

# En empirisk studie av utenforskap blant unge i Norge

av

Aksel Kvanne Vestfossen



---

Masteroppgaven er levert for å fullføre graden

**Master i Samfunnsøkonomi**

Universitetet i Bergen, Institutt for økonomi  
1. september 2015

---

## Forord

Min interesse for samfunnsøkonomi begynte våren 2009 da jeg tok ECON130. Dette var i tiden rett etter finanskrisen hadde inntruffet og samfunnsøkonomi var dagsaktuelt. Jeg hoppet etterhvert over fra å studere matematikk til å studere samfunnsøkonomi. Interessen innen samfunnsøkonomi har vært forskjellig, men en rød tråd gjennom hele studiet har vært statistikk og økonometri.

Jeg vil takke veileder Kjell Vaage for tålmodighet og for å ha hjulpet å finne et spennende og utfordrende tema for masteroppgaven min, i tillegg for å ha tilrettelagt et datasett til meg. Det må også rettes en spesiell takk til veileder Arnstein Mykletun som har kommet med mange gode innspill og veiledning underveis. Jeg vil også takke Bridianne O'Dea for god hjelp i startfasen.

Takker også Magnus, Håvard, Vidar, Karl Ove, Arthur og Magne for korrekturlesing og forslag til forbedringer. Jeg vil også takke resten av mine medstudenter på instituttet for gode samtaler på lunsjrommet og utenfor instituttet.

Jeg vil dessuten takke familien og kjæresten min for støtte gjennom studiene og skriving av masteroppgaven.



---

Aksel Kvamme Vestfossen, Bergen, 1. september 2015

## Sammendrag

---

### En empirisk studie av utenforskap i Norge

av

**Aksel Kvamme Vestfossen, Master i samfunnsøkonomi**

Universitetet i Bergen, 2015

Veileder: Kjell Vaage

Biveileder: Arnstein Mykletun

---

Unge som faller utenfor samfunnet, i form av at de ikke er i arbeid, i utdanning eller er på opplæring (NEET; engelsk forkortelse for Neither in Employment, Education or Training) er et samfunnsproblem. Tidligere forskning viser at arbeidsledighet i ung alder fører til økt risiko for arbeidsledighet og lav lønn senere i livet (Arulampalam, Gregg og Gregory, 2001). NEET-gruppen er sammensatt og består av langtidsledige, passive stønadsmottakere og økonomisk inaktive personer. Formålet med oppgaven er å finne omfanget av NEET i Norge ved bruk av registerdata, og å se på utviklingen over tid og beskrive omfanget, og se på hvilke faktorer som disponerer for denne tilstanden. Det er også av interesse å undersøke om NEET er en overgangstilstand eller absorberende tilstand.

Jeg har benyttet meg av paneldata fra FD-trygd (ForløpsData trygd), i fra dette har jeg ved hjelp av inntekts-, utdannings- og tryggedata vurdert om en person faller innenfor oppgavens definisjon av NEET for et gitt år. Jeg har sett på omfanget av NEET for de over 20 år som er født i perioden 1977 til 1984 i perioden 2000 til 2008.

Omtrent 9 % av 23- og 24-åringer er NEET mellom 2000 og 2008. Det er generelt flere kvinner som er i denne tilstanden enn menn, likevel indikerer resultatene mine at NEET er et økende problem for menn. NEET-andelen varierer stort med alder, hvor det var flest menn og kvinner som var NEET som 20-åring. I snitt er de som er NEET det i litt over to år, kvinner ser ut til være i denne tilstanden noe lengre enn menn. Risikofaktorer som ser ut til å øke NEET-risikoen er frafall i videregående opplæring og historikk som trygde- og sosialbruker. Resultatene indikerer at NEET-risikoen er større for de som fullfører allmennfaglig studieretning enn de som fullfører yrkesfag, i forhold til frafall.

Statistikkprogrammet Stata/IC 13.1 er brukt for den statistiske analysen. Oppgaven er skrevet i L<sup>A</sup>T<sub>E</sub>X.

# Innhold

<b>Forord</b>	<b>I</b>
<b>Sammendrag</b>	<b>II</b>
<b>1 Innledning</b>	<b>1</b>
<b>2 Definisjon av NEET</b>	<b>6</b>
2.1 Arbeidsmarkedstilknytning . . . . .	7
2.1.1 Arbeidsmarkedstilknytningsmodell . . . . .	10
2.1.2 Kritikk av arbeidsmarkedstilknytningsmodell . . . . .	11
2.2 Frafall og forsinket skolegang - definisjonsavklaring . . . . .	13
2.3 NEET og sykepenger . . . . .	14
2.4 NEET og attføring og rehabilitering . . . . .	14
2.5 Svakheter ved NEET-definisjon i oppgaven . . . . .	15
2.6 En liten oppsummering . . . . .	16
<b>3 Utvalgt relevant forskning</b>	<b>17</b>
3.1 Nordiske studier . . . . .	17
3.2 Andre internasjonale studier . . . . .	20
3.3 Sammendrag og kritikk . . . . .	21
<b>4 Beskrivelse av datamaterialet</b>	<b>24</b>
4.1 Datamaterialet . . . . .	24
4.2 utfordringer . . . . .	25
4.2.1 Utdanningsdata . . . . .	25
4.3 Utvalget . . . . .	26
4.4 Arbeidsmarkedstilknytning og gjennomstrømming i vgo . . . . .	27
<b>5 Deskriptive funn</b>	<b>33</b>
5.1 Utviklingen i NEET over tid . . . . .	33
5.2 NEET - et overgangsfenomen eller en absorberende tilstand? . . . . .	38
5.2.1 NEET: Hvor kommer de fra? . . . . .	41
5.2.2 Overgang til jobb, utdanning eller tiltak . . . . .	42

<b>6</b>	<b>Økonometriske metoder</b>	<b>44</b>
6.1	Paneldata og regresjon . . . . .	44
6.1.1	Lineære paneldatamodeller . . . . .	44
6.1.2	Fast Effekt-modeller . . . . .	46
6.1.3	Tilfeldig effekt-modeller . . . . .	46
6.2	Korrelert tilfeldig effekt-modeller . . . . .	47
6.3	Valg av modell . . . . .	48
<b>7</b>	<b>Regresjonsanalyse</b>	<b>50</b>
7.1	Sannsynligheten for å være NEET . . . . .	52
7.1.1	Utdanning . . . . .	55
7.1.2	Velferdsordninger . . . . .	55
7.1.3	Demografiske forskjeller . . . . .	57
7.1.4	Testing av antagelser og FE-testen . . . . .	57
7.2	Hvor robust er resultatene? . . . . .	61
7.2.1	Valg av modell . . . . .	61
7.2.2	Endogenitet . . . . .	61
<b>8</b>	<b>Oppsummering og konklusjoner</b>	<b>62</b>
8.1	Hovedfunn . . . . .	62
8.2	Konklusjon: Har jeg funnet opp kruttet på nytt? . . . . .	63
8.3	Veien videre . . . . .	64
<b>A</b>	<b>Vedlegg</b>	<b>70</b>
A.1	Institusjonelle forhold . . . . .	70
A.1.1	Velferdsstaten . . . . .	70
A.2	Økonometriske metoder . . . . .	73
A.2.1	Tilfeldig Effekt-modeller . . . . .	73
A.2.2	Binære modeller . . . . .	77
A.3	Utfyllende regresjonsresultater . . . . .	81
A.4	Utviklingen i grunnbeløpet og antall helt ledige . . . . .	89

## Figurer

1	Estimat av NEET fra OECD . . . . .	3
---	------------------------------------	---

2	Internasjonal sammenligning . . . . .	4
3	Arbeidsmarkedstilknytning i andeler syv år etter avsluttet ungdomsskole i perioden 2000 til 2007. . . . .	27
4	Arbeidsmarkedstilknytning i 2006, syv år etter avsluttet ungdomsskole delt inn etter kjønn og fullføring av vgo. . . . .	29
5	Gjennomstrømming i vgo fordelt på allmennfaglig og yrkesfaglig utdanning. . . . .	31
6	NEET (%) etter år . . . . .	33
7	NEET etter alder . . . . .	35
8	NEET dekomponert . . . . .	37
9	Antall år som NEET . . . . .	38
10	Forløpene til personer som er NEET i 2004 i perioden 2000-2003. . . . .	41
11	Overganger til forskjellige tilstander for personer som var NEET i 2004 i perioden 2005 til 2008. . . . .	42
12	QQ-plott over fordelingen til de uobserverte individeffektene ( $\alpha_i$ ) for modell (3). . . . .	58
13	QQ-plott over fordelingen til residualene ( $u_{it}$ ) for modell (4). . . . .	59
14	QQ-plott av $\epsilon_{it} = \alpha_i + u_{it}$ for modell (3) (t.v.) og modell (4). . . . .	84
15	Antall helt ledige i alderen 20 til 24 år . . . . .	90

## Tabeller

1	Tabell med litteraturoversikt . . . . .	23
2	Antall individer per kohort . . . . .	26
3	Overgangsmatrise . . . . .	39
4	Variabelliste . . . . .	
5	LPM regresjonsresultater . . . . .	54
6	FE-test . . . . .	60
7	Oppsummering for forklaringsvariabler (gjennomsnittsverdi). . . . .	81
8	Lineær sannsynlighetsmodell for sannsynligheten for å være NEET syv år etter avsluttet ungdomsskole. . . . .	82
9	Korrelert tilfeldig effekt probit-modell for sannsynligheten for å være i tilstanden NEET. . . . .	84
10	Korrelert tilfeldig effekt-modell for sannsynligheten for å være i tilstanden NEET, lagget 3 år. . . . .	87

11	Grunnbeløpet pr. år . . . . .	89
----	-------------------------------	----

# 1 Innledning

I denne oppgaven vil jeg ta for meg inaktivitet blant ungdom, unge som ikke går på skole, ikke er i jobb og ikke deltar på arbeidsrettede tiltak. Jeg vil kartlegge omfanget av denne inaktiviteten blant ungdom, og jeg vil, med hjelp av økonometriske metoder, finne frem til faktorer som kan disponere for, og lede fram til eller føre ungdom inn i en slik tilstand av inaktivitet. Men først vil jeg ta opp noen negative virkninger som ung inaktivitet kan ha for den enkelte, for familien og for samfunnet, og på den måten forklare hvorfor dette emnet er verdt det omfattende arbeidet som en masteroppgave er.

Problemet med utenforskap er selvsagt størst for det enkelte individet som er rammet, men også for familien og samfunnet representerer dette et problem. Slik ledighet, blant personer som egentlig er frisk nok til å jobbe, er et produksjonstap og med det et velferdstap for samfunnet. Det er også kjent at lav utdanning og arbeidsledighet i ung alder gir økt risiko for arbeidsledighet og lavere inntekt senere i livet (Arulampalam et al., 2001).

Det er én årsak som merker seg ut, og som er mye omtalt i media, og det er frafall i videregående skole. Om lag 95 til 97 prosent av de som fullfører ungdomskolen begynner på videregående utdanning samme året, men etter fem år har bare rundt to av tre fullført denne utdanningen. Bare om lag tre prosent av de som ikke hadde fullført etter 5 år, fullførte det på 6. året (Falch, Johannesen og Strøm, 2009). Det var en betydelig større prosentandel av de som ikke hadde fullført videregående som var registrert på en eller annen offentlig støtteordning senere i livet, enn de som hadde fullført (Falch et al., 2009).

Hva skjer så med de som faller fra videregående opplæring (vgo)? Noen kommer ut i arbeid, andre ikke. Noen kommer ikke inn i arbeidslivet på grunn av kroniske sykdommer eller psykiske problemer. Andre velger å ta seg et friår, jobbe frivillig o.l.. Hva kan så de som de har falt fra vgo og ikke jobber, gjøre? Det eksisterer forskjellige typer tiltak for å få ungdom ut i arbeid, enten i form av arbeidsrettede tiltak eller utdanning som tiltak. Disse tiltakene tilbys i hovedsak av NAV og Oppfølgingstjenesten. Oppfølgingstjenesten<sup>1</sup> har ansvar for å følge opp ungdom som har rett til vgo, men som ikke er i utdanning eller arbeid. Motivasjonen bak denne tjenesten er å hjelpe ungdom til å få studiekompetanse,

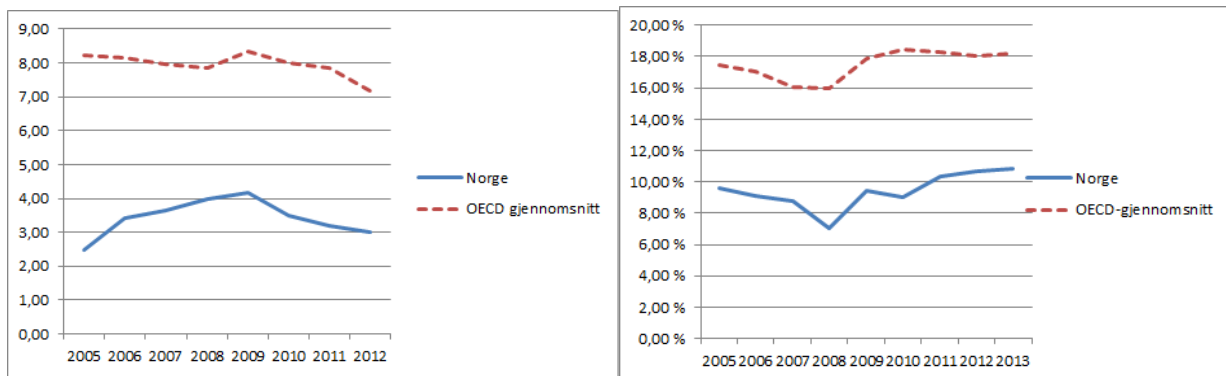
---

<sup>1</sup>Alle unge har rett til å ta tre års heltids videregående utdanning. Hele retten må tas ut i løpet av en sammenhengende periode på fem år, innen utgangen av det året de fyller 21 år (Kunnskapsdepartementet, 2010).



yrkeskompetanse/fagbrev eller annen kompetanse (Kunnskapsdepartementet, 2015). En gruppe som er særlig interessant å se på er de som verken er i utdanning, i jobb eller på tiltak/opplæring. Disse kalles gjerne for NEET (Neither in Education, Employment or Training). Fordelen med å følge denne gruppen mot det å følge ungdomsarbeidsledigheten er at NEET-gruppen inkluderer de ungdommene som er registrert arbeidsledige, i tillegg til å inkludere de som er inaktive. Disse personene er inaktive i den forstand at de ikke søker jobb, og dermed ikke tilfredsstillter kravene for å være definert som arbeidssøkende. NEET som et mål kan dessuten gi oss en økt forståelse av de som står utenfor arbeidsmarkedet, enn bare ved å se ungdomsarbeidsledighet eller frafalls- eller tiltaksstatistikken. NEET som et sammenfattende mål kan kanskje best beskrives ved historien om de fire blinde mennene som står og kjenner på hver sin del av en elefant. Ingen av de beskriver elefanten, dette gjør heller ikke hvert av enkeltmålene. NEET-begrepet kan forhåpentligvis tilby et samlende mål på utenforskap.

Det finnes flere organisasjoner som følger utviklingen i antallet personer som er NEET. I Norge er det SSB som beregner andelen personer som er utenfor. SSB bruker da Arbeidskraftsundersøkelsen (AKU) (Bø og Vigran, 2014). I Europa og resten av verden brukes lignende spørreundersøkelser (National Labour Force Survey, NLFS) for å beregne andelen personer som verken er i jobb, utdanning eller på tiltak. Eurostat beregner andelen unge som er uten jobb eller er inaktive for landene i EU. OECD fører også en oversikt over NEET for mange land, hvor de henter data fra Eurostat og andre lands NLFS, som tilsvarer AKU (OECD, 2014). OECD følger utviklingen av det de kaller for «ungdomsinaktivitet», altså NEET-kategorien for de mellom 15 og 24 år, hvor de deler inn ytterlige i ungdom i alderen 15 til 19 og 20 til 24. Jeg kommer i neste kapittel til å presentere en eksakt definisjon av NEET basert på registerdata. Figuren under er med for å gi et førsteinntrykk av omfanget av NEET i Norge i forhold til OECD-gjennomsnittet. Andelen er beregnet ved hjelp av AKU for Norge og AKU for andre land.



(a) NEET (%) i alder 15-19 år.

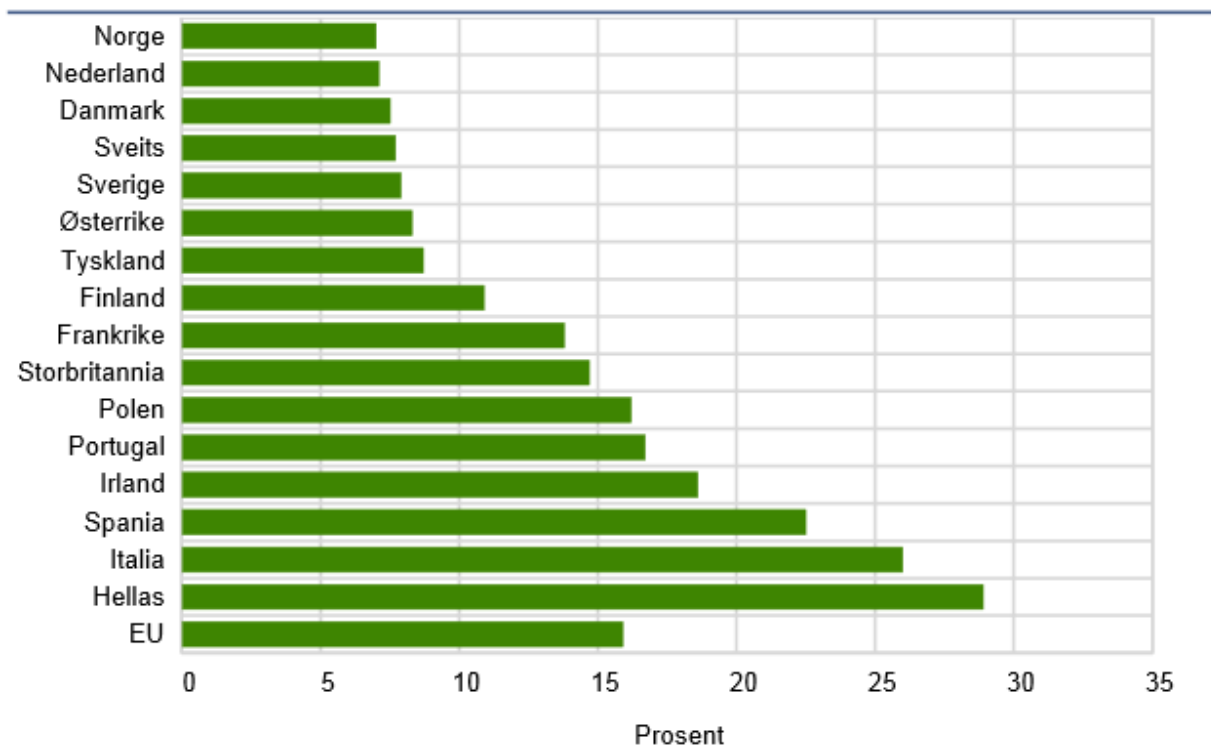
(b) NEET (%) i alder 20-24 år.

Figur 1: Utviklingen i andel NEET i Norge sammenlignet med OECD-gjennomsnittet.<sup>1</sup>

Vi ser at andelen tenåringer som er NEET i Norge er betydelig lavere enn gjennomsnittet i OECD, og ligger stabilt mellom 3 og 4 prosent i perioden 2005 til 2012. Vi ser videre at prosentandelen NEET i Norge er omtrent halvparten så stor som OECD-gjennomsnittet både for de mellom 15 og 19 år og de mellom 20 og 24 år. Andelen NEET mellom 15 og 19 er økende frem til 2009 for så å avta etter dette. En mulig grunn til denne nedgangen etter 2009, kan være at flere unge tar utdanning i usikre tider, noe tiden etter finanskrisen i 2008 kan sies å være. Andelen som er NEET i alderen 20 til 24 ser ut til å synke frem mot finanskrisen i 2008, for så å øke etter dette. Denne økningen kan kanskje sees i sammenheng med at det i dårligere tider er vanskeligere å få jobb, men det er vanskelig å si nøyaktig hva som gav denne økningen. Fra figur 1 finner vi at andelen NEET mellom 15 og 24 er på rundt 6 prosent. Bø og Vigran (2014) finner at ca. 5 prosent av de mellom 15 og 24 var NEET i 2009. Det er altså forskjeller i hva de forskjellige forfatterne anslår NEET-andelen til å være, noe som kanskje kommer av forskjellige definisjoner.

Andelen unge som er NEET er et stort problem i flere land. I figur 2 kan vi se en sammenligning av andelen NEET i forskjellige europeiske land.

<sup>1</sup>Data hentet fra <http://data.oecd.org/youthinac/youth-not-in-education-or-employment-neet.htm>



Kilde: Eurostat

Figur 2: Internasjonal sammenligning av ungdom mellom 15 og 29 år som er NEET i 2013 (Bø og Vigran, 2014).

Vi ser fra figuren over at omtrent hver fjerde italiener og greker mellom 15 og 29 verken var i jobb eller utdanning i 2013. Norge og Nederland er de landene med lavest andel NEET, med rundt 7 og 8 prosent. Vi ser at Finland er det av de nordiske landene med høyest andel som er NEET, med en andel på over 10 prosent i 2013.

I denne oppgaven vil jeg starte med å konstruere en definisjon for NEET og beskrive omfanget av gruppen og drøfte hva som kjennetegner medlemmene i den. Relevant forskning tar jeg opp i kapittel 3. Jeg vil deretter analysere økonometrisk hvilke faktorer som ser ut til å disponere for tilstanden NEET. Hvor mye betyr det å ha droppet ut av utdanning? Er det forskjell på menn og kvinner? Er det forskjell på å ha gått yrkesfaglig og allmennfaglig retning? Hvor mye påvirker tidligere bruk av sosial- og trygdetjenester risikoen for NEET? Jeg vil også undersøke om NEET er en forbigående tilstand eller er det en tilstand man ikke kommer ut av.

Jeg benytter meg av registerdata fra SSB som er koblet med utdanningsdata og inntektsdata. Jeg observerer etnisk norske personer født i 1977 til 1984 i perioden 1998 til

2008. Jeg har altså anledning til å observere personer i alderen 14 til 31. Jeg begrenser meg til å se på hvilken tilstand personene er i fra de er 20 år. Grunner til dette diskuterer jeg i kapittel 2.2.

## 2 Definisjon av NEET

På bakgrunn av diskusjonen i innledningskapittelet vil punktene under avgjøre hvem som er kvalifisert til å være NEET i et gitt år:

1. Personer med avkortet utdanning eller problemer i utdanningsløpet:
  - (a) Personer som har falt fra videregående opplæring (vgo).
  - (b) Personer som er forsinket i vgo.
  - (c) Personer som ikke har startet på høyere utdanning.
2. Personer som ikke er i jobb.
3. Personer som ikke er på tiltak.

En person er da NEET hvis en av delkategoriene under 1) er tilfredsstilt og 2 og 3 er oppfylt. Resten av dette kapittelet kommer til å svare på når en person ikke er i utdanning, jobb eller på tiltak.

Det finnes ingen standard definisjon på hvordan man skal kartlegge NEET i registerdata. Det finnes en del artikler som undersøker NEET-gruppen ved hjelp av registerdata, men ingen av dem bruker samme definisjon (Bø og Vigran, 2014; Bäckman, Jakobsen, Lorentzen, Österbacka og Dahl, 2011; Falch og Nyhus, 2011; Grødem, Skevik og Strand, 2014).

NEET-begrepet har den styrken at det kan løse problemet med at de øvrige målene på utenforskap i offentlig sektor er utdaterte og har sårbarheter. Vi har skolevesen, derav frafall. Vi har helsevesen, derav diagnoser og helserelevante trygdeordninger. Vi har arbeidsmarked, og derav arbeidsledighetstrygd. Hver av disse sektorene ser på bare på en del av problemet med utenforskap, men ingen sammenfatter det og ser på det samlede problemet med utenforskap. Dette gjør NEET-begrepet, siden det sammenfatter alle disse målene under ett. Barth og von Simson (2012) påpeker at andelen arbeidsledige kan være høyere enn andelen NEET. Arbeidsledighetsandelen kan f.eks. reduseres gjennom at arbeidssøkere trekker seg ut av arbeidsstyrken. Siden NEET-andelen beregnes ut i fra det totale antallet i de respektive aldersgruppene, er NEET sannsynligvis et bedre mål på utenforskap (Grødem et al., 2014). Dette fordi endringen i NEET-andelen hele tiden kommer fra at folk enten går inn eller ut av denne tilstanden, og ikke forsvinner ut av «statistikken».

Det finnes forskjellige måter å se på NEET-begrepet, alt etter hva slags fokus man har. Skal man lage en internasjonal sammenligning, er det kanskje best å bruke definisjonen til OECD og SSB (se Bø og Vigran (2014)). OECD kaller NEET for en indikator på «skole-til-jobb-overgangen». En høy NEET-andel tyder, i følge OECD, på at mange personer som nettopp er ferdig med, eller falt fra utdanning har vanskeligheter med å finne arbeid. Det kan være flere grunner til dette. Mange jobber krever tidligere jobberfaring eller høyere utdanning, slik at personer som faller fra vgo eller ikke tar høyere utdanning sliter med å komme inn arbeidslivet. Hvis andelen NEET skal være en indikator på friksjon i overgangen mellom skole og arbeid, er det ikke alle grupper som er like relevant å inkludere i NEET-begrepet. Er det f.eks. relevant å inkludere de som av selvvalgte grunner ikke er arbeid, utdanning eller på tiltak; de som tar et friår og reiser eller jobber frivillig? Er det de hjemmeværende mødrene eller de som pga. lyter og kronisk sykdom ikke kan jobbe eller ta utdanning? Jeg vil si at de som vi bør fange opp med NEET-begrepet ikke er disse, men heller de som har en ufrivillig periode hvor de ikke jobber etter endt utdanning. Fra dette perspektivet kan man da stille spørsmål ved hvem som skal inkluderes i NEET-begrepet. Skal f.eks. de langtidssykemeldte eller de uføretrygdete inkluderes? De langtidssykemeldte må ha vært i arbeid før sykeperioden, så disse har i realiteten en arbeidstilknytning. Det er mer uklart når det kommer til de uføretrygdete. Det er grunner til å inkludere de i NEET-begrepet, og grunner til å ikke inkludere de. En grunn for å inkludere de unge uføretrygdete er at de i likhet med de arbeidsledige og de økonomisk inaktive er utenfor. Det kan derfor virke merkelig at OECD inkluderer de uføretrygdete, og kaller NEET for et mål på friksjonen i overgangen mellom skole og arbeid. Uføretrygdete kommer sannsynligvis aldri ut i arbeid, og OECD overestimerer derfor andelen som opplever et ufrivillig opphold mellom endt skolegang og starten på arbeidslivet. Jeg har valgt å ekskludere de som er uføretrygdete fra NEET nettopp på grunn av dette.

Et problem er hvordan vi skal operasjonalisere disse kategoriene i denne oppgaven i forhold til de registerdata vi har til disposisjon. F.eks. når er en person i jobb eller ikke når vi benytter oss av registerdata. Neste delkapittel beskriver dette.

## 2.1 Arbeidsmarkedstilknytning

Hos SSB, OECD og Eurostat brukes det et standard sett med spørsmål for å avdekke om en person er sysselsatt. SSB har arbeidskraftsundersøkelsen (AKU) og man har tilsvarende

arbeidskraftsundersøkelser internasjonalt. Disse spørsmålene fanger opp småjobber helt ned til arbeid som varer i én time i uken (Bø og Vigran, 2014). For registerdata har man ikke anledning til å spørre personen om hva slags arbeid personen har. Man må derfor f.eks. bruke inntektsdata, antall dager i jobb registrert i løpet av ett år eller arbeidsledighetsdata. Jeg har benyttet meg av inntektsdata, nærmere bestemt yrkesinntekt. Yrkesinntekt er inntekter fra lønn pluss netto næringsinntekt (f.eks. selvstendig næringsdrivende) minus arbeidsledighetstrygd. Disse dataene gir forhåpentligvis et godt mål på den inntekt som en person får fra arbeid. En annen grunn til at jeg benytter meg av inntektsdata, er at det ser ut til å være en del avvik mellom inntekt og antall dager i jobb. Folk kan være registrert som å ha vært i jobb hele året, men ha for lav inntekt til at dette kan stemme. Hvis jeg bare bruker arbeidsledighetsdata kan jeg ikke finne de som er økonomisk inaktive. De økonomisk inaktive har lav inntekt i tillegg til å ikke være registrert som arbeidsledige. Fordelen med å benytte seg av registerdata er at man til forskjell ved spørreundersøkelser ikke støter på problemet med ikke-respons (Bäckman et al., 2011).

Det er vanskelig å avgjøre hvor stor inntekt en person må ha i et gitt år for at vi skal kunne si at denne personen er i arbeid. Det finnes forskjellige tilnærminger. Én tilnærming er å si at en person som ikke har noen inntekt i et år, da ikke har vært i arbeid. Denne tilnærmingen fanger bare de med lavest arbeidsmarkedstilknytning. Personer som står oppført med noe inntekt i løpet av et år, blir ikke fanget opp, selv om de i realiteten kan ha jobbet mindre enn en måned av året.

