



*Magne Bråthen, Jørn Ivar Hamre og  
Trond Pedersen*

## **Evaluering av ordinære arbeidsmarkedstiltak**

Beskrivende analyse av deltakerne i  
2002 og forslag til ny evaluerings-  
metode

## **Forord**

Arbeidet er utført på oppdrag fra Arbeids- og sosialdepartementet. Takk til Hege Torp, Institutt for samfunnsforskning og til Oddbjørn Raaum og Knut Rød ved Stiftelsen Frischsenteret for samfunnsøkonomisk forskning for nyttige kommentarer og innspill. Gjenstående feil og mangler står forfatterne selv for.

## **Sammendrag**

Notatet omhandler SSB sitt forslag til opplegg for en årlig evaluering av ordinære arbeidsmarkedstiltak.

Opplegget er todelt. Første del vil bestå av en deskriptiv analyse av tiltaksdeltakere og helt ledige.

Tallmaterialet er ment som et supplement til den løpende statistikken. Del to utgjøres av ulike effektanalyser av ordinære tiltak, etter et opplegg som implementerer nyere forskning på feltet.

Vi vil beregne endringen i gjennomsnittlig arbeidstakerandel etter ett år for personer som deltok på ordinære arbeidsmarkedstiltak i fjerde kvartal sett i forhold til om de ikke deltok på tiltak. Til beregningene trengs en sammenligningsgruppe av personer som skal representere deltakernes forløp om de ikke hadde gått på tiltak. Metoden for å danne de enkelte sammenligningsgruppene vil være matching ved hjelp av en såkalt propensity score. Vi modellerer individenes "valg" mellom enten å delta på arbeidskraftsopplæring, lønnstilskudd, arbeidspraksis eller ikke delta på tiltak i en flervalgsmodell med relevante forklaringsvariable separat for mest mulig like undergrupperinger.

# Innhold

<b>1. Innledning</b> .....	<b>4</b>
<b>2. Effektevaluering av ordinære arbeidsmarkedstiltak til nå</b> .....	<b>6</b>
2.1 SSBs registerbaserte effektevalueringer av ordinære tiltak.....	6
2.2 Nyere effektanalyser av ordinære tiltak .....	8
<b>3. Metodiske utfordringer ved evalueringsstudier</b> .....	<b>10</b>
3.1 Effekten av hva ?.....	10
3.2 Definere tiltaksgruppen.....	11
3.3 Gruppering av tiltak .....	14
3.4 Definere sammenlikningsgruppe - matching .....	14
3.5 Valg av suksesskriterium .....	17
<b>4. Forslag til opplegg for en årlig, beskrivende analyse av deltakere på ordinære arbeidsmarkedstiltak</b> .....	<b>18</b>
4.1 Sammensetningen av tiltaksdeltakere og sammenlikningsgruppen av helt ledige.....	18
4.2 Tilpasning på arbeidsmarkedet målt ett år etter .....	20
<b>5. Forslag til opplegg for årlige effektevalueringer av ordinære arbeidsmarkedstiltak</b> .....	<b>23</b>
5.1 Separate analyser av tiltaksgrupper.....	23
5.2 Evaluering av effekten på tvers av tiltakene .....	24
<b>6. Forslag til tilleggsprosjekter</b> .....	<b>25</b>
6.1 Sammenligning med SSB sin tidligere anvendte metode .....	25
6.2 Utvidelse av suksesskriteriet .....	25
6.3 Evaluering av hele tiltakspakken.....	25
<b>Referanser</b> .....	<b>27</b>
<b>Vedlegg A</b> .....	<b>29</b>
A.1 Mer formell beskrivelse av metoden i kapittel 5.1 og 5.2 .....	29
A.2 Mer formell beskrivelse av metoden i kapittel 6.3 .....	30

# 1. Innledning

Den aktive arbeidsmarkedspolitikken har en helt sentral plass i den økonomiske politikken i Norge. Begrunnelsen er at den skal bidra til et mer effektivt og vel fungerende arbeidsmarked og derved redusert strukturell arbeidsledighet. Myndighetenes dimensjonering og sammensetning av de forskjellige ordinære tiltakene er avhengig av en rekke faktorer, hvor effektevalueringer er ansett som et viktig hjelpemiddel. I Stortingsmelding nr. 19 (Et velfungerende arbeidsmarked) heter det blant annet:

"En har de senere årene fått mye forskningsbasert kunnskap om hensiktsmessig utforming av tiltak. Tiltakene bør være av begrenset varighet og godtgjørelsen ved tiltaksdeltakelse bør være lavere enn den arbeidssøkeren kan få i ordinært arbeid. På denne måten unngås at deltakerne blir «låst inne» i tiltak ved at deltakelse i tiltak oppfattes som bedre enn å komme i ordinært arbeid. Det offentlige bør ikke fullfinansiere tiltakene. Egenandel for arbeidsgivere skal bidra til å sikre at arbeidsgiverne er interessert i å gi tiltaksdeltakerne reelle arbeidsoppgaver. Omfang og bruk av ulike typer tiltak overfor den enkelte arbeidssøker eller yrkeshemmede skal vurderes i forhold til bidraget tiltaket kan gi for å bedre muligheten til å komme i jobb. "

Statistisk sentralbyrå (SSB) foretok i perioden 1997 til 2000 løpende resultatanalyser av ordinære arbeidsmarkedstiltak, basert på informasjonen fra administrative registre. Formålet var å måle effekten av tiltaksdeltakelse på sannsynligheten for å komme jobb. Analysen omfattet alle typer ordinære tiltak, og ble gjennomført på samme måte fra år til år for å gi sammenlignbare resultater. Utgangspunktet for SSBs involvering i den løpende tiltaksevalueringen var basert på at det skulle benyttes et evalueringsopplegg som forskerne på dette tidspunktet anså som mest velegnet. Som følge av at datagrunnlaget kun var registerbasert ble kostnadene ved å utarbeide en slik analyse relativt lave og produksjonstiden kort.

I notatet "Systematisk evaluering av aktive arbeidsmarkedstiltak" av Raaum, Røed og Torp (2003), blir det gitt en vurdering av analyseopplegget som lå til grunn for den løpende evalueringen som SSB utførte. Her konkluderes det med at verdien av den type evalueringer er begrenset. Dette er i hovedsak begrunnet med følgende tre momenter:

- De statistiske metodene er for enkle. Siden 1997 har det vært store fremskritt innenfor fagfeltet, både nasjonalt og internasjonalt, når det gjelder behandling av ulike seleksjonseffekter.
- Gruppeinndelingen er for grov. Effektene er evaluert på svært heterogene grupper. Det er for eksempel ikke tatt tilstrekkelig hensyn til alder, arbeidsmarkedserfaring eller dagpengerettigheter.
- Fokus på korttidseffekter på et gitt tidspunkt kan gi misvisende indikasjoner på "de sanne" effektene av tiltak. Suksessopptak et halvt år senere er nødvendigvis ikke representativ for totaleffekten et tiltak har hatt.

Arbeids- og sosialdepartementet (ASD) har gitt SSB i oppdrag å gjøre en vurdering av hvilken type årlige analyser av ordinære tiltak basert på registerdata som kan gjennomføres. Med bakgrunn i forskning på feltet belyser denne rapporten ulike metodiske problemstillinger knyttet til systematiske effektevalueringer, og det skisseres et mulig opplegg. Notatet er utformet med tanke på at den skal utgjøre en del av beslutningsgrunnlaget for å avgjøre innretningen og omfanget av de videre systematiske analysene.

Opplegget til analyse vil være todelt. Den første delen består av en deskriptiv analyse som beskriver hvem som utgjør gruppene av tiltakdeltakere og helt ledige, samt tabeller over deltakernes status på

arbeidsmarkedet en tid etter de avsluttet tiltaket. Tabellene er ment som et supplement til den løpende statistikken fra Aetat og SSB.

Del to utgjøres av ulike effektanalyser av ordinære tiltak, etter et opplegg som implementerer nyere forskning på feltet. I de tidligere resultatanalysene utført av SSB ble en modell benyttet for å belyse en rekke problemstillinger. Det ble beregnet gjennomsnittseffekten av å delta på ordinære arbeidsmarkedstiltak sammenlignet med det å ikke delta, effekter av enkelt tiltak, forskjeller i effekten tiltakene imellom, samt forskjeller i tiltakseffekten over tid. Sammenlignet med denne fremgangsmåten argumenteres det i rapporten for å utføre enkeltanalyser av hver problemstilling, for å oppfylle de metodiske krav nyere forskning på feltet setter.

I kapittel 2 utdypes nærmere bakgrunnen for prosjektet. Det blir gitt en gjennomgang av de effektanalysene som SSB har utført til nå, og vi beskriver kort enkelte nyere effektanalyser av ordinære tiltak. Kapittel 3 omhandler metodiske sider ved evaluering av arbeidsmarkedstiltak. Vi kommer inn på valg av tiltaks- og sammenligningsgruppe, suksesskriteriet, hva og hvordan vi måler effekten, hvilke antagelser vi må gjøre og hvor realistiske de er. Forslag til beskrivende analyse presenteres i kapittel 4, med tall fra 4. kvartal 2002. I kapittel 5 skisserer vi et konkret forslag til metode for tiltaksevaluering. Forslag til tilleggsanalyse hvor vi sammenligner resultater basert på metoden skissert i kapittel 5 med tilsvarende resultater basert på metodene brukt i SSB sine tidligere evalueringer av arbeidsmarkedstiltak er bl.a. beskrevet i kapittel 6.

## **2. Effektevaluering av ordinære arbeidsmarkedstiltak til nå**

Aktive arbeidsmarkedstiltak kan påvirke flere sider ved arbeidsmarkedet. I Torp (2000) gis det en oversikt over ulike effekter. De direkte effektene er i stor grad knyttet opp mot de som deltar, og gjør seg for eksempel gjeldende i form av økt kompetanse og at mulighetene for å få jobb endres. Disse positive effektene kan forsterkes eller motvirkes av indirekte effekter. To eksempler kan være at tiltaksdeltakere fortrenger arbeidsløse arbeidssøkere, eller at adferden til aktørene i arbeidsmarkedet endres ved at det tas hensyn til sannsynligheten for bli tilbudt tiltaksplass. I tillegg kan omfanget av arbeidsmarkedstiltakene også påvirke hvilken vekt lønnsfastsetterne legger på å forebygge arbeidsledighet. I analysene til SSB har effektmålet vært knyttet til deltakernes muligheter for å få jobb. Det er med andre ord kun sett på partielle effekter av arbeidsmarkedstiltakene.

Selve analyseopplegget for denne effektevalueringen kan kort oppsummeres som følger. Målet var å undersøke i hvilken grad sannsynligheten for å gå over i jobb blir endret som følge av deltakelse på tiltak. Det ble med andre ord kun sett på partielle effekter av arbeidsmarkedstiltakene. For å besvare et slikt spørsmål etablertes to grupper, hvorav den ene bestod av personer som avsluttet tiltaksdeltakelsen seks måneder før evalueringstidspunktet, og den andre av en sammenlikningsgruppe med personer registrert som helt ledige. Ved å kontrollere for kjennetegn som kunne tenkes å påvirke tilpasningen på arbeidsmarkedet, ble effekten av tiltaket beregnet ved se på forskjellene i tilpasningen for de to gruppene.

I dette kapittelet presenteres effektanalyser gjennomført av SSB i perioden 1997 - 2000, samt en gjennomgang av analyser som er utført av norske forskningsmiljøer i den senere tid hvor tiltakseffekten for deltakerne måles.

### **2.1 SSBs registerbaserte effektevalueringer av ordinære tiltak**

SSB publiserte i perioden 1997 - 2000 resultater fra fire registerbaserte evalueringer av ordinære arbeidsmarkedstiltak. Formålet var å undersøke i hvilken grad arbeidsmarkedstiltakene bidrar til å oppfylle hovedmålsettingen om at deltakerne skal få bedret sine jobbmuligheter. Bakgrunnen for undersøkelsene ble lagt ved et forprosjekt som ble utført i Statistisk sentralbyrå i 1996 og publisert i Vassnes (1996). Her ble det vurdert i hvilken grad data fra administrative kilder kunne erstatte data fra spørreskjema ved evaluering av arbeidsmarkedstiltak. Resultatene viser at opplegget for evaluering av arbeidsmarkedstiltak basert på registerdata hovedsakelig gir samme resultat som tilsvarende undersøkelse med spørreskjema.

Sentralt i prosjektet var å få fram:

- en analyse som omfattet alle ordinære tiltak,
- aktuell statistikk innenfor en relativt kort produksjonstid,
- sammenliknbare resultater på tvers av tiltakene,
- sammenliknbare resultater over flere år.

Evalueringen skulle omfatte beregninger av gjennomsnittseffekten av å delta på et ordinært tiltak, effekten av å delta på en definert hovedgruppe av tiltak, og effekten av deltakelse på enkelt tiltak.

#### **Datagrunnlaget**

En rekke administrative kilder ble benyttet i analysen. For å identifisere tiltaksdeltakere og konstruere en sammenlikningsgruppe ble Aetats SOFA-søkerregister over personer registrert som ledige av arbeidskontorene benyttet. Registeret ble også benyttet for å konstruere ledighets- og tiltakshistorikk. Informasjon om arbeidsforhold ble hentet fra Rikstrygdeverkets arbeidstakerregister. I tillegg ble det innhentet informasjon fra registeret over personer høyeste utdanning, fødelandsregisteret m.m.

### **Gruppering av ordinære tiltak**

Følgende hovedgrupper av tiltak var inkludert i evalueringsopplegget:

- *Offentlige sysselsettingstiltak*
- *Lønnstilskudd for arbeidsgivere*
- *Praksisplasser med fadderordning*
- *Vikarplasser*
- *Arbeidsmarkedsopplæring (AMO)*
- *Jobbklubber*

I perioden evalueringene ble foretatt trappet myndighetene ned offentlige sysselsettingstiltak og vikarplasser i så stor grad at de ikke inngikk i effektevalueringene for 1999 og 2000.

### **Avgrensning av populasjon og utvalg**

Populasjonen i undersøkelsene bestod av en tiltaksgruppe av deltakere og en sammenlikningsgruppe.

Tiltaksgruppen ble definert som :

- Personer som avsluttet et ordinært arbeidsmarkedstiltak i løpet av april det aktuelle året.

Sammenlikningsgruppen ble definert som :

- Personer som i slutten av april det aktuelle året var registrert som helt ledige arbeidssøkere og som ikke var registrert på noen form for tiltak i slutten av mai samme år.

Utvalget ble videre selektert ved at permitterte ble tatt ut av sammenlikningsgruppen, og de som var registrert som yrkeshemmede eller deltar på attføringstiltak det aktuelle året ble tatt ut av begge gruppene.

### **Suksesskriterium**

Med bakgrunn i at det primære målet med tiltakene er å få flest mulig ut i arbeid, ble det vurdert som mest nærliggende å benytte sysselsetting som suksesskriterium i effektevalueringene. Informasjon om sysselsetting ble kun hentet fra arbeidstakerregisteret. Arbeidsforhold som selvstendig næringsdrivende eller vernepliktige var ekskludert fra undersøkelsen, fordi produksjonstiden på registrene var for lang.

### **Evalueringstidspunkt**

Evalueringstidspunktet var satt 6 måneder etter tiltaket ble avsluttet. Dette blir videre omtalt som korttidseffekten av å delta på tiltak. Det ble i Bråthen og Landfald (1999) foretatt en analyse av effekten av tiltak målt også 30 måneder etter de ble avsluttet.

### **Resultater**

Tabell 1 samler beregningene fra de fire årene analysen ble foretatt. Den gir estimert jobbsannsynlighet for en referanseperson med gitte bakgrunnsvariable<sup>1</sup> som henholdsvis var helt ledig ved utgangen av april, eller som avsluttet et ordinært tiltak i en av de seks hovedgruppene på samme tidspunkt.

Hovedkonklusjonen for alle de fire evalueringene var at deltakerne på et ordinært arbeidsmarkedstiltak i gjennomsnitt hadde en høyere estimert sannsynlighet for å komme i jobb seks måneder senere enn sammenlikningsgruppen av helt ledige. Tiltakseffekten, det vil si differensen i jobbsannsynlighet mellom de to gruppene, varierte mellom 11 til 14 prosentpoeng i de fire analysene. Tiltakene som

---

<sup>1</sup> Referansepersonen i analysene er i utgangspunktet mann, 30-49 år, ugift, ikke innvandrere, har Videregående skole 11-12 år, 4 års yrkeserfaring, 5 mnd ledighetserfaring, 1 mnd tiltakserfaring, 1 mnd delvis sysselsatt, 0 mnd erfaring attføringstiltak, har ikke vært yrkeshemmet siste 12 mnd, har industri som yrkesbakgrunn og er bosatt i Oslo.



inngikk i gruppen lønnstilskudd var de som hadde størst effekt, mens praksisplasser med fadderordning hadde liten eller ingen effekt på jobsannsynligheten.

**Tabell 2.1. Beregnede jobsannsynligheter for personer som avsluttet et ordinært arbeidsmarkedstiltak i april for årene 1996 til 1999.**

Tiltaksgruppe	Referanse-	Offentlige	Lønns-	Praksispl. m.		Jobb-	
Tiltaksår	person	sys. tiltak	tilskudd	fadderordning	AMO	klubber	Vikarplasser
<b>Jobsannsynligheter beregnet 6 måneder etter avsluttet tiltak</b>							
1996	0,27	0,35	0,58	0,31	0,37	0,42	0,53
1997	0,29	0,49	0,66	0,37	0,39	0,43	0,49
1998	0,32	0,64	0,61	0,34*	0,40	0,45	0,55
1999	0,25	-	0,58	0,27*	0,35	0,42	-
<b>Jobsannsynligheter beregnet 30 måneder etter avsluttet tiltak</b>							
1996	0,46	0,52	0,65	0,54	0,55	0,56	0,67

Jobsannsynligheter merket med \* indikerer at estimatet ikke er signifikant på 5 prosent nivå.

## 2.2 Nyere effektanalyser av ordinære tiltak

Hovedmønsteret når det gjelder effekten av de ulike tiltaksgruppene presentert i forrige avsnitt, kan gjenfinnes både i nasjonale og internasjonal litteratur som beregner gjennomsnittseffekten for alle som faktisk deltar. Som nevnt i Raaum et al (2002) er det for eksempel i ferd med å etablere seg en slags enighet i den internasjonale forskningslitteraturen om at lønnstilskudd til private bedrifter som tar imot ledige er tiltaket med størst effekt for deltakerne, og at tiltak i offentlig sektor er det minst effektive jmfør Martin og Grubb (2001), Calmfors et al (2002) og Schmidt et al (2001). Det advares imidlertid mot å trekke for bastante konklusjoner av flere grunner. Et argument som nevnes er de ulike seleksjonsmekanismer inn til disse tiltakstypene, som neppe er luket helt ut av de effektevalueringene som ligger til grunn for disse vurderingene.

Både nasjonalt og internasjonalt har forskningen på feltet vært særlig fokusert på å utarbeide metoder som fanger opp seleksjonsskjevheter. Det er et velkjent problem i analyser av denne typen at det kan være systematiske forskjeller mellom tiltaksdeltakere og sammenlikningsgruppen av ikke- deltakere som vi ikke kan avsløre ved våre beskrivelser, enten fordi datamaterialet ikke er tilstrekkelig eller fordi enkelte ulikheter kan være vriene å observere. Seleksjonsskjevheter kan oppstå dersom disse uobserverbare kjennetegnene ved personene både har betydning for om de blir valgt ut til å delta, og i tillegg påvirker tilpasningen på arbeidsmarkedet.

I notatet "Systematisk evaluering av aktive arbeidsmarkedstiltak" er det gitt en vurdering av SSBs tidligere analyseopplegg. Her argumenteres det for at modellen som er valgt i for liten grad tar hensyn til seleksjonsmekanismer. Enkelte nyere norske studier, som Hardoy (2003), Raaum, Torp og Zhang (2002a, b), benytter empiriske strategier hvor seleksjonsproblematikken blir behandlet med en mer systematisk måte. Utgangspunktet er en antagelse om at de uobserverbare kjennetegnene som både påvirker rekrutteringen til tiltak og eventuell suksess på arbeidsmarkedet har sammenheng med observerbare kjennetegn. Sammenlikningsgruppen av ikke-deltakere som skal representere tiltaksdeltakernes forløp dersom de ikke hadde gått på tiltak, etableres ved å beregne såkalte matchingestimatorer. En nærmere beskrivelse av både hvilke forutsetninger som ligger til grunn for metodene, samt ulike fremgangsmåter som kan benyttes er gitt i kapittel 3.

Av resultatene fra de nevnte norske studiene, bekrefter for øvrig Raaum et al (2002a) resultatene fra tidligere analyser basert på både eksperimentelle og ikke-eksperimentelle data når det gjelder arbeidsmarkedsopplæring (AMO). Her ble den gjennomsnittlige effekten på arbeidsinntekten som følge deltakelse på tiltaket beregnet for deltakerne i perioden 1991-1996. Analysene viste at deltakelse har en signifikant positiv effekt på arbeidsinntekten for deltakere med rett på dagpenger, dvs personer som har sysselsettingserfaring rett forutfor ledighetsperioden. For nykommere i arbeidsmarkedet, de som ikke har rett til dagpenger, er det liten eller ingen inntektseffekt. I Raaum et al (2002b) er analysen knyttet opp mot å måle om tiltakenes positive effekt på inntekten faktisk overstiger de direkte kostnadene knyttet til å gjennomføre tiltakene. Resultatene viser at effektene faktisk oversteg disse, kostnadene for populasjonen av deltakere med rett til dagpenger. Analysen gjort av Hardoy (2003) antyder at tiltakene rette spesielt mot ungdom i liten grad har effekt på senere tilpasning på arbeidsmarkedet.

### 3. Metodiske utfordringer ved evalueringsstudier

Forutsetningen om lik tiltakseffekt for alle individer, ev. betinget på et sett av forklaringsvariable i tradisjonelle regresjonsanalysen skaper problemer. Vi kan få skjeve effekter dersom deltakere sammenlignes med ikke-deltakere som har ingen eller svært liten sjanse for å havne på tiltak. Dette blir i litteraturen ofte kalt mangel på "common support". Ulik sammensetning av de to gruppene kan videre føre til feil vekting av sammenligningsgruppen, med gale gjennomsnittseffekter som resultat.

En metodisk tilnærming som forsøker å fjerne disse to feilkildene, går ut på å matche deltakere med ikke-deltakere slik at sammenligningsgruppen er "likest" mulig deltakerne i betydningen samme fordeling av observerbare forklaringsvariable. Ulike matching metoder er mye brukt i forskningsarbeider om evaluering av arbeidsmarkedstiltak i inn- og utland. Problemer med uobserverbar heterogenitet er også til stede i matchingstudier, men rikelig med data og god matching prosedyre kan bidra til å redusere problemet.

Tidligere forsøkte vi å evaluere tiltakene samlet, enkelt tiltak og forskjellen i effekt tiltakene imellom ved hjelp av en og samme modell. I og med at seleksjonsprosessen til ulike typer tiltak kan være forskjellig, vil separate analyser for hver tiltakstype kunne være nødvendig for å analysere de tre problemene over tilfredstillende.

Videre i kapitlet ser vi nærmere på viktige metodiske problemstillinger som må drøftes i forbindelse med tiltaksevaluering.

#### 3.1 Effekten av hva ?

Målet med den typen effektanalyse vi legger opp til her er å måle forskjellen i suksess for de som har deltatt på (et) tiltak, i gjennomsnitt, i forhold til dersom de som deltok ikke hadde vært på tiltak(et). Denne typen mål er det mye brukt i evalueringslitteraturen. På engelsk betegnes ofte målet med ATET av "average treatment effect on the treated". Målet sier noe om gjennomsnittet for de som **har deltatt**. Målet sier ikke noe om hvordan de som ikke er på tiltak ville ha gjort det dersom de var på tiltak. Målet sier heller ikke noe om effekten dersom omfanget av personer på tiltak endres. Målet sier heller ikke noe om hvilken suksess deltakerne på et annet tiltak ville ha hatt dersom de i stedet hadde gått på det evaluerte tiltaket.

I de tidligere rapportene fra SSB sammenlignet vi jobbsannsynligheter for referansepersoner som hadde vært på tiltak med for referansepersoner som ikke hadde vært på tiltak. Referansepersonen var en tenkt person som for hver variabel fikk tilordnet den mest hyppige verdi<sup>2</sup>. Slik sett ble referanseperson konkret og intuitiv, men ikke nødvendigvis representativ for gjennomsnittet av de som deltok på tiltak. Manglende representativitet av tiltaksdeltakerne er årsaken til at vi ønsker å endre det vi måler i de nye analysene.

Om vi ønsker å måle effekten av å ha gått på et av tiltakene i en tiltaksgruppe sett i forhold til fravær av tiltak bør en analysere tiltaksgruppen separat, i og med at seleksjonsprosessen til ulike typer tiltak kan være forskjellig. Egne analyser av mer homogene undergrupper, som ungdom/voksne, menn/kvinner og personer med/uten dagpengerettigheter kan også være en fordel. Årsaken er at individer i ulike undergrupper kan i gjennomsnitt ha forskjellige motiver for å velge de enkelte tiltak, og fordi saksbehandlerne på arbeidskontorene kan vurdere ulike undergrupper forskjellig. Bakgrunnen for bruk av undergrupper er for det første at de bakenforliggende effektene av de observerbare forklaringsvariable kan være forskjellig for ulike undergrupper. Dessuten fanger

---

<sup>2</sup> Referansepersonen i SSB sine tidligere studier er ugift, i alderen 30-49 år, og med videregående skole nivå II (11-12 år) som høyeste fullførte. Videre er referansepersonen ikke innvandrere, har fire års yrkeserfaring fra industri, fem måneders ledighetserfaring, en måneds tiltakserfaring, og en måneds erfaring som delvis sysselsatt. Referansepersonen har ikke vært yrkeshemmet, og er bosatt i Oslo.

estimerte effekter av observerbare forklaringsvariable også opp evt. uobserverbare variable av betydning for tiltaksdeltakelse dersom er korrelert de observerbare forklaringsvariablene. De ovennevnte korrelasjonene kan være forskjellig for f.eks. ungdom/voksne, menn/kvinner og for personer med/uten dagpengerettigheter. Det er grunn til å tro at sannsynligheten for å komme på tiltak er korrelert med egenskaper som har betydning for senere jobbsannsynlighet.

Sammenligning av effekten av ulike tiltaksgrupper opp mot hverandre krever separate matching studier av hver enkelt kombinasjon pga. potensiell ulik seleksjon til de ulike tiltaksgruppene og fordi fordelingen av forklaringsvariable potensiell kan være ulik. Det blir derfor generelt ikke riktig å sette effekten av tiltak sett i forhold til ikke-deltakelse for de ulike tiltaksgruppene opp mot hverandre. Bakgrunnen er at de som går på ulike grupper av tiltak i gjennomsnitt kan være forskjellig.

Ønsker vi imidlertid å måle effekten av tiltakene og systemet for rekruttering inn på tiltak samlet kan det være tilstrekkelig å ha en modell som dekker alle tiltakene.

### 3.2 Definere tiltaksgruppen

I SSBs tidligere evalueringsstudier ble tiltaksdeltakere identifisert ut fra personer som avsluttet et tiltak i april de ulike årene. Sammenligningsgruppen ble konstruert ut fra populasjonen av personer registrert som helt ledige på samme tidspunkt. Det tas altså utgangspunkt i tidspunktet for utstrømming fra tiltak. Ved å ikke inkludere tidsrommet hvor deltakerne går på tiltak i resultatperioden, det vil si tiden fra uttrekkstidspunktet til evalueringstidspunktet, kan det argumenteres for at avgrensningen førte til en overvurdering av beregnede tiltakseffekten.

Resonnementet går på at en arbeidsledig som ikke deltar på tiltak i større grad vil søke jobb i tidsrommet for tiltaket enn en tiltaksdeltaker. For å forbedre populasjonsavgrensningen av vil SSB definere tiltaksgruppen ut fra starttidspunktet av tiltaket, dvs. deltakere til de som påbegynte et ordinært tiltak i løpet av en periode.

I dette avsnittet vil vi gi enkelte tabeller for å illustrere mulig størrelse på tiltaksgruppene ved en slik tilnærming. Når vi senere i kapittel 4 gir eksempler på beskrivende analyser, vil de ta utgangspunkt i populasjonen definert her.

Lengden på uttrekksperioden beskrevet under er henholdsvis en måned og et kvartal. Tabellene nedenfor gir tall for personer som påbegynte et ordinært tiltak i løpet av oktober<sup>3</sup> 2002, samt personer som påbegynte et tiltak i fjerde kvartal<sup>4</sup> 2002. Personer vurdert som yrkeshemmede av Aetat er utelatt.

**Tabell 3.1 Personer som påbegynte et ordinært tiltak i løpet av oktober etter tiltakstype. 2002. Antall og prosent.**

Arbeidssøkerstatus	Antall	Prosent
I alt	3 084	100
1. Lønntilskudd	305	9,9
2. Arbeidspraksis	858	27,8
3. Opplæring	1 904	61,7
4. Midlertidige sysselsettingstiltak	17	0,6

<sup>3</sup> Personer registrert på et ordinært tiltak i oktober 2002, men ikke i september samme år.

<sup>4</sup> Personer registrert på et ordinært tiltak i oktober 2002, men ikke i september samme år, personer registrert på et ordinært tiltak i november 2002, men ikke i september og oktober samme år, personer registrert på et ordinært tiltak i desember 2002, men ikke i september, oktober samme år.

**Tabell 3.2 Personer som påbegynte et ordinært tiltak i løpet av oktober etter tiltakstype, fin inndeling. 2002. Antall og prosent.**

Arbeidssøkerstatus	Antall	Prosent
I alt	3 084	100
1. Lønnstilskudd	305	9,9
2. Arbeidspraksis i ordinær virksomhet	858	27,8
3. Arbeidsmarkedsopplæring (AMO)	1 868	60,6
3. Arbeidsmarkedsopplæring (AMO), enkeltplasser	14	0,5
3. Jobbklubb	22	0,7
4. Sysselsettingstiltak for langtidsledige	17	0,6

**Tabell 3.3 Personer som påbegynte et ordinært tiltak i løpet av 4. kvartal etter tiltakstype. 2002. Antall og prosent.**

Arbeidssøkerstatus	Antall	Prosent
I alt	7 538	100,0
1. Lønnstilskudd	924	12,3
2. Arbeidspraksis	2 682	35,6
3. Opplæring	3 900	51,7
4. Midlertidige sysselsettingstiltak	31	0,4
Andre arbeidssøkere	1	0,0

**Tabell 3.4 Personer som påbegynte et ordinært tiltak i løpet av 4. kvartal etter tiltakstype, fin inndeling. 2002. Antall og prosent.**

Arbeidssøkerstatus	Antall	Prosent
I alt	7 538	100,0
1. Lønnstilskudd	924	12,3
2. Arbeidspraksis i ordinær virksomhet	2 682	35,6
3. Arbeidsmarkedsopplæring (AMO)	3 790	50,3
3. Arbeidsmarkedsopplæring (AMO), enkeltplasser	61	0,8
3. Jobbklubb	49	0,7
4. Sysselsettingstiltak for langtidsledige	28	0,4
4. Vikarplass	3	0,0
5. Nye plasser institusjonelle tiltak	1	0,0

Valget av lengden på uttrekksperioden har flere effekter. Desto kortere uttrekksperiode, jo større er sannsynligheten for målefeil som følge størrelsen på utvalget. Som tabellene ovenfor viser, vil det å benytte en måned mer enn halvere gruppen tiltaksdeltakere. Bakdelen med en lengre uttrekksperiode, er at den øker sannsynligheten for skjevheter i utvalget når det gjelder forskjelle i de uobserverbare kjennetegnene som både påvirker rekrutteringen til tiltak, og den senere tilpasningen på arbeidsmarkedet.

SSB velger å avgrense populasjonen av tiltaksdeltakere til de som påbegynte et ordinært tiltak 4. kvartal 2002.

En annen avgrensning av populasjonen av både tiltaksdeltakere og sammenligningsgruppene er å fjerne personer over 59 år, alternativt utelukke de over 54 år. I tiltaksgruppen er det kun 276 personer over 54 år blant de som påbegynte ordinære tiltak i 4. kvartal 2004. De er trolig en selektert gruppe som skiller seg ut fra de yngre, bl.a. ved at de nærmer seg (førtids)pensjonsalder, og at de kanskje har større problem med å få jobb. Dessuten øker andelen uførepensjonister vesentlig med alderen over 50 år. I sammenligningsgruppen er det en god del personer over 54 år. Det kan være uheldig om disse personene matches mot andre langt yngre personer i tiltaksgruppen som har lav sannsynlighet for å havne på tiltak.

Vi vil også vurdere andre avgrensninger av populasjonen av tiltaksdeltakere og sammenligningsgruppen etter hvert som vi blir mer kjent med datamaterialet om vi mener det er nødvendig for å få en god evaluering. Eksempel på dette er kan være personer som begynner i førstegangstjeneste i søkeperioden, minimumsalder i populasjonene på 16 år eller ev. krav om at personene har vært registrert arbeidsledig et vist antall uker før 4. kvartal 2002. Slike populasjonsavgrensninger fører til at vi ikke evaluerer effekten av ordinære arbeidsmarkedstiltak for samtlige personer som begynte på tiltak i 4. kvartal.

### 3.3 Gruppering av tiltak

Nedenfor følger en gruppering av tiltakene rettet mot ordinære arbeidssøkere, som blir benyttet i Stortingsmelding 19.

- **Lønnstilskudd** kan i en begrenset periode gis til arbeidsgivere som ansetter arbeidssøkere som har særlige problemer med å komme inn på arbeidsmarkedet på ordinære lønns- og arbeidsvilkår med sikte på varig arbeidsforhold.
- **Arbeidspraksis** bidrar til å prøve ut og styrke arbeidssøkernes muligheter på arbeidsmarkedet og styrke deltakernes muligheter til å komme i arbeid eller utdanning. Arbeidspraksis foregår i ordinær eller i skjermet virksomhet i form av tilrettelagt arbeidstrening med oppfølging.
- **Arbeidsmarkedsopplæring (AMO)** bidrar til at arbeidssøkere kvalifiseres til ledige jobber. Tiltaket kan også brukes for å hindre utstøting av arbeidstakere som står i fare for å falle ut av arbeidslivet. Opplæring og oppfølging i jobbklubber mv. faller inn under dette tiltaket.
- **Midlertidige sysselsettingstiltak** gir personer som står i fare for å bli varig utestengt fra arbeidslivet eller som har problemer med å etablere seg på arbeidsmarkedet pga manglende arbeidserfaring. Tiltaket gis i form av sysselsettingstiltak for langtidsledige, sysselsettingstiltak for yrkeshemmede eller vikar plass.
- **Bedriftsintern opplæring** styrker kompetansen til ansatte i bedrifter som har omstillings- eller strukturproblemer som er av omfattende karakter eller støtter rekruttering til ledige stillinger som er vanskelige å besette. Bedriftsintern opplæring inneholder teori, eventuelt kombinert med praktisk trening.

Målsettingen for analysen er at den skal være så bred som mulig, dvs. dekke flest mulig ordinære tiltakene. Det tas utgangspunkt i grupperingen ovenfor. Som tabellene i foregående avsnitt viste er antall deltakere på midlertidige sysselsettingstiltak og bedriftsintern opplæring er lavt. I tillegg har vi jobbklubber er kortvarige tiltak som skiller seg fra andre AMO kurs, og vil derfor bli holdt utenfor denne gruppen i tiltaksevalueringene. SSBs analyse vil derfor begrense seg til å omfatte Lønnstilskudd, Arbeidspraksis i ordinær virksomhet og Arbeidsmarkedsopplæring (AMO) uten jobbklubber.

### 3.4 Definere sammenlikningsgruppe - matching

Målet med den typen effektanalyse vi legger opp til her er å måle forskjellen i suksess for de som deltar på (et) tiltak, i gjennomsnitt, i forhold til dersom de ikke hadde vært på tiltak(et). I og med at det kontrafaktiske utfallet ikke er observerbart, defineres som før nevnt en sammenlikningsgruppe av personer som skal representere deltakernes forløp om de ikke hadde gått på tiltak. Sammenlikningsgruppen avgrenses til personer som var registrert helt arbeidsledig i september samme år som deltakerne er på tiltak, i tillegg til at de ikke ha deltatt arbeidsmarkedstiltak i fjerde kvartal. Sammenlikningsgruppen må være representativ for de som er på tiltak for at evalueringen skal bli korrekt. Det er et velkjent problem i analyser av denne typen at det kan være systematiske forskjeller mellom tiltaksdeltakere og sammenlikningsgruppen av ikke - deltakere som vi ikke kan avsløre ved våre beskrivelser, simpelthen fordi enkelte ulikheter kan være vriene å observere eller de er uobserverbare. Seleksjonsproblemer kan oppstå dersom det er spesielle uobserverbare kjennetegn ved personene som har betydning for om personene blir valgt ut til å delta på et tiltak og som også påvirker jobsannsynligheten. Hvis en hadde hatt randomisert trekking av tiltaksdeltakere og sammenlikningsgruppe ville man unngå systematiske forskjeller gruppene imellom, noe som ville vært et godt utgangspunkt for forventningsrette effektestimater.

I fravær av randomisert trekking av tiltaksdeltakere, vil vi forsøke å løse identifikasjonsproblemet på en annen måte. Innenfor forskningen på feltet er det gjort en rekke fremskritt når det gjelder å behandle slike seleksjonsproblemer. Det er i den senere tid vært fokusert på bruk av matching. Dette er en metode for å danne en sammenlikningsgruppe på ikke-eksperimentelle data. Strategien går kort sagt ut på å betinge utvalget av tiltaksdeltakere og personer i sammenlikningsgruppen på et sett av observerbare kjennetegn. Dette for å fange opp forskjeller som kan påvirke rekrutteringen til tiltaket som evalueres, og som igjen kan innvirke på beregningen av effekten ved tiltaksdeltakelse.

En utfordring med matching prosedyren er å definere hvilke observerbare kjennetegn som påvirker rekruttering til tiltak<sup>5</sup>. For eksempel er trekking av deltakere og ikke-deltakere fra samme lokale arbeidsmarked ønskelig slik at de har relativt like muligheter til å oppnå suksess.

Dersom deltakerne har karakteristika som er uobserverbare for SSB, og som gir deltakerne fordeler (ulempen) når det gjelder å oppnå suksess sammenlignet med kontrollgruppen, vil det kunne føre til skjeve estimater. Det vil si at effekten av tiltakene vil kunne overvurderes (undervurderes).

Utvalgsskjevhet kan følge både av uobserverbar heterogenitet mellom deltakere og ikke deltakere, og ved at forskjeller i observerte egenskaper/kjennetegnene har ulik fordeling mellom de to gruppene, jfr Raaum et al (2002a). Riktig matching tar bort sistnevnte evalueringsskjevhet, mens den første består. Rikt datamateriale og god matchingprosedyre kan redusere ev. skjevhet knyttet til uobserverbar heterogenitet.

Matching metoden bygger på en forutsetning om at vi har et sett observerbare variable,  $X$ , slik at kontrollert for dem, så er de potensielle suksessutfallene uavhengig av tiltaksindikatoren. Dette blir kalt conditional independence assumption i engelskspråklig litteratur, med forkortelsen CIA. I tillegg må vi forutsette at det finns tiltaksdeltakere og personer i sammenlikningsgruppen for alle verdier av de observerbare variablene  $X$ , dvs. at alle har en sannsynlighet for å ha vært tiltaksdeltaker betinget på forklaringsvariablene  $X$ , som ligger strengt mellom 0 og 1.

Det vil som før nevnt bli benyttet resultater og erfaringer fra andre norske forskningsprosjekter for å velge en metode som er mest hensiktsmessig gitt datamaterialet tilgjengelig. I nyere litteratur på området er propensity score matching<sup>6</sup> en mye brukt metode. Hver tiltaksdeltaker blir her forsøkt matchet med en ikke-deltaker med observerte egenskaper som gir personen en sannsynlighet for å havne på tiltak som er tettest opptil deltakeren selv. Vi får da en sammenlikningsgruppe med fordeling av kjennetegn som påvirker tiltaksdeltakelse som er svært lik den en finner hos de som deltar.

Propensity score matching er eksempelvis benyttet i de norske analysene til Raaum et al (2002b) og Hardoy (2003), og i analysen til Larsson (2003) som er basert på svenske data. Fremgangsmåtene er imidlertid noe forskjellige.

Propensity score matching er blitt en populær metode, men nødvendigvis ikke en enkel metode for å få forventningsrette estimater i praksis. Smith og Todd (2005) finner at flere matching estimatorer på tverrsnittsdata er svært sensitive for valg av variable brukt til å estimere propensity scorene og valg av subgruppering. De skriver at mulige forklaringer på sensitiviteten bl.a. er at tiltaksdeltakere ikke er matchet med sammenlikningspersoner fra samme lokale arbeidsmarked.

I en ny litteraturoversikt av Imbens (2004) gjennomgås ulike metoder for effektestimering av gjennomsnittlig behandling (tiltak) gitt eksogenitet. Han skiller mellom regresjonsestimatorer, matching estimatorer, ulike propensity score metoder og kombinasjoner av metodene over. Debatten

---

<sup>5</sup> Ikke inkluder variable som følger av selve tiltaksdeltakelsen, jamfør Heckman et. al(1999, s. 1956).

<sup>6</sup> Metoden er velegnet dersom en har et rikt datamateriale, slik at direkte matching (dvs. kobling) på alle forklaringsvariablene gir relativt få treff.



om de praktiske fordelene ved de ulike metodene pågår fortsatt, og han kommer ikke med en konklusjon om hvilken type metode som er best.

### **Plausibiliteten av matching antagelsen for datamaterialet**

Forutsetningen i matchingstudier om betinget uavhengighet(CIA) er ikke testbar. Vi vil derfor forsøke å sannsynliggjøre troverdigheten av antagelsen. Avgjørende for troverdigheten av vår forutsetning er at vi har rikelig med relevante forklaringsvariable som vi kan betinge på, slik at der er troverdig at potensielt suksessutfall er uavhengig av tiltaksindikatoren, gitt forklaringsvariablene.

Personens utdanningsnivå, tidligere arbeidserfaring og tidligere arbeidsledighetserfaringer er trolig viktige observerbare faktorer for å bestemme om et individ vil delta på tiltak og tiltakstype. Disse faktorene vil sannsynligvis også påvirke framtidig utfall på arbeidsmarkedet, Jamfør Larrson (2003). For at betinget uavhengighetsforutsetningen skal være plausibel bør derfor faktorene være med i beregningene av tiltakssannsynlighetene.

Andre tilgjengelige faktorer som ifølge litteraturen bør være med for å gjøre CIA troverdig er eksempelvis:

En indikator om personer har dagpengerettigheter påvirker også trolig insentivene for å delta på tiltak og for å komme i jobb. Flere studier har med en slik variabel enten i propensity score beregningene eller via separate analyser for personer med og uten dagpengerettigheter.

Indikatorer for det lokale arbeidsmarkedet er også viktig å ha med i matchingen. Viktigheten av å unngå geografisk mismatch mellom tiltaksdeltakere og personer fra sammenligningsgruppen poengteres bl.a. i Heckman et al(1997) artikkel. De påpeker også at detaljert inntektshistorikk også er viktig for at CIA er plausibel.

Sammenlignet med mange andre norske og utenlandske studier er våre data gode, slik at vi mener CIA-antagelsen vil være plausibel i vår studie. Vi tror at vi har tilstrekkelig med avgjørende informasjon til å rettfærdiggjøre CIA-antagelsen.

Enkelte svenske studier har imidlertid også informasjon om hvor parat den ledige er for å ta arbeid ifølge saksbehandler på arbeidskontoret. Det informasjon vi ikke har, men som kunne ha vært nyttig i evalueringsoyemed.

Raaum et al (2002b) ser de ikke på deltakelse i et spesifikt tiltak som utfall av et enkelt binært valg eller som en seleksjonsprosess med to gjensidig utelukkende utfall. Deltakere og ikke deltakere på det evaluerte arbeidsmarkedstiltaket matches ved sannsynligheten for a) deltakelse i tiltaket som evalueres b) deltakelse i andre tiltak c) forlater ledighetsregisteret, alt som alternativ til d) være i ledighetsregisteret. De mener at matching av utvalget på to av disse sannsynlighetene gjør CIA mer plausibel. De senere årene er multinomisk propensity-score matching brukt en del i evalueringsslitteraturen.

Flere studier foretar separate analyser av enkelte undergrupper av populasjonen. Slike subgrupperinger kan være skille mellom unge og voksne, personer med og uten dagpengerettigheter, og mellom menn og kvinner. Dette vil gjøre undergruppene mer homogene enn totalen. Til gjengjeld blir det mange undergrupper å analysere separat, med tilhørende ekstra ressursbruk. En av hovedinnvendingene mot SSB sitt tidligere analyseopplegg var at vi analyserte for grove heterogene grupper samlet. Vi vil derfor foreta separate analyser for mer homogene undergrupper.

### 3.5 Valg av suksesskriterium

Avsnittet nedenfor er hentet fra Stortingsmelding nr. 19 (Et velfungerende arbeidsmarked)

"Bruken av arbeidsmarkedstiltak skal baseres på individuell vurdering, tiltakskjeding og helhetstenkning. Tiltaksstrukturen og betydelig frihet i regelverket gir det enkelte Aetat lokal mulighet til å tilpasse tiltakene til brukernes behov. Rett valg av tiltak skal fremme rask overgang til jobb. Selv om hovedparten av tiltakene kan brukes både av ordinære arbeidssøkere og yrkeshemmede, så er yrkeshemmede tilgodesett med både høyere tilskuddssatser og lengre varighet på tiltakene. De særlige tiltakene for yrkeshemmede skal sikre et tilbud som gjenspeiler den store variasjonen i behovet blant personene i denne gruppen. Det vises til nærmere omtale i avsnitt 8.5."

Det er nærliggende å tro at myndighetene ønsker effektanalyser hvor forskjeller i jobbsannsynligheter er et ønsket suksesskriterium. Det er imidlertid problemer beheftet med hvor representativt dette kriteriet er. For enkelte tiltak kan formålet være todelt. Et eksempel er arbeidsmarkedstiltak hvor målgruppen er for ungdom. I tillegg til å forbedre sannsynligheten for å lykkes på arbeidsmarkedet, er målsettingen også å øke deltakernes mulighet for å komme ordinær utdanning. I en effektanalyse av arbeidsmarkedstiltak for ungdom Hardoy (2003) defineres **suksess** i ut ifra om personen er **sysselsatt eller i utdanning** på måletidspunktet.

Målet om kort produksjonstid gjør at vi i praksis velger sannsynligheten for å være registrert arbeidstaker i registeret i november året etter som suksesskriterium og ev. sannsynligheten for å være arbeidstaker i november to år etter. Indirekte ser vi da på det å bli selvstendig næringsdrivende ikke som suksess. Denne utilfredstillende avgjørelsen er tatt fordi vi ikke har data om selvstendig næringsdrivende tidlig nok.

Et tilleggskriterium vi vil ha er om personen har fått arbeidsinntekt registrert i lønn- og trekkoppgaveregisteret for å få en ekstra kontroll på at det er reelle arbeidsforhold vi finner.

#### ***Evalueringstidspunkt***

Vi ønsker å benytte samme referansetidspunkt i november som den vanlig sysselsettingsstatistikken benytter. Evalueringstidspunktene blir da november året etter tiltaksdeltakelse og november to år etter. Valg av samme referansetidspunkt som den vanlige sysselsettingsstatistikken er gjort for å spare kostnader. Dessuten er den sysselsettingsstatistikken klar utpå våren, slik at resultatene blir mest mulig oppdatert for rapportering på sensommeren.

## **4. Forslag til opplegg for en årlig, beskrivende analyse av deltakere på ordinære arbeidsmarkedstiltak**

En registerbasert effektevaluering av ordinære tiltak krever at det settes sammen informasjon fra en rekke administrative kilder. Dette gir også muligheter for å produsere tabeller som beskriver sider ved sammensetningen av gruppen tiltakdeltakere og de helt ledige, som ikke er tilgjengelig i den løpende statistikken til Aetat og SSB. De tidligere analysene fra SSB inneholdt deskriptiv statistikk for tiltaksdeltakerne og sammenligningsgruppen av helt ledige etter en rekke kjennetegn. Disse kjennetegnene inngikk også som forklaringsvariable i selve effektevalueringen.

I notatet "Systematisk evaluering av aktive arbeidsmarkedstiltak" (Raaum, Røed Torp 2003) hevdes det at analysen benytter et svært begrenset sett av observerbare kjennetegn som nok ikke forklarer tiltaksdeltakelse i særlig grad. Et moment i utarbeidelsen av et nytt opplegg vil bli å bestemme hvilke variable som skal inngå for å:

1. definere den aktuelle populasjonen,
2. benyttes som forklaringsvariable,
3. definere suksesskriterium (angi status på arbeidsmarkedet).

I denne fasen av arbeidet vil vi skissere noen tabellforslag. Den beskrivende analysen i det ferdige evalueringsopplegget vil i stor grad være knyttet opp mot hvilke forklaringsvariable som kommer til inngå i modellen(e). Populasjon som benyttes består av deltakere som påbegynte et ordinært tiltak 4. kvartal 2002 (se kapittel 3.2). Tabellene inneholder også tall over personer registrert som helt ledige i september, og som ikke påbegynte et ordinært tiltak i løpet av fjerde kvartal 2002. Det er blant disse sammenligningsgruppen vil bli trukket. I tillegg til å fungere som en illustrasjon på hva den beskrivende analysen vil inneholde, vil tallgrunlaget også være til hjelp i arbeidet med å utforme den endelige evalueringsmodellen.

Hvis vi ser bort fra selve effektevalueringen, er det også mulige å lage alle de tabellene som her presenteres for alle tiltaksdeltakere på et gitt tidspunkt, samt alle personer registrert som helt ledige. Dette som et supplement til den løpende statistikken.

### **4.1 Sammensetningen av tiltaksdeltakere og sammenligningsgruppen av helt ledige**

#### **4.1.1 Demografiske kjennetegn**

I SSBs analyser ble tiltaksdeltakere og sammenligningsgruppe fordelt på kjønn, alder og bosted, i tillegg til sivil status, innvandringsbakgrunn, utdanning. Nedenfor gis en tilsvarende tabell over tiltaksdeltakere som påbegynte et ordinært tiltak i fjerde kvartal 2002 etter ulike demografiske kjennetegn, og personer som var registrert som helt ledige ved kvartalets inngang..

**Tabell 4.1 Personer som påbegynte et ordinært tiltak i løpet av 4. kvartal 2002 og helt ledige ved inngangen til 4. kvartal 2002 etter et utvalg demografiske kjennetegn på uttrekkstidspunktet. Antall og prosent.**

Demografiske kjennetegn	Deltakere		Helt ledige	
	Antall	Prosent	Antall	Prosent
<b>I alt</b>	<b>7 538</b>	<b>100</b>	<b>62 646</b>	<b>100,0</b>
<b>Alder</b>	<b>7 538</b>	<b>100,0</b>	<b>62 646</b>	<b>100,0</b>
Under 20 år	1 662	22,0	1 247	2,0
20-24 år	1 317	17,5	8 210	13,1
25-39 år	2 920	38,7	29 804	47,6
40-54 år	1 363	18,1	15 282	24,4
55-66 år	274	3,6	7 957	12,7
67 år og mer	2	0,0	146	0,2
<b>Region</b>	<b>7 538</b>	<b>100,0</b>	<b>62 646</b>	<b>100,0</b>
Oslo og Akershus	1 391	18,5	14 465	23,1
Hedemark og Oppland	624	8,3	3 740	6,0
Sør-Østlandet	1 531	20,3	11 304	18,0
Agder og Rogaland	1 181	15,7	9 253	14,8
Vestlandet	1 045	13,9	10 438	16,7
Trøndelag	631	8,4	5 953	9,5
Nord-Norge	1 030	13,7	7 030	11,2
Uoppgitt	105	1,4	463	0,7
<b>Kvinner</b>	<b>3 483</b>	<b>46,2</b>	<b>28 121</b>	<b>44,9</b>
<b>Gift</b>	<b>2 351</b>	<b>31,2</b>	<b>21 324</b>	<b>34,0</b>
<b>Innvandrere</b>	<b>2 005</b>	<b>26,6</b>	<b>8 540</b>	<b>13,6</b>
<b>Utdanningsnivå</b>	<b>7 538</b>	<b>100,0</b>	<b>62 646</b>	<b>100,0</b>
Uoppgitt/Ingen	1 070	14,2	3 522	5,6
Grunnskole	1 431	19,0	8 616	13,8
VGI	2 876	38,2	22 408	35,8
VGII	1 275	16,9	15 436	24,6
Påbygging VG	176	2,3	1 873	3,0
Univ./Høgskole	710	9,4	10 791	17,2

For personer i aldersgruppen under 25 år, innvandrere og de med utdanning lavere enn VGII er andelen tiltaksdeltakere klart høyere enn andelen helt ledige pr. 4. kvartal 2002. Disse undergruppene er med andre ord overrepresenterte når det gjelder tiltaksdeltakelse.

#### 4.1.2 Historikk på arbeidsmarkedet

Viktigheten av å benytte kjennetegn som beskriver historikken på arbeidsmarkedet i en effektanalyse av denne typen, er fremhevet i en rekke studier bl.a. Heckman et al (1997). I de tidligere analysene fra SSB ble ulike sider tidligere arbeidsmarkedserfaring forsøkt belyst og benyttet som forklaringsvariable i modellen. Ved hjelp av informasjon fra Arbeidstakerregisteret, ble det konstruert en variabel som anga yrkeserfaring siste ti år fra uttrekkstidspunktet. Denne variabelen tok imidlertid kun hensyn til om en person var registrert med et aktivt arbeidstakerforhold på et gitt tidspunkt i året, men ikke hvor stor andel av året vedkommende hadde vært sysselsatt eller omfanget av arbeidsforholdet (heltid, deltid). Det vil derfor bli lagt vekt på å konstruere en forklaringsvariabel som fanger opp yrkeserfaringen på en mer nøyaktig måte, og presentere tall for deltakere og ledige.

På samme måte antar vi også at tiltaks- og ledighetshistorikk er en sentral variabel som påvirker både rekrutteringen til tiltak og tilpasningen på arbeidsmarkedet. I de tidligere undersøkelsene ble disse variablene konstruert ved å benytte informasjon fra Aetats register over registrert ledige 12 måneder før uttrekkstidspunktet. Historikkvariabelen begrenset seg altså til kun å gjelde siste året. Det kan argumenteres for at dette er en for kort periode til å gi et godt nok bilde av både tiltaks- og ledighetserfaringen.

Andre historikk variable som det kan være naturlig å inkludere, omfatter erfaring med attføring, rehabilitering og uførepensjon i forkant av ledighets og tiltaksperiode.

#### 4.2 Tilpasning på arbeidsmarkedet målt ett år etter

Ved å koble på informasjon fra Arbeidstakerregisteret og Aetats register over personer registrert som helt ledige ble i SSBs tidligere analyser tilpasningen gruppert i statusene jobb, ledig, på tiltak, og annet<sup>7</sup>. Tabell 4.2 angir en liknende tabell for gruppen av tiltaksdeltakere deltakere og helt ledige. Informasjonen er her hentet fra det samme datamaterialet som ligger til grunn for SSBs registerbaserte sysselsettingsstatistikk. Dette betyr at i tillegg til Arbeidstakerregisteret innhentes også informasjon fra andre registerkilder som skattemyndighetenes Lønns- og trekkoppgaveregister, sivil- og verneplikts registeret. Klassifiseringen av personer som sysselsatte er altså basert på en rekke ulike kilder, og det er bygget opp et system i SSB for en samlet utnyttning av disse. Kategorien *ordinære tiltak* begrenser seg her til kun å omfatte arbeidsmarkedsopplæring, mens *annet* kan være selvstendigvirksomhet, utdanning, ulike former for trygdeordninger eller at personene er utenfor arbeidsstyrken.

**Tabell 4.2 Personer som påbegynte et ordinært tiltak i løpet av 4. kvartal 2002 og helt ledige ved inngangen til 4. kvartal 2002 etter arbeidsmarkedsstatus. 4.kvartal 2003. Antall og prosent.**

Arbeidsmarkedsstatus	Deltakere		Helt ledige	
	Antall	Prosent	Antall	Prosent
<b>I alt</b>	<b>7 538</b>	<b>100,0</b>	<b>62 646</b>	<b>100,0</b>
Lønnstakere	3 261	43,3	22 972	36,7
Ordinære tiltak (arbeidsmarkedsopplæring)	458	6,1	2 237	3,6
Helt ledige	1 019	13,5	18 704	29,9
Annet	2 800	37,1	18 733	29,9

Suksesskriteriet det legges opp til å benytte i de videre effektanalysene vil være knyttet opp til om personer er i jobb eller ikke. Tabell 4.3 gir andelen av tiltaksdeltakerne som er registrert som

<sup>7</sup> Se rapport 2000/22, [http://www.ssb.no/emner/06/90/rapp\\_200022/rapp\\_200022.pdf](http://www.ssb.no/emner/06/90/rapp_200022/rapp_200022.pdf).

lønnstakere i den ordinære sysselsettingsstatistikken fordelt på de ulike tiltakene. Selv om dette gir en indikasjon på i hvor stor grad deltakerne kommer over i lønnet arbeid ett år etter tiltaket startet, sier det ikke noe om effekten av selve tiltaksdeltakelsen. Når det gjelder hovedgruppene av tiltak er lønnstilskudd tiltaket med størst andel lønnstakere målt ett år etter tiltaksstart. Andelen lønnstakere av deltakerer på arbeidspraksis og AMO er tilnærmet den samme.

**Tabell 4.3 Personer som påbegynte et ordinært tiltak i løpet av 4. kvartal 2002 etter tiltakstype og sysselsettingsstatus 4. kvartal 2003. Antall og prosent.**

<b>Arbeidssøkerstatus</b>	<b>Antall</b>	<b>Andelen lønnstakere 4.kv 2003</b>
<b>I alt</b>	<b>7 538</b>	<b>43,3</b>
1. Lønnstilskudd	924	65,3
2. Arbeidspraksis i ordinær virksomhet	2 682	39,1
3. Arbeidsmarkedsopplæring (AMO)	3 790	40,3
3. Arbeidsmarkedsopplæring (AMO), enkeltplasser	61	65,6
3. Jobbklubb	49	40,8
4. Sysselsettingstiltak for langtidsledige	28	67,9
4. Vikarplass	3	100,0
5. Nye plasser institusjonelle tiltak	1	0,0

I tabell 4.3 gis det andelen lønnstakere målt ett år etter uttrekkstidspunktet for både deltakergruppen og de helt ledige, fordelt på de samme demografiske kjennetegnene som tidligere. Antall personer på jobbklubb er trolig lavt fordi vi betinger på at personene ikke er på tiltak i september. Jobbklubb er ofte siste tiltak i en kjede av tiltak. Jobbklubb-deltakere i 4. kvartal som var på et annet tiltak i september faller derfor utenfor vår avgrensning av tiltaksdeltakerpopulasjonen.

Andelen lønnstakere 4. kvartal 2003 var høyere for tiltaksdeltakere 4. kvartal 2002 enn for helt arbeidsledige ved kvartalets inngang, jmfør tabell 4.4. Dette gjelder total sett, for de ulike aldersgruppene, regioner, kjønn, innvandrere og for ulike utdanningsnivå.

**Tabell 4.4 Personer som påbegynte et ordinært tiltak i løpet av 4. kvartal 2002 og helt ledige ved inngangen til 4. kvartal 2002 etter utvalgte demografiske kjennetegn på uttrekkstidspunktet og sysselsettingsstatus 4. kvartal 2003. Antall og prosent.**

Demografiske kjennetegn	Deltakere		Helt Ledige	
	Antall	Andelen Lønnstakere 4.kv 2003	Antall	Andelen lønnstakere 4.kv 2003
<b>I alt</b>	<b>7 538</b>	<b>43,3</b>	<b>62 646</b>	<b>36,7</b>
<b>Alder</b>	<b>7 538</b>	<b>43,3</b>	<b>62 646</b>	<b>36,7</b>
Under 20 år	1 662	35,3	1 247	41,9
20-24 år	1 317	46,5	8 210	46,7
25-39 år	2 920	47,2	29 804	40,1
40-54 år	1 363	42,4	15 282	34,3
55-66 år	274	39,1	7 957	17,9
67 år og mer	2	0,0	146	8,2
<b>Region</b>	<b>7 538</b>	<b>43,3</b>	<b>62 646</b>	<b>36,7</b>
Oslo og Akershus	1 391	37,9	14 465	32,1
Hedemark og Oppland	624	47,1	3 740	39,7
Sør-Østlandet	1 531	41,9	11 304	36,0
Agder og Rogaland	1 181	42,4	9 253	36,4
Vestlandet	1 045	47,6	10 438	39,3
Trøndelag	631	44,7	5 953	39,3
Nord-Norge	1 030	49,9	7 030	42,2
Uoppgitt	105	4,8	463	1,1
<b>Kvinner</b>	<b>3 483</b>	<b>43,2</b>	<b>28 121</b>	<b>39,4</b>
<b>Gift</b>	<b>2 351</b>	<b>47,4</b>	<b>21 324</b>	<b>35,0</b>
<b>Innvandrere</b>	<b>2 005</b>	<b>40,8</b>	<b>8 540</b>	<b>28,8</b>
<b>Utdanningsnivå</b>	<b>7 538</b>	<b>43,3</b>	<b>62 646</b>	<b>36,7</b>
Uoppgitt/Ingen	1 070	40,3	3 522	25,3
Grunnskole	1 431	31,0	8 616	28,0
VGI	2 876	42,9	22 408	33,0
VGII	1 275	54,2	15 436	43,9
Påbygging VG	176	53,4	1 873	40,9
Univ./Høgskole	710	51,7	10 791	43,8

## 5. Forslag til opplegg for årlige effektevalueringer av ordinære arbeidsmarkedstiltak

SSB vil evaluere effekten av å delta på ordinære arbeidsmarkedstiltak for personer som begynte på et tiltak i løpet av fjerde kvartal år t. Sammenligningsgrupper vil bli trukket fra massen av personer registrert som helt arbeidsledige i september år t, og som i tillegg ikke påbegynte et ordinært arbeidsmarkedstiltak i løpet fjerde kvartal. Metoden for å danne de enkelte sammenligningsgruppene er basert på propensity score matching.

Som suksesskriterium benyttes om personen er kommet i jobb på et evalueringstidspunkt satt til november året etter, og resultatene presenteres i form av beregnede jobbsannsynligheter. Kommende års analyser vil det også være aktuelt å følge opp de opprinnelige utvalgene av tiltaksdeltakere og (matchede) sammenligningsgrupper, og måle tiltakseffektene to år senere. Resultatene vil gi uttrykk for gjennomsnittseffekten av det ordinære tiltaket på den beregnede jobbsannsynligheten *for de som faktisk deltok*, sammenlignet med om de ikke skulle ha deltatt.

SSBs forslag til årlige effektevalueringer består av en hoveddel og et tillegg. Hoveddelen inneholder separate analyser av de tre definerte hovedgruppene av ordinære arbeidsmarkedstiltak: arbeidsmarkedsopplæring (AMO), lønnstilskudd og arbeidspraksis. Effekten av tiltaksdeltakelse for de som deltok blir målt i forhold til fravær av tiltak, både for ordinære tiltakene samlet og for de tre tiltaksgruppene. Tillegget består av sammenligninger av effektene av de tre tiltaksgruppene opp mot hverandre.

I dette kapitlet gis det en nærmere gjennomgang av analysene. En mer formell beskrivelse av metoden er gjengitt i vedlegg A.

### 5.1 Separate analyser av tiltaksgrupper

Den første delen av den årlige analysen omfatter effektanalyser av hver enkelt av de tre hovedgruppene av ordinære arbeidsmarkedstiltak samt en beregning av den gjennomsnittelige effekten av deltakelse på et av de tre tiltaksgruppene. I samråd med forskere anser vi følgende analysemetode som mest hensiktsmessig.

Vi modellerer individenes "valg" mellom enten å delta på AMO, lønnstilskudd, arbeidspraksis eller ikke delta på tiltak, i en multinomisk logitmodell med relevante tilgjengelige forklaringsvariable. Sannsynligheten for å delta på hver av de tre tiltaksgruppene blir beregnet gitt de observerbare kjennetegnene til hver person (propensity score). Dette gjøres separat for relevante undergrupper.<sup>8</sup>

Personer i en tiltaksgruppe og tilsvarende masse av ikke-deltakere tas ut dersom de mangler overlapp i propensity score. På bakgrunn av de beregnede sannsynlighetene (propensity score) blir så sammenligningsgruppene plukket ut. Hver tiltaksdeltaker matches med én person fra massen av ikke-deltakere som har likest propensity score. Matchingen gjøres separat for hver tiltaksgruppe og for mest mulig like undergrupper.<sup>9</sup>

De matchede parene er utgangspunkt for beregning av andelen arbeidstakere på evalueringstidspunktet. Dette gjøres både for tiltaksdeltakere og for matchede ikke-deltakere, enten for de tre tiltaksgruppene hver for seg eller samlet. Vi får beregnet en gjennomsnittseffekt av de tre ordinære tiltaksgruppene for de som har deltatt dersom beregningene gjøres samlet.

<sup>8</sup> Åtte relevante undergrupperinger kan være kjønn krysset med alder (ungdom/voksen) og dagpengerettigheter (med/uten).

<sup>9</sup> Med 8 undergrupperinger blir det til sammen 24 tiltaksgrupper og tilhørende masser av ikke-deltakere hvor vi fjerner individer med manglende overlapp og matcher deltakere og ikke-deltakere.



Differansen i arbeidstakerprosenten mellom tiltaksdeltakere og sammenligningsgruppen rapporteres, ved siden av varianser og noen kvalitetsindikatorer.

Metoden er utviklet av Lechner (2001) og Imbens (2000). Tilsvarende metode på svenske data er anvendt av Sianesi (2001). En lignende matching metode på norske data er anvendt av Raaum et al (2002b). De benytter en flervalgmodell der individene kan "velge" mellom a) deltakelse i tiltaket som evalueres b) deltakelse i annet tiltak c) forlate ledighetsregisteret, alt som alternativ til d) være i ledighetsregisteret.

## 5.2 Evaluering av effekten på tvers av tiltakene

Her sammenligner vi effektene av de tre tiltaksgruppene opp mot hverandre.

Vi evaluerer endringen i gjennomsnittlig arbeidstakerprosent for individene som faktisk gjennomførte et av tiltakene i en bestemt tiltaksgruppe dersom de i stedet hadde gått på et tiltak i en annen bestemt tiltaksgruppe.

Analysen beskrevet i 5.1 gir den estimerte effekten av tiltaksdeltakelse ved å sammenligne beregnet jobbsannsynlighet etter å ha deltatt på et tiltak med jobbsannsynligheten uten tiltaksdeltakelse. For å analysere forskjeller i effekt tiltakene imellom tar vi utgangspunkt i de samme modellene og individenes beregnede propensity score for de tre tiltaksgruppene. Sammenligningsgruppene som tiltaksdeltakerne matches mot blir imidlertid ulike. Her benytter vi hver av de to andre tiltaksgruppene som sammenligningsgruppe avhengig av hvilke effekter som beregnes.

Etter å ha fjernet personer i tiltak- sammenligningsgruppen med manglende overlapp, matches hver tiltaksdeltaker i den analyserte tiltaksgruppen med den deltakeren fra sammenligningsgruppen med likest propensity score. Differansen i arbeidstakerprosenten mellom tiltaksdeltakerne og sammenligningsgruppen rapporteres.

Med tre grupper av tiltak gir det oss 6 kombinasjoner å analysere fordi disse effektene ikke er symmetriske.<sup>10</sup>

---

<sup>10</sup> Med 8 undergrupperinger blir det til sammen 48 tiltaksgrupper og tilhørende sammenligningsgruppe hvor vi fjerner individer med manglende overlapp og matcher deltakere og individ fra sammenligningsgruppen.

## 6. Forslag til tilleggsprosjekter

I dette kapitlet skisseres kort tre tilleggsprosjekter. I det første tilleggsprosjektet vil vi sammenligne metoden beskrevet i kapittel 5 med SSB sin tidligere anvendte metode. I det andre tilleggsprosjektet vurderer vi andre suksesskriterier, mens vi i det tredje leggsprosjektet i større grad vil måle effekten av tiltakene og selve rekrutteringssystemet samlet.

### 6.1 Sammenligning med SSB sin tidligere anvendte metode

Forskjellen mellom analyseopplegget som presenteres i denne rapporten og det som lå til grunn for SSB sine tidligere analyser, er i hovedsak valg av uttrekkstidspunkt (innstrømming på tiltak vs. utstrømming fra tiltak) og estimeringsmetode. Estimeringsmetoden beskrevet i denne rapporten går ut på å definere sammenlikningsgruppene ved hjelp av matching teknikker, for deretter å foreta ikke-parametriske gjennomsnittsberegninger. I tillegg til at vi her foreslår estimering av (multinomiske) logitmodeller *separat* for 8 mer homogene undergrupper i stedet for samlet. Tidligere anvendte SSB logistisk regresjon for å beregne jobbsannsynligheten forklart ved en tiltaksindikator og en rekke andre forklaringsvariable. Her ønsker vi å se på eventuelle resultatforskjeller som skyldes ulik estimeringsmetode. Dette vil kunne gi en indikasjon på hvor mye den metodiske tilnærmingen har å si for resultatene. Vi vil derfor foreta en effektevaluering basert på logistisk regresjon, som i SSB sine tidligere evalueringstudier, men med noen av de samme avgrensningene og valgene som i metoden beskrevet i dette notatet. Vi tenker oss ikke disse sammenligningene som en del av den årlige standardiserte analysen. Denne analysen foretar vi en gang andre halvår 2005, etter at den årlige standardiserte tiltaksevalueringen er ferdig.

### 6.2 Utvidelse av suksesskriteriet

I 2006 kan vi foreta en analyse av tiltaksdeltakerne i 4. kvartal 2004 lik den skissert i kapitel 5, hvor suksesskriteriet utvides til å inneholde status som både lønntakere og selvstendig næringsdrivende. Årsaken til at eventuelle sysselsettingsforhold som selvstendig næringsdrivende, ikke kan inkluderes i analysene utført våren 2005, skyldes at kilden for opplysninger om næringsvirksomhet er Selvangivelsesregisteret. Produksjonstiden for dette registeret er slik at opplysninger om 2004, først foreligger ved inngangen av 2006.

For enkelte tiltak og deltakergrupper vil utdanning være en status som det er naturlig å inkludere i suksesskriteriet ved siden av sysselsetting. Produksjonstiden for disse opplysningene er også så lang at en slik analyse av tiltaksdeltakerne i 4. kvartal 2004 først kan foretas våren 2006.

### 6.3 Evaluering av hele tiltakspakken

Vi kan analysere effekten av hele tiltakspakken for de som deltar på tiltak.

Når vi modellerer rekrutteringen til tiltak for ulike grupper av personer og tiltak samlet, vil vi i praksis måle effekten av tiltakene og selve rekrutteringssystemet samlet. Dette kan være en interessant analyse som ikke trengs å gjøres hvert år.

Vi foretar en logistisk regresjonsanalyse for å beregne sannsynligheten for å havne på tiltak gitt relevante forklaringsvariable. Sannsynligheten for å delta på tiltak blir beregnet gitt de observerbare kjennetegnene til hver person (propensity score). Personer i tiltaksgruppen og massen av ikke-deltakere tas ut dersom de mangler overlapp i propensity score. På bakgrunn av de beregnede sannsynlighetene (propensity score) blir så sammenlikningsgruppene plukket ut. Hver tiltaksdeltaker matches med én person fra massen av ikke-deltakere som har likest propensity score. De matchede parene er utgangspunkt for beregning av andelen arbeidstakere på evalueringstidspunktet både for

tiltaksdeltakere og sammenligningsgruppen. Differansen i arbeidstakerprosenten mellom tiltaksdeltakere og sammenligningsgruppen rapporteres, sammen med varianser og noen kvalitetsindikatorer.

## Referanser

- Abedie, A. og Imbens, G. (2002): "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects". NBER technical working paper no 283, 2002.
- Bråthen, M. og Landfald, Ø. (1999): Evaluering av ordinære arbeidsmarkedstiltak 1999. Dokumentasjon og analyse av effekter på kort og lang sikt. Rapporter 1999/21, Statistisk sentralbyrå, Oslo.
- Calmfors, L., Forslund, A. og Hemström, M. (2002): "Does active labour market policy work? Lessons from the Swedish experiences", Working paper 2002:4, IFAU, Uppsala.
- Hardoy, Inés(2003): A study of the effect of labour market programs for youths in Norway using propensity score matching. Manuskript, EALE2003.
- Heckman, J., Ichimura, H. and Todd, P., 1997: Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *Review of Economic Studies* 64 4, pp. 605–654.
- Heckmann et al. (1999); The Economics and Econometrics of Active Labour Market Programmes. Trykt i O. Ashenfelter og D.Card, (red.): *Handbook of Labor Economics Vol. III*, North-Holland, Amsterdam.
- Imbens, Guido W.(2000): The role of the propensity score in estimating dose-responses functions. *Biometrika* 87, pp. 706–710.
- Imbens, Guido W. (2004): Nonparametric estimation of average treatment effects under exogeneity: A review, *Economics and Statistics (2004), February*.
- Larsson, Laura (2003): Evaluation of Swedish Youth Labor Market Programs, *Journal of Human Resources*, XXXVIII nr. 4, 2003.
- Lechner (2001): Identification and estimation of causal effects of multiple treatments under conditional independence assumption. In Lechner, M. og Pfeiffer, F. (Eds.): *Econometric Evaluation of Labour Market Policies*, Zew Economic studies 13, Physica-Verlag, 2001.
- Martin og Grubb (2001): What works and for whom: a review of OECD countries' experiences with active labour market policies. Working paper 2001:14, IFAU, Uppsala.
- Raaum, O. og Torp, H. (2002): "Labour market training in Norway - effect on earnings", *Labour Economics*, Vol. 9, 207-247.
- Raaum, Oddbjørn, Røed, Knut og Torp, Hege (2002a): Riktig satsing i arbeidsmarkedspolitikken ?, *Norsk Økonomisk Tidsskrift* nr. 2, 2002
- Raaum, Oddbjørn, Torp, Hege og Zhang, Tao (2002b): Business cycles and the impact of labour market programs. Memorandum 14/2002. Department of Economics, University of Oslo.
- Raaum, Oddbjørn, Torp, Hege og Zhang, Tao (2002c): Do individual programme effects exceeds the costs? Norwegian evidence on long run effects of labour market training. Memorandum 15/2002. Department of Economics, University of Oslo.
- Raaum, Oddbjørn, Røed, Knut og Torp, Hege (2003): Systematisk evaluering av aktive arbeidsmarkedstiltak, upublisert notat datert 3.2.2003 laget på oppdrag for AAD.

- Rosenbaum, P. and Rubin, D., (1983): The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, pp. 41–55.
- Røed, K., Torp, H., Tuveng, I. og Zhang, T. (2000): Hvem vil og hvem får delta? Analyser av rekruttering og utvelgelse av deltakere til arbeidsmarkedstiltak i Norge på 1990-tallet. Rapport 4/2000, Frischsenteret.
- Schmidt, C. M., Zimmermann, K. F., Fertig, M. og Kluge, J. (2001): *Perspektiven der Arbeitsmarktpolitik. Internationaler Vergleich und Empfehlungen für Deutschland.* Springer.
- Sianesi, B. (2001): Differential effects of Swedish active labour market programmes for unemployed adults during the 1990s. Working paper W01/25, Institute for Fiscal Studies, London.
- Smith, J. A. og Todd, P. E. (2005): Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* 125 (2005) 305-353.
- Stortingsmelding nr. 19 (2003-2004): Et velfungerende arbeidsmarked. Arbeids- og administrasjonsdepartementet.
- Torp, H. (2000): Rekruttering, seleksjon og effektevaluering.  
Kapitel 2 i rapporten Røed, K., Torp, H., Tuveng, I. og Zhang, T. (2000): Hvem vil og hvem får delta? Analyser av rekruttering og utvelgelse av deltakere til arbeidsmarkedstiltak i Norge på 1990-tallet. Rapport 4/2000, Frischsenteret.
- Vassnes, E. (1996): Evaluering av arbeidsmarkedstiltak - bruk av registerdata. Notater 1996/17. Statistisk sentralbyrå, Oslo

## Vedlegg A

### A.1 Mer formell beskrivelse av metoden i kapittel 5.1 og 5.2

I kapitel 5.1 ønsker vi å evaluere effekten på jobbsannsynligheten av tre grupper ordinære arbeidsmarkedstiltak separat sett i forhold til å ikke ha deltatt på tiltak. Vi benytter en flervalgsmodell utviklet av Lechner (2001) og Imbens (2000). De generaliserer potensielt utfall tilnærmingen for tilfellet med et tiltak, beskrevet i Rosenbaum og Rubin's (1983) artikkel, til tilfellet med flere tiltak. I en multinomisk logit modell modelleres individenes "valg" mellom 0) ikke å delta på tiltak, 1) Arbeidsmarkedsopplæring, 2) lønnstilskudd og 3) arbeidspraksis ved hjelp av et sett forklaringsvariable.<sup>11</sup> Et sett av potensielle utfall(suksess) er assosiert med hver av de fire valgene:  $Y^0, Y^1, Y^2, Y^3$ , der  $Y_i^k$  betegner utfallet for individ  $i$  hvis individet fikk behandling  $k$ . La tiltaksindikatoren  $S \in \{0,1,2,3\}$  betegne faktisk behandling ev. mangel på sådan, slik at  $S_i=k$  dersom individ  $i$  fikk behandling  $k$ . Da står 0 for mangel på tiltak, mens verdiene en til tre hhv. karakteriserer de tre nummererte tiltaksgruppene beskrevet over. Siden hvert individ bare kan få en behandling vil individets potensielle utfall under de 3 andre behandlingene ev. mangel på sådan, være uobserverbare kontrafaktiske utfall. Vi er derfor nødt til å gjøre en identifiserende forutsetning for å overkomme det fundamentale problemet med ikke observerbare kontrafaktiske utfall. I tråd med Lechner (2001) gjør vi en utvidet antagelse om betinget uavhengighet (CIA) som tillater oss å identifisere de nødvendige kontrafaktiske utfallene. For et underrom  $\chi$  i rommet av forklaringsvariable forutsetter vi at:

$$Y^l \perp\!\!\!\perp S | X=x, S \in \{l,m\} \text{ og } 0 < P^j(x) < 1 \text{ holder for } \forall x \in \chi \text{ og } \forall j=m,l$$

Da er  $\theta_0^{ml} = E(Y^m - Y^l | S = m)$  identifiserbar.

Ifølge Lechner (2001) følger det da også at

hvis  $Y^l \perp\!\!\!\perp S | P^{l|ml}(X) = P^{l|ml}(x), S \in \{l,m\} \text{ og } 0 < P^j(x) < 1 \text{ for } \forall x \in \chi \text{ og } \forall j=m,l$ , der

$$P^{l|ml}(x) = P(S=l | S \in \{l,m\}, X=x), \text{ så følger det at } \theta_0^{ml} = E(Y^m - Y^l | S = m) \text{ er identifiserbar.}$$

Hovedimplikasjonen av dette er at antall dimensjoner i matchingen kan reduseres til en også for flervalgsmodeller.

Gangen i analysen er at vi benytter en multinomisk valgmodell til å forklare utfallet at behandlingsindikatoren ved hjelp av et sett av forklaringsvariable. Ut ifra estimatene i en multinomisk logitmodell kan vi beregne propensity score for hver behandling for hvert individ  $(\hat{P}_N^0(x), \hat{P}_N^1(x), \hat{P}_N^2(x), \hat{P}_N^3(x))$ . For hvert individ benytter vi de til å predikere den betingete valgsannsynligheten for behandling  $l$  gitt at enten behandling  $l$  eller  $m$  blir valgt:

$$\hat{P}_N^{l|ml}(x) = \frac{\hat{P}_N^l(x)}{\hat{P}_N^l(x) + \hat{P}_N^m(x)}, \text{ der } m, l \in \{0,1,2,3\} \text{ og } m \neq l$$

De predikerte betingete sannsynlighetene benyttes til å fjerne personer i tiltaksgruppen som analyseres og sammenligningsgruppen der det ikke er *common support*. Deretter benytter vi de predikerte betingete sannsynlighetene til å matche personer i tiltaksgruppe,  $l$ , som analyseres med personer i sammenligningsgruppen,  $m$ , som i kapitel 5.1 er de som ikke gikk på tiltak ( $m=0$ ). I kapitel 5.2 vil sammenligningsgruppen  $m$  være en av de to andre tiltaksgruppene. Alternativt benyttes både  $\hat{P}_N^l(x)$  og  $\hat{P}_N^m(x)$  i matchingen ved å minimere Mahalanobis distanse.

<sup>11</sup> De andre gruppene av tiltak er holdt utenfor analysen siden det er svært få som deltar på dem. Strengt tatt modellerer vi derfor ikke alle valg av tiltak individene står overfor.

De matchede parene er utgangspunkt for beregning av andelen arbeidstakere på evalueringstidspunktet både for tiltaksdeltakere og for sammenligningsgruppen. Differansen i arbeidstakerprosenten mellom tiltaksdeltakere og sammenligningsgruppen rapporteres, ved siden av varianser og noen kvalitetsindikatorer.

Vi kan få fram effekten av de ordinære tiltakene samlet for de som deltok på dem. Da tar vi utgangspunkt i samtlige matchede par fra de ulike undergruppene for alle tiltaksgruppene i fra analysen i kapittel 5.1, og beregner andelen arbeidstakere på evalueringstidspunktet for tiltaksdeltakerne samlet og for ikke-deltakerne samlet.

## A.2 Mer formell beskrivelse av metoden i kapittel 6.3

I kapittel 6.3 ønsker vi å evaluere effekten av hele "tiltaks pakken", dvs. effekten for alle som deltar på ordinære tiltak. Vi måler da effekten av tiltakene og selve rekrutteringssystemet samlet. Formelt kan dette skrives som:

$$E(Y_1 - Y_0 | D=1) = E(Y_1 | D=1) - E(Y_0 | D=1),$$

der  $D$  er en binær variabel som er lik 1 dersom en person har vært på tiltak og verdien 0 dersom personen ikke har vært på tiltak.  $Y_1$  er potensielt utfall dersom en person har vært på tiltaket, mens  $Y_0$  er potensielt utfall dersom en person ikke har vært på tiltaket. Problemet er at (det kontrafaktiske) ikke-tiltaksutfallet er uobserverbart for de som har vært på tiltak. En annen person kan benyttes til å beregne det kontrafaktiske utfallet slik at  $E(Y_0 | D=1) = E(Y_0 | D=0)$ . Med andre ