendiks A: Estimering av 6.1 med lineær fast effekt-estimator

Fullstendig utvalg									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0035	0.0002	0.0000	-0.0022	0.0003	0.0000	-0.0068	0.0004	0.0000
Offentlig sektor	-0.0067	0.0017	0.0001	-0.0178	0.0024	0.0000	-0.0082	0.0025	0.0010
Off.sekt*ledighet	-0.0016	0.0004	0.0000	-0.0007	0.0005	0.1497	0.0008	0.0005	0.1602
Alder	-0.0111	0.0005	0.0000	-0.0049	0.0006	0.0000	-0.0235	0.0007	0.0000
Alder ²	0.0002	0.0000	0.0000	0.0001	0.0000	0.0000	0.0003	0.0000	0.0000
Utdanning	0.0006	0.0004	0.0807	-0.0007	0.0005	0.1477	0.0003	0.0005	0.5438
Inntekt	0.0016	0.0000	0.0000	0.0007	0.0000	0.0000	0.0061	0.0002	0.0000
Inntekt ²	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Deltid	-0.0574	0.0010	0.0000	-0.0550	0.0021	0.0000	-0.0424	0.0012	0.0000
Gift	0.0037	0.0010	0.0002	0.0016	0.0012	0.1869	0.0082	0.0017	0.0000
Barn under 11	-0.0158	0.0007	0.0000	0.0004	0.0008	0.6307	-0.0297	0.0011	0.0000
Gift*barnu11	0.0066	0.0008	0.0000	0.0031	0.0009	0.0010	0.0071	0.0012	0.0000
Konstant	0.2767	0.0107	0.0000	0.1134	0.0135	0.0000	0.5833	0.0168	0.0000
R ² -within	0.008			0.006			0.014		
F-test	1966			755			1984		
Individer	556011			278862			277149		
Observasjoner	4146868			2113812			2033056		
Stabilt utvalg (pensjonsgivende inntekt alle år de inngår i utvalget)									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0029	0.0002	0.0000	-0.0017	0.0003	0.0000	-0.0059	0.0004	0.0000
Offentlig sektor	-0.0048	0.0019	0.0120	-0.0134	0.0025	0.0000	-0.0078	0.0029	0.0059
Off.sekt*ledighet	-0.0020	0.0004	0.0000	-0.0008	0.0005	0.0938	0.0004	0.0006	0.5123
Alder	-0.0100	0.0005	0.0000	-0.0041	0.0006	0.0000	-0.0227	0.0008	0.0000
Alder ²	0.0002	0.0000	0.0000	0.0001	0.0000	0.0000	0.0003	0.0000	0.0000
Utdanning	0.0014	0.0004	0.0010	0.0002	0.0006	0.6674	0.0011	0.0006	0.0885
Inntekt	0.0011	0.0000	0.0000	0.0004	0.0000	0.0000	0.0049	0.0002	0.0000
Inntekt ²	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Deltid	-0.0515	0.0012	0.0000	-0.0452	0.0023	0.0000	-0.0408	0.0014	0.0000
Gift	0.0032	0.0011	0.0026	0.0008	0.0012	0.5371	0.0083	0.0019	0.0000
Barn under 11	-0.0144	0.0007	0.0000	0.0001	0.0009	0.8626	-0.0314	0.0013	0.0000
Gift*barnu11	0.0066	0.0008	0.0000	0.0032	0.0010	0.0007	0.0075	0.0014	0.0000
Konstant	0.2324	0.0115	0.0000	0.0785	0.0141	0.0000	0.5457	0.0188	0.0000
R ² -within	0.007			0.005			0.011		
F-test	1511			618			1236		
Individer	439815			233558			206257		
Observasjoner	3526804			1888730			1638074		

Appendiks B: Estimering av 6.2 med MKM

Fullstendig utvalg									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0306	0.0018	0.0000	-0.0300	0.0024	0.0000	-0.0312	0.0027	0.0000
Offentlig sektor	-0.1287	0.0102	0.0000	-0.1549	0.0218	0.0000	-0.1113	0.0130	0.0000
Off.sekt*ledighet	-0.0015	0.0027	0.5815	-0.0025	0.0058	0.6724	-0.0009	0.0035	0.8029
Alder	-0.0505	0.0032	0.0000	-0.0139	0.0053	0.0091	-0.0698	0.0041	0.0000
Alder ²	0.0007	0.0000	0.0000	0.0003	0.0001	0.0000	0.0008	0.0000	0.0000
Utdanning	-0.0153	0.0006	0.0000	-0.0225	0.0010	0.0000	-0.0122	0.0008	0.0000
Inntekt	-0.0094	0.0004	0.0000	-0.0095	0.0004	0.0000	-0.0114	0.0008	0.0000
Inntekt ²	0.0000	0.0000	0.0002	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
Deltid	0.1271	0.0062	0.0000	0.2042	0.0167	0.0000	0.1117	0.0069	0.0000
Gift	-0.0763	0.0045	0.0000	-0.1199	0.0073	0.0000	-0.0523	0.0057	0.0000
Barn under 11	-0.0084	0.0039	0.0312	-0.0347	0.0063	0.0000	0.0007	0.0051	0.8901
Gift*barnu11	0.0039	0.0045	0.3847	0.0279	0.0073	0.0001	-0.0069	0.0058	0.2372
Gradert	0.8566	0.0035	0.0000	0.9064	0.0058	0.0000	0.8342	0.0044	0.0000
Konstant	4.6871	0.0690	0.0000	3.8998	0.1126	0.0000	5.1370	0.0879	0.0000
R ²	0.079			0.084			0.076		
F-test	5076			2214			2890		
Observasjoner	691291			266734			424557		
Stabilt utvalg (pensjonsgivende inntekt alle år de inngår i utvalget)									
	Alle			Menn			Kvinner		
	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z	Koeff.	St. feil	P>z
Arbeidsledighet	-0.0295	0.0020	0.0000	-0.0281	0.0026	0.0000	-0.0314	0.0030	0.0000
Offentlig sektor	-0.0784	0.0112	0.0000	-0.1424	0.0235	0.0000	-0.0604	0.0144	0.0000
Off.sekt*ledighet	-0.0055	0.0030	0.0676	-0.0015	0.0063	0.8075	-0.0046	0.0039	0.2381
Alder	-0.0977	0.0035	0.0000	-0.0523	0.0057	0.0000	-0.1229	0.0045	0.0000
Alder ²	0.0012	0.0000	0.0000	0.0007	0.0001	0.0000	0.0015	0.0001	0.0000
Utdanning	-0.0142	0.0007	0.0000	-0.0219	0.0012	0.0000	-0.0115	0.0009	0.0000
Inntekt	-0.0057	0.0003	0.0000	-0.0060	0.0003	0.0000	-0.0060	0.0005	0.0000
Inntekt ²	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Deltid	0.1302	0.0072	0.0000	0.2084	0.0193	0.0000	0.1159	0.0079	0.0000
Gift	-0.0600	0.0049	0.0000	-0.1013	0.0079	0.0000	-0.0377	0.0063	0.0000
Barn under 11	0.0063	0.0043	0.1439	-0.0175	0.0067	0.0092	0.0156	0.0058	0.0068
Gift*barnu11	-0.0083	0.0050	0.0957	0.0104	0.0078	0.1833	-0.0136	0.0066	0.0384
Gradert	0.9323	0.0038	0.0000	0.9758	0.0062	0.0000	0.9079	0.0049	0.0000
Konstant	5.4348	0.0750	0.0000	4.4730	0.1202	0.0000	5.9897	0.0964	0.0000
R ²	0.096			0.096			0.095		
F-test	4907			2103			2804		
Observasjoner	557320			223080			334240		

Appendiks C: Utregning av vendepunkter

Modell 6.1 er estimert som en logistisk fordelt sannsynlighetsfunksjon Λ . Forenklet kan vi illustrere sannsynligheten for sykefravær for et individ i på følgende måte:

$$P_i = \Lambda(\beta X_i' + \delta_1 \text{alder}_i + \delta_2 \text{alder}_i^2) \quad (\text{C.1})$$

, hvor X_i' er en vektor som inneholder alle andre forklaringsvariabler enn alder. Ved hjelp av kjerneregelen kan vi finne marginaleffekten til alder slik:

$$\frac{\partial P_i}{\partial \text{alder}_i} = \Lambda'(\beta X_i' + \delta_1 \text{alder}_i + \delta_2 \text{alder}_i^2)(\delta_1 + 2\delta_2 \text{alder}_i) \quad (\text{C.2})$$

Den førsteordensderiverte er et uttrykk for stigningstallet til (C.1). Vi setter uttrykket ovenfor lik null for å finne vendepunktet:

$$\text{alder}_i = -\frac{\delta_1}{2\delta_2} \quad (\text{C.3})$$

Fra den andreordensderiverte av (C.1) kan det enkelt vises at fortegnet på koeffisienten δ_2 bestemmer om vendepunktet i (C.3) er ett topp- eller bunnpunkt. Ved å bruke koeffisientene for kvinne fra logitregresjonen i det fullstendige utvalget får vi følgende bunnpunkt:

$$\text{alder}_i = -\frac{-0,1673}{2(0,0021)} = 39,8 \quad (\text{C.4})$$

Tilsvarende utregninger gir ingen meningsfulle resultater for menn.