En annen tilnærming er å si hvilken inntekt som er bærekraftig, altså at inntekten greier å forsørge en person gjennom et år. Denne tilnærmingen bruker Grødem et al. (2014). Her sier de at hvis man har en inntekt på under 2 ganger grunnbeløpet ( $G$ ), så er inntekten for lav til å forsørge seg selv. Grunnbeløpet var pr. 1. mai 2014 på 88 370 kr (NAV, 2014b).

En annen måte å vurdere arbeidsmarkedstilknytningen til en person på, er beskrevet av Bäckman et al. (2011). Denne tilnærmingen ligner på Grødem et al. (2014) sin, men er mer dynamisk, siden den vurderer hvilken tilstand en person er i ved tid  $t$  ut ifra tid  $t$ ,  $t - 1$  og  $t - 2$ . Modellen til Bäckman et al. (2011) innbefatter flere gjensidig utelukkende kategorier, f.eks. kjernearbeidsstyrken (Core labour force), som består av de personene som har en høy nok inntekt til å forsørge seg selv i ett år. Denne grensen er satt til 3,5 ganger Nordic Base Amount (NBA). Forfatterne definerer NBA som 25% av medianbruttoinntekt. Neste kategori er den ustabile arbeidsstyrken (Unstable Labour force), som består av personer

som er i etableringsfasen, de som jobber deltid og de som har svakere arbeidsmarkedstilknøyning. Studenter er også en kategori. Så kommer de kategoriene hvor man har så svak arbeidsmarkedtilknøyning at Bäckman et al. (2011) vurderer det til at man kan være utenfor jobb. Dette er de langtidssykemeldte, langtidsarbeidsledige og de uføre. I tillegg er det de som har alternative inntektskilder og de som er økonomisk inaktive. Forfatterne bruker denne modellen til å avgjøre hvilken tilstand et individ er i syv år etter at det hadde begynt på vgo, altså hvilken tilstand man er i når man er ca. 23 år gammel. For en person som f.eks. er født i 1977 vil hans tilstand i år 2000 vurderes ut i fra 1998, 1999 og 2000.

Det kan stilles spørsmål om de som er uføretrygdet eller de langtidssykemeldte i virkeligheten er kategorier som faller innenfor NEET-konseptet. Dette avhenger av hvordan man ser på NEET. Er NEET de unge som står utenfor arbeidslivet i tillegg til at de ikke tar utdanning og er på tiltak, og dermed står i risiko for dårligere arbeidsutsikter? Er vi interessert i å inkludere de personene som er allerede i ung alder har kroniske sykdommer eller lyter som ikke gjør dem i stand til å jobbe, og dermed kan ha krav på uføretrygd? Flesteparten av de som er uføretrygdet har kanskje aldri fullført vgo og har lav eller ingen arbeidsevne. Det virker unaturlig å inkludere uføretrygdede i NEET-gruppen, i hvert fall hvis NEET skal beskrive problemer i overgangen mellom skole og jobb. Å inkludere de langtidssykemeldte som NEET-gruppe diskuteres senere, i kapittel 2.3.

I denne oppgaven kommer jeg til å bruke arbeidsmarkedstilknøyningsmodellen til Bäckman et al. (2011) til to forskjellige analyser. Den første for å se hvilken tilstand en person er i 7 år etter påbegynt vgo (ca. 23 år gammel). Men siden jeg ikke har informasjon om når en person begynte på vgo, bruker jeg syv år etter avsluttet ungdomsskole. Dette blir relativt like mål, siden nesten alle begynte på vgo samme år som de avsluttet ungdomsskolen (Falch et al., 2009). Den andre analysen bruker jeg for å avdekke arbeidsmarkedstilknøyningen for et individ for hvert år fra personen er 20 år gammel, ikke bare syv år etter avsluttet ungdomsskole. Slik at arbeidsstatus som 20-åring blir vurdert ut ifra tilstand som 18-, 19- og 20-åring. Med fødselskohortene 1977-1984 og observasjonsperioden 1998-2008 rekker den yngste kohorten å bli 24 år, mens den eldste kohorten rekker å bli 31.



## 2.1.1 Arbeidsmarkedstilknytningsmodell

### Kjernerarbeidstyrken

Arbeidsinntekt på minst 3,5 G i minst to av tre påfølgende år. Arbeidsinntekt mellom 0,5 og 3,5 G i maks ett av de tre årene. Ingen uføretrygd.

### Ustabil arbeidsstyrke Omfatter flere underkategorier:

#### *I re- eller etableringsfasen*

Arbeidsinntekt på minst 3,5 G i det siste av tre påfølgende år. Arbeidsinntekt lavere enn 0,5 G i det første og andre året. Ingen uføretrygd eller tidsbegrenset uførepensjon.

#### *Deltidsjobb*

Arbeidsinntekt mellom 0,5 G og 3,5 G i minst to av tre år, mindre enn 180 dager i et gitt år som arbeidsledig. Under 0,5 G eller over 3,5 G i maks ett av de tre årene.

#### *I det marginale arbeidsmarkedet*

Arbeidsinntekt mellom 0,5 G og 3,5 G i det første av tre år og under 0,5 G i de to påfølgende årene.

### Studenter

Arbeidsinntekt over 0 i de to første årene. Arbeidsinntekt mindre enn 3,5 G og i høyere/lavere utdanning det tredje året.

### Langtidsarbeidsledig

Arbeidsmarkedsinntekt lavere enn 0,5 G og minst 180 dager registrert som arbeidsledig hvert år i to av tre år. NEET-kategori.

### Langtidssykemeldte

Minst 90 dagers sykefravær i minst to av tre år.<sup>3</sup>

### Uføretrygdet

Registrert med uføretrygd eller tidsbegrenset uføretrygd med overføringer fra folketrygden på minst 1 G og arbeidsinntekt på mindre en 0,5 G i minst to av tre år.

### **Økonomisk inaktive**

Personer som ikke tilhører noen av kategoriene over og som har en arbeidsinntekt lavere enn 0,5 G i to av tre år. Eller inntekt lavere enn 0,5 G i tre påfølgende år. NEET-kategori.

Fra kategoriene over kan vi nå avdekke hvilke grupper som er i jobb eller ikke. De som er i jobb er de som er i kjernearbeidsstyrken og i den ustabile arbeidsstyrken, mens studenter, uføretrygdede og langtidssykemeldte er andre kategorier som heller ikke gir mening å inkludere i NEET-begrepet. Jeg inkluderer jobb-kategoriene og de andre kategoriene for å kunne se hvordan de forskjellige gruppene står i forhold til hverandre i størrelse. Det er også av interesse å se hvordan disse kategoriene endrer seg over tid.

Jeg har skiftet ut NBA med grunnbeløpet, siden de i den opprinnelige modellen som Bäckman et al. (2011) benytter seg av brukes price base amount (PBA), som benyttes for eksempelvis å beregne pensjon (Bäckman et al., 2011). På denne måten tilsvarer PBA det norske grunnbeløpet. I 2014 var 3,5 G lik 309 295.

Jeg har ikke benyttet meg av alle kategoriene nevnt i Bäckman et al. (2011). Jeg har utelatt pensjonistkategorien, siden det ikke eksisterer pensjonister i datamaterialet mitt. Jeg har også valgt å slå sammen NEET-kategoriene; alternative inntektskilder og økonomisk inaktive og kalt disse for økonomisk inaktive. Jeg har valgt å slå disse sammen fordi kategoriene er nokså like. Dette vil bidra å lette tolkningen av NEET-kategoriene, og gjøre det enklere å tolke endringene over tid for NEET-kategoriene.

### **2.1.2 Kritikk av arbeidsmarkedstilknytningsmodell**

Når jeg benytter meg av modellen i boks 2.1.1 på datamaterialet er det ikke alle endringer i tilstand modellen greier å fange opp. Modellen har vanskelig for å identifisere tilstanden til et individ som har forskjellige arbeidstilknytninger tre år på rad. F.eks. har modellen vanskelig å bestemme hvilken arbeidstilknytning en person er i hvis personen første året er arbeidsledig store deler av året, for så å være i jobb andre året, for så å være mye sykemeldt

<sup>3</sup>Det kan diskuteres om dette er NEET-kategori. Jeg diskuterer dette i kapittel 2.3.

det tredje året. Dette er uheldig siden mange unge har en varierende arbeidstilknytning, og jeg får da ikke med de med varierende arbeidsmarkedstilknytning. Jeg har derfor forsøkt å forbedre modellen til Bäckman et al. (2011). Jeg har undersøkt forløpene som ikke blir fanget opp av den opprinnelige modellen og har funnet en del forløp som går igjen. Underkategorier som «re- eller etableringsfasen» og «marginal arbeidsmarkedstilknytning» er kategorier som er tenkt å plukke opp personer som har en varierende arbeidsmarkedstilknytning. For mine data er inntektsgrensene kanskje for restriktive, og greier f.eks. ikke å indentifisere personer som har mellom 0.5G og 1G i inntekt og har mer enn 360 dager (180 + 180 dager) som arbeidsledig i løpet av tre år. Selv om de har en inntekt høyere enn den i artikkelen til Bäckman et al. (2011), vil disse individene likevel kunne kategoriseres som langtidsledige.

Personer som hopper mellom forskjellige arbeidstilknytning tre år på rad er vanskeligere å kategorisere. F.eks. hvilken kategori bør en person som det første året jobber fulltid (over 3,5G i inntekt), for så å være sykemeldt i mer enn 90 dager i det neste året, for så å ha en middels inntekt (f.eks. 2G i inntekt) det tredje året. Hvilket forløp beskriver best mulig denne personen? Det kan fremstå som om denne personen er på vei tilbake inn i arbeidslivet igjen, og bør derfor kategoriseres i «re- eller etableringsfasen». En annet forløp kan være at personen er student første året, har høy inntekt andre året og er arbeidsledig store deler av det tredje året. Hvilken kategori beskriver best mulig denne personen sitt forløp? Personen kan kategoriseres som arbeidsledig, men da kan man spørre seg hvorfor ikke de to foregående årene hvor personen gjorde noe «meningsfullt», altså å jobbe eller ta utdanning, tas med i beregningen. Jeg vil påstå at den kategorien som best beskriver denne personen er i «det marginale arbeidsmarked». Dette fordi personen gjør noe meningsfullt de to foregående årene, hvor personen jobber og tar utdanning.

Modellen plukker heller ikke opp personer som utfører førstegangstjeneste og hjemmeverende mødre. Det kan derimot settes spørsmål ved om modellen er god til å plukke opp personer som utfører ideelt arbeid eller er tar seg et friår. Personer som utfører ideelt arbeid eller tar seg et friår skal ideelt sett plukkes opp av kategorier som «i re- eller etableringsfasen» eller «marginal arbeidstilknytning», disse kategoriene beskriver personer som har en viss form for arbeidstilknytning. Dette forutsetter likevel at denne perioden med reiser, ideelt arbeid eller lignende ikke varer i lengre enn 1-2 år. Fra figur 3 ser vi at disse kategoriene er relativt små. Det kan derimot settes spørsmål ved hvor mange som utfører ideelt arbeid eller tar seg et friår når de er 23 år gamle, og at disse kategoriene er så små for denne aldersgruppen.

Det kan også settes spørsmål ved om 3,5 G er for høyt beløp for å indikere om en person er i fullt arbeid. Dette kan bidra til at for mange av de unge blir klassifisert som å jobbe deltid, eller faller i de andre svakere arbeidsmarkedskategoriene.

## 2.2 Frafall og forsinket skolegang - definisjonsavklaring

Utdanningsvariablene i datasettet består av normert antall år høyeste oppnådde utdanning har tatt og hvilken type utdanning dette er. Jeg kan derfor ikke skille mellom personer som faller tidlig fra i videregående utdanning, eksempelvis første år, eller personer som faller seint fra, f.eks. tredje år. Normert tidsbruk for fullføring av videregående utdanning for allmennfaglig studieretning er 3 år og for yrkesfaglig utdanning 4 år. Det viser seg at store deler av de som tar videregående utdanning bruker mer enn normert tid på å fullføre (Falch, Borge, Lujala, Nyhus og Strøm, 2010). I oppgaven defineres frafall som de som bare har ungdomsskole som høyeste fullførte utdanning. Det er høy sannsynlighet for at disse har begynt på videregående utdanning, men så falt fra og derfor ikke har oppnådd en høyere utdanning enn ungdomsskolen. Dette fordi om lag 95 til 97 prosent av hver kohort starter videregående utdanning samme år som de avslutter ungdomsskolen (Falch et al., 2009). På samme grunnlag kan jeg heller ikke avdekke om en person har droppet ut av høyere utdanning, disse vil stå oppført med vgo som høyeste fullførte utdanning. Forsinket skolegang defineres som det å bruke mer enn fem år på videregående utdanning. Et mulig problem med registrering av forsinket skolegang er personer som bruker flere år enn observasjonsvinduet til å fullført. For mine data vil disse individene være registrert med frafall i vgo.

Siden jeg bare har informasjon om høyeste fullførte utdanning er det vanskelig å skille mellom om en person er NEET eller om personen er i utdanning for de under 18 år. Dette fordi personene kan ha liten inntekt fordi de er i utdanning eller fordi de har falt fra utdanningen og ikke arbeider. Derfor blir det også vanskelig vurdere tilstanden til en 18- og 19-åring, siden arbeidsstatus for disse aldersgruppene vurderes ut ifra tilstanden de er i som 16-, 17-, 18- og 19-åring.

## 2.3 NEET og sykepenger

Det kan diskuteres hvilke grupper som skal inkluderes i NEET-definisjonen. Bäckman et al. (2011) inkluderer personer som har mer enn 90 dager med sykefravær i NEET-kategorien. Dette gir ikke nødvendigvis mening for en norsk definisjon, siden man må ha en arbeidsgiver for å få rett til sykepenger. Det kan da argumenteres for at disse personene ikke har en perifer arbeidsmarkedstilknytning, og derfor ikke er NEET. Det viser seg dessuten at de som blir klassifisert som langtidssykemeldte i mitt utvalg har en rimelig høy medianinntekt, ca. 119 000 kroner, noe som gir mening siden man blir kompensert fullt ut for inntekten man hadde i forveien av sykdomsperioden. I denne oppgaven kommer jeg i motsetning til Bäckman et al. (2011) å ekskludere de langtidssykemeldte som en NEET-kategori.

## 2.4 NEET og attføring og rehabilitering

Personer som mottar ytelser som et resultat av svekkede arbeidsevner, kan få tilbud om forskjellige arbeidsfremmende tiltak, eller medisinsk rehabilitering hvis grunnen til at de ikke jobber er medisinsk. I realiteten viser det seg at mange av som får stønad til å være på arbeidsfremmende tiltak, er passive stønadsmottakere, se f.eks. Mandal, Jakobsen, Jensen og Osborg (2015). Passive stønadsmottakere er i realiteten en NEET-kategori, så lenge de ikke er i jobb eller i utdanning, for de er realiteten heller ikke «på» tiltak; de bare mottar stønad. Spørsmålet er da hvordan jeg kan skille mellom de «passive» og «aktive» stønadsmottakerne i mine data. Jeg har valgt å se på hvilke attføringskoder og arbeidsledighetskoder personer er registrert med. En person er en aktiv stønadsmottaker hvis han i løpet av et år er registrert med<sup>1</sup>:

- Arbeidssøkerkode:
  - Ordinær tiltakdeltager.
  - Yrkeshemmede på tiltak.
  - Arbeidssøker i jobb, skole eller som mottar etablerertilskudd.

---

<sup>1</sup>Se Akselsen, Lien og Sivertstøl (2007) eller <https://www.nav.no/no/NAV+og+samfunn/Statistikk/Arbeidssokere+og+stillinger++statistikk/Tiltaksdeltakere> for detaljer.

Passive stønadsmottakere har jeg definert som at de får utbetalt attføring- eller rehabiliteringspenger, men er ikke registrert med arbeidssøkerkodene nevnt over.

## 2.5 Svakheter ved NEET-definisjon i oppgaven

Ved å definere NEET som jeg gjør fanger jeg opp mange personer som av legitime grunner ikke er i jobb eller tar utdanning, f.eks. pga. udiagnostiserte psykiske lidelser. Mange rekker ikke å få innvilget uføretrygd før de er i 30-årene, selv om en aldri har hatt noen stabil arbeidssituasjon (Olsen et al., 2013, s. 42). Siden de eldste personene i datasettet bare rekker å bli 31 år, blir nok andelen NEET sannsynligvis høyere enn hva den da burde vært. Jeg fanger i tillegg opp hjemmeværende mødre og personer som utfører førstegangstjeneste. Quintini og Martin (2006) finner at omtrent halvparten av unge menn som er NEET i Sverige, utfører obligatorisk militærtjeneste eller reiser i utlandet. Siden jeg ikke kan utelukke de som utfører førstegangstjeneste i mine data kan da NEET-andelen være for høye for menn. Både hjemmeværende mødre og personer som utfører førstegangstjeneste er i realiteten ikke NEET, mitt fokus er på de som kunne ha jobbet eller tatt utdanning, men som ikke gjør det. NEET-begrepet består i realiteten av en rekke forskjellige undergrupper, og noen av disse består av personer som er svært sårbare. Hyggen (2013) og Furlong (2006) deler inn i følgende hovedkategorier:

- Personene som er korttids- og langtidsledige. De korttidsledige er ikke nødvendigvis i en sårbar situasjon. Det er dog slik at tidligere arbeidsledighet gir økt sjanse for fremtidig arbeidsledighet (Arulampalam et al., 2001).
- Unge som ikke søker jobber eller ikke er i utdanning. Disse er gjerne en del av den gruppen som kalles for kjerne-NEET (Serracant, 2014).
- Personer som aktivt søker arbeid eller utdanning, men som venter med å jobbe eller ta utdanning fordi de mulighetene som presenteres for dem ikke oppleves å være i tråd med deres ønsker.
- Personer som tar et friår for å reise eller fokusere på aktiviteter som musikk o.l.

Vi ser at NEET-gruppen består av flere forskjellige undergrupper, med hver sine utfordringer, og hvor enkelte av undergruppene er sårbare med tanke på fremtidig arbeid. Tiltak for å

redusere andelen NEET har ikke nødvendigvis en like god effekt for alle undergruppene, siden hver undergruppe har sine egne utfordringer (Serracant, 2014). En annen ulempe ved bruk av NEET-begrepet er mangelen på en klar definisjon, noe som gjør det vanskelig å gjøre internasjonale sammenligninger (Furlong, 2006).

I motsetning til slik Hyggen (2013) fremstiller svakhetene ved NEET-definisjonen, kan arbeidsmarkedstilknytningsmodellen som jeg benytter meg av hjelpe å redusere størrelsen på enkelte av de ovennevnte hovedkategoriene. Når man observerer hvert individ i tre påfølgende år, reduseres problemet med grupper bestående av personer som tar seg et friår, jobber frivillig et år eller de som venter på riktig jobb eller utdanning.

## 2.6 En liten oppsummering

Hovedfokuset i denne oppgaven er på de personene som ikke blir fanget opp av de ordinære målene for utenforskap, altså de som ikke nødvendigvis har for svak utdanning eller svake ferdigheter som gjør det vanskelig å få jobb. Jeg bruker arbeidsmarkedstilknytningsmodellen til Bäckman et al. (2011) for å avdekke om en person er i jobb. Om en person er i utdanning avdekkes ved hjelp av utdanningsvariabler som blir beskrevet senere, og om en person er på tiltak blir definert som nevnt i kapittel 2.4.

1. Personer med avkortet utdanning eller problemer i utdanningsløpet:
  - (a) Personer som har falt fra videregående opplæring (vgo).
  - (b) Personer som er forsinket i vgo.
  - (c) Personer som ikke har startet på høyere utdanning.
2. Personer som ikke er i jobb:
  - (a) Personer som er langtidsarbeidsledige.
  - (b) Personer som er økonomisk inaktive.
3. Personer som er passive stønadsmottakere.

En person er da NEET hvis en av delkategoriene under 1) er tilfredsstilt og 2 og 3 er oppfylt.

Arbeidsmarkedstilknytningsmodellen vurderes ut ifra et treårs perspektiv. Personer som er økonomisk inaktive har så lav inntekt i 2 eller 3 av 3 år at de sannsynligvis ikke har anledning til å forsørge seg selv med egen inntekt. Utdannings- og tiltaksstatus vurderes årlig. Slik at om en person er NEET i 2003, vurderes jobbstatus ut ifra årene 2001, 2002 og 2003. Mens utdannings- og tiltaksstatus vurderes bare for 2003.

### **3 Utvalgt relevant forskning**

NEET er opprinnelig et sosiologisk begrep, og det er innen dette fagfeltet vi finner de fleste studiene gjort på området. Det er også utført mye forskning på områder som grenser opp mot NEET, som ungdomsarbeidsledighet og frafall, men i denne gjennomgangen kommer jeg til å begrense meg til noen artikler som spesifikt ser på NEET. Dette for å gi et inntrykk av den forskning som foregår på området. Som vi etterhvert skal se, tilbyr forskjellige artikler forskjellige definisjoner av NEET-begrepet. Flere benytter seg av registerdata og arbeidskraftsundersøkelser. Det meste av forskningen på området er rent deskriptive studier, altså at det ikke gjøres noen statistiske analyser som f.eks. regresjoner o.l..

#### **3.1 Nordiske studier**

Grødem et al. (2014) bruker registerdata. Det er data fra NAV kombinert med befolkningsstatistikk fra SSB. I dette materialet skiller først forfatterne ut en gruppe som de kaller registerbasert NEET. Den består av personer som har tre kjennetegn: de er ikke registrert som lønnstakere i november hvert år, de er ikke i formell utdanning i oktober hvert år og de deltar ikke i arbeidsmarkedstiltak i november hvert år. Fra denne gruppen skiller så forfatterne ut en undergruppe som består av personer som, i tillegg til kjennetegnene nevnt over, ikke mottar offentlig støtte: sosialhjelp, uførepensjon, attførings- eller rehabiliteringspenger. I tillegg til dette en inntekt som ikke overstiger 2 G. Denne gruppen kaller forfatterne for familieforsørget NEET. Dette navnet har forfatterne valgt fordi personene som inngår i gruppen trolig ikke har høy nok inntekt til å forsørge seg selv. I perioden 2000



til 2009 utgjorde gruppen omtrent 5 prosent av alle unge mellom 18 og 30 år, mens den førstnevnte gruppen, registerbasert NEET, utgjorde 13 prosent av samme aldersgruppe. Forskerne beskriver bare deskriptivt denne gruppen, og ser ikke på hvilke faktorer som kan disponere for tilstanden. Det argumenteres også for at det offentlige i større grad enn før fanger opp unge som ikke er i arbeid og utdanning. Dette kommer av at det i perioden skjer en økning i bruken av helserelaterte ytelser for unge. Forfatterne stiller også spørsmål ved om økningen i bruken av helserelaterte ytelser skyldes dårligere helse blant unge, eller om disse ytelsene tilbys som en løsning på problemet med unge voksne som faller utenfor arbeidsmarkedet.

Bäckman et al. (2011) sammenligner landene i Skandinavia når det gjelder NEET. Forfatterne benytter seg av samme arbeidsmarkedstilknytningsmodell som meg for å avdekke om en person er NEET i et gitt år eller ikke. I denne analysen inkluderer NEET de som er langtidsarbeidsledige, langtidssykemeldte, uførepensjonister, personer med alternative inntektskilder og de økonomisk inaktive. For den norske analysen benytter de registerdata fra FD-trygd, og for perioden 1995 til 2007 finner de at ca. 6 prosent av unge mellom 20 og 24 år er NEET. Målsettingen til studien er å se på hvordan organiseringen av yrkesutdanningen påvirker variasjonen til «suksessraten» i «skole-til-jobb-overgangen» mellom de skandinaviske landene. Forfatterne estimerer risikoen for å bli NEET med hjelp av en lineær sannsynlighetsmodell med fast-effekt for fylke. Fast-effekten skal være en grov korreksjon for arbeidsmarkedseffekter. Frafall fra vgo er en viktig risikofaktor i alle de skandinaviske landene. Forskerne ser på risikoen for å være NEET syv år etter at personen begynte på vgo. Den er høyere for de elevene som falt ut av vgo, særlig de som falt ut i det første eller andre året. Denne sammenhengen er sterkere for Norge enn for de andre skandinaviske landene. I Norge er risikoen for å være NEET høyest blant de som har falt ut av allmennfag, men i de andre nordiske landene er denne risikoen størst for de som falt ut av yrkesfag. Forfatterne har en a priori-hypotese om at mangel på lærlingplasser kan ha betydning for at det er større risiko for å være NEET i Norge enn i de andre skandinaviske landene, men dette finner de ikke beviser for. Et funn er at kjønnsforskjellene er store for Norge; kvinner har større risiko for å være NEET enn menn. Kjønnsforskjellene er ubetydelige i de andre landene.

Bø og Vigran (2014) ser på utviklingen i NEET over tid for unge under 30 år. Det benyttes

både tall fra AKU, samt data fra SSBs «System for persondata» (se artikkel for detaljer). For den registerbaserte NEET-definisjonen legger forfatterne til grunn sysselsettings-, utdannings- og tiltaksinformasjon fra 4. kvartal hvert år. Forfatterne ser først på utviklingen i NEET ved bruk AKU, og ser så på samme utvikling ved bruk av registerdata. Forfatterne finner at rundt 7 % av befolkningen mellom 15 og 29 år er NEET i 2013 når de bruker AKU. Generelt finner de at andelen NEET blir høyere når de bruker registerdata. Andelen NEET er i 2012 på rundt 11 %. Forskerne nevner at det er enkelte utfordringer ved bruk av registerdata. F.eks. har de problemer med å fange opp alle de som er i opplæring, jf. de som er på tiltak. De finner også at tallene på de som er i arbeid, utdanning eller i opplæring er noe lavere ved bruk av registerdata enn ved bruk av AKU. Dette fører igjen til at andelen NEET blir større ved å bruke registerdata. Forfatterne ser bare på omfanget av NEET deskriptivt, men gjør en tilsvarende analyse som jeg har gjort, hvor jeg ser på hvilke tilstander et individ som har vært NEET hopper over i i årene etter. De konkluderer med at personer som er NEET i et gitt år, ser ut til å ha mer varige problemer med å komme inn i arbeidsmarkedet.

Falch og Nyhus (2011) ser på hvilken betydning fullføring av vgo har på sysselsetting. Det benyttes registerdata fra FD-trygd som er koblet med utdanningsdata og andre databaser. Forfatterne benytter seg av en multinomisk logit-modell, altså hvor den avhengige variabelen kan ha flere enn to diskrete utfall. De deler avhengig variabel inn i de gjensidig utelukkende tilstandene «i utdanning», «sysselsatt i hel stilling», «sysselsatt i deltidsstilling» og et lignende begrep til NEET, det å være «ikke-i-jobb ikke-i-utdanning», eller inaktiv som forfatterne kaller det. Forskjellen på NEET og inaktivitet er de som er på tiltak eller opplæring. Forfatterne påpeker at det finnes flere uobserverbare faktorer som påvirker sysselsetting og som kan være korrelert med sannsynligheten for å fullføre vgo. De argumenterer for at god oppfølging hjemmefra bidrar til å øke sannsynligheten for å være i jobb eller utdanning og å fullføre vgo. Derfor inkluderes foreldrenes utdanningsbakgrunn, som forfatterne mener vil fange opp noe av denne sammenhengen. Et av funnene er at fullføring av vgo reduserer sannsynligheten for inaktivitet med 14,5 prosentpoeng. I tillegg gjennomfører forfatterne en separat analyse hvor de forsøker å korrigere for utelatte forklaringsfaktorer ved å sammenligne to ganske like grupper, de som «nesten» fullførte mot de som «så vidt» fullførte. Gruppen som «så vidt» gjennomførte er definert med et «lavt» karaktersnitt på vitnemålet fra vgo. De som «nesten» fullførte hadde et relativt godt karaktersnitt fra grunnskolen og hadde minst vært tre år i vgo, men ikke fullført. Resultatene fra denne analysen gir

sterkere effekt av det å ha fullført vgo, enn analysen som ikke brukte «sammenliknbare» grupper. Forfatterne understreker at siden analysen deres begrenser seg til de som startet på allmennfaglig studieretning, så kan ikke de uten videre sammenligne sine resultater med en analyse basert på alle individer.

Barth og von Simson (2012) undersøker konjunkturfølsomheten til arbeidsledighet og NEET og benytter seg av kvartalvise data fra arbeidskraftsundersøkelsen (AKU). De har benyttet seg av tidsrekkeregresjon der arbeidsledighet og NEET forklares ut ifra avvik fra BNP-trend med lineær trend. Lineær trend inkluderes for å korrigere for det at arbeidsledighet/NEET og konjunkturer trender samtidig. Forfatterne finner at ungdomsarbeidsledigheten er svært konjunkturfølsom. Det vil si at når det er dårlige tider og lite jobber, trekker unge seg i større grad ut av arbeidsstyrken. Det argumenteres også for at ungdommer i utdanning representerer en reservearbeidsstyrke som entrer arbeidsmarkedet i gode tider, men trekker seg ut i dårlige tider. Dermed blir ungdomsarbeidsledigheten mindre konjunkturfølsom hvis man trekker ut studenter og elever. Et annet funn er at konjunkturfølsomheten blant den yngste NEET-gruppen (16 til 19 år) er nesten fraværende, mens den er større for de eldre NEET (20 til 24 år). Forfatterne mener at det ligger andre forhold enn muligheter i arbeidsmarkedet til grunn for at de yngste er NEET, og mener at det er forhold som helse og motivasjon som styrer de yngste til utenforskap. Forfatterne finner også at det har skjedd en økning i andelen NEET som er arbeidsufør, både for menn og kvinner mellom 16 og 19 år. Andelen hjemmевærende NEET har sunket i hele perioden fra 1972 til 2012, både for menn og kvinner.

### **3.2 Andre internasjonale studier**

Furlong (2006) benytter seg av Scottish School Leavers Survey, hvor personer ble bedt om å svare på en spørreundersøkelse rundt ett år etter de gikk på det siste året av den obligatoriske utdanningen, jfr. grunnskolen. De spurte består av de som var 17 år i november 2003. Omtrent 10 prosent av mennene og 9 prosent av kvinnene var NEET. Forfatterne argumenterer for at NEET-gruppen er ekstremt heterogen, og forskjellige tiltak for å redusere andelen NEET derfor bare har begrenset effekt. Et interessant funn er at storparten av de som var NEET svarte at de ikke hadde funnet en passende jobb, etterfulgt av de som svarte at de ikke hadde bestemt seg for hvilken jobb eller utdanning de skulle ta. En annen stor

gruppe var de som svarte at at det ikke er noen skikkelige jobber eller utdanning der de bor. De tre viktigste årsakene til at de som var NEET ikke var i jobb eller utdanning kom fra valg de selv hadde tatt, fremfor mangel på muligheter til å jobbe eller ta utdanning. Av de mer «alvorlige» faktorene var «personlige problemer», hvor det var flest kvinner. Blant de som oppga at dårlig helse eller uførhet hindret dem å jobbe eller ta utdanning var flertallet kvinner. Forfatterne fant også at de som ikke var NEET i mindre grad hadde skulket eller blitt utvist, kom i fra mer privilegerte familier, og at det var høyere sannsynlighet for at foreldrene hadde akademiske grader og mindre sjanse for å være arbeidsledige blant de som ikke var NEET.

Quintini og Martin (2006) benytter seg av databasen til OECD, og fokuserer på situasjonen til unge i arbeidsmarkedet og ser på flere aspekter som påvirker dette. Et av funnene er at NEET-status først og fremst er en overgangsfase for de fleste unge personer, men at i land som Italia og Hellas tilbrakte mellom 20 og 30 prosent av unge fem etterfølgende år som NEET. Forfatterne finner også at NEET-andelene er høye i de fleste øst-europeiske land, og spesielt høy i f.eks. Tyrkia, hvor nesten 50 prosent av alle mellom 20 og 24 er NEET. Blant de yngste (15 til 19 år) har Norge blant de laveste andelene NEET, med under 5 prosent i 2003. Andel unge (20 til 24 år) NEET i Norge i 2003 ligger på 10 prosent.

### 3.3 Sammendrag og kritikk

Som det sees fra de to foregående delkapitlene, brukes det et spekter av forskjellige definisjoner for NEET. F.eks. benytter Grødem et al. (2014) og Bø og Vigran (2014) seg av informasjon fra enten én spesifikk måned eller ett kvartal av året for å beregne andelen som er NEET i et gitt år. Svakheten ved dette er at personer som ikke er i jobb i november eller i 4. kvartal i et gitt år fint kan ha vært i arbeid over lengre perioder tidligere i året. Det samme gjelder for utdanning og tiltak. Så personer kan feilaktig bli kategorisert som NEET, på tross av at de har jobbet, tatt utdanning eller vært på tiltak store deler av året. Modellen som jeg benytter meg av benytter seg dessuten av informasjon fra tre påfølgende år, og skal da i prinsippet være mindre utsatt for endringer i arbeidstilknytning fra år til år (Bäckman et al., 2011). Ellers brukes det arbeidskrafts- eller spørreundersøkelser (se Bø og Vigran (2014), Barth og von Simson (2012), Furlong (2006) eller Quintini og Martin

(2006)) for å definere NEET. Artikkelen til Falch og Nyhus (2011) analyserer en nærliggende gruppe til NEET, de ser på de som verken jobber eller tar utdanning.

Flere av artiklene inkluderer dessuten grupper med individer som jeg mener ikke er relevant for NEET-begrepet. F.eks. inkluderer alle artiklene nevnt i dette kapitlet uføretrygdede i NEET-kategorien. Jeg mener som tidligere nevnt at dette er en gruppe som ikke bør inkluderes i NEET-begrepet. Alle artiklene unntatt Grødem et al. (2014), inkluderer personer som får sykepengen i NEET-kategorien.

Tabell 1 viser en oversikt over de forskjellige artiklene med oversikt over hvilket fagfelt de kommer fra, anslag på andelen NEET, hvilke type data som benyttes og om det gjøres deskriptive eller statistiske analyser.

	Forfattere	Definisjon	% NEET	Type data	Type modeller
Sosiologi	Grødem et al. (2014)	Se kapittel 4.1	13 % registerbasert 5 % familieforsørget	Registerdata fra NAV og SSB	Deskriptiv
	Furlong (2006)	Nevnes ikke spesifikt.	10 % av menn 9 % av kvinner	Spørreundersøkelse	Deskriptiv
	Bäckman et al. (2011)	Se delkapittel 2.1.1	6 % av 20-24 år	Registerdata fra SSB	Deskriptiv og LPM.
Økonomi	Falch og Nyhus (2011)	ikke-i-jobb ikke-i-utdanning		Registerdata fra SSB	Multinomisk logit
	Barth og von Simson (2012)	AKU	10 % 20-24 år	AKU	Tidsrekkerregresjon
	Quintini og Martin (2006)	Definisjon tilsvarende AKU	5 % for 15-19 år 10 % for 20-24 år.	OECD, Eurostat	Deskriptivt
	Bø og Vigran (2014)	AKU	11 % av 15-29 år.	AKU og SSBs register	Deskriptivt

Tabell 1: En kort oversikt over de forskjellige artiklene, delt inn etter fagfelt osv.

## 4 Beskrivelse av datamaterialet

I dette kapittelet presenterer jeg datamaterialet for denne oppgaven. Det blir også en diskusjon rundt forskjellige utfordringer i datasettet. I tillegg presenteres noen viktige årsaker til at folk er NEET, f.eks. gjennomstrømning i vgo. Jeg kommer også til å presentere hvordan de forskjellige arbeidsmarkedstilknytningene står i forhold til hverandre og utvikler seg over tid.

### 4.1 Datamaterialet

Data er hentet fra ForløpsDatabase-Trygd (FD-Trygd) hos Statistisk Sentralbyrå. Dette er en database som består av data om personers trygdeforhold, sosialhjelp, samt bakgrunnsopplysninger om familieforhold, bostedskommune m.m.. Her kan det registreres hvordan en person over tid beveger seg mellom tilstander som arbeid og trygd, og mellom forskjellige trygdestatuser og sosialhjelpsordninger. Datamaterialet er anonymisert, men ved hjelp av unike identifikasjonsnummer kan data fra forskjellige databaser kobles sammen.

I tillegg har jeg fått opplysninger om hvert individs utdanningsløp fra Nasjonal utdanningsdatabase (NUDB). Her finnes opplysninger om hvor mange års utdanning en person har, og hvilket utdanningsvalg en person har tatt. Opplysningene herfra strekker seg helt fra personer som ikke har noen form for utdanning til personer med doktorgrader og universitetsutdanninger.

Jeg følger etnisk norske personer født i perioden 1977 til 1984 fra 1998 til 2008. Jeg har valgt ut disse fødselskohortene fordi de blir gamle nok til å bli ferdig med høyere utdanning og komme seg ut i arbeidslivet. Datasettet består av totalt 373 451 individer. Personene født i den tidsperioden jeg ser på begynner på videregående utdanning både før og etter Reform 94. Reform 94 gav 16 til 19 åringer rett til videregående utdanning i ett av tre prioriterte grunnkurs. Reformen reduserte også antall linjer man kunne velge til 13 grunnkurs og et antall videregående kurs I og II (Kirke-, utdannings- og forskningsdepartementet, 1996). Utdanningsløpet for de som er født før 1982 vil sannsynligvis være mer uoversiktlig, men det som er av interesse for oppgaven, er å se om disse har fullført videregående opplæring, om de har falt fra eller er forsinket i vgo. Det er to variabler knyttet til utdanningsløpet til

individene i datasettet. Den ene er høyeste fullførte utdanning, oppgitt i antall år. 10 år betyr at personen har fullført grunnskolen, 13 eller 14 betyr at personen har fullført videregående opplæring. I tillegg har vi opplysninger om hvilken utdanning en person har tatt, f.eks. hvilken videregående studieretning en person har gått på, eller hvilken universitets- eller høyskoleutdanning en person har tatt. Denne variabelen bruker jeg hovedsakelig for å se om en person har tatt yrkesfaglig eller allmennfaglig utdanning. Jeg har selv laget en variabel som i hvert år viser om en person er i utdanning eller ikke. Jeg har gitt de som går på vgo opptil 5 år på å fullføre vgo.

Andre variabler av interesse er inntekt, forskjellige sosial- og trygdeordninger som sosialhjelp, arbeidsledighetstrygd/dagpenger og andre overføringer fra folketrygden. Inntektsvariabelen som jeg bruker i min analyse er yrkesinntekt, som er lønnsinntekt i tillegg til næringsinntekt fratrukket arbeidsledighetstrygd.

I tillegg inneholder datasettet variabler om attføring og rehabilitering, hvor jeg har opplysninger om antall dager en person får attførings- eller rehabiliteringspenger, i tillegg til hva slags tiltak personen er på og hvor stor hvor stor uføregraden er. Datasettet inneholder også variabler om tidsbegrenset uføretrygd og varig uføretrygd, hvor jeg har informasjon om antall dager i snitt i løpet av en måned et individ har vært tidsbegrenset eller varig uføretrygdet, i tillegg til uføregrad.

## 4.2 utfordringer

En del av variablene har manglet informasjon i et eller flere år. Dette har resultert i at disse individene har blitt fjernet fra datasettet. Jeg har gjort dette for å skape et balansert panel.

### 4.2.1 Utdanningsdata

Flere individer manglet informasjon om høyeste fullførte utdanning i antall år eller utdanningsløpet sitt. I tillegg var det personer som hadde et utdanningsløp som virket umulig, personer som tok universitetsutdanning når de var i en alder som tilsa at burde vært i videregående opplæring. Totalt fjernet jeg 27 594 personer som hadde mangelfull informasjon, eller ufullstendige/«feile» forløp i utdanningsvariablene. Jeg fjernet dessuten 1644 personer som var registret uten noen informasjon om utdanning.



### 4.3 Utvalget

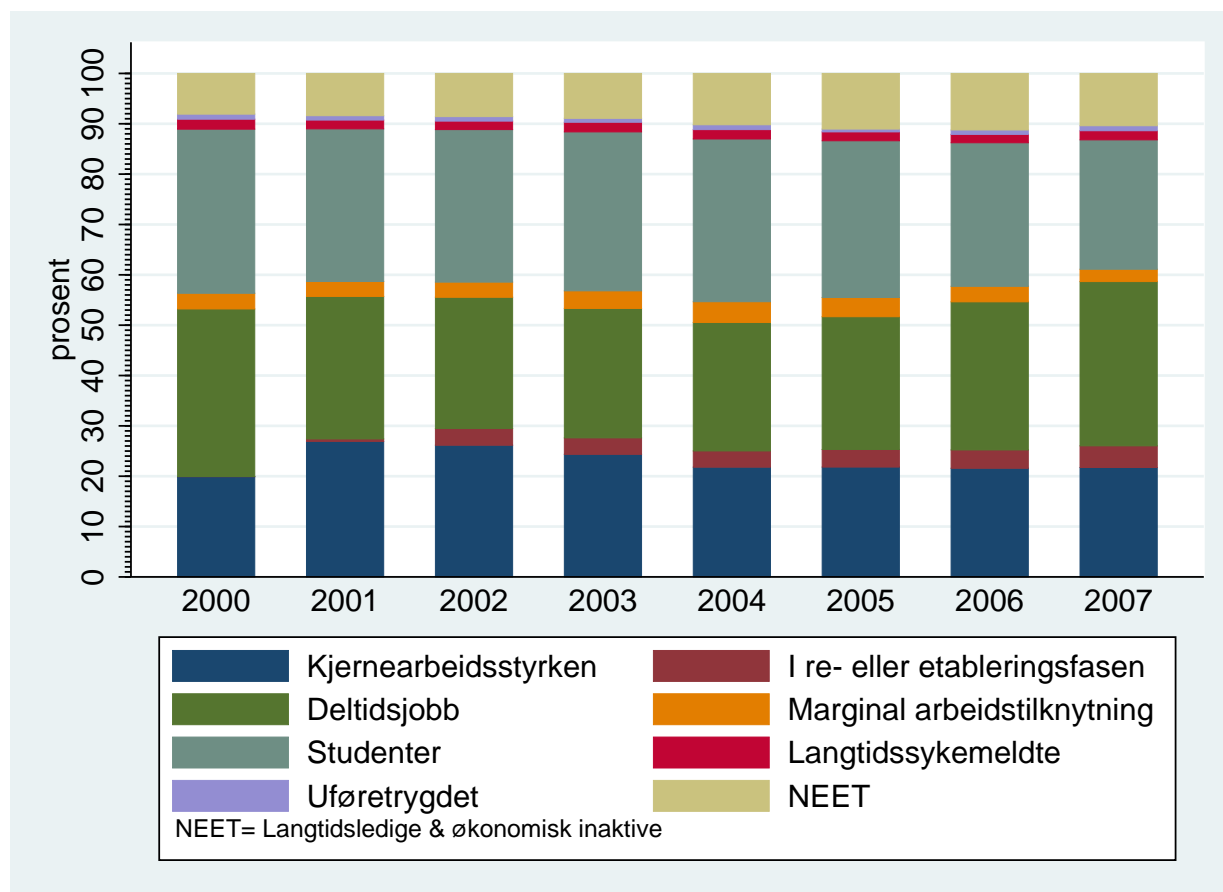
Utvalget mitt består av 344 213 individer etter at jeg fjernet individer med mangelfull informasjon.

Fødselsår	Kjønn		Total
	Kvinne	Mann	
1977	20 714	22 044	42 758
1978	20 745	22 401	43 146
1979	20 848	22 310	43 158
1980	20 580	22 324	42 904
1981	20 823	21 956	42 779
1982	20 965	22 728	43 693
1983	20 466	22 088	42 554
1984	20 860	22 361	43 221
Total	166 001	178 212	344 213

Tabell 2: Antall individer per kohort

Tabell 2 viser at kohortstørrelsen er høyest for 1982, mens den er lavest for 1983. Vi ser også at det er generelt et lite flertall av menn i hver kohort.

## 4.4 Arbeidsmarkedstilknytning og gjennomstrømming i vgo

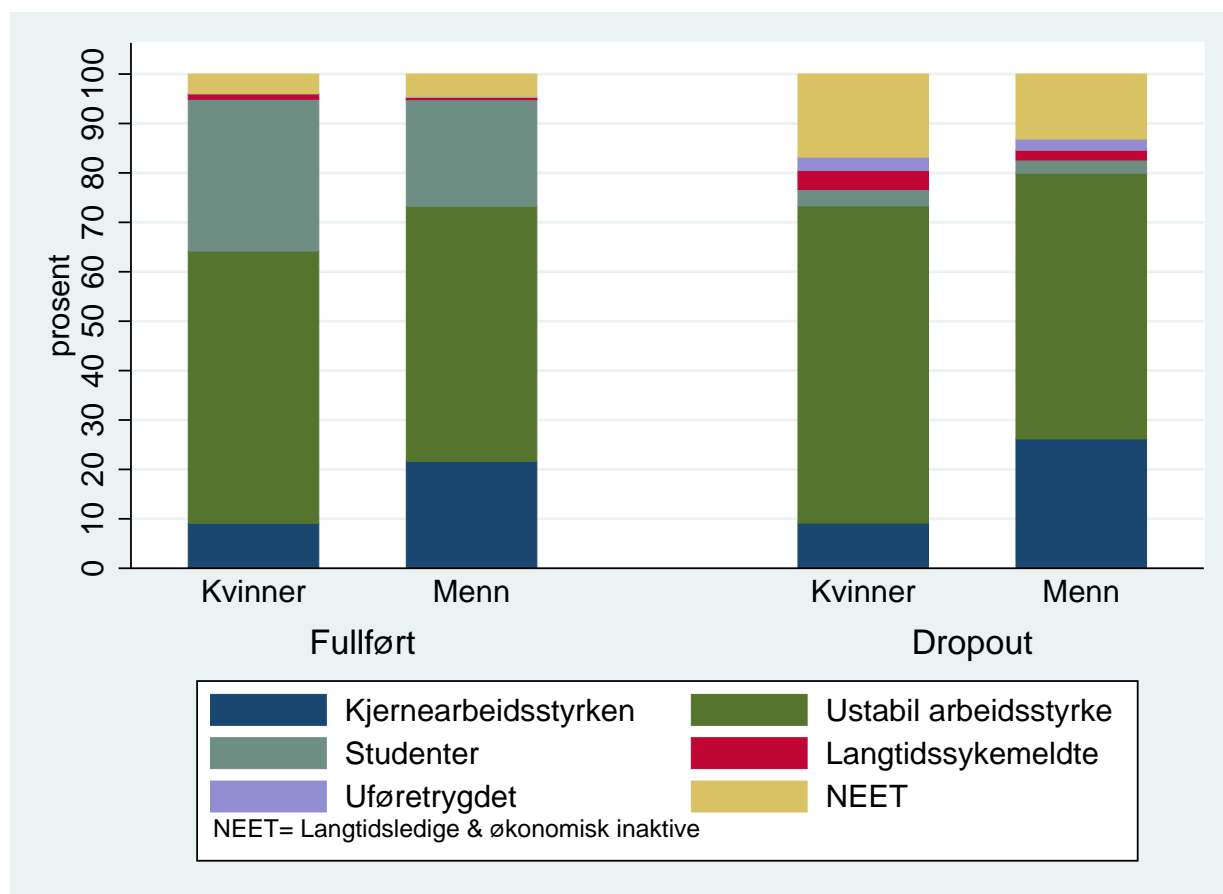


Figur 3: Arbeidsmarkedstilknytning i andeler syv år etter avsluttet ungdomsskole i perioden 2000 til 2007.

Figur 3 viser fordelingen av de forskjellige arbeidsmarkedstilknytningene presentert i tabell 2.1.1 syv år etter avsluttet grunnskole, når personene er 23 år. År 2000 er altså det året de som er født i 1977 er 23 år, mens i 2007 er de som er født i 1984 23 år. De to største kategoriene er de som jobber deltid eller er studenter. Det gir mening at de fleste 23-åringene enten studerer eller jobber deltid og ikke har hatt to av tre påfølgende år med «høy» inntekt, slik som det kreves for å tilhøre kjernearbeidsstyrken. En annen mulig grunn til at personer som er klassifisert til å være i deltidsjobb, egentlig jobber fulltid, men at inntektsgrensene er for høye til at de blir klassifisert som i fulltidsjobb. Vi ser at andelen som jobber deltid reduseres med omtrent en fjerdedel i fra 2000 til 2001, mens kjernearbeidsstyrken øker i samme periode. Man skulle kanskje egentlig forvente at det skjedde en nedgang i disse kategoriene, siden IT-boblen sprakk i 2001 (Cappelen og Eika, 2010), likevel viser ikke

modellen noen reduksjon i disse kategoriene. Dette kan komme av at modellen virker med et visst tidsetterslep. Slik at hvis det var nedgangskonjunktur i 2001 og 2002, begynner ikke dette å vise seg før i 2003 eller 2004 siden modellen baserer tilstanden i 2003 ut i fra 2001, 2002 og 2003. Det kan virke som om det er det som skjer, siden det skjer en reduksjon i deltidsjobbkategorien og kjernearbeidsstyrken i 2003 og 2004. På samme tid ser vi en økning i studentkategorien, noe som er konsistent med litteraturen på området som tilsier at personer i større grad velger å ta utdanning fremfor å jobbe i dårlige tider (Barth og von Simson, 2012). Vi ser at fra 2004 til 2007 ligger kjernearbeidsstyrken stabilt på rundt 25 %, mens det skjer en økning i andelen deltidsjobbende på bekostning av studentkategorien. Fra 2003 frem mot 2005 skjer det også en økning i kategorien «marginal arbeidstilknytning». Dette er personer som er på vei ut av arbeidsmarkedet og inn i arbeidsledighet eller økonomisk inaktivitet. Det virker rimelig at det er flere av personer med dette forløpet i nedgangskonjunktur. For 2000 og 2001 er det nesten ingen personer som er «i re- eller etableringsfasen», dette kommer sannsynligvis av problemer med operasjonaliseringen av disse gruppene. Modellen kan ha problemer med å plukke opp denne kategorien for de to årene. For de senere årene ligger den jevnt på rundt 5 %. NEET-gruppen varierer rundt ti prosent i perioden, NEET-andelen var størst i 2005 og lavest i 2001. Rundt 10 % av alle 23-åringer faller innenfor min definisjon av NEET.

Figur 4 har jeg valgt å ta med for å kunne sammenligne med tilsvarende figur i artikkelen til Bäckman et al. (2011, s. 28). Bäckman et al. (2011) ser på de som gikk på yrkesfaglig retning. Jeg har ikke anledning til å dele inn på samme måte, fordi jeg bare har informasjon om den høyeste fullførte utdanningen til en person. Jeg kan dermed ikke trekke ut informasjon om hvilket linjevalg en person som droppet ut av vgo gjorde.

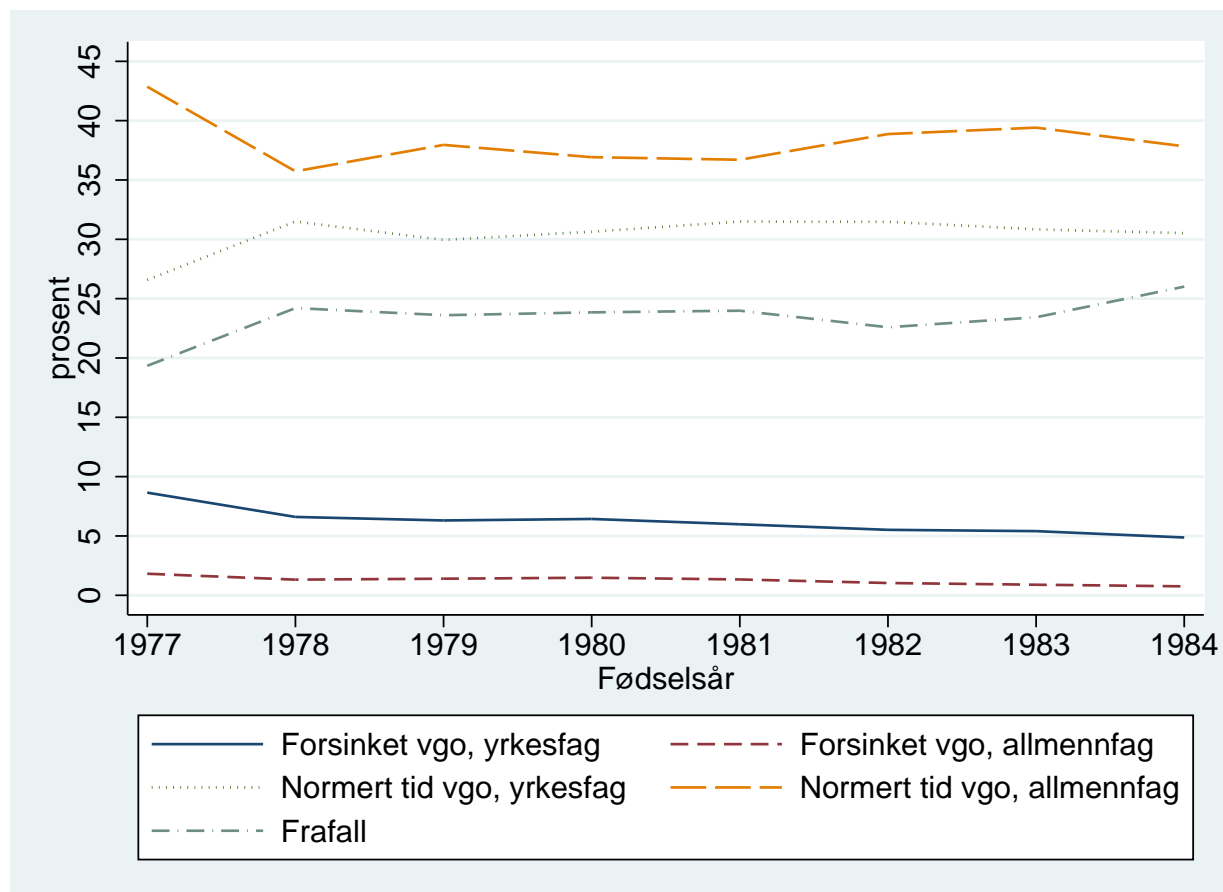


Figur 4: Arbeidsmarkedstilknytning i 2006, syv år etter avsluttet ungdomsskole delt inn etter kjønn og fullføring av vgo.

Vi ser fra figur 4 at den «ustabile arbeidsstyrken» som består av deltidsjobbende, personer som er i re- eller etableringsfasen og personer som har en marginal arbeidstilknytning, er den desidert største kategorien. Til forskjell fra tilsvarende figur i artikkelen til Bäckman et al. (2011) er kjernerarbeidsstyrken betydelig mindre, hvor Bäckman et al. (2011) finner at ca. 60 % av menn som fullførte vgo var i kjernerarbeidsstyrken syv år etter påbegynt vgo. Tilsvarende finner jeg at 20 % av menn var i kjernerarbeidsstyrken. Bäckman et al. (2011) begrenser seg, som tidligere nevnt, til å se på de som valgte yrkesfaglig linje. Det

er vanskelig å vite hva forskjellen i størrelsen kommer av, den kan komme av forskjellig operasjonalisering. Jeg har valgt å bruke Grunnbeløpet, mens Bäckman et al. (2011) belager seg på et beløp som utgjør 25 % av median bruttoinntekt for hvert år. Dette beløpet kaller de for Nordic Base Amount (NBA), det kan da hende at NBA er lavere enn grunnbeløpet. Og flere personer tilfredsstiller inntektsgrensen i NBA enn i G. Vi ser at studentkategorien er større for min figur, noe som korresponderer med at personer som velger allmennfaglig studieretning i større grad tar høyere utdanning enn personer som har valgt yrkesfag. Andre interessante funn er at kategoriene langtidssykemeldte og uføretrygdete er større for de som droppet ut av vgo. Det er flere kvinner som er langtidssykemeldte og uføretrygdete. Vi ser også at kategorien uføretrygdete er fraværende blant de som fullførte, dette er kanskje ikke så overraskende siden størsteparten av de som er uføretrygdete i en alder av 23 oftere er psykisk utviklingshemmet, har medfødte misdannelser eller kromosomavvik (f.eks. Downs syndrom) (Brage og Thune, 2015). Vi ser også at det er langt flere av de som faller fra vgo som er NEET. Og kvinner som er NEET er i flertall. Mens for de som fullførte er det omtrent like mange menn som kvinner som er NEET. Andre interessant funn ved tabellen er at det ser ut til å være langt flere langtidssykemeldte og uføretrygdete blant de som har falt fra vgo. Det ser også ut til å være flere kvinner som er NEET eller langtidssykemeldte blant de som falt fra, i forhold til mennene som falt fra vgo.

Vi har nå sett på hvordan arbeidsmarkedstilknytningene og NEET står i forhold til hverandre. Det kan være interessant å se hvordan videregående utdanning varierer med fødselskohortene. Falch og Nyhus (2011) og Bäckman et al. (2011) finner at frafall er en viktig prediktor for å være NEET. Det kan derfor være interessant å se på utviklingen i frafall for de forskjellige kullene. I tillegg kan det være interessant å se hvor mange som er forsinket eller fullførte på normert tid.



Figur 5: Gjennomstrømming i vgo fordelt på allmennfaglig og yrkesfaglig utdanning.

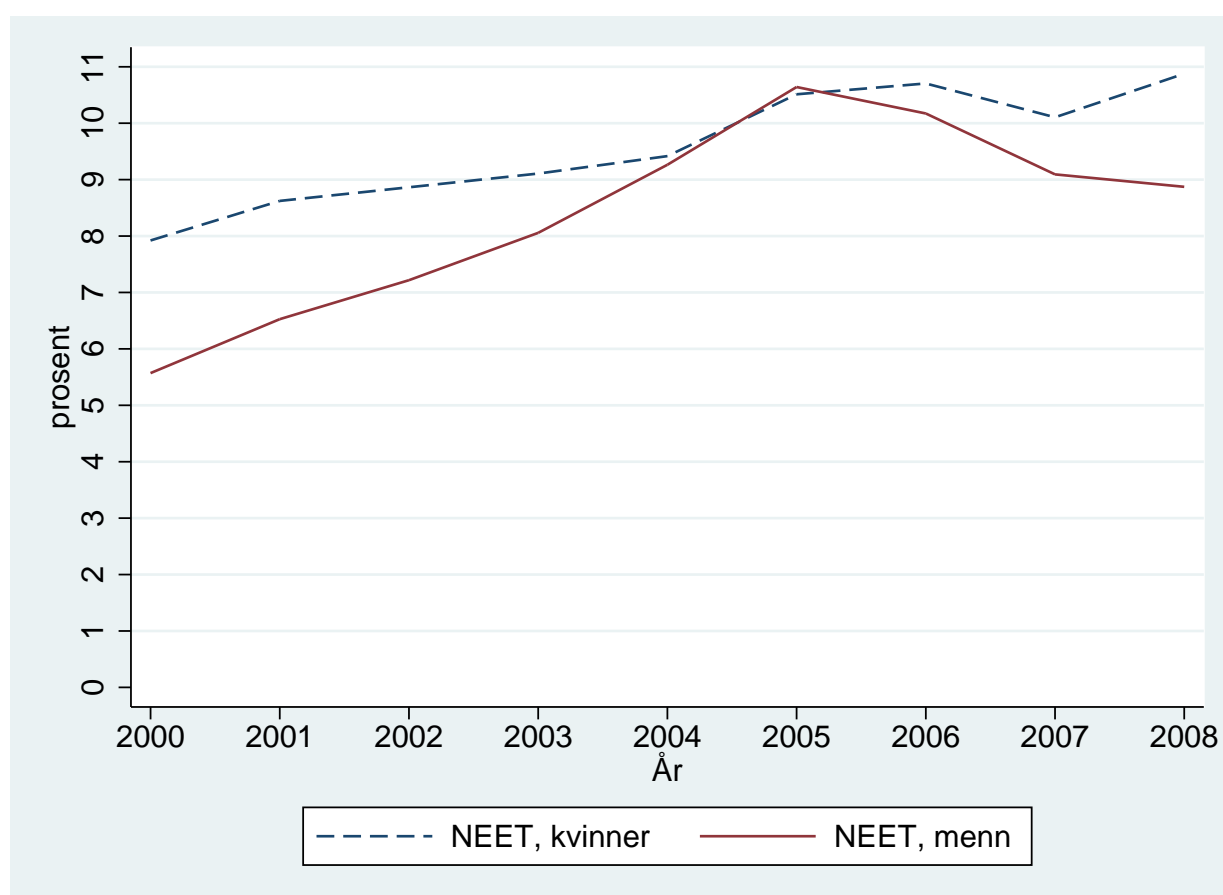
Vi ser i figur 5 at frafallet varierer mellom 20 og 25 prosent for alle fødselskohortene. Et interessant funn er imidlertid den utviklingen vi finner blant de eldste fødselskohortene. 1978-kohorten var den første kohorten som ble omfattet av reform 94. Vi ser en markant nedgang i de som gjennomførte allmennfaglig studieretning på normert tid fra 1977- til 1978-kohorten. Det kunne vært interessant å se på hvordan frafall fordelte seg mellom yrkesfag og allmennfag, men det har jeg dessverre ikke anledning til, siden jeg ikke har data på det. Vi ser og at det er flere som fullfører vgo på normert tid på allmennfag enn

på yrkesfag hvor gjennomstrømningen ligger på mellom 35 og 45 prosent. I motsetning ser vi en økning i andelen som gjennomførte yrkesfag på normert tid og de som falt fra vgo for de samme fødselskohortene. Vi ser også at det er en liten nedgang i andelen som er forsinket i vgo blant de yrkes- og allmennfaglige fra 1977- til 1978-kohorten. Totalt ligger gjennomføringandelen på rundt 70 prosent, noe som stemmer overens med tilsvarende tall fra SSB (SSB, 2014a).

## 5 Deskriptive funn

Første delen av kapittelet kommer til å ta for seg forskjellige aspekt ved utviklingen av NEET for personer i alderen 20 til 31. Andre del av kapittelet jeg til å gjøre en deskriptiv forløp- og overgangsanalyse over hvilke tilstander personer beveger seg mellom.

### 5.1 Utviklingen i NEET over tid

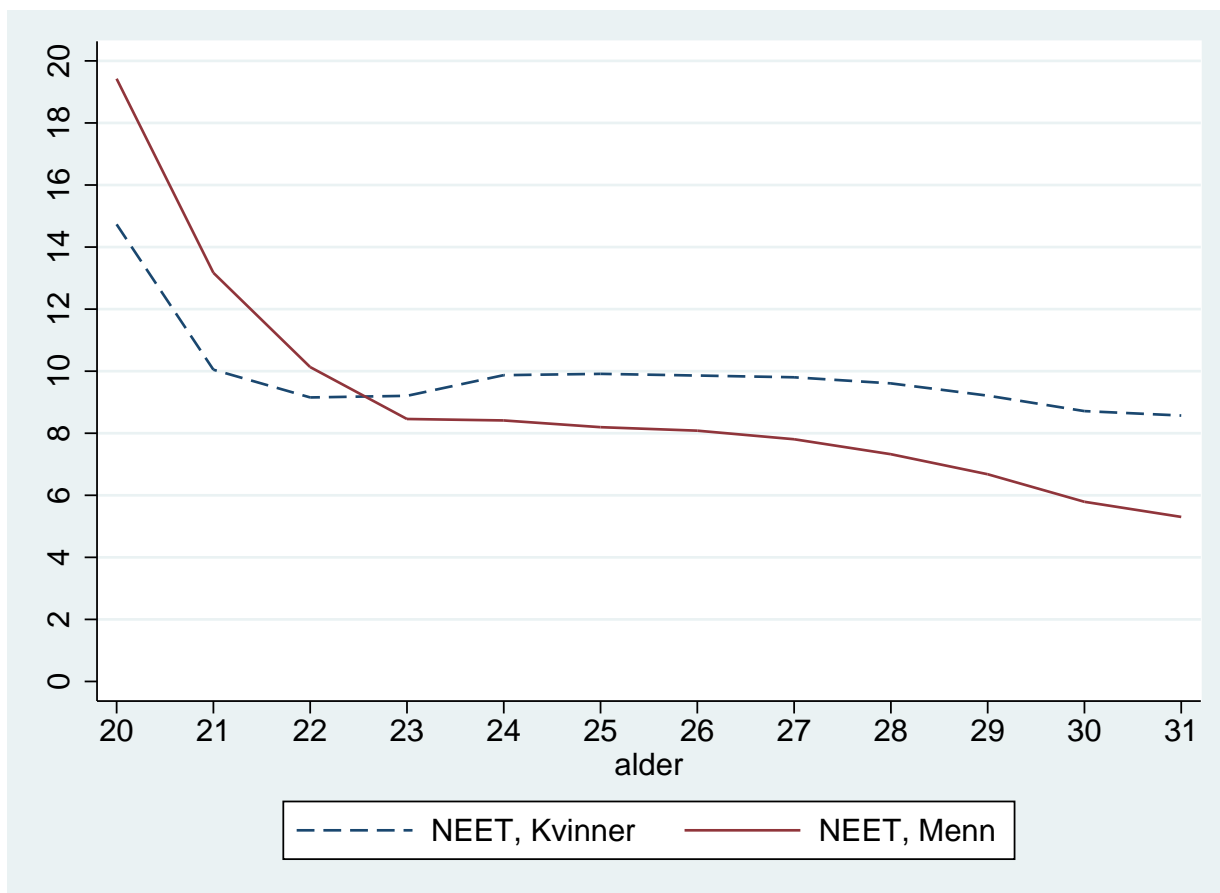


Figur 6: NEET (%) i perioden 1998 til 2008 for 23- og 24-åringer.

I figuren over ser vi utviklingen i andelen menn og kvinner som er NEET. Andelen er beregnet for 23- og 24-åringer fordi dette aldersspennet finnes gjennom hele perioden. Andelen NEET varierer i perioden på mellom rundt 6 og 10 prosent. Det er generelt flere kvinner som er NEET enn menn. I 2005 er det marginalt flere menn enn kvinner som er NEET. Det



er større variasjon i andelen menn som er NEET enn for kvinner. I perioden 2000 til 2005 øker andelen menn som er NEET nesten til det doble. Er det slik at det er blitt vanskeligere å tre inn i arbeidsmarkedet for unge menn i denne perioden? Noe av økningen i NEET kan muligens komme som følge av arbeidsinnvandring. En del av arbeidsinnvandrerne kan ha tatt ufaglærte jobber som f.eks. var vanlig å ta for personer som har falt fra. Arbeidsinnvandrerne stiller kanskje med sterkere kvalifikasjoner enn de som norske som har falt fra vgo. I 2001 sprakk IT-boblen og i 2001 og 2002 var det nedgangskonjunktur (Cappelen og Eika, 2010), og vi ser fra figur 15 (i vedlegget) at antall helt ledige stiger frem til 2003 for så å synke. Fra 2003 til 2005 er økningen større enn før 2003. En mulig grunn til dette kan være måten arbeidsmarkedstilknytningsmodellen min fungerer, den virker med et tidsetterslep. De «dårlige» tidene begynner sannsynligvis først å vise seg i arbeidsmarkedstilknytningsmodellen min i 2002 eller 2003, fordi jeg vurderer tilstanden ved tidspunkt  $t$  ut i fra  $t$ ,  $t - 1$  og  $t - 2$ . Barth og von Simson (2012) finner at NEET er konjunkturfølsomt. I nedgangskonjunktur øker altså andelen NEET. Dette forklarer muligens noe av økningen i andelen NEET i 2003 til 2005. Fra 2005 synker andelen menn som er NEET. I 2004 ble ordningen med tidsbegrenset uføretrygd innført, denne ordningen var særlig rettet mot unge (Bragstad, 2009). Noen av nedgangen etter 2005 kan derfor forklares av at noen av de som var NEET etter min definisjon, fikk tidsbegrenset uføretrygd. I antall var det 259 644 tilfeller av NEET blant de mellom 20 og 31 i perioden 2000 til 2008, fordelt på 103 500 personer. De fleste som da var NEET var det da sannsynligvis mer enn én gang. Både Bäckman et al. (2011) og Grødem et al. (2014) finner en lignende utviklingen i andelen som er NEET. Bäckman et al. (2011) finner at en lavere andel er NEET gjennom samme periode. Dette kan komme av forskjellige definisjoner og at vi ser på forskjellige alderspenn. Grødem et al. (2014) finner imidlertid at flest personer er registerbasert NEET i 2004.



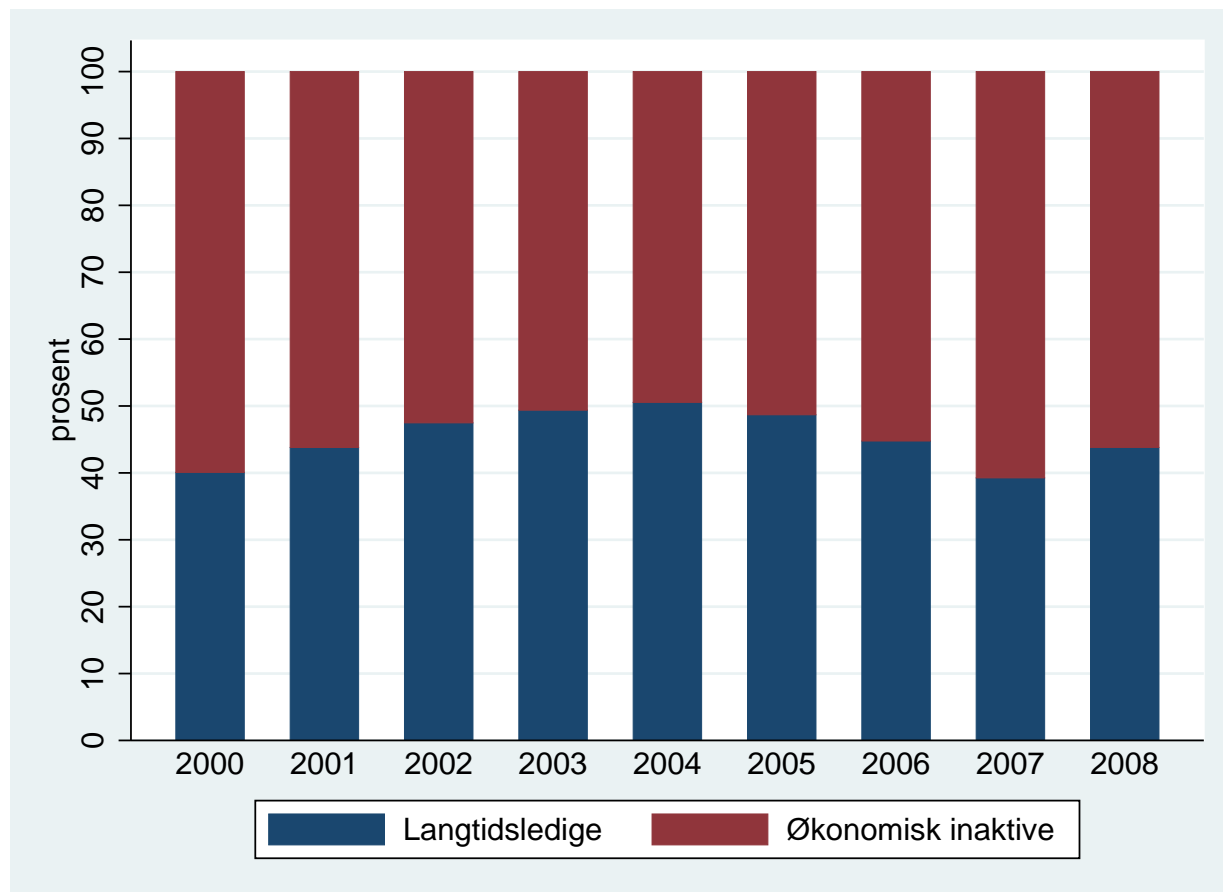
Figur 7: NEET for menn og kvinner etter alder.

Figur 7 viser hvordan andelen NEET endrer seg med alder. Vi ser at andelen NEET er størst for 20-åringene. Hvor ca. 1/5 menn er NEET når de er 20 år. Grunnen til at vi ser en opphopning for menn rundt denne alderen, kan være at mange menn utfører førstegangstjenesten. Førstegangstjenesten kan utføres fra man er 18 år. Man blir som regel ikke innkalt når man er i videregående utdanning. Personer som avlegger førstegangstjenesten er i prinsippet ikke NEET, fordi de gjør en jobb, eller er under en slags utdanning. Grunnen til at vi ser en opphopning blant 20 år gamle kvinner, kan være kvinner som har falt fra vgo som ennå ikke har fått seg arbeid. Vi ser fra figuren at andelen menn som er NEET er større enn for kvinnene frem til 23-årsalderen, hvor andelen kvinner som er NEET blir større. Bø og Vigran (2014) finner også at det flere kvinner enn menn som er NEET i alderen 25 til 29. At andelen kvinner som er NEET blir større etter denne alderen kan være fordi kvinner i større grad begynner å få unger etter denne alderen. Det kan argumenteres for at hjemmeværende mødre ikke er NEET; de oppdrar småbarn og gjør noe «funksjonelt» i

motsetning til andre unge som ikke jobber, tar utdanning eller er på tiltak. I prinsippet kunne vi da luket ut de hjemmeværende småbarnsmødrene og funnet at andelen kvinner som er NEET blir lavere.

Grødem et al. (2014) finner ikke samme utvikling i NEET etter alder. De finner at andelen NEET stiger fra 18- til 19-åringer, for så stabilisere seg for de eldre. Hvorfor de finner en annen utvikling enn hva jeg finner er vanskelig å vite. En mulig grunn kan være at Grødem et al. (2014) ser på situasjonen i 2009, altså rett etter finanskrisen inntraff. Arbeidsmarkedet for de over 20 kan da være vanskeligere enn perioden jeg ser på. Barth og von Simson (2012) finner at de som er NEET og mellom 20 og 24 år er mer konjunkturfølsomme enn den yngre NEET-gruppen. Det at min definisjon av NEET er forskjellig fra definisjonen til Grødem et al. (2014) bidrar også til at vi ser forskjeller.

Hvordan er sammensetningen av NEET-gruppen? Er det de som er langtidsledige eller er det personer som eventuelt støttes med penger hjemmefra eller jobber svart o.l.?

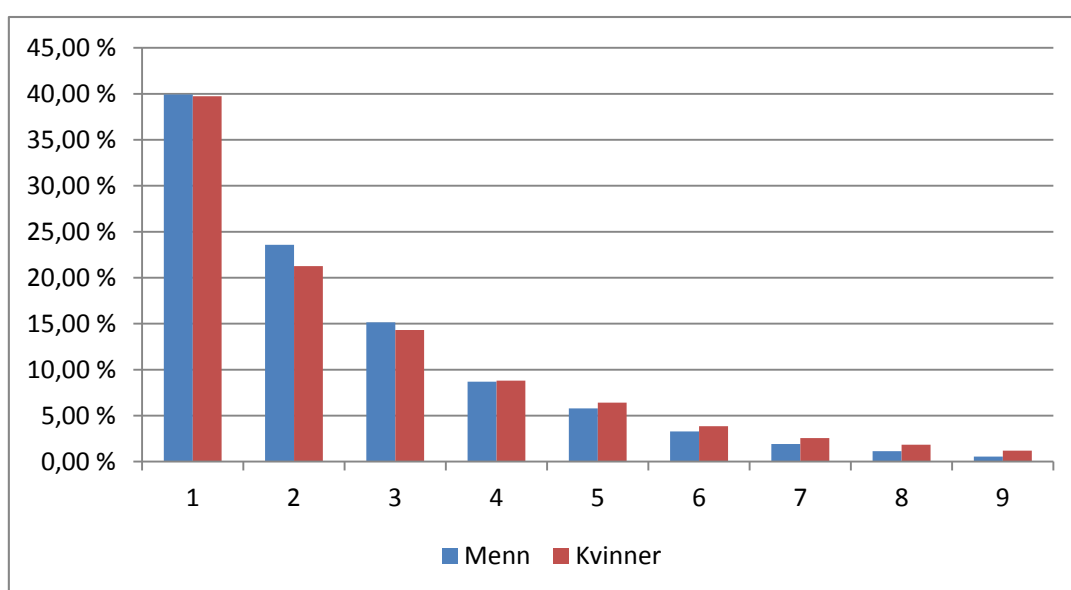


Figur 8: NEET dekomponert i forskjellige kategorier.

Figur 8 viser utviklingen i de forskjellige NEET-kategoriene over tid for 23- og 24-åringer. Den største gruppen er de langtidsledige, etterfulgt av de som er økonomisk inaktive. Den største NEET-gruppen er også de som har den laveste arbeidstilknytningen. De som er økonomisk inaktive er mest sannsynlig personer som i stor grad finansieres hjemmefra eller jobber svart. En av fordelene ved NEET-begrepet er at den som mål fanger opp arbeidsledigheten samtidig til de som står helt utenfor arbeidsmarkedet. Andelen langtidsledige øker i perioden 2000 til 2004, for å så å reduseres etter dette og så øke fra 2007 til 2008. Som tidligere nevnt kan grunnen til at vi ser en økning i andelen langtidsledige komme fra nedgangskonjunktoren i 2001 og 2002, grunnen til at denne økningen varer helt til 2004 kan komme fra operasjonaliseringen av disse kategoriene. Reduksjonen i langtidsledige etter 2004 kan komme av at en del av de får tidsbegrenset uførepensjon.

## 5.2 NEET - et overgangsfenomen eller en absorberende tilstand?

Er det slik at personer er NEET over lang tid og forblir der eller det slik at personer i er i den tilstanden over et par år og så kommer seg ut av den? Dette delkapittelet kommer til å forsøke å svare på dette ved å se på hvor lenge de som er NEET er i denne tilstanden og hvilke tilstander de eventuelt går over til.



Figur 9: Antall år som NEET fordelt på kjønn.

Figur 9 viser hvor mange år de som er NEET tilbringer i denne tilstanden. Vi ser at det er flest som er NEET i ett år, andelen er så synkende etter antall år. Medianantallet i år er ca. 2 år for menn og kvinner, hvor medianen for kvinner er litt høyere. Det er generelt lite kjønnsforskjeller. Vi ser at menn er mer representert for de er som NEET i 2 til 4 år, mens kvinner ser ut til å være noe lengre NEET enn menn. Det er omtrent dobbelt så mange kvinner som er NEET i hele tidsvinduet (9 år) som menn. Dette kan være en indikator på at kvinner som først har vært NEET har det vanskeligere å komme tilbake i arbeidslivet

enn menn. Raffe (2003) finner óg at kvinner forblir NEET lengre enn menn. Mulige grunner til at kvinner forblir lengre i denne tilstanden kan være feilkilder, som hjemmeværende mødre. En annen mulig grunn kan være at kvinner i større grad enn menn blir forsørget av sine samboere.

Et av spørsmålene i begynnelsen av dette kapittelet var om NEET var en tilstand som man var i et par år og så kom ut av, eller om det var en absorberende tilstand. Uføretrygd blir av og til omtalt som en absorberende tilstand (Haugen, 2013). Hvis vi da sammenligner hvor mange som forlater NEET-tilstanden, med hvor mange som fra ett år til det neste ikke lengre har uføretrygdstatus, kan vi få en indikasjon på om NEET er en absorberende tilstand.

		Tilstand <sub>t+1</sub>					
		1	2	3	4	5	6
Tilstand <sub>t</sub>	1.Jobb	91.54	4.14	3.54	0.67	0.07	0.03
	2.NEET	28.98	60.63	4.79	3.59	1.12	0.88
	3.I utdanning	24.16	2.29	73.17	0.36	0.01	0.00
	4.På tiltak	43.74	34.81	10.72	9.72	0.36	0.66
	5.Tidsbegrenset ufør	0.83	0.16	0.02	0.04	94.28	4.67
	6.Full uførepensjon	0.13	0.01	0.00	0.00	0.01	99.85
Andel i hver kategori		65.58	9.60	22.07	0.99	0.40	1.36

Tabell 3: Overgangsfrekvenser for forskjellige tilstander

Tabell 3 viser andelen som enten forblir i samme tilstand eller gikk over til en annen året etter. Generelt så ser vi at majoriteten av de som var i en tilstand et gitt år, forblir i samme tilstand året etter. Rundt 90 % av de som var i jobb, forblir i jobb året etter, mens rundt 4 % går over til å være NEET. Flesteparten av de som var NEET i et gitt år var i samme tilstand året etter, mens rundt 4 % gikk over i utdanning. Andelene som over på tiltak, tidsbegrenset (TU) og full uføretrygd var nærmest fraværende. Rundt 30 % av de som var NEET gikk over i jobb året etter. Rundt 4 % gikk over fra å være NEET, til å være på tiltak eller utdanning. Flesteparten av de som var under utdanning, forblir under utdanning året etter. Den nest mest vanlige overgangen fra å være under utdanning, er å gå over til jobb. Rundt 2 % av de som var i utdanning, gikk over til å være NEET. For de som var på tiltak er det annerledes, overgangene var mer jevnt spredd mellom de andre tilstandene. Rundt 44 % av de som var på tiltak gikk over i jobb året etter. Mens andelen som gikk over til å være NEET var på rundt 35 %. Rundt 11 % gikk over i utdanning og rundt 10 % gikk

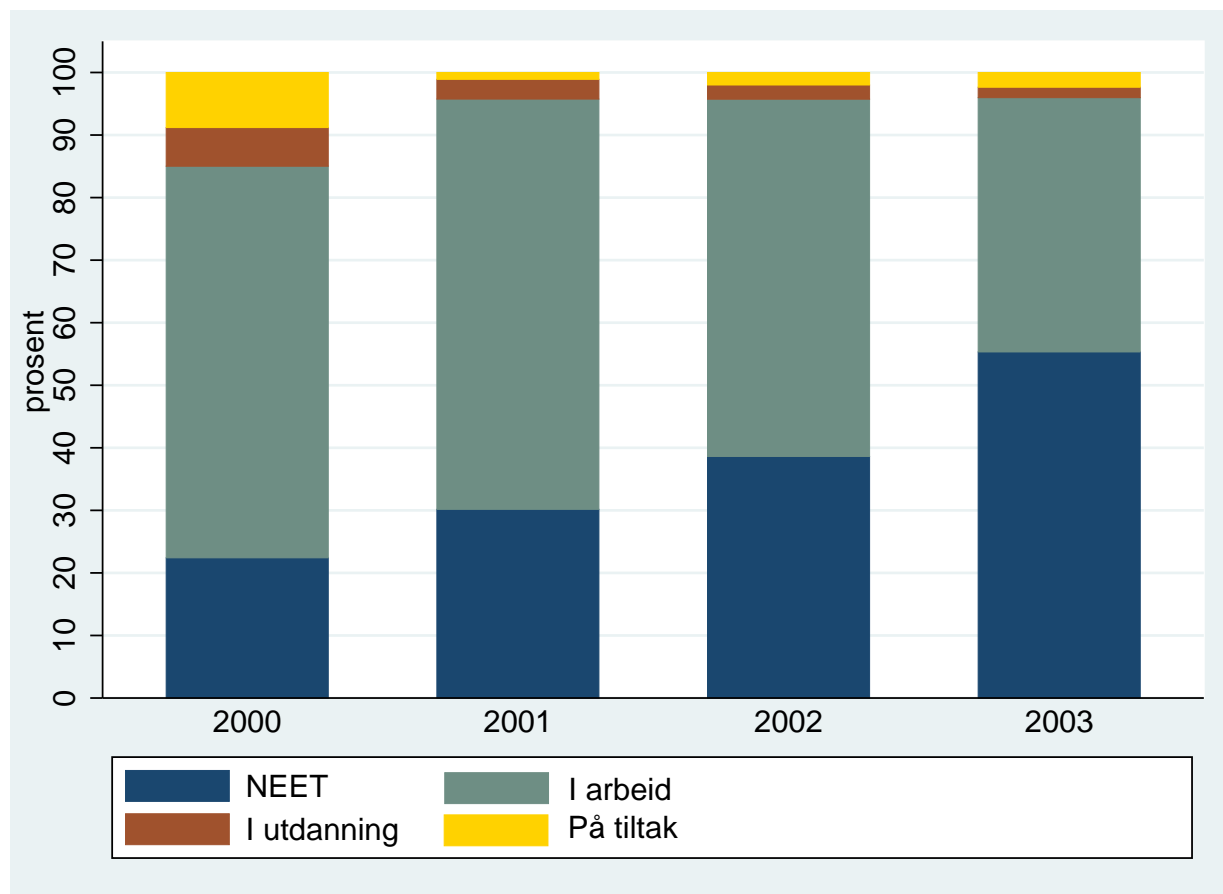
over på tiltak. Blant de som mottar TU, forblir nesten alle i samme tilstand året etter. Den nest mest vanlige overgangen blant de som mottok TU, var å gå over på full uføretrygd. Vi ser at nesten alle som fikk full uføretrygd et år, også mottok det året etter. Det var ingen som gikk over i utdanning og ingen som gikk over på tiltak, blant som mottok uføretrygd året før.

Alt i alt, er NEET-gruppen klart mer mobil enn de som mottok TU og full uføretrygd. Det var langt flere av de som var NEET som gikk over til andre tilstander året etter, enn for de som mottok TU og full uføretrygd.

Hvordan ser bildet ut hvis vi utvider tidsperspektivet. At vi ser lengre enn ett år frem tid, og i tillegg ser hvilke tilstander de som er NEET i et gitt år kommer fra? Resten av kapittelet kommer til å svare på det.

### 5.2.1 NEET: Hvor kommer de fra?

Et interessant spørsmål er hvilken tilstand de som er NEET kommer fra. Har de tidligere vært i utdanning eller i jobb eller vært på tiltak? Figur 10 viser forløpet til de som var NEET som 24-åring i 2004 <sup>1</sup>. Figur 10 viser at flesteparten som var NEET i 2004 hadde en



Figur 10: Forløpene til personer som er NEET i 2004 i perioden 2000-2003.

historikk som NEET. Omtrent 50 % av de som var NEET i 2004, var det også året før, for så å reduseres til litt under 40 % to år før. Vi ser også at omtrent 40 % av de som var NEET i 2004 var i arbeid året før. Denne andelen øker desto lengre tilbake i tid vi går. Det er relativt få av de som er NEET som i umiddelbar fortid har vært i utdanning. Vi ser at rundt 5 % av de som var NEET i 2004 var under utdanning i 2000, dette kan være personer som fullførte videregående på fire år. Under fem prosent av de var NEET i 2004 var på tiltak året før, denne andelen holder seg stabil tilbake 3 år i tid. Fire år før ser vi at rundt

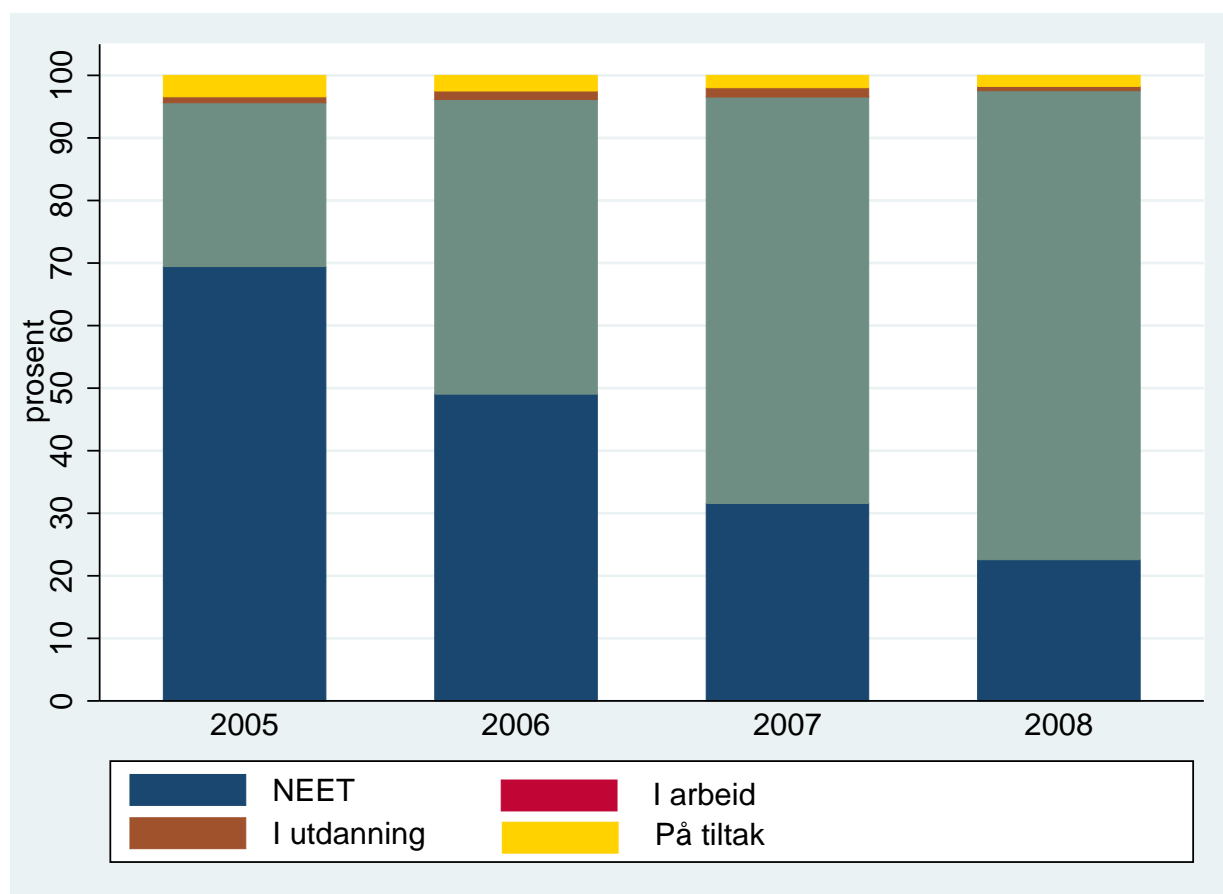
<sup>1</sup>Andre år som «ankerpunkt» viser omtrent samme forløp for jobb, utdanning eller tiltak.



10 % av de som var NEET som 24-åring var på tiltak. En mulig grunn til at denne andelen er høy, kan være at man hadde bare rett på attføring- og rehabiliteringspenger i fire år og de som da var på tiltak i 2000 hadde brukt opp denne rettigheten i 2004.

### 5.2.2 Overgang til jobb, utdanning eller tiltak

Mange personer vil ha en NEET-status flere år på rad. På tross av dette vil noen av de som er NEET i et gitt år gå over i arbeid, utdanning eller på tiltak. I figur 11 ser vi hvilke tilstander de som var NEET som 24-åring i 2004 «hopper» over i perioden 2005 til 2008 <sup>1</sup>.



Figur 11: Overganger til forskjellige tilstander for personer som var NEET i 2004 i perioden 2005 til 2008.

Vi ser fra figur 11 at omtrent 70 % av de som var NEET i 2004 også var det i 2005. Litt over 20 % av de som var NEET i 2004 var det også i 2008. Rundt 25 % av de som var

<sup>1</sup>Andre år som «ankerpunkt» viser omtrent samme overganger til jobb, utdanning eller tiltak.

NEET i 2004 hadde gått over i jobb året etter, og nesten 75 % av de som var NEET i «basisåret» var kommet seg i arbeid i 2008. Figuren viser også at få av de var NEET i 2004 gikk over til å være i utdanning senere. Vi ser at en større andel av de som var NEET i 2004 går over på tiltak enn de som kom fra tiltak før 2004. Likevel er som tidligere nevnt over 60 % også NEET i 2005. Dette kan tyde på at NAV og oppfølgingstjenesten ikke er flinke nok til å plukke opp de som er NEET. Det kan hende at de får utbetalt stønad, men at de er passive stønadsmottakere, siden det er en av NEET-gruppene. Det er dessuten interessant at en veldig liten andel av de som var NEET går over i utdanning. En andel av de som er på tiltak som innebærer at de tar utdanning, men de har jeg ikke anledning til å skille ut med mine data. Slik at andelen som reelt er under utdanning, er større enn det jeg finner. Overgangene jeg finner kan også tyde på at de som er NEET for så å komme seg ut av denne tilstanden gjør det selv, og ikke med hjelp av NAV eller andre tjenester. Hvis det skulle vært slik at arbeidsfremmende tiltak eller lignende ordninger var den måten flest av de som var NEET kom seg tilbake i jobb på, skulle endringen i andelen på tiltak fra 2005 til 2006 vært større. Likevel går flesteparten av de som var NEET i 2005 og skiftet tilstand til 2006 over i jobb, mens andelen som gikk over på tiltak reduseres noe fra året før.

Bø og Vigran (2014) finner at 77 % av de som var NEET i 2008 var sysselsatt i 2012. Jeg finner omtrent det samme, rundt 75 % av de som var NEET i 2004 var i arbeid i 2008. I motsetning til meg finner de at i overkant av 10 % var i utdanning fire år senere. For mine data var nærmest ingen i utdanning fire år senere. Dette kan komme av at jeg ser på andre år enn Bø og Vigran (2014) eller av andre faktorer. Det kan også komme av de ser på flere aldersgrupper enn meg, det kan f.eks. være slik at de som er yngre enn 23 år og er NEET i større grad tar utdanning senere. Det kan også komme av arbeidsmarkedet etter 2008 har blitt vanskeligere for de yngre, på grunn av finanskrisen, og at de yngre i større grad velger å ta høyere utdanning. Dette stemmer i så fall overens med hva Barth og von Simson (2012) finner, at studenter en mobil arbeidsstyrke som trekker seg inn i arbeidsmarkedet i gode tider, og ut av det i dårlige tider.

## 6 Økonometriske metoder

Dette kapitlet kommer til å presentere relevante økonometriske metoder for problemstillingen i denne oppgaven. Jeg kommer i hovedsak til å basere meg på Verbeek (2012) og Wooldridge (2012).

### 6.1 Paneldata og regresjon

Paneldata baserer seg på at vi følger samme individ, firma e.l. over tid, hvor disse individene eller firmaene trekkes tilfeldig fra en populasjon på et gitt tidspunkt. Siden vi kan følge samme individet over tid, kan man i følge Verbeek (2012) ved bruk av paneldata estimere mer kompliserte og realistiske modeller enn ved tidsseriedata og tverrsnittsdata. Paneldata gjør det også mulig å kontrollere for variabler som man ikke kan observere eller måle, slik som evner, motivasjon og lignende. Dette kalles gjerne for uobservert heterogenitet; i mitt tilfelle kaller jeg det for uobserverte individeffekter. Hvordan vi kontrollerer for uobserverte individeffekter, diskuteres senere i kapitlet.

#### 6.1.1 Lineære paneldatamodeller

En standard regresjonsmodell med individdimensjon  $i$  og tidsdimensjon  $t$  kan skrives som

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \dots + \beta_k x_{itk} + \epsilon_{it} = \beta_0 + \mathbf{x}'_{it} \boldsymbol{\beta} + \epsilon_{it}. \quad (1)$$

hvor  $y_{it}$  er fenomenet som skal forklares, og  $\mathbf{x}'_{it}$  er den transponerte vektoren av en rekke forskjellige forklaringsvariabler,  $\boldsymbol{\beta}$  er en vektor av parametere som beskriver forklaringsvariablene sitt bidrag til  $y_{it}$  og  $\epsilon_{it}$  er det stokastiske feilleddet. For å finne  $\hat{\boldsymbol{\beta}}$ , kan vi bruke Minste Kvadraters Metode (MKM). Denne metoden minimerer de kvadrerte avvikene mellom predikert verdi og observert verdi. MKM-estimatoren kan skrives som

$$\hat{\boldsymbol{\beta}}_{\text{MKM}} = \frac{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})(y_{it} - \bar{y})}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})^2}. \quad (2)$$

---

$$\bar{x} = \frac{1}{NT} \sum_{n=1}^N \sum_{t=1}^T x_{it} \text{ og } \bar{y} = \frac{1}{NT} \sum_{n=1}^N \sum_{t=1}^T y_{it}.$$

Estimatoren over består av forholdet mellom samvariasjonen mellom forklaringsvariabel og avhengig variabel og variansen til forklaringsvariabelen.  $\hat{\beta}$  er et mål på effekten forklaringsvariabel  $x$  har på avhengig variabel  $y$ .

Vi antar også at

$$E\{\epsilon_{it}\} = 0 \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (3a)$$

$$E\{\epsilon_{it} | \mathbf{x}_{it}\} = 0 \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (3b)$$

$$\text{Var}\{\epsilon_{it}\} = \sigma^2 \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (3c)$$

$$\text{cov}\{\epsilon_{it}, \epsilon_{js}\} = 0 \quad i, j = 1, \dots, N, \quad t, s = 1, \dots, T \quad t \neq s \quad (3d)$$

Hvis antagelsene over holder, kan det vises at  $\beta_{\text{MKM}}$  vil være den beste lineære forventningsrette estimatoren (BLUE, engelsk: Best Linear Unbiased Estimator).

For paneldata gjelder antagelsene (3a) til (3c) fortsatt. Wooldridge (2012, s. 433) påpeker at det kan hende at vi bryter antagelse (3d) om uavhengighet, siden en observasjon på et individ på et tidspunkt kan påvirke størrelsen på observasjonen til samme individ på et annet tidspunkt. Dette fører til at estimatorene blir ineffisient med MKM på paneldata. En vanlig metode for paneldata er å inkludere den uobserverte individeffekten som et ledd i det stokastiske feilleddet

$$\epsilon_{it} = \alpha_i + u_{it} \quad u_{it} \sim \text{NID}(0, \sigma_u^2), \quad (4)$$

hvor variabelen  $\alpha_i$  fanger uobserverte individeffekter som er konstant over tid og som påvirker  $y_{it}$ . Vi får da modellen

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + \alpha_i + u_{it}. \quad (5)$$

Minste kvadraters metode krever at

$$E\{\mathbf{x}'_{it}\epsilon_{it}\} = 0. \quad (6)$$

Denne antagelsen regnes som ganske streng, og man kan derfor bruke den mindre strenge antagelsen om at den uobserverte individeffekten  $\alpha_i$  er korrelert med vektoren av forkla-

ringsvariabler  $\mathbf{x}_{it}$ . Matematisk kan vi skrive dette som

$$E\{\mathbf{x}'_{it}\alpha_i\} \neq 0. \quad (7)$$

Hvilken antagelser vi gjør om  $\alpha_i$  resulterer i forskjellige modeller. *Fast Effekt*-modellen tillater at den individspesifikke effekten kan korrelere med forklaringsvariablene, men det er ingen antagelse om at den individspesifikke effekten *må* korrelere med forklaringsvariablene. Fast-effekt-modellen modellerer  $\alpha_i$  som en deterministisk komponent, og ikke som en del av feilledet. Hvis vi antar at  $\alpha_i$  er stokastisk og «oppfører» seg som  $u_{it}$  ender vi opp med *Tilfeldig Effekt*-modellen.

### 6.1.2 Fast Effekt-modeller

Fast Effekt-modeller (FE-modeller, engelsk: Fixed Effect) fokuserer bare på variasjon over tid hos individet og ikke på variasjon på tvers av individer. FE-modellen transformerer vekk  $\alpha_i$  fra modellen ved å trekke fra individgjennomsnittene. Når vi gjør dette, fjerner vi også effekten av alle tidskonstante observerbare variabler. Dette er problematisk fordi flere av forklaringsvariablene som er viktig for analysen min, er tidskonstant. Hvis vi kan anta at alle forklaringsvariablene er strengt eksogene, altså at det ikke er noen samvariasjon mellom feilledet  $u_{it}$  og  $x_{it}$ , så vil FE-estimatoren være forventningsrett og konsistent.

### 6.1.3 Tilfeldig effekt-modeller

Tilfeldig Effekt-modeller (RE-modell, engelsk: Random Effect) antar at den uobserverte individeffekten er tilfeldig fordelt på tvers av individer. Matematisk blir dette at ligning 4 også inneholder en antagelse om at  $\alpha_i \sim \text{NID}(0, \sigma_\alpha^2)$ . Denne modellen inkluderer også variasjon over tid hos individene, men benytter seg også av variasjon mellom individene. Dette betyr at vi kan se hvilken effekt tidskonstante variabler har på den avhengige variabelen. Antagelsene for RE-modeller er mer restriktive enn for FE-modeller, siden vi antar  $\alpha_i$  ikke samvarierer med noen av forklaringsvariablene over alle tidsperiodene. RE-estimatoren er forventningsrett og konsistent hvis vi kan anta at det ikke er noen samvariasjon mellom den uobserverte individeffekten og alle forklaringsvariablene i tillegg til at det ikke er noen samvariasjon mellom det stokastiske feilledet  $u_{it}$  og alle forklaringsvariabler. En RE-modell

er mer effisient enn FE-modellen hvis antagelsene over holder, siden vi nå både utnytter informasjon over tid hos individene i tillegg til informasjon på tvers av individer.

## 6.2 Korrelert tilfeldig effekt-modeller

Diskusjonen her følger Wooldridge (2012, s. 479) og Hirano (2010). Korrelert tilfeldig effekt-modeller (CRE-modeller, engelsk Correlated Random Effects approach) er en mindre restriktiv modell enn «vanlige» RE-modeller. Dette fordi den i stedet for å anta at det ikke er noen korrelasjon mellom den uobserverte individeffekten og forklaringsvariablene, så forsøker den å modellere korrelasjon mellom den uobserverte individeffekten og forklaringsvariablene. Dette gjør den gjennom å inkludere individgjennomsnittene for å kontrollere for nivåforskjeller mellom individene. Vi antar det er lineær sammenheng mellom  $\alpha_i$  og gruppegjennomsnittet av en forklaringsvariabel  $x_{itk}$ . Matematisk blir dette

$$\alpha_i = a + \gamma \bar{\mathbf{x}}_i' + r_i, \quad (8)$$

hvor vi antar at  $r_i$  og  $\mathbf{x}_{it}$  er ukorrelert, hvor  $\bar{x}_{ik} = \frac{1}{T} \sum_{t=1}^T x_{itk}$ . Dette fører igjen til at

$$E\{\bar{x}_{ik}|r_i\} = 0 \quad \forall i, k, \quad (9)$$

dette fordi at  $\bar{x}_{ik}$  er en lineær funksjon av  $x_{itk}$ . Modellen blir da

$$y_{it} = a + \mathbf{x}_{it}'\boldsymbol{\beta} + \bar{\mathbf{x}}_i'\boldsymbol{\gamma} + r_i + u_{it}. \quad (10)$$

Vi ser at ligningen over ligner på RE-modellen med unntak fra  $\bar{\mathbf{x}}_i'$  som er vektoren bestående av gjennomsnittsnivået til forklaringsvariablene og  $\boldsymbol{\gamma}$  er vektoren bestående av tilhørende regresjonsparametere. Ligning 10 gir RE-estimatoren hvis  $\boldsymbol{\gamma}$  er null. Antagelsen om at det ikke er samvariasjon mellom den uobserverte individeffekten og alle forklaringsvariablene holder hvis vi bytter  $\alpha_i$  med  $r_i$ , hvor  $r_i$  er et tidskonstant uobservert ledd. Vi antar også at det ikke er noen korrelasjon mellom det stokastiske feilleddet  $u_{it}$  og alle forklaringsvariablene. Dette fører i essens til at vi kan betrakte ligningen over som en RE-modell. Men et viktig spørsmål er da hvordan  $\hat{\beta}_{\text{CRE}}$  ser ut. Wooldridge (2010, s. 273) argumenterer for at  $\hat{\beta}_{\text{CRE}} = \hat{\beta}_{\text{FE}}$ . En av fordelene med å bruke denne modellen fremfor å bruk RE-modellen, er at vi slipper å gjøre noen antagelser om  $\alpha_i$ . Hvis det er samvariasjon mellom de uobserverte individeffektene og forklaringsvariablene, får vi ikke forventningskjevne eller inkonsistente

estimatorer. Det virker ikke urimelig at f.eks. evner påvirker både sannsynligheten for å falle fra vgo eller å være NEET. Antagelse 17i (i vedlegget) virker kanskje derfor urimelig. En annen fordel med CRE-modeller mot FE-modeller er at vi kan inkludere tidskonstante forklaringsvariabler, dette er en fordel i min analyse siden mange av forklaringsvariablene er tidskonstante.

Wooldridge (2012, s. 480) argumenterer for at vi ved å bruke CRE-modeller gir oss en måte å velge mellom RE-modeller og FE-modeller. Wooldridge (2013) påpeker at Hausman-testen gir feil antall frihetsgrader når man bruker CRE-modellen. Som nevnt gir blir CRE-estimatoren lik RE-estimatoren om  $\gamma$  er lik null. Ved estimering av ligning 10 får vi  $\hat{\gamma}_{\text{CRE}}$  og tilhørende standardfeil. Vi har da hypotesesituasjonen

$$H_0 : \gamma = 0 \qquad H_a : \gamma \neq 0.$$

Testobservatoren er en  $t$ -test

$$T = \frac{\hat{\gamma}_{\text{CRE}} - 0}{\text{SE}\{\hat{\gamma}_{\text{CRE}}\}} \sim t(N - K). \qquad (11)$$

Under nullhypotesen er testobservatoren  $t$ -fordelt med  $N - K$  frihetsgrader. Det bør brukes cluster-robuste standardfeil (Wooldridge, 2012). Videre i oppgaven kaller jeg denne testen for den regresjonsbaserte FE-testen.

### 6.3 Valg av modell

For min økonometriske analyse kommer jeg til å benytte meg av lineære sannsynlighetsmodeller estimert med minste kvadraters metode (LPM, engelsk: Linear Probability Model), tilfeldig effekt-LPM (RE-LPM) og korrelert tilfeldig effekt-LPM (CRE-LPM). Det kan settes spørsmål ved om antagelsene for de LPM og RE-LPM holder. Som tidligere nevnt virker det ikke urimelig at uobserverte individspesifikke faktorer som evner og motivasjon påvirker både forklaringsvariablene og avhengig variabel. LPM- og RE-estimatorene vil da bli forventningskjevne og inkonsistente. En fordel ved å bruke LPM og RE-modeller er at man kan estimere effekten tidskonstante variabler har på sannsynligheten for å være NEET, og som tidligere nevnt er mange interessante forklaringsvariabler tidskonstante.

CRE-LPM blir et kompromiss mellom RE- og FE-modeller, og kan brukes til å sammenligne RE-modeller med FE-modeller. Testen nevnt i forrige delkapittel kan brukes for å teste om RE- eller FE-modellen er den korrekte modellen.



## 7 Regresjonsanalyse

Frem til nå har jeg bare sett deskriptivt på NEET-fenomenet. En del av problemstillingen er hvilke faktorer som disponerer for NEET-tilstanden. Fordelen ved å benytte seg av regresjonsanalyse er at vi, ideelt sett, kan identifisere årsakssammenhenger. I tillegg kan vi se om faktorer systematisk påvirker utfallet, og at påvirkningen, eller sammenhenger, ikke er tilfeldig. Multippel regresjon har den fordelen at vi kan måle partielle effekter, som er effekten av en forklaringsvariabel kontrollert for variasjonen fra de andre forklaringsvariablene. For min problemstilling og mine data er kan det være vanskelig å avdekke årsakssammenhenger, fordi forklaringsvariablene ikke er manipulert gjennom et eksperimentelt design, eller at datamaterialet ikke er resultat av et naturlig eksperiment. Resultatene fra regresjonsanalysen viser derfor ikke rene kausale sammenhenger, men heller hvilken samvariasjon det er mellom egne, foreldres og myndighetenes valg og sannsynligheten for å falle innenfor oppgavens definisjon av NEET.

Jeg kommer også til å skille mellom *økonomisk signifikans* og *statistisk signifikans*, hvor statistisk signifikans bestemmes ved en hypotesetest. Økonomisk signifikans, og kalt praktisk signifikans, bestemmes av størrelsen og tegnet til koeffisienten. Wooldridge (2012, s. 128) argumenterer for at hvis man bare fokuserer på statistisk signifikans, kan dette føre til at vi feilaktig konkluderer at en forklaringsvariabel er viktig selv om effekten er lav. Hvilken grense man skal sette for hva som er økonomisk signifikant avhenger av problemstilling og enheten til forklaringsvariabelen. Jeg har valgt at hvis risikoen endrer seg med mindre enn ett promillepoeng, alt annet like, så er effekten ikke økonomisk signifikant.

Jeg kommer til å benytte meg av et signifikansnivå på én prosent. Dette kommer av at desto større utvalgene er, desto mindre blir standardfeilene og desto større sjanse er det å forkaste nullhypotesen. Dette betyr at sjansen for *Type II-feil*<sup>1</sup> er liten i store utvalg. Ved å velge et lavere signifikansnivå reduseres også den *statistiske styrken*<sup>2</sup> til hypotesetesten (Verbeek, 2012, s. 31). Vi ser likevel at de fleste effektene er statistisk signifikante.

---

<sup>1</sup>Type II-feil: Sannsynligheten for å beholde nullhypotesen, når den alternative hypotesen er sann.

<sup>2</sup>Statistisk styrke: Sannsynligheten for å forkaste nullhypotesen, når den alternative hypotesen er sann.

<i>Variabelnavn</i>	<i>Ligning</i>	<i>Definisjon</i>
NEET	12, 13, 14	Dummy-variabel definert i boks 2.6 (1=NEET).
Fullført allmennfag	12, 13, 14	indikatorvariabel (1=allmennfag, 2=yrkesfag. Referansegruppe er frafall (0=fracfall).
Fullført yrkesfag	12, 13, 14	indikatorvariabel (1=allmennfag, 2=yrkesfag. Referansegruppe er frafall (0=fracfall).
Arbeidssøker året før	12, 13, 14	Dummy-variabel om person er arbeidssøker det foregående året. Referansegruppe er ikke arbeidssøkende
Arbeidssøker to år før	12, 13, 14	Dummy-variabel om person er arbeidssøker de foregående to årene. Referansegruppe er ikke arbeidssøkende
Arbeidssøker tre år før	12, 13, 14	Dummy-variabel om person er arbeidssøker de foregående 3 årene. Referansegruppe er ikke arbeidssøkende
Fikk sosialhjelp året før	12, 13, 14	Dummy-variabel som indikerer om en person fikk sosialhjelp det foregående året (1=mottatt sosialhjelp).
Fikk sosialhjelp to år før	12, 13, 14	Dummy-variabel som indikerer om en person fikk sosialhjelp de to foregående årene (1=mottatt sosialhjelp).
Fikk sosialhjelp 3 år før	12, 13, 14	Dummy-variabel som indikerer om en person fikk sosialhjelp de 3 foregående årene (1=mottatt sosialhjelp).
Passiv stønadsmottaker året før	12, 13, 14	Faktorvariabel (=1, 2=aktiv tiltaksdeltager) referansegruppe er de som ikke fikk noen stønad.
Aktiv tiltaksdeltager året før	12, 13, 14	Faktorvariabel (=2, 1=passiv stønadsmottaker) Referansegruppe er de som ikke fikk noen stønad.
Passiv stønadsmottaker to år før	12, 13, 14	Faktorvariabel, sum av de to foregående årene. (=1, 2=aktiv tiltaksdeltager) Referansegruppe er de som ikke fikk noen stønad.
Aktiv tiltaksdeltager to år før	12, 13, 14	Faktorvariabel, sum av de to foregående årene. (=2, 1=passiv stønadsmottaker) Referansegruppe er de som ikke fikk noen stønad.
Passiv stønadsmottaker 3 år før	12, 13, 14	Faktorvariabel, sum av de 3 foregående årene. (=1, 2=aktiv tiltaksdeltager) Referansegruppe er de som ikke fikk noen stønad.
Aktiv tiltaksdeltager 3 år før	12, 13, 14	Faktorvariabel, sum av de 3 foregående årene. (=2, 1=passiv stønadsmottaker) Referansegruppe er de som ikke fikk noen stønad.
Mann=1, kvinne=0	12, 13, 14	Dummy-variabel som indikerer kjønn.
Forsørget av én forelder	12, 13, 14	Dummy-variabel som indikerer om en person er registrert med én eller to foreldre (1= én forelder, 0= to foreldre).
Antall søsken	12, 13, 14	Variabel som indikerer det maksimale antall søsken under 18 år (0=referanseverdi).
Gj.snitt antall ganger arbeidssøkende	14	Gjennomsnittlig antall ganger som arbeidssøkende pr. person.
Gj.snitt antall ganger mottatt sosialhjelp	14	Gjennomsnittlig antall ganger mottatt sosialhjelp pr. person.
Gj.snitt antall ganger på tiltak	14	Gjennomsnittlig antall ganger på tiltak pr. person.

Tabell 4: Oversikt over forklaringsvariabler i regresjonsanalysen.

## 7.1 Sannsynligheten for å være NEET

Jeg kommer til å benytte meg av henholdsvis tre typer lineære sannsynlighetsmodeller. To av disse modellene kontrollerer for uobserverte individeffekter. Disse er tilfeldig effekt- og korrelert tilfeldig effekt-modeller. Minste kvadraters metode (MKM) kontrollerer ikke for uobserverte individeffekter. Det er av interesse å se om det å kontrollere for disse effektene fører til annerledes resultater.

$$P(\text{NEET})_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}'_{i(t-s)}\boldsymbol{\gamma} + \mathbf{z}'_i\boldsymbol{\delta} + \epsilon_{it} \quad (12)$$

Ligning 12 viser MKM-modellen for sannsynligheten for individ  $i$  ved tid  $t$  å være NEET.  $\mathbf{x}'_{it}$  er vektoren bestående av tidsvariende forklaringsvariabler,  $\mathbf{z}'_i$  er en vektor bestående av tidskonstante forklaringsvariabler og  $\mathbf{x}'_{i(t-s)}$  er vektoren bestående av «laggede» tidsvarierende forklaringsvariabler. Jeg introduserer aggregert historikk for noen forklaringsvariabler, fordi det kan være interessant å se hvordan forløp som sosial- eller trygdebruker påvirker NEET-risikoen.  $\boldsymbol{\beta}$ ,  $\boldsymbol{\delta}$  og  $\boldsymbol{\gamma}$  er vektorer bestående av regresjonsparametere for forklaringsvariablene. Variablene som er tidsvarierende er om personen er arbeidssøker, om en person har mottatt sosialhjelp eller om personen er passiv stønadsmottaker eller aktiv tiltaksdeltaker. De tidskonstante variablene relaterer seg til utdanning, f.eks. frafall/fullføring, altså de som fullfører vgo og hvilken linje de fullførte. Andre tidskonstante variabler er antall søsken, om en person er forsørget av én forelder og kjønn. Jeg inkluderer også årsvariabler og aldersvariabler for å kontrollere for at endringer i andre forklaringsvariabler ikke kommer fra endring i alder eller år. Lineære sannsynlighetsmodeller estimert ved MKM har heller ikke homoskedastiske feilledd, slik som vist i kapittel A.2.2.1. Hypotesetesting av regresjonskoeffisientene ved bruk av MKM er ikke valid når man har heteroskedastiske feilledd (Wooldridge, 2012, s. 259), og estimatoren er ikke lengre den mest effisiente, fordi det kan finnes andre estimatører med lavere varians. Bruk av robuste standardfeil vil redusere problemet med heteroskedastisitet (White, 1980).

Som vi så i den deskriptive delen, endret andelen som er NEET seg veldig med alder. Ved å studere hvordan tilstanden endrer seg over tid for individet, kan jeg utnytte mer informasjon enn ved å se på en spesifikk aldersgruppe. Ved å se på mer enn en aldersgruppe kan jeg også utnytte paneldatastrukturen i datamaterialet mitt og forhåpentligvis estimere mer realistiske modeller enn ved tversnittsanalyser (Verbeek, 2012). Paneldatamodeller kan

benytte seg av variasjon over tid hos individ, så vel som variasjon mellom individ. Estimering av regresjonene i kolonne (1) og (2) i tabell 5 baserer seg på estimeringsligning 12 for lineære sannsynlighetsmodeller. For Tilfeldig Effekt-modellen (RE) bruker jeg estimeringsligning

$$P(\text{NEET})_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}'_{i(t-s)}\boldsymbol{\gamma} + \mathbf{z}'_i\boldsymbol{\delta} + \alpha_i + u_{it}. \quad (13)$$

Ligning 13 ligner på ligning 12. Eneste forskjellen er at ligningen over inkluderer den uobserverte individsesifikke komponenten  $\alpha_i$ . Denne individsesifikke komponenten fanger opp alle uobserverte faktorer som er konstant over tid.

Estimeringsligningen for Korrelert tilfeldig effekt-modellen (CRE) blir

$$P(\text{NEET})_{it} = a + \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{x}'_{i(t-s)}\boldsymbol{\gamma} + \bar{\mathbf{x}}'_i\boldsymbol{\lambda} + \mathbf{z}'_i\boldsymbol{\delta} + r_i + u_{it}. \quad (14)$$

Ligning 14 ligner på ligning 13, bortsett fra  $\bar{\mathbf{x}}'_i$  som er gjennomsnittet over tid til de tidsvarierende forklaringsvariablene.  $\boldsymbol{\beta}$  blir i denne ligningen lik estimatoren vi hadde fått om vi hadde brukt en FE-modell.

Det kan derimot settes spørsmåltegn ved om forutsetningene for tilfeldig effekt-modellen holder. Forutsetningen om at det ikke skal være noen kovariasjon mellom de uobserverte individeffektene og forklaringsvariablene er en av forutsetningene som virker urimelig. Dette av grunner som er nevnt over. Tilfeldig effekt-estimatoren blir da inkonsistent. Korrelert tilfeldig effekt-modellene har ikke like restriktive antagelser som RE-modellen, førstnevnte modell har også en fordel over fast effekt-modellen, og at den inkluderer variasjon mellom individ.

Jeg har nevnt at jeg ser på aggregert historikk for noen av forklaringsvariablene. Det betyr at forklaringsvariabel *Arbeidssøker to år før* betyr at personen har vært arbeidsledig/arbeidssøkende de to foregående årene. *Fikk sosialhjelp to år før* betyr mottatt sosialhjelp de to foregående årene. Det samme gjelder forklaringsvariablene *passiv stønadsmottaker* og *aktiv tiltaksdeltaker*. Alle tolkninger av koeffisientene er gjennomsnittlig endring i NEET-risiko når vi endrer forklaringsvariabelen fra 0 til 1, alt annet like.

Tabell 5: Lineære sannsynlighetsmodeller for sannsynligheten for å være i tilstanden NEET.

	LPM (23år)	LPM (2)	RE (3)	CRE (4)
	$\beta$ / SE	$\beta$ / SE	$\beta$ / SE	$\beta$ / SE
<i>Utdanning</i>				
Fullført allmennfag	-0.06930*** (0.001563)	-0.06577*** (0.000566)	-0.08859*** (0.000990)	-0.02772*** (0.000957)
Fullført yrkesfag	-0.08073*** (0.001572)	-0.08039*** (0.000567)	-0.10268*** (0.000975)	-0.04850*** (0.000929)
<i>Velferdsordninger</i>				
Arbeidssøker året før	0.06342*** (0.002243)	0.06960*** (0.000861)	0.05398*** (0.000791)	0.04226*** (0.000797)
Arbeidssøker to år før	0.01863*** (0.003918)	0.02249*** (0.001591)	0.01329*** (0.001365)	0.00829*** (0.001362)
Arbeidssøker 3 år før	0.01935*** (0.005246)	0.04222*** (0.002126)	0.02492*** (0.002033)	0.01487*** (0.002032)
Fikk sosialhjelp året før	0.14555*** (0.005513)	0.16383*** (0.002062)	0.09589*** (0.001896)	0.04258*** (0.001928)
Fikk sosialhjelp to år før	0.06645*** (0.009368)	0.07637*** (0.003390)	0.06118*** (0.002867)	0.03411*** (0.002877)
Fikk sosialhjelp 3 år før	0.12933*** (0.009201)	0.11908*** (0.003378)	0.08939*** (0.003292)	0.03179*** (0.003300)
Passiv stønadsmottaker året før	0.21179*** (0.008394)	0.21487*** (0.002972)	0.13877*** (0.002732)	0.11568*** (0.002750)
Aktiv tiltaksdeltager året før	0.13399*** (0.007960)	0.15077*** (0.002889)	0.12620*** (0.002773)	0.10027*** (0.002785)
Passiv stønadsmottaker to år før	0.07684*** (0.015588)	0.04111*** (0.005213)	0.04068*** (0.004272)	0.03199*** (0.004243)
Aktiv tiltaksdeltager to år før	0.04933** (0.020914)	0.03726*** (0.007747)	0.05668*** (0.007211)	0.04985*** (0.007190)
Passiv stønadsmottaker 3 år før	0.11938*** (0.020809)	0.08843*** (0.006862)	0.07786*** (0.006700)	0.06476*** (0.006683)
Aktiv tiltaksdeltager 3 år før	0.15707*** (0.033886)	0.12457*** (0.015560)	0.11321*** (0.014018)	0.10743*** (0.013924)
<i>Demografi</i>				
Mann=1, kvinne=0	-0.01310*** (0.001004)	-0.00957*** (0.000370)	-0.00848*** (0.000621)	-0.00730*** (0.000590)
Forsørget av én forelder	0.05507*** (0.004087)	0.04359*** (0.001089)	0.05818*** (0.001928)	0.01797*** (0.001809)
Antall søsken=1	-0.00698*** (0.001214)	-0.00671*** (0.000470)	-0.00748*** (0.000783)	-0.00778*** (0.000746)
Antall søsken=2	-0.00924*** (0.001544)	-0.01059*** (0.000618)	-0.01105*** (0.001010)	-0.01218*** (0.000968)
Antall søsken=3	-0.01142*** (0.002646)	-0.01015*** (0.001095)	-0.00934*** (0.001763)	-0.01317*** (0.001681)
Observasjoner	296673	2385483	2385483	2385483
Antall personer			330351	330351
adjusted $R^2$	0.158	0.150	0.145	0.177
$\rho$			0.256	0.254
Gjennomsnittlig predikert sannsynlighet	0.092	0.099	0.099	0.100
LR $\chi^2$	54		72814.534	101820.867
Prob > $\chi^2$	0.000	0.000	0.000	0.000

\* p &lt; 0.1, \*\* p &lt; 0.05, \*\*\* p &lt; 0.01. Robuste standardfeil i parentes.

*Merknad:* Inkludert årsdummyer og aldersdummyer, se tabell 8 i vedlegget for alle regresjonsresultatene.

Kolonne (1) og (2) er Minste Kvadraters lineære sannsynlighetsmodeller (LPM), mens kolonne (3) og (4) er henholdsvis tilfeldig effekt-LPM og korrelert tilfeldig effekt-LPM. Jeg kommer til å kalle kolonne (1) for modell (1), kolonne (2) for modell (2) osv. For modell (1) ser jeg på NEET-risikoen syv år etter avsluttet ungdomsskole, mens for modell (2), (3) og (4) ser jeg på NEET-risikoen for alle aldersgrupper.

Vi ser at antall observasjoner er lavere for modell (1). Dette kommer av at jeg NEET-risikoen syv år etter avsluttet ungdomsskole, når personene er rundt 23 år gamle. Antallet blir derfor lavere enn når jeg ser på alle aldersgrupper.

Generelt ligger forklaringskraften på mellom 15 % og 18 %. Altså forklares mellom 15 % og 18 % av variasjonen i NEET av variasjonen i modellen. Modellen med høyest forklaringskraft er modell (4) som er korrelert tilfeldig effekt-modellen. Modellen med lavest forklaringskraft er tilfeldig effekt-modellen (modell (3)).

### 7.1.1 Utdanning

For utdanningsvariabelen er det sterke negative effekter av det å ha fullført vgo mot det å ha falt fra. For de som fullførte allmennfag var det 6,9 prosentpoeng (modell (1)) NEET-risiko mot de som falt fra. Effekten er sterkest for modell (3). Reduksjonen i NEET-risiko er sterkere for de som fullførte yrkesfag. Her varierer effektene mellom 10,2 og 4,8 prosentpoeng redusert NEET-risiko for de som fullførte yrkesfag mot de som falt fra. Det kan virke som om det er vanskeligere å komme seg inn i arbeidsmarkedet for de som har fullført allmennfag enn for yrkesfag. Dette kan komme av at de som har gått på yrkesfag kommer i kontakt med en potensiell arbeidsgiver i læretiden sin og derfor har tidligere hatt kontakt med arbeidsmarkedet enn de som går på allmennfag. En mulig feilkilde kan være de som velger å ta seg noen år fri etter vgo til å jobbe frivillig eller reise o.l, i større grad kommer fra allmennfaglig studieretning. Denne gruppen er, som tidligere diskutert, ikke en NEET-gruppe.

### 7.1.2 Velferdsordninger

Når vi ser på historikk som arbeidssøkende er alle koeffisientene statistisk og økonomisk signifikant, for alle modellene. Det ser ut til å være økt NEET-risiko assosiert med det å

ha et forløp som arbeidssøkende. Effekten er sterkest for året før og svekkes noe for de to foregående årene, for så å styrkes noe for de tre foregående årene. Effektene svekkes noe når vi kontrollerer for uobserverte individeffekter. En mulig grunn til at vi ser en økt NEET-risiko for de som har et forløp som arbeidsledig/arbeidssøkende, kan være at det en «scarring»-effekt assosiert med det å være arbeidsledig (Arulampalam et al., 2001).

Det å ha hatt en historikk med sosialhjelp er assosiert med en økt NEET-risiko. For modell (1) er det rundt 14 prosentpoeng økt NEET-risiko hvis personen har mottatt sosialhjelp året før. Det gir mening at personer som har hatt behov for sosialhjelp oftere har en så perifer arbeidstilknytning over tid at de fanges opp av oppgavens definisjon av NEET. Effektene av det å ha mottatt sosialhjelp de tre foregående årene styrkes relativt til det å ha mottatt sosialhjelp de to foregående årene.

For det å ha en historikk som aktiv tiltaksdeltager eller passiv stønadsmottaker er generelt økt NEET-risiko assosiert med å være i disse tilstandene. NEET-risikoen assosiert med en historikk som passiv stønadsmottaker er større enn for de som ikke mottok noen stønad. Det å være passiv stønadsmottaker året før øker NEET-risikoen med rundt 21 prosentpoeng (modell (1) og (2)). Det gir god mening at dette er tilfellet, siden de fleste av de som er NEET er det i to år, hvis man da er passiv stønadsmottaker et år er man det gjerne året etter også. Vi ser at effektene svekkes noe når vi kontrollerer for uobserverte individeffekter. En mulig grunn til dette kan være at evner og motivasjon spiller en rolle, og at vi derfor overestimerer effekten av det å være passiv stønadsmottager ved å ikke kontrollere for disse uobserverte individeffektene. Effekten svekkes noe for de som var passive stønadsmottagere de to foregående årene mot det å være det året før eller de tre foregående årene. For modell (1) er effekten av å være passiv stønadsmottager de to foregående årene ikke statistisk signifikant.

NEET-risikoen assosiert med det å være aktiv tiltaksdeltager er større enn for de som ikke de ikke mottok atferds- og rehabiliteringspenger. Det ser ut til å være generelt høyere NEET-risiko for de var passive stønadsmottakere mot det å ikke motta noen støtte, enn for de var aktive tiltaksdeltakere. Det at det er generelt høyere NEET-risiko for de som har hatt behov for hjelp av det offentlige for å komme seg i arbeid, kan tyde på at disse personene i snitt har vanskeligere for takle det å være i arbeid, enn for de som ikke mottok noen slik hjelp.

### 7.1.3 Demografiske forskjeller

For alle modellene er det høyere NEET-risiko assosiert med det å være kvinne. Effekten er sterkest syv år etter avsluttet ungdomsskole (modell (1)), hvor det 1,3 prosentpoeng lavere NEET-risiko for menn. Denne effekten reduseres når vi ser på alle aldersgrupper og kontrollerer for uobserverte individeffekter. Det er vanskelig å vite om kjønnsforskjellen i NEET-risiko kommer fra hjemmевærende mødre, som tidligere diskutert ikke er en NEET-gruppe. Det kan også tenkes at kjønnsvariabelen plukker opp effekter av lønnsforskjeller. Og at vi derfor overestimerer effekten av kjønn.

Det ser ut til å være en økt NEET-risiko assosiert med det å være bosatt hos en av foreldrene som 18-åring, mot det å være bosatt hos begge. Dette stemmer med tidligere forskning (Schoon, 2014). Det kan være at én forelder har mindre tid og ressurser til å hjelpe barna sine, og at disse barna igjen derfor har mindre sjanser til å komme seg inn i arbeidsmarkedet eller i utdanning. For modell (1) øker NEET-risikoen av det å være bosatt hos én av foreldrene mot begge med rundt 5 prosentpoeng.

Vi ser at NEET-risikoen reduseres etter jo flere søsken man har, mot det å ikke ha noen. Effekten styrkes noe desto flere søsken man har, men for modell (1) er effektene ikke statistisk signifikant forskjellig fra hverandre. Og det å ha 2 eller 3 søsken er ikke statistisk signifikant forskjellig fra hverandre for noen av modellene. I prinsippet så handler det da om forskjellen mellom å være enebarn mot det å ha én søster eller bror.

### 7.1.4 Testing av antagelser og FE-testen

Når vi kontrollerer for uobserverte individeffekter, slik som vi gjør for modell (3) og (4), endrer noen av effektene seg. Noen av effektene blir sterkere, mens andre blir svakere. Som tidligere nevnt kan det tenkes at evner og motivasjon er negativt korrelert med fullføring av vgo. Hvis vi ikke kontrollerer for disse uobserverte individeffektene underestimerer vi effekten av fullføring av vgo. Som vi ser fra modell (3) styrkes effekten av fullføring av vgo når vi kontrollerer for uobserverte individeffekter, dette kan indikere at hypotesen min om sammenhengen mellom uobserverte individeffekter stemmer. Det er likevel uklart om det er

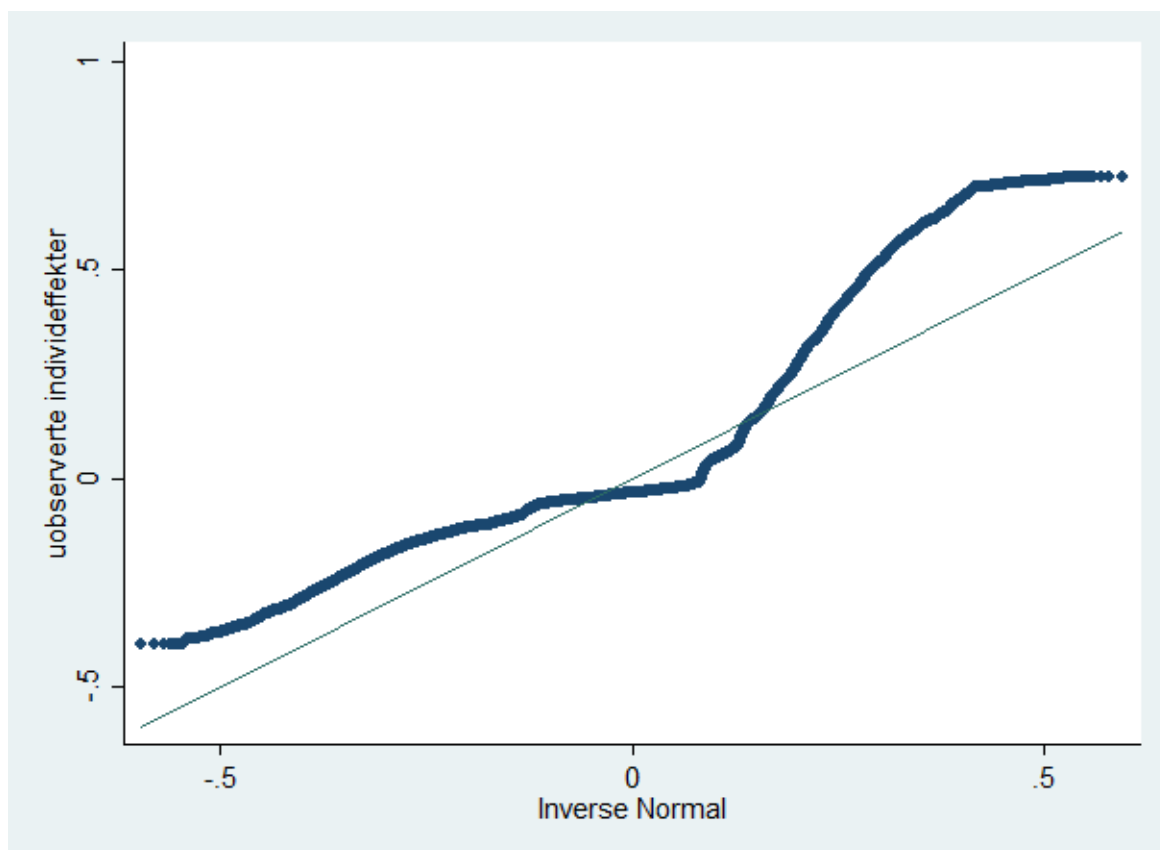


dette som bidrar, siden det sannsynligvis er flere faktorer som fanges opp av feilleddet  $\alpha_i$ .

$$\rho = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2}, \quad (15)$$

viser den andelen av den totale variansen som kommer fra  $\alpha_i$  (StataCorp., 2014). For modell (3) og (4) ligger denne andelen på rundt 25 %. Hvis denne hadde vært lik null, hadde RE-estimatoren vært MKM-estimatoren. P-verdien ( $\text{prob} > \chi^2$ ) fra likelihood ratio-testen, hvor man tester om  $\rho$  er statistisk signifikant forskjellig fra null, viser at RE-estimatoren for modell (3) og (4) er signifikant forskjellig fra MKM-estimatoren. Denne testen er i prinsippet lik Breusch Pagan-testen (Rodríguez og Elo, 2003).

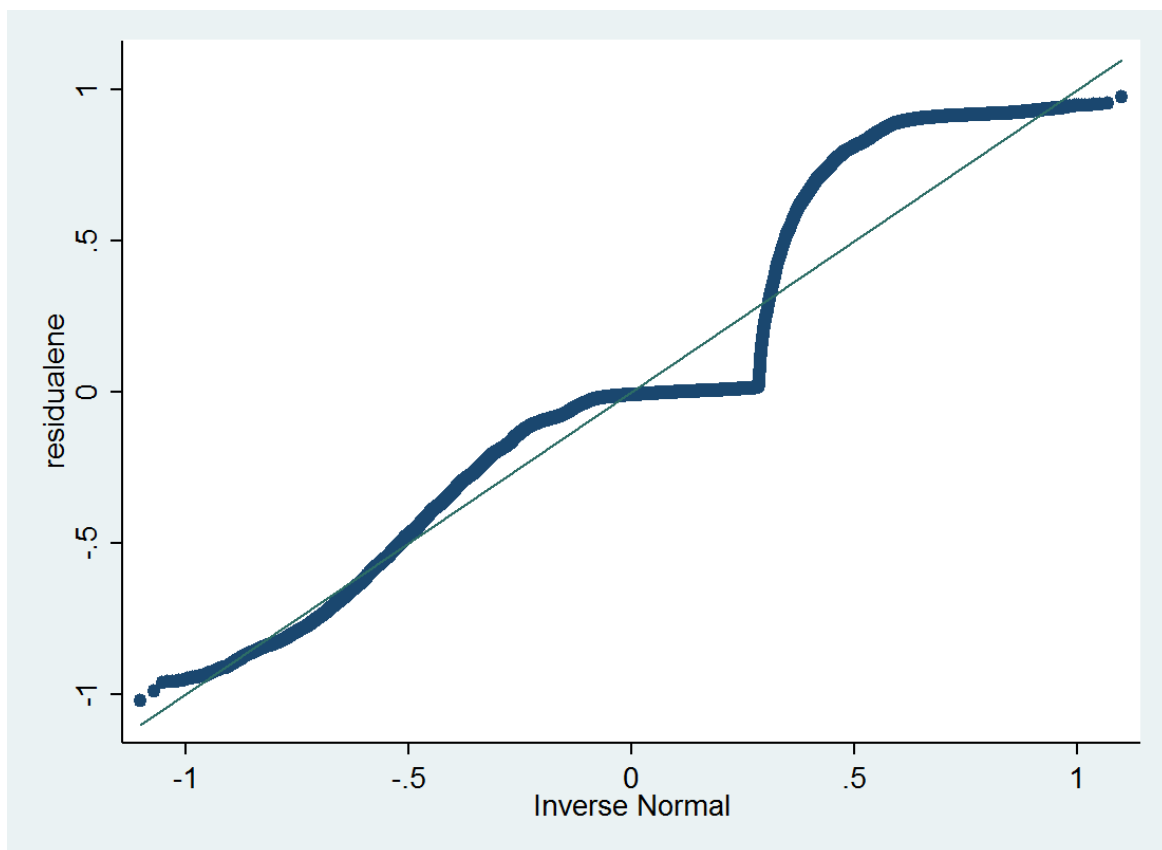
En antagelse om tilfeldig effekt-modellen var at  $\alpha_i \sim \text{NID}(0, \sigma_\alpha^2)$ . Denne antagelsen kan vi sjekke om stemmer ved hjelp av se om kvantilene til  $\alpha_i$  stemmer overens med kvantilene til normalfordelingen



Figur 12: QQ-plott over fordelingen til de uobserverte individeffektene ( $\alpha_i$ ) for modell (3).

Figur 12 viser at kvantilene til de uobserverte ikke perfekt følger en normalfordeling. Vi

ser at det er store avvik for de øverste kvantilene av  $\alpha_i$ . Dette kan tyde på de uobserverte individeffektene ikke er normalfordelt, og at residualene ikke oppfører seg på en tilfeldig måte. Hypotesetesting av RE-koeffisientene krever normalfordelte  $\alpha_i$ . Slik at det kan stilles spørsmål ved hypotesetesting av regresjonskoeffisienter for RE-modellen. Korrelert tilfeldig effekt-modellen setter ingen krav til fordelingen til  $\alpha_i$ , men det er fortsatt et krav om at residualene skal være normalfordelt. Figur 13 viser om dette er tilfelle.



Figur 13: QQ-plott over fordelingen til residualene ( $u_{it}$ ) for modell (4).

Figur 13 viser at det også kan settes spørsmål om residualene er normalfordelt for den korrelerte tilfeldig effekt-modellen. Vi ser at for lave residualer ligner kvantilene til residualene på normalfordelte kvantiler. Normalitetsantagelsen for residualene kan også trekkes i tvil for CRE-modellen, men residualene enn de uobserverte individeffektene for RE-modellen. Figur 14 i kapittel A.3 i vedlegget viser at det «totale» feilleddet til CRE-modellen er mer normalfordelt enn for RE-modellen, siden kvantilene ligger tettere innpå linjen for CRE-modellen.

Ved bruk av korrelert tilfeldig effekt-modeller (CRE) kontrollerer vi for variasjonen mellom individer ved å inkludere gjennomsnittet for de tidsvarierende forklaringsvariablene. CRE-modellen har mindre restriktive antagelser enn RE-modellen, og CRE-estimatoren er derfor konsistent og forventningsrett hvis antagelsen om uavhengighet mellom forklaringsvariablene og den uobserverte individeffekten ikke stemmer. Hvis gjennomsnittsverdien til de tidsvarierende forklaringsvariablene er statistisk signifikant, er dette en indikasjon på at FE-modellen er den korrekte spesifikasjonen. I følge Schmidheiny (2014) tester vi antagelsen om uavhengighet mellom de uobserverte individeffektene og forklaringsvariablene, når vi tester simultant om alle koeffisientene til gjennomsnittsvriablene er lik null.

Tabell 6: Regresjonsbasert FE-test

	CRE (4) $\beta$ / SE
Gjn.snitt antall ganger som arbeidssøkende	0.1301*** (0.00263)
Gjn.snitt antall ganger mottatt sosialhjelp	0.3981*** (0.00462)
Gjn.snitt antall ganger på tiltak	0.1421*** (0.00377)
Gjennomsnittsalder	-0.0077*** (0.00018)
JHT $\chi^2$	19844.083
p-verdi	0.000

\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Robuste standardfeil i parentes.

*Merknad:* JHT: Joint Hypothesis Test

Tabell 6 viser at alle koeffisientene er statistisk signifikant, både simultant (JHT  $\chi^2$ ) og individuelt. Dette tyder på at antagelsen om uavhengighet mellom de uobserverte individeffektene og forklaringsvariablene er brutt, og at RE-estimatoren er inkonsistent og forventningskjev. FE-estimatoren derfor den korrekte, og som det ble diskutert i kapittel 6.2 er  $\hat{\beta}_{CRE} = \hat{\beta}_{FE}$ . Dette indikerer at korrelert tilfeldig effekt-modellen er den modellen som sannsynligvis gir de mest korrekte størrelsene på effektene.

## 7.2 Hvor robust er resultatene?

Jeg har gjennom de to siste kapitlene diskutert forskjellige økonomiske utfordringer som kan tenkes å påvirke resultatene fra kapittel 7.1. Jeg kommer i følgende delkapittel til å diskutere hvordan valg av regresjonsmodell og endogenitet kan påvirke resultatene fra regresjonsanalysen.

### 7.2.1 Valg av modell

Pischke og Angrist (2009, s. 107) påpeker at effektene fra lineære sannsynlighetsmodeller (LPM) blir ganske lik marginaeffektene man får hvis man bruker ikke-lineære metoder, f.eks. probit og logit. Kumulativ standard normalfordeling, probit, vil f.eks. være noenlunde lineær for predikerte sannsynligheter på rundt 0,5. Hvis de predikerte sannsynlighetene ligger nærmere 0 eller 1 vil avvikene i effektstørrelsene være større. Som nevnt i kapittel A.2.2.2 sørger probit-transformasjonen for at sannsynlighetene holder seg innenfor 0 og 1, noe som man ikke er sikret mot hvis man estimerer en lineær sannsynlighetsmodell ved hjelp av minste kvadraters metode (MKM).

Alternativet til å bruke lineære sannsynlighetsmodeller blir å bruke ikke-lineære modeller, tabell 9 viser marginaeffektene fra en korrelert tilfeldig effekt probit-modell. Tabell 9 viser at de fleste marginaeffektene er kraftig redusert i forhold til modell (4) i tabell 5. Som nevnt i avsnittet over er det ikke urimelig at det er forskjeller i marginaeffektene når effektene er såpass lave, siden vi for lave effekter overestimerer effektene ved bruk av lineære sannsynlighetsmodeller.

### 7.2.2 Endogenitet

Det er sannsynligvis langt flere faktorer enn de som er tilgjengelig eller som det ikke er mulig å observere som påvirker utfallet, og modellene som jeg har kjørt greier ikke å kontrollere for uobserverte faktorer som varierer over tid. Jeg har ikke korrigert for regionale forskjeller, fordi jeg ikke har informasjon om region i mitt datasett. Bäckman et al. (2011) hevder at det er overraskende at det er høy NEET-risiko i Norge blant de som faller fra vgo. Dette kan skyldes at mange unge personer fra andre skandinaviske land flytter til Norge og tar

ufaglærte jobber. Jobber som sannsynligvis passer til personer som har falt fra vgo. Falch og Nyhus (2011) finner likevel at effekten av regionale forskjeller er små.

Det kan stilles spørsmål ved hvor gode forklaringsvariabler arbeidssøker 1 og 2 år før er. Som vi husker definisjonen fra arbeidsmarkedstilknytningsmodellen blir jobbstatus i år  $t$  bestemt ut i fra år  $t$ ,  $t - 1$  og  $t - 2$ . For NEET-kategorien «langtidsledige» var et av kravene at de skulle ha mer enn 180 dager som arbeidsledig i to av tre år. Derav kan vi få et endogenitetsproblem som følge av omvendt kausalitet mellom NEET i år  $t$  og det å være arbeidssøker i  $t - 1$  og  $t - 2$ . Tabell 10 viser resultatene når jeg ekskluderer det å være arbeidssøker i år  $t - 1$  og  $t - 2$ , og bare ser på hvilken betydning det har å være arbeidssøker i  $t - 3$  har for NEET-risikoen. Det ser da ut til å være lavere NEET-risiko for de som var arbeidssøkende tre år før. Dette kan bety at vi har et simultanitetsproblem, og MKM-estimatorene er inkonsistente og forventningsskjeve (Wooldridge, 2012). Det samme gjelder for paneldatamodellene, siden en av antagelse for disse er at det ikke skal være avhengighet over tid i  $u$ .

## 8 Oppsummering og konklusjoner

Formålet med denne oppgaven var å finne en definisjon for NEET og beskrive omfanget av dette over tid. Jeg skulle så se økonometrisk på hvilke faktorer som disponerer for NEET, i tillegg å se om NEET er en absorberende tilstand. Datamaterialet er hentet fra FD-trygd og koblet sammen med utdanningsdata fra Nasjonal Utdanningsdatabase.

Den empiriske strategien har vært å analysere hvordan frafall/fullføring i vgo kan påvirke sannsynligheten for å være NEET. Der har dessuten vært av interesse å se hvordan historikk som trygde- eller sosialmottaker kan forklare NEET-risikoen, og hvordan familiekarakteristika påvirker NEET-risikoen.

### 8.1 Hovedfunn

Jeg finner at det har skjedd en økning over tid i andelen menn som er NEET. For 23- og 24-åringer øker denne andelen med over 80 % i perioden 2000 til 2005, fra ca. 6 til 9 %. Andelen kvinner som er NEET øker også i samme periode, men veksten er ikke like sterk.

Resultatene mine indikerer at NEET er aldersavhengig, det var klart flest som var NEET når de var rundt 20 år. Omtrent 8 % av kvinner i 31-årsalderen var NEET, ca. 6 % for menn. Personer som var NEET i et gitt år hadde oftere et forløp som NEET og forblir NEET i snitt i 2 år. Flesteparten av de som var NEET kommer seg likevel tilbake i arbeid. Resultatene fra forløp- og overgangsanalysen (kapittel 5.2) indikerer også at en veldig liten andel av de som var NEET tok utdanning etter å ha vært NEET.

Alle effekter fra regresjonsanalysen er både statistisk og økonomisk signifikante. Det ser ut til å være høy NEET-risiko blant de som falt fra vgo. Vi finner også at NEET-risikoen er lavere for de som fullførte yrkesfag mot de som falt fra, enn for de som fullførte allmennfag mot de som falt fra. Når jeg har sett på andre faktorer som disponerer for NEET, finner jeg at de som har hatt en historikk med sosialhjelp har høyere NEET-risiko. Det at det er økt NEET-risiko for de som mottar sosialhjelp er ikke overraskende, siden sosialhjelp som regel bare innvilges når en person har ingen andre måter å tjene til livsopphold. Resultatene indikerer at det å tidligere ha vært arbeidsledig gir høyere NEET-risiko. Det gir også mening at antall dager som arbeidsledig gir økt NEET-risiko. Nettopp fordi de som er arbeidsledig er en av NEET-gruppene. Resultatene indikerer også at det å blitt forsørget av én forelder er assosiert høyere NEET-risiko.

## 8.2 Konklusjon: Har jeg funnet opp kruttet på nytt?

Motivasjonen for å introdusere begrepet NEET, var å tilby et samlet mål på utenforskap, noe som verken arbeidsledighets- eller frafallstatistikken kan. Spørsmålet er, når vi nå har sett på hvordan forskjellige tilstander disponerer for NEET, om vi har fått noe ny innsikt i utenforskap. En av styrkene ved NEET-begrepet er at det også inkluderer de som er helt utenfor arbeidsmarkedet. Spørsmålet er om disse oppfører seg annerledes enn de som f.eks. er arbeidsledig. I prinsippet skulle man tro at de som var økonomisk inaktive har en annen økonomisk atferd enn de som er arbeidsledig. De som er arbeidsledig skulle i prinsippet ha en større villighet til å komme seg i arbeid, men arbeidsledighet er, forhåpentligvis, i de fleste tilfeller en ufrivillig tilstand. Det kan virke som om forskjellen mellom disse to gruppene da er det å søke hjelp fra det offentlig, mot det å få økonomisk støtte hjemmefra eller andre steder.

Generelt finner andre studier, både de som studerer NEET, arbeidsledighet og frafall,

lignende effekter til meg. Slik som Falch og Nyhus (2011) og Bäckman et al. (2011) finner jeg at fullføring av vgo er assosiert med lavere NEET-risiko, mot frafall i vgo. Falch et al. (2010) finner at de som fullførte vgo i mindre grad mottok offentlige stønader eller var arbeidssøkere. Effektene av det å ha fullført vgo ser ut til å være sterkere for de som mottar offentlig stønad i forhold de som er arbeidssøkere. Falch et al. (2010) finner at de som er arbeidssøkere i større grad mottar stønad. Resultatene fra regresjonsanalysen i denne oppgaven indikerer óg at dette er tilfellet, hvor det ser ut til at personer som har en historikk med å motta ytelser fra det offentlige har høyere NEET-risiko. Alt i alt, ser det ut som om det foreligger de samme risikofaktorene for de som er NEET som for de som er arbeidsledige.

Teoretisk sett tilbyr NEET som et mål på utenforskap mindre svakheter enn f.eks. ungdomsarbeidsledighet. Dette avhenger av hvilken statistikk man bruker for å beregne arbeidsledighetsandelen. Arbeidskraftsundersøkelsen (AKU) gir det mest dekkende bildet på arbeidsledighet, fordi den inkluderer personer som ikke er registrert som arbeidssøkere. Det er også ut ifra AKU at SSB beregner andelen NEET. Andre mål på arbeidsledighet beregnes ut i fra størrelsen på arbeidsstyrken (summen av arbeidsledige og sysselsatte). Det er velkjent at unge trekker seg ut av arbeidsmarkedet i dårlige tider, og derfor ikke lengre er en del av arbeidsstyrken (Barth og von Simson, 2012). Andelen NEET beregnes ut fra alle personer i de relevante aldersgruppene, og endres ikke av at folk trer ut av statistikken.

### 8.3 Veien videre

En naturlig oppfølging til denne oppgaven blir å gjøre det samme for nyere data, og i tillegg å se på den yngste gruppen som kan være NEET, de mellom 15 og 18 år. Da må man ha mer nøyaktige utdanningsdata, slik at man kan avdekke på hvilket år en person har falt fra, i tillegg til å se hvilken studieretning en person har falt fra. Det vil også være ønskelig å finne måter å løse problemet med feilkilder som personer som utfører førstegangstjeneste og hjemmeværende mødre.

En annen problemstilling kan være å se hvilke konsekvenser det å være NEET i ung alder har å si for inntekt, arbeidsledighet og utdanning senere i livet.

En annen interessant problemstilling vil være å se på «sosial dynamikk», som er hvordan f.eks. foreldres erfaring med arbeidsledighet eller det å være NEET, påvirker NEET-risikoen

til barna. Eller f.eks. hvordan andelen NEET i skoleklassen eller nabolaget påvirker ens egen NEET-risiko.



## Referanser

- Akselsen, A., Lien, S. og Sivertstøl, Ø. (2007). *FD-trygd variabelliste*. Hentet 15.5.2015 fra [http://www.ssb.no/a/publikasjoner/pdf/notat\\_200717/notat\\_200717.pdf](http://www.ssb.no/a/publikasjoner/pdf/notat_200717/notat_200717.pdf)
- Arulampalam, W., Gregg, P. og Gregory, M. (2001). Unemployment scarring. *The Economic Journal*, 111(475), 577–584.
- Bø, T.P. og Vigran, Å. (2014). Ungdom som verken er i arbeid eller utdanning. *SSB-rapport*, 37.
- Bäckman, O., Jakobsen, V., Lorentzen, T., Österbacka, E. og Dahl, E. (2011). Dropping out in Scandinavia: Social exclusion and labour market attachment among upper secondary school dropouts in Denmark, Finland, Norway and Sweden. (*Arbetsrapport/Institutet för Framtidsstudier*, 8.
- Barth, E. og von Simson, K. (2012). Ungdomsarbeidsledighet og konjunkturer. *Økonomiske analyser*(5), 38–44.
- Brage, S. og Thune, O. (2015). Ung uførhet og psykisk sykdom. *Arbeid og Velferd*, 1.
- Bragstad, T. (2009). Tidsbegrenset uførestønad - evaluering fire år etter innføring. *NAV-rapport*(3).
- Breusch, T.S. og Pagan, A.R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The Review of Economic Studies*, 239–253.
- Cappelen, Å. og Eika, T. (2010). 2000-tallet: Fra optimisme til krise. *Rapporter*(14).
- Falch, T., Borge, L.-E., Lujala, P., Nyhus, O.H. og Strøm, B. (2010). Årsaker til og konsekvenser av manglende fullføring av videregående opplæring. *SØF-rapport*, 10(3).
- Falch, T., Johannesen, A. og Strøm, B. (2009). Kostnader av frafall i videregående opplæring. *Trondheim: Senter for økonomisk forskning AS.*, 15, 2012.
- Falch, T. og Nyhus, O.H. (2011). Betydningen av fullført videregående opplæring for sysselsetting og inaktivitet blant unge voksne. *Senter for Økonomisk Forskning*, 01.
- Försäkringskassan. (2014). *Sjuk- och aktivitetsersättning*. Hentet 18.8.2015 fra [http://www.forsakringskassan.se/statistik/sjuk/sjukersattningaktivitetsersattning/sjukochaktivitetsersattning/sjakantalmotlan/sjukers\\_akters\\_offstat2](http://www.forsakringskassan.se/statistik/sjuk/sjukersattningaktivitetsersattning/sjukochaktivitetsersattning/sjakantalmotlan/sjukers_akters_offstat2)
- Furlong, A. (2006). Not a very NEET solution representing problematic labour market transitions among early school-leavers. *Work, employment & society*, 20(3), 553–569.
- Grødem, A., Skevik, R.A.N. og Strand, A.H. (2014). Unge mottakere av helserelevante

- ytelser: Fordelingen mellom offentlig og familiebasert forsørgelse av unge NEET. *FAFO-rapport*, 37.
- Haugen, R. (2013). *Gradert uførepensjon: En empirisk analyse av pensjonstilgangen og utviklingen i årene etter* (Upublisert masteroppgave). Universitetet i Bergen.
- Hirano, K. (2010). *Lecture note 10: Panel data part II*. Forelesningsnotat i Economics 696, Topics in Econometrics. Hentet 9.4.2015 fra [http://www.u.arizona.edu/~hirano/696\\_2010/ln10.pdf](http://www.u.arizona.edu/~hirano/696_2010/ln10.pdf) (Forelesningsnotat)
- Hyggen, C. (2013). Unge i og utenfor arbeidsmarkedet i norden. *Søkelys på arbeidslivet*, 40(4), 357–378.
- Kirke-, utdannings- og forskningsdepartementet. (1996). *Videregående opplæring etter reform 94*. Hentet 18.2.2015 fra <https://www.regjeringen.no/nb/dokumenter/Vidaregaande-opplaring-etter-Reform-94/id87404/>
- Kunnskapsdepartementet. (2010). *Rett til videregående opplæring for ungdom (16-24 år)*. Hentet 19.6.2015 fra <https://www.regjeringen.no/no/tema/utdanning/grunnopplaring/videregaende-opplaring/Rett-til-videregaende-opplaring-for-ungdom-16-24-ar/id600783/>
- Kunnskapsdepartementet. (2015). *Oppfølgingstjenesten*. Hentet 5.3.2015 fra [https://utdanning.no/tema/hjelp\\_og\\_veiledning\\_nyttig\\_informasjon/oppfolgingstjenesten](https://utdanning.no/tema/hjelp_og_veiledning_nyttig_informasjon/oppfolgingstjenesten)
- Mandal, R., Jakobsen, H.O., Jensen, C. og Osborg, S.O. (2015). *Hvordan fungerer arbeidsavklaringspenger som ytelse og ordning?* (Teknisk rapport). Trondheim: SINTEF. (Forskningsrapport)
- NAV. (2014a). *Arbeidsavklaringspenger (aap)*. Hentet 3.12.2014 fra <https://www.nav.no/no/Person/Arbeid/Sykmeldt%2C+arbeidsavklaringspenger+og+yrkesskade/Arbeidsavklaringspenger+%28AAP%29.217377.cms>
- NAV. (2014b). *Grunnbeløpet i folketrygden*. Hentet 21.4.2014 fra <https://www.nav.no/no/NAV+og+samfunn/Kontakt+NAV/Utbetalinger/Grunnbelopet+i+folketrygden>
- NAV. (2014c). *Økonomisk stønad (økonomisk sosialhjelp)*. Hentet 15.7.2015 fra <https://www.nav.no/no/Person/Flere+tema/Sosiale+tjenester/%C3%98konomisk+st%C3%B8nad+%28%C3%B8konomisk+sosialhjelp%29.87469.cms#chapter-2>
- NAV. (2014d). *Mottakere av arbeidsavklaringspenger. januar-oktober 2014. tidsserie (pdf)* (Teknisk rapport). Forfatter. Hentet 4.12.2014 fra <https://www.nav.no/no/NAV+og+samfunn/Statistikk/AAP+nedsatt+arbeidsevne+og+uforepensjon++statistikk/Arbeidsavklaringspenger>

- NAV. (2014e). *Sykepenger til arbeidstakere*. Hentet 4.12.2014 fra <https://www.nav.no/no/Person/Arbeid/Sykmeldt%2C+arbeidsavklaringspenger+og+yrkesskade/Sykepenger/Sykepenger+til+arbeidstakere#chapter-1>
- NAV. (2015). *Mottakere av uføretrygd, etter kjønn og alder. pr. 31.03.2006-2015* \*). *antall*. Hentet 15.7.2015 fra [https://www.nav.no/no/NAV+og+samfunn/Statistikk/AAP+nedsatt+arbeidsevne+og+uforetrygd+-+statistikk/Tabeller/Mottakere+av+uf%C3%B8retrygd%2C+etter+kj%C3%B8nn+og+alder.+Pr.+31.03.2006-2015+\\*%29.+Antall.414281.cms](https://www.nav.no/no/NAV+og+samfunn/Statistikk/AAP+nedsatt+arbeidsevne+og+uforetrygd+-+statistikk/Tabeller/Mottakere+av+uf%C3%B8retrygd%2C+etter+kj%C3%B8nn+og+alder.+Pr.+31.03.2006-2015+*%29.+Antall.414281.cms)
- Nygaard, H. (2015). *Utviklingen i sykefraværet, 4. kvartal 2014*. Hentet 06.7.2015 fra [https://www.nav.no/no/NAV+og+samfunn/Statistikk/Sykefravar+-+statistikk/Sykefravar/\\_attachment/409858?\\_download=true&\\_ts=14c08a13970](https://www.nav.no/no/NAV+og+samfunn/Statistikk/Sykefravar+-+statistikk/Sykefravar/_attachment/409858?_download=true&_ts=14c08a13970)
- OECD. (2014). *Education at a glance: OECD indicators* (Teknisk rapport). Forfatter. Hentet fra <http://dx.doi.org/10.1787/eag-2014-en>
- Olsen, T., Tägtström, J., Bremberg, S., Mykletun, A., Olofsson, J., Arnardóttir, J.R., ... others (2013). *For det som vokser: Unge, psykisk uhelse og tidlig uførepensjonering i norden. en antologi*. Nordens välfärdscenter.
- Pischke, J.S. og Angrist, J.D. (2009). *Mostly harmless econometrics*. Princeton, NJ: Princeton University Press.
- Quintini, G. og Martin, S. (2006). Starting well or losing their way?: the position of youth in the labour market in oecd countries. *OECD social, employment and migration working papers*, 39(8).
- Raffe, D. (2003). *Young people not in education, employment or training*. Centre for Educational Sociology, University of Edinburgh.
- Rodríguez, G. og Elo, I. (2003). Intra-class correlation in random-effects models for binary data. *The Stata Journal*, 3(1), 32–46.
- Schmidheiny, K. (2014). *Panel data: Fixed and random effects*. Short Guide to Microeconomics. Hentet 27.8.2015 fra <http://www.schmidheiny.name/teaching/panel2up.pdf> (Forelesningsnotat)
- Schoon, I. (2014). *Parental worklessness and the experience of neet among their offspring. evidence for the longitudinal study of young people in england (lsype)*. Hentet 4.08.2015 fra <http://www.llcsjournal.org/index.php/llcs/article/download/279/296>
- Serracant, P. (2014). A Brute Indicator for a NEET Case: Genesis and Evolution of a Problematic Concept and Results from an Alternative Indicator. *Social Indicators*

- Research*, 117(2), 401–419.
- Socialstyrelsen. (2014). *Økonomisk bistånd*. Hentet 15.7.2015 fra <http://www.socialstyrelsen.se/statistik/statistikdatabas/ekonomisktbistand>
- SSB. (2014a). *Gjennomstrømming i videregående opplæring, 2008-2013*. Hentet 8.4.2015 fra <http://www.ssb.no/utdanning/statistikker/vgogjen/aar/2014-06-19#content>
- SSB. (2014b). *Økonomisk sosialhjelp*. Hentet 15.7.2015 fra <https://www.ssb.no/soshjelpk/>
- StataCorp. (2014). *xtreg — fixed-, between-, and random-effects and population-averaged linear model*. Hentet 20.5.2015 fra <http://www.stata.com/manuals13/xtxtreg.pdf>
- Verbeek, M. (2012). *A guide to modern econometrics* (4. utg.). John Wiley & Sons.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, 817–838.
- Wooldridge, J.M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- Wooldridge, J.M. (2012). *Introductory econometrics: A modern approach* (5. utg.). Cengage Learning.
- Wooldridge, J.M. (2013). *Correlated random effects panel data models*. IZA Summer School in Labor Economics. Hentet 5.6.2015 fra [http://www.iza.org/conference\\_files/SUMS\\_2013/slides\\_1\\_linear\\_iza.pdf](http://www.iza.org/conference_files/SUMS_2013/slides_1_linear_iza.pdf) (Forelesningsnotat)

# A Vedlegg

## A.1 Institusjonelle forhold

Gitt de forskjellige sosial- og trygdeordningene som finner i Norge, kan de diskuteres i hvilken grad de som er på disse ordningene hører til NEET-begrepet. Jeg har valgt å utelate de som mottar sykepenger og de uføretrygdede. Å ekskludere disse kategoriene fører til at mine estimater av NEET-andelen er lavere enn andre analyser (se: Bäckman et al. (2011), Grødem et al. (2014), Bø og Vigran (2014) o.l.). Jeg omtaler bare de velferdsordningene som jeg mener er relevant for NEET-begrepet; som sykepenger, uføretrygd, sosialhjelp og attførings- og rehabiliteringspenger.

### A.1.1 Velferdsstaten

I Norge har vi offentlige ordninger som garanterer landets borgere hjelp hvis de skulle få problemer med helsen eller ikke har nok midler til livsopphold. Systemet innebærer at alle borgere har rett på grunnleggende velferdsgoder som arbeidsledighetstrygd, sykepenger, alderspensjon, barnetrygd, foreldrepermisjon og lignende ordninger. Noen ordninger er behovsprøvd, og brukes som siste utvei, f.eks. sosialhjelp. Velferdsordningene er i stor grad finansiert med skatter.

**Sykepenger** får man når man er arbeidsufør på grunn egen sykdom eller skade. Man kan først benytte seg av egenmelding, hvor man må melde fra til arbeidsgiver at man er syk eller skadet, uten å fremvise legeerklæring. Man kan ha opp til 8 sammenhengende egenmeldingsdager 3 ganger i løpet av et år, hvis bedriften man jobber i er tilknyttet IA-avtalen<sup>5</sup>. Hvis bedriften man jobber i ikke er tilknyttet IA-avtalen, så kan man ha 3 egenmeldingsdager 4 ganger per år. Om man er syk utover egenmeldingsdagene, må man kunne legge frem en legeerklæring. Man har krav på sykepenger hvis man har vært i jobb i minst i fire uker. Arbeidsgiver betaler sykepengene de første 16 kalenderdagene av sykefraværet. Etter dette er det folketrygden som betaler. Størrelsen på utbetalingen av

---

<sup>5</sup>Samarbeidsavtale mellom forskjellige etater og departementer og arbeidsgiver/arbeidstaker med intensjon om å redusere sykefravær.

sykepenger er tilsvarende den gjennomsnittlige ukeseinntekten din rett før man ble sykmeldt. Man kompenseres fullt ut for det tap man har i inntekt når man er syk, og man har rett på sykepenger i opp til ett år. Hvis man ikke er tilbake i jobb etter ett år, kan personen har rett på andre ytelser, som arbeidsavklaringspenger (NAV, 2014e). Fjerde kvartal 2014 var sykefraværet på rundt 6 %. Det nesten dobbelt så mange kvinner som menn som var representert i denne statistikken (Nygaard, 2015). I Sverige var sykefraværet for 2014 på 10,6 prosent. Disse tallene lar seg ikke uten videre sammenligne nøyaktig siden det for Norge bare registreres sykefraværet for fjerde kvartal, mens det i Sverige anslås for hele året.

**Arbeidsavklaringspenger/yrkesrettet attføring og medisinsk rehabilitering** er et begrep brukt om ytelser som har til hensikt å få personer tilbake i arbeid. Denne ytelsen får man dersom man har sykdom eller skader gir redusert arbeidsevne. Yrkesrettet attføring og medisinsk rehabilitering ble erstattet av arbeidsavklaringspenger i 2010. For å få denne ytelsen forutsetter det at man må gå på arbeidsrettede tiltak, få medisinsk behandling eller gå på andre tiltak fra NAV (NAV, 2014a). I oktober 2014 var det totalt 4,6 % av befolkningen som mottok arbeidsavklaringspenger (NAV, 2014d). Medisinsk attføring/rehabilitering er en ytelse for personer som på grunn av skade eller sykdom har fått nedsatt funksjonsevnen. Arbeidsfremmende tiltak under medisinsk attføring kan f.eks. være fysioterapi. Innføringen av den nye ordningen med arbeidsavklaringspenger skulle gi tettere oppfølging av mottakerne. Intensjonen var da at færre skulle ende opp som passive stønadsmottakere. I rapporten fra Mandal et al. (2015) finner de at unge brukere uten alvorlig sykdom får for lett tilgang på AAP, og at disse ofte blir passive stønadsmottakere.

**Uføretrygd** er en ordning som sikrer inntekt til en person som er varig ufør. Uføretrygden beregnes ut fra gjennomsnittsinntekten i de beste tre av de fem siste årene før man ble ufør. Sykdom eller skade må være hovedårsaken til at inntekten er nedsatt. Andre typer tiltak eller behandlinger må være forsøkt før uføretrygden innvilges. Det er vanlig å dele inn i full og gradert uføretrygd. Gradert uføretrygd betyr at man har en restarbeidsevne. Full uføretrygd får man når det forventes at det ikke foreligger noen restarbeidsevne. Man kan ved siden av både full og gradert uføretrygd tjene opp til 0,4 G, som i 2014 tilsvarte ca. 35 000 kroner. Uføretrygden har preg av å være en absorberende tilstand: man kommer inn i den, men ikke ut av den (Haugen, 2013). I 2014 var det 307 349 som mottok uføretrygd, totalt 9,4 prosent av befolkningen. Flertallet av disse var kvinner (NAV, 2015). I Sverige

var det 354 818 personer som fikk uføretrygd i desember 2014 (Försäkringskassan, 2014).

**Tidsbegrenset uføretrygd** var en ordning som ble innført i 2004, med hensikt å redusere det økende antallet personer som fikk varig uføretrygd. Tanken bak denne ordningen var at man var usikker på en persons fremtidige arbeidsevne. Myndighetene ønsket en endring i holdningene med tanke på den varige inntektskilden som kommer med uføretrygd, og trodde at det var mulig etter en periode på uføretrygd å tilbakeføre folk i arbeid (Bragstad, 2009). Personer på tidsbegrenset uføretrygd måtte dessuten følge en oppfølgingsplan i samarbeid med NAV. Kravene for varighet av sykdom, og lengden på forutgående behandling, var mindre streng for den tidsbegrensede varianten enn varig uføretrygd. Ordningen ble avskaffet i 2010, og ble erstattet av ordningen med arbeidsavklaringspenger. NAV-rapporten av Bragstad (2009) viser at over 30 % av de som var mottakere av tidsbegrenset uføretrygd i 2004 var gått over på varig uføretrygd i 2008. Denne ytelsen ble erstattet av arbeidsavklaringspenger i 2010.

**Sosialhjelp** er en ordning som skal sikre at personer har nok midler til et forsvarlig livsopphold, når personen ikke har noen andre muligheter til å forsørge seg selv. Andre forsørgingsalternativer kan være oppsparte midler, andre stønadsordninger eller inntektsgivende arbeid. NAV vurderer individuelt størrelsen på stønaden. Stønaden blir ofte gitt i kombinasjon med veiledning om hvordan å forsørge seg selv (NAV, 2014c). Det var 125 407 personer som mottok sosialhjelp i 2014, og det var nesten dobbelt så mange menn som kvinner som mottok sosialhjelp (SSB, 2014b). I Sverige var antall sosialhjelpsmottakere på 270 300 i 2014, med et lite flertall av menn blant mottakerne (Socialstyrelsen, 2014). Siden sosialhjelp bare utbetales når det ikke er andre forsørgingsmuligheter, kan det tenkes at de som mottar sosialhjelp gjerne også samtidig er i NEET-kategorien. Sosialhjelp kan derfor være en sterk indikator på NEET.

## A.2 Økonometriske metoder

### A.2.1 Tilfeldig Effekt-modeller

Et ankepunkt mot å bruke Tilfeldig Effekt-modeller er at de har mer restriktive antagelser enn Fast Effekt-modeller (FE-modell). Det kan tenkes at det er samvariasjon mellom den uobserverte individeffekten og forklaringsvariablene, men når viktige forklaringsvariabler, slik som utdanning, er tidskonstante, kan vi som tidligere nevnt ikke bruke FE-modeller. Spørsmålet er da om det rimelig å anta om de forskjellige forklaringsvariablene ikke i seg selv er et resultat av valg personer har gjort på grunnlag av individuelle preferanser eller evner, som fanges i  $\alpha_i$ . Det er ikke utenkelig at f.eks. evner påvirker muligheten til å gjennomføre vgo, og at antagelsen om at det ikke skal være samvariasjon mellom uobserverte individeffekter (f.eks. evner/motivasjon) og forklaringsvariabler derfor ikke er realistisk. I følge modellen skal all korrelasjon i feilleddene over tid kommer fra  $\alpha_i$ . Siden feilleddene er seriekorrelerte blir standardfeil beregnet med MKM feil. Det vanlige er da å benytte seg av GLS (engelsk: Generalized Least Squares), denne metoden benyttes når det er en viss grad av korrelasjon mellom observasjonene, eller når man har heteroskedastiske feilledd.

RE-modellen kan skrives som

$$y_{it} = \beta_0 + \mathbf{x}'_{it}\boldsymbol{\beta} + \alpha_i + u_{it}, \quad u_{it} \sim \text{NID}(0, \sigma_u^2) \quad \alpha_i \sim \text{NID}(0, \sigma_\alpha^2). \quad (16)$$

Vi antar følgende om  $\alpha_i$  og  $u_{it}$

$$E\{u_{it}\} = 0 \quad (17a)$$

$$E\{\alpha_i\} = 0 \quad (17b)$$

$$E\{u_{it}^2\} = \sigma_u^2 \quad (17c)$$

$$E\{\alpha_i^2\} = \sigma_\alpha^2 \quad (17d)$$

$$E\{u_{it}u_{js}\} = 0 \quad \forall t \neq s \text{ og } i \neq j \quad (17e)$$

$$E\{\alpha_i\alpha_j\} = 0 \quad \forall i \neq j \quad (17f)$$

$$E\{u_{it}\alpha_i\} = 0 \quad (17g)$$

$$E\{u_{it}\mathbf{x}_{it}\} = 0 \quad (17h)$$

$$E\{\alpha_i\mathbf{x}_{it}\} = 0 \quad (17i)$$



Ved bruk av RE-modeller antar vi at det er ikke er heteroskedastiske feilledd, både for  $\alpha_i$  og  $u_{it}$ . Vi antar at det er uavhengighet mellom feilleddene og at det ikke er seriekorrelasjon i feilleddene. I tillegg til antagelsen om at forklaringsvariablene er strengt eksogene.

Vi definerer så summen av feilleddene til å være

$$w_{it} = \alpha_i + u_{it} \quad (18)$$

$w_{it}$  har følgende egenskaper

$$\begin{aligned} E\{w_{it}w_{is}\} &= E\{(\alpha_i + u_{it})(\alpha_i + u_{is})\} \\ &= E\{\alpha_i\alpha_i + \alpha_i u_{is} + u_{it}\alpha_i + u_{it}u_{is}\} \\ &= E\{\alpha_i^2\} + E\{\alpha_i u_{is}\} + E\{u_{it}\alpha_i\} + E\{u_{it}u_{is}\} \\ &= E\{\alpha_i^2\} + 0 + 0 + 0 \\ &= \sigma_\alpha^2 \end{aligned} \quad (19)$$

Det er altså en samvariasjon over tid hos individet som kommer fra den individspesifikke faktoren  $\alpha_i$ .

Den totale variansen til feilleddet  $w_{it}$  blir

$$\begin{aligned} E\{w_{it}^2\} &= E\{(\alpha_i + u_{it})(\alpha_i + u_{it})\} \\ &= E\{\alpha_i^2 + 2\alpha_i u_{it} + u_{it}^2\} \\ &= E\{\alpha_i^2\} + 2E\{\alpha_i u_{it}\} + E\{u_{it}^2\} \\ &= \sigma_\alpha^2 + \sigma_u^2 \end{aligned} \quad (20)$$

Korrelasjonen i feilleddet over tid for samme individ blir

$$\rho = \text{corr}(w_{it}, w_{i,t-s}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_u^2} \quad s \neq t. \quad (21)$$

Dette bryter antagelse 3d. MKM er ikke lenger den mest effisiente estimatoren og vi må finne en mer effisient estimator som utnytter korrelasjonsstrukturen i dataene.

For  $T$  observasjoner av individ  $i$  så har vi

$$\mathbf{w}_i = \begin{bmatrix} w_{i1} \\ w_{i2} \\ \vdots \\ w_{iT} \end{bmatrix} \quad \text{og} \quad E\{\mathbf{w}_i \mathbf{w}_i'\} = \text{Var}\{\mathbf{w}_i\} = \begin{bmatrix} \sigma_\alpha^2 + \sigma_u^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_u^2 & \cdots & \sigma_\alpha^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 + \sigma_u^2 \end{bmatrix} = \mathbf{\Omega}.$$

Vi ser her at kovariansmatrisen  $\mathbf{\Omega}$  for RE-modellen er forskjellig fra kovariansmatrisen for vanlig MKM hvor alle andre plasser enn diagonalen ville vært null. Matematisk blir dette at  $\mathbf{\Omega} \neq \sigma_w^2 \mathbf{I}_k$ . MKM-estimatoren er ineffisient og vi vil ha forventningskjev varians p.g.a. seriekorrelasjon forårsaket av  $\alpha_i$ . Løsningen på dette problemet er å bruke GLS (Generaliserte Minste Kvadraters Metode, engelsk: Generalized Least Squares). Prinsippet med GLS er å transformere den opprinnelige ligningen til en ligning der de klassiske Gauss-Markov vilkårene er oppfylt. I RE-tilfellet kan det vises at denne transformasjonen vil være

$$y_{it} - \vartheta \bar{y}_i = \beta_0(1 - \vartheta) + \beta(\mathbf{x}'_{it} - \vartheta \bar{x}_i) + (w_{it} - \vartheta \bar{w}_{it}), \quad (22)$$

hvor  $\vartheta$  er lik

$$\vartheta = 1 - \sqrt{\psi} = 1 - \frac{\sigma_u}{\sqrt{\sigma_u^2 + T\sigma_\alpha^2}}. \quad (23)$$

$\vartheta$  er vektingen, og beskriver andelen av individuelle gjennomsnitt som blir fjernet fra de originale data-ene.

Tilfeldig effekt-estimatoren kan skrives som

$$\hat{\beta}_{\text{RE}} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)(x_{it} - \bar{x}_i) + \psi T \sum_{i=1}^N (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y})}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 + \psi T \sum_{i=1}^N (\bar{x}_i - \bar{x})^2}. \quad (24)$$

La oss nå sammenligne  $\hat{\beta}_{\text{RE}}$  med  $\hat{\beta}_{\text{FE}}$

$$\hat{\beta}_{\text{FE}} = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - \bar{y}_i)(x_{it} - \bar{x}_i)}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2}. \quad (25)$$

Vi ser at hvis  $\psi$  fra ligning 24 er null, så blir RE-estimatoren lik FE-estimatoren. Dette skjer når  $T$  går mot uendelig, nevneren i ligning 23 blir da uendelig stor. Når det ikke er noen individuell variasjon i feilledet ( $\sigma_\alpha^2 = 0$ ), så blir  $\psi$  lik én og RE-estimatoren blir

---


$$\bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}, \quad \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$$

MKM-estimatoren. Vi kan teste om RE-estimatoren er lik MKM-estimatoren, dette gjør vi ved å anvende en *Breusch-Pagan* Lagrange Multiplier test:

### Breusch-Pagan Lagrange Multiplier test

Denne testen bruker vi for å sjekke om det eksisterer uobserverte individeffekter. Hypotesene er

$$H_0 : \sigma_\alpha^2 = 0 \qquad H_a : \sigma_\alpha^2 \neq 0.$$

Under nullhypotesen eksisterer det ikke uobserverte individeffekter, og MKM er da konsistent og den meste effisiente estimatoren (Wooldridge, 2010, s. 264).

Testobservatoren er en Lagrange Multiplier (LM) teststatistikk og kan skrives som

$$LM = \frac{NT}{2(T-1)} \left[ \frac{\sum_{i=1}^N (\sum_{t=1}^T \hat{e}_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{e}_{it}^2} - 1 \right]^2 \sim \chi^2(1), \quad (26)$$

hvor  $\hat{e}_{it} = y_{it} - \hat{y}_{it}$  og  $\hat{y}_{it}$  er den predikerte verdien til den avhengige variabelen. Testobservatoren er *asymptotisk*  $\chi^2$ -fordelt under nullhypotesen med én frihetsgrad (Breusch og Pagan, 1980). Asymptotisk fordelt betyr at testobservatoren sin fordeling konvergerer mot en spesifikk fordeling når antall observasjoner øker. Høy verdi av testobservatoren betyr typisk at vi forkaster nullhypotesen. Verbeek (2012, s. 392) påpeker i en fotnote at testen i de fleste tilfeller forkaster nullhypotesen, og man bør utvise forsiktighet i tolkningen av konklusjonen til testen.

For RE-modeller med binær avhengig variabel, gjør Stata en tilsvarende test. Testobservatoren er en likelihood-ratio test. Prinsippet er at den tester en modell med restriksjoner (med likelihood  $\tilde{\theta}$ ) mot en uten restriksjoner (med likelihood  $\hat{\theta}$ ). I vårt tilfelle blir dette å teste for tilstedeværelsen av den tilfeldige effekten  $\alpha_i$ . Hypotesene er

$$H_0 : \rho = 0 \qquad H_a : \rho \neq 0.$$

Testobservatoren kan skrives som

$$\text{LR} = 2[\log L(\hat{\theta}) - \log L(\tilde{\theta})] \sim \chi^2(0; 1), \quad (27)$$

hvor testobservatoren er tilnærmet fordelt under nullhypotesen som en  $\chi^2$ -fordeling med null og én frihetsgrad, som er den tilnærmingen Stata bruker (Rodríguez og Elo, 2003).

Den mest effisiente estimatoren er den som har minst varians. Variansen til RE-estimatoren er lik

$$\text{Var}\{\hat{\beta}_{\text{RE}}\} = \frac{\sigma_u^2}{\sum_{i=1}^N \sum_{t=1}^T (x_i - \bar{x}_i)^2 + \psi T \sum_{i=1}^N (\bar{x}_i - \bar{x})^2} \quad (28)$$

$$\text{Var}\{\hat{\beta}_{\text{FE}}\} = \frac{\sigma_u^2}{\sum_{i=1}^N \sum_{t=1}^T (x_i - \bar{x}_i)^2} \quad (29)$$

Vi ser fra ligningen over at  $\text{Var}\{\hat{\beta}_{\text{RE}}\} < \text{Var}\{\hat{\beta}_{\text{FE}}\}$  så lenge  $\psi$  er større enn null. Dette betyr at RE-estimatoren er mer effisient enn FE-modellen. Dermed foretrekkes RE-modellen hvis antagelse 17a til 17i gjelder.

## A.2.2 Binære modeller

Ofte kan det være interessant å se hvilke forskjellige faktorer som påvirker om en person er i en bestemt tilstand eller ikke. Når det eksisterer bare to tilstander kaller vi dette for binære modeller, f.eks. NEET eller ikke.

### A.2.2.1 Lineære sannsynlighetsmodeller

Sannsynligheten for at individ  $i$  er NEET i periode  $t$ , kan estimeres ved å bruke en lineær binær modell, hvor den avhengige variabelen er 0 hvis personen ikke er NEET og 1 hvis personen er NEET. Vi vil så forklare den avhengige variabelen ut i fra en rekke bakgrunnsvariabler gitt i modellen

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \epsilon_{it}. \quad (30)$$

Vi antar at  $E\{\epsilon_{it}|\mathbf{x}_{it}\} = 0$  og får at

$$E\{y_{it}|\mathbf{x}_{it}\} = E\{\mathbf{x}'_{it}\boldsymbol{\beta} + \epsilon_{it}\} = E\{\mathbf{x}'_{it}\boldsymbol{\beta}\} + 0 = E\{\mathbf{x}'_{it}\boldsymbol{\beta}\} = \mathbf{x}'_{it}\boldsymbol{\beta}. \quad (31)$$

$y_{it}$  kan tolkes som en sannsynlighet og skal da ligge i intervallet  $[0, 1]$ . Wooldridge (2012, s. 241) bruker et eksempel med hvilke faktorer som påvirker jobbdeltagelse for kvinner, og finner at det å ha et ekstra barn reduserer sannsynligheten for jobbdeltagelse med 0,262, slik at det å få fire ekstra barn reduserer sannsynligheten for jobbdeltagelse med  $0,262 \cdot 4 = 1,048$ . Tolkningen av dette resultatet gir altså ikke mening. Wooldridge (2012, s. 241) påpeker at 96 % av utvalget hadde enten ingen eller ett barn, slik at man kanskje skal begrense seg til disse utfallene når man tolker estimatene.

Et annet problem er at vi bryter antagelsen om at variansen til  $y$  ikke avhenger av  $\mathbf{x}'$ . Vi vet at  $y$  er Bernoulli-fordelt og variansen til  $y$  kan da skrives som

$$\text{Var}\{y_{it}|\mathbf{x}_{it}\} = P\{y_{it} = 1|\mathbf{x}_{it}\}(1 - P\{y_{it} = 1|\mathbf{x}_{it}\}) = \mathbf{x}'_{it}\boldsymbol{\beta}(1 - \mathbf{x}'_{it}\boldsymbol{\beta}). \quad (32)$$

Vi ser at variansen til  $y$  avhenger av  $\mathbf{x}$  og  $\boldsymbol{\beta}$  og vi har derfor heteroskedastisitet. Heteroskedastisitet gir forventningskjevne standardfeil og kan føre til at vi feilaktig får signifikante regresjonskoeffisienter, eller at vi feilaktig ikke får signifikante regresjonskoeffisienter. Siden vi får forventningskjevne standardfeil vil også estimatoren være ineffisient. Heteroskedastisitet kan også gi ikke-normalfordelte feilledd, og testobservatoren blir er ikke da lengre  $t$ -fordelt. For å korrigere for forventningskjevne standardfeil kan vi benytte oss av heteroskedastisk-robuste metoder, som også kalles robuste standardfeil (White, 1980).

### A.2.2.2 Probit

En måte som sikrer at  $y_{it}$  ligger i intervallet mellom 0 og 1, er å anta at

$$P\{y_{it} = 1|\mathbf{x}_{it}\} = F(\mathbf{x}'_{it}\boldsymbol{\beta}), \quad (33)$$

hvor  $F(\mathbf{x}'_{it}\boldsymbol{\beta})$  bare tar verdier mellom 0 og 1.  $F(\cdot)$  har egenskapen at  $F(-\infty) = 0$  og  $F(+\infty) = 1$ .

Et eksempel på en funksjon  $F$  som møter disse kriteriene, er den kumulative standard

normalfordelingsfunksjonen

$$F(\mathbf{x}'_{it}\boldsymbol{\beta}) = \Phi(\mathbf{x}'_{it}\boldsymbol{\beta}) = \int_{-\infty}^{\mathbf{x}'_{it}\boldsymbol{\beta}} \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}t^2\} dt, \quad (34)$$

Vi har da at

$$P\{y_{it} = 1 | \mathbf{x}'_{it}\boldsymbol{\beta}\} = \Phi(\mathbf{x}'_{it}\boldsymbol{\beta}) \quad (35)$$

Ligning 35 er kjent som **probit-modellen**. Et problem med probit-modellen er at det er vanskelig å tolke noe annet enn fortegnet til koeffisientene (Verbeek, 2012, s. 208). For å si noe om effekten fra en av forklaringsvariablene på den avhengige variabelen må man derfor beregne marginaleffekter. Marginaleffekter for en kontinuerlig variabel  $x_{ik}$  er definert som

$$\frac{\partial \Phi(\mathbf{x}'_{it}\boldsymbol{\beta})}{\partial x_{ik}} = \phi(\mathbf{x}'_{it}\boldsymbol{\beta})\beta_k \quad (36)$$

Ligning 36 viser sannsynligheten for at den avhengige variabelen tar verdien 1 når vi øker  $x_{ik}$  med én enhet.

Marginaleffekten for en binær variabel er definert som

$$P\{y_{it} = 1 | \mathbf{x}'_{it}, x_{itk} = 1\} - P\{y_{it} = 1 | \mathbf{x}'_{it}, x_{itk} = 0\} \quad (37)$$

Ligningen over forteller oss hvordan den predikerte sannsynligheten endrer seg når vi endrer den binære uavhengige variabelen fra 0 til 1, gitt at vi holder alle andre forklaringsvariabler konstante. Hvis vi nå finner denne differansen for alle individene i datasettet og tar gjennomsnittet, så får vi gjennomsnittlig marginaleffekt (AME, engelsk: Average Marginal Effect)

$$AME = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T P\{y_{it} = 1 | \mathbf{x}'_{it}, x_{itk} = 1\} - P\{y_{it} = 1 | \mathbf{x}'_{it}, x_{itk} = 0\} \quad (38)$$

### A.2.2.3 Tilfeldig effekt probit-modeller

Siden NEET er en binær variabel, altså at den enten tar verdien 0 hvis ikke NEET eller 1 hvis NEET, vil vi ikke kunne benytte oss av «vanlig» RE-modeller. Å benytte seg av probit som funksjonsform sikrer at sannsynlighetene holder seg mellom null og én.

Vi antar at

$$\alpha_i \sim \text{NID}(0, \sigma_\alpha^2) \quad (39)$$

$$u_{it} \sim \text{NID}(0, 1 - \sigma_\alpha^2) \quad (40)$$

Likelihood-funksjonen for individ  $i$  ved tid  $t$  kan skrives som, gitt at vi antar at alle  $e_{it}$  er uavhengig

$$f(y_{it} | \mathbf{x}_{it}, \alpha_i, \boldsymbol{\beta}) = \Phi \left( \frac{\mathbf{x}'_{it} \boldsymbol{\beta} + \alpha_i}{\sqrt{1 - \sigma_\alpha^2}} \right) \quad \text{hvis } y_{it} = 1 \quad (41)$$

$$= 1 - \Phi \left( \frac{\mathbf{x}'_{it} \boldsymbol{\beta} + \alpha_i}{\sqrt{1 - \sigma_\alpha^2}} \right) \quad \text{hvis } y_{it} = 0, \quad (42)$$

hvor  $\Phi$  er den kumulative sannsynlighetsfordelingen til standard normalfordelingen.

$\alpha_i$  er normalfordelt og sannsynlighetsfordelingen kan da skrives som

$$f(\alpha_i) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp \left[ -\frac{1}{2} \frac{\alpha_i^2}{\sigma_\alpha^2} \right]. \quad (43)$$

Gitt at vi antar at  $u_{it}$  er uavhengig over tid og på tvers av individer kan simultansannsynligheten for alle individer på tvers av tid skrives som

$$\begin{aligned} f(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \boldsymbol{\beta}) &= \int_{-\infty}^{\infty} f(y_{i1}, \dots, y_{iT} | x_{i1}, \dots, x_{iT}, \alpha_i, \boldsymbol{\beta}) f(\alpha_i) d\alpha_i \\ &= \int_{-\infty}^{\infty} \left[ \prod_t f(y_{it} | \mathbf{x}_{it}, \alpha_i, \boldsymbol{\beta}) \right] f(\alpha_i) d\alpha_i. \end{aligned} \quad (44)$$

Ligning 44 har ingen analytisk løsning, men lar seg løse numerisk (Verbeek, 2012, s. 422).

### A.3 Utfyllende regresjonsresultater

Tabell 7: Oppsummering for forklaringsvariabler (gjennomsnittsverdi).

	Gjennomsnitt
NEET	9.77%
<i>Utdanning</i>	
Fullført allmennfag	44.01%
Fullført yrkesfag	31.99%
<i>Velferd</i>	
Arbeidsøker året før	12.88%
Arbeidsøker de to foregående årene	5.91%
Arbeidsøker de 3 foregående årene	2.87%
Fikk sosialhjelp året før	5.34%
Fikk sosialhjelp de to foregående årene	3.42%
Fikk sosialhjelp de 3 foregående årene	2.38%
Passiv stønadsmottaker året før	1.49%
Aktiv tiltaktdeltager året før	1.37%
Passiv stønadsmottaker de to foregående årene	0.60%
Aktiv tiltaktdeltager de to foregående årene	0.22%
Passiv stønadsmottaker de 3 foregående årene	0.24%
Aktiv stønadsmottaker de 3 foregående årene	0.04%
<i>Demografi</i>	
Mann=1, kvinne=0	51.73%
Forsørget av én forelder	4.77%
Antall søsken	0.68
alder (år)	22.50
Observasjoner	3097917



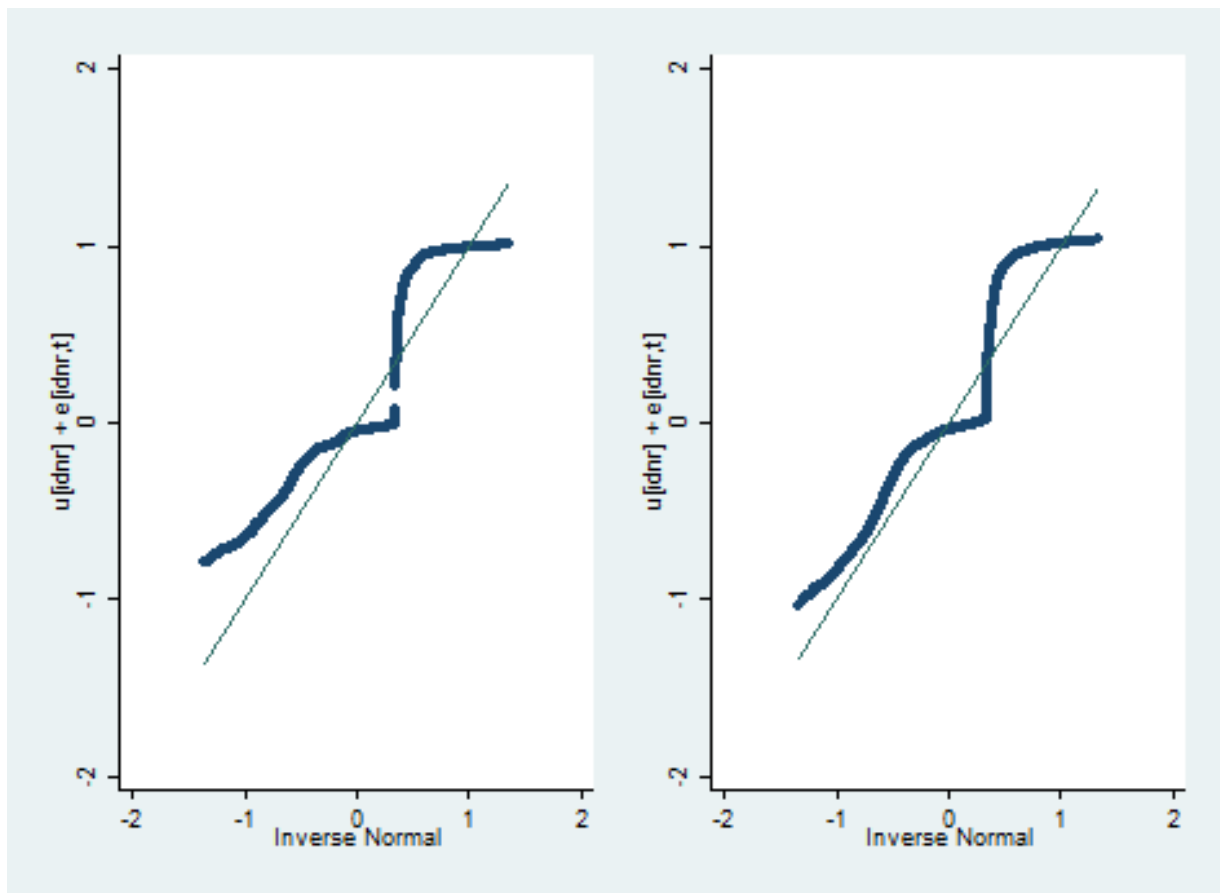
Tabell 8: Lineær sannsynlighetsmodell for sannsynligheten for å være NEET syv år etter avsluttet ungdomsskole.

	LPM (23år)	LPM (2)	RE (3)	CRE (4)
	$\beta$	$\beta$	$\beta$	$\beta$
<i>Utdanning</i>				
Fullført allmennfag	-0.06930***	-0.06577***	-0.08859***	-0.02772***
Fullført yrkesfag	-0.08073***	-0.08039***	-0.10268***	-0.04850***
Arbeidssøker året før	0.06342***	0.06960***	0.05398***	0.04226***
Arbeidssøker to år før	0.01863***	0.02249***	0.01329***	0.00829***
Arbeidssøker 3 år før	0.01935***	0.04222***	0.02492***	0.01487***
Gjn.snittlig antall ganger som arbeidssøker				0.13007***
Fikk sosialhjelp året før	0.14555***	0.16383***	0.09589***	0.04258***
Fikk sosialhjelp to år før	0.06645***	0.07637***	0.06118***	0.03411***
Fikk sosialhjelp 3 år før	0.12933***	0.11908***	0.08939***	0.03179***
Gjennomsnittlig antall ganger med sosialhjelp				0.39813***
Passiv stønadsmottaker året før	0.21179***	0.21487***	0.13877***	0.11568***
Aktiv tiltaksdeltager året før	0.13399***	0.15077***	0.12620***	0.10027***
Passiv stønadsmottaker to år før	0.07684***	0.04111***	0.04068***	0.03199***
Aktiv tiltaksdeltager to år før	0.04933**	0.03726***	0.05668***	0.04985***
Passiv stønadsmottaker 3 år før	0.11938***	0.08843***	0.07786***	0.06476***
Aktiv tiltaksdeltager 3 år før	0.15707***	0.12457***	0.11321***	0.10743***
Gjennomsnittlig antall ganger på tiltak				0.14214***
<i>Demografi</i>				
Mann=1, kvinne=0	-0.01310***	-0.00957***	-0.00848***	-0.00730***
Forsørget av én forelder	0.05507***	0.04359***	0.05818***	0.01797***
<i>tidsdummyer</i>				
=2001	0.00000	0.00000	0.00000	0.00000
=2002	0.00728***	0.00623***	0.00673***	0.00106*
=2003	0.01557***	0.01709***	0.01728***	0.00672***
=2004	0.02252***	0.02388***	0.02499***	0.01066***
=2005	0.03167***	0.03486***	0.03630***	0.01715***
=2006	0.03574***	0.03479***	0.03575***	0.01117***

=2007	0.03342***	0.03085***	0.03079***	0.00003
=2008		0.03843***	0.03727***	0.00000
<i>aldersdummyer</i>				
=21		-0.07261***	-0.06939***	-0.06154***
=22		-0.09922***	-0.09475***	-0.08088***
=23		-0.10953***	-0.10538***	-0.08664***
=24		-0.10995***	-0.10658***	-0.08303***
=25		-0.11293***	-0.10958***	-0.08180***
=26		-0.11664***	-0.11369***	-0.08141***
=27		-0.12122***	-0.11905***	-0.08197***
=28		-0.12536***	-0.12390***	-0.08193***
=29		-0.12889***	-0.12804***	-0.08089***
=30		-0.13358***	-0.13405***	-0.08136***
=31		-0.13636***	-0.13864***	-0.08044***
Gjennomsnittalder				-0.00772***
Konstantledd	0.09771***	0.20050***	0.22470***	0.31478***
Observasjoner	296673	2385483	2385483	2385483
Antall personer			330351	330351
adjusted $R^2$	0.158	0.150	0.145	0.177
$\rho$			0.256	0.254
Gjennomsnittlig predikert sannsynlighet	0.092	0.099	0.099	0.100
LR $\chi^2$			72814.534	101820.867
Prob > $\chi^2$	0.000	0.000	0.000	0.000

\* $p < 0.10$ , \*\*  $p < 0.51$ , \*\*\*  $p < 0.01$ . Robuste standardfeil i parentes.

*Merknad:* Kontrollert for antall søsken.



Figur 14: QQ-plott av  $\epsilon_{it} = \alpha_i + u_{it}$  for modell (3) (t.v.) og modell (4).

Tabell 9: Korrelert tilfeldig effekt probit-modell for sannsynligheten for å være i tilstanden NEET.

	(1) NEET	
	$\beta$	AME
<i>Utdanning</i>		
Fullført allmennfag	-0.16823***	-0.01250***
Fullført yrkesfag	-0.42411***	-0.02746***
<i>Velferdsordninger</i>		
Arbeidssøker året før	0.32802***	0.02155***
Arbeidssøker to år før	0.02338***	0.00154***

Arbeidssøker 3 år før	0.02865***	0.00188***
Gjennomsnittlig antall ganger som arbeidssøker	1.77178***	***
Fikk sosialhjelp året før	0.14991***	0.00985***
Fikk sosialhjelp to år før	0.10202***	0.00670***
Fikk sosialhjelp 3 år før	0.14114***	0.00927***
Gjennomsnittlig antall ganger med sosialhjelp	1.71110***	***
Passiv stønadsmottaker året før	0.54185***	0.04749***
Aktiv tiltaksdeltager året før	0.52341***	0.04539***
Passiv stønadsmottaker to år før	0.14044***	0.00995***
Aktiv tiltaksdeltager to år før	0.18917***	0.01377***
Passiv stønadsmottaker 3 år før	0.23053***	0.01718***
Aktiv tiltaksdeltager 3 år før	0.43082***	0.03592***
Gjennomsnittlig antall ganger på tiltak	1.34390***	***
<i>Demografi</i>		
Mann=1, kvinne=0	-0.09433***	-0.00620***
Forsørget av én forelder	0.20863***	0.01371***
<i>aldersdummyer</i>		
=21	-0.47570***	0.00000***
=22	-0.65720***	0.00000***
=23	-0.71992***	0.00000***
=24	-0.68497***	0.00000***
=25	-0.67556***	0.00000***
=26	-0.66992***	0.00000***
=27	-0.67822***	0.00000***
=28	-0.68593***	0.00000***
=29	-0.68478***	0.00000***
=30	-0.69737***	0.00000***
=31	-0.69199***	0.00000***
Gjennomsnittalder	-0.08899***	***
<i>tidsdummyer</i>		
=2001	0.00000	0.00000
=2002	-0.01597**	0.00000**
=2003	0.02903***	0.00000***
=2004	0.06162***	0.00000***

=2005	0.13256***	0.00000***
=2006	0.07969***	0.00000***
=2007	-0.03374***	0.00000***
=2008	0.00000	0.00000
Konstantledd	0.23602***	***
lnsig2u	0.06534***	***
Observasjoner	2385483	
Antall personer	330351	
McFadden's $R^2$	0.126	
$\rho$	0.516	
Gjennomsnittlig predikert sannsynlighet	0.057	
LR $\chi^2$	106725.838	
Prob > $\chi^2$	0.000	

\* $p < 0.10$ , \*\*  $p < 0.51$ , \*\*\*  $p < 0.01$ . Standardfeil i parentes.

*Merknad:* Kontrollert for antall søsken. AME= Average Marginal Effect. Det ser ut som om STATA ikke har villet beregne marginaleffektene av tids- og aldersdummyene.

Tabell 10: Korrelert tilfeldig effekt-modell for sannsynligheten for å være i tilstanden NEET, lagget 3 år.

	CRE (1) $\beta$
<i>Utdanning</i>	
Fullført allmennfag	-0.02774***
Fullført yrkesfag	-0.04834***
<i>Velferdsordninger</i>	
Arbeidssøker 3 år før	-0.02104***
Gjennomsnittlig antall dager arbeidledig	0.21691***
Fikk sosialhjelp året før	0.05493***
Fikk sosialhjelp to år før	0.03261***
Fikk sosialhjelp 3 år før	0.03036***
Gjennomsnittlig antall ganger med sosialhjelp	0.38559***
Passiv stønadsmottaker året før	0.06839***
Aktiv tiltaksdeltager året før	0.01108***
Passiv stønadsmottaker to år før	0.18068***
Aktiv tiltaksdeltager to år før	0.06883***
Passiv stønadsmottaker 3 år før	0.00676
Aktiv tiltaksdeltager 3 år før	0.01389***
Gjennomsnittlig antall ganger på tiltak	0.09928***
<i>Demografi</i>	
Mann=1, kvinne=0	-0.00694***
Forsørget av én forelder	0.01793***
<i>aldersdummyer</i>	
=20	0.00000
=21	-0.05810***
=22	-0.07568***
=23	-0.08002***
=24	-0.07632***
=25	-0.07562***
=26	-0.07519***

=27	-0.07545***
=28	-0.07533***
=29	-0.07427***
=30	-0.07457***
=31	-0.07361***
Gjennomsnittalder	-0.00855***
<i>tidsdummyer</i>	
=2001	0.00000
=2002	0.00124**
=2003	0.00988***
=2004	0.01607***
=2005	0.02206***
=2006	0.01520***
=2007	0.00216***
=2008	0.00000
Konstantledd	0.32434***
Observasjoner	2385483
Antall personer	330351
adjusted $R^2$	0.176
$\rho$	0.253
Gjennomsnittlig predikert sannsynlighet	0.100
LR $\chi^2$	98913.738
Prob > $\chi^2$	0.000

\* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . Standardfeil i parentes.

*Merknad:* Kontrollert for antall søsken.

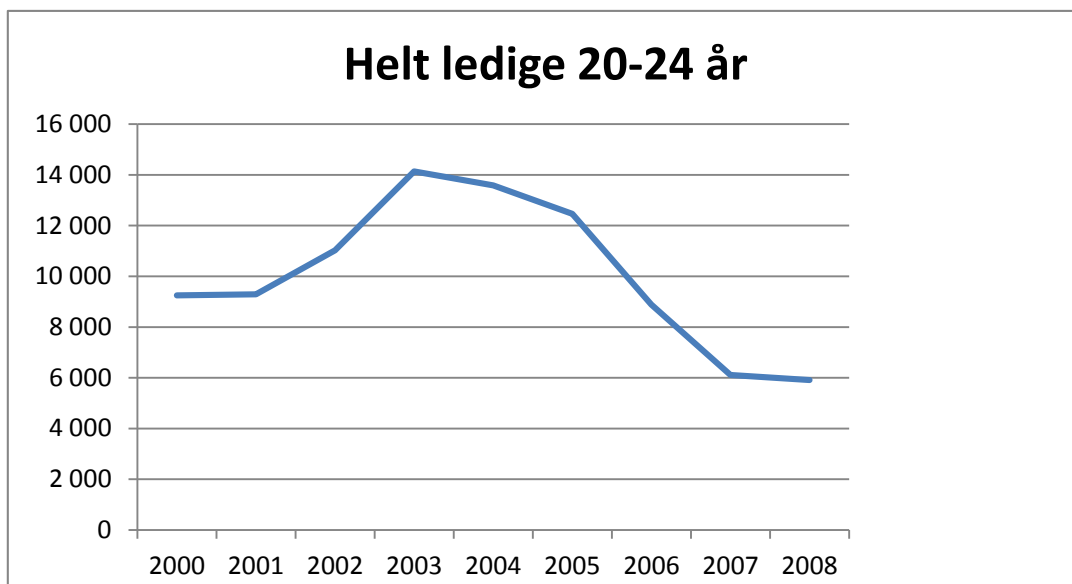
## A.4 Utviklingen i grunnbeløpet og antall helt ledige

År	Grunnbeløp
2008	70 256
2007	66 812
2006	62 892
2005	60 600
2004	58 779
2003	56 861
2002	54 861
2001	51 360
2000	49 090
1999	46 950
1998	45 370
Gjennomsnitt	56 712

Tabell 11: Grunnbeløpet pr. år

Tallene er hentet fra <https://www.nav.no/no/NAV+og+samfunn/Kontakt+NAV/Utbetalinger/Grunnbeløpet+i+folketrygden>.





Figur 15: Antall helt ledige i alderen 20 til 24 år

Tallene er hentet fra [https://www.nav.no/no/NAV+og+samfunn/Statistikk/Arbeidssokere+og+stillinger+--+statistikk/Historisk+statistikk/\\_attachment/409333?\\_download=true&\\_ts=14bea26d748](https://www.nav.no/no/NAV+og+samfunn/Statistikk/Arbeidssokere+og+stillinger+--+statistikk/Historisk+statistikk/_attachment/409333?_download=true&_ts=14bea26d748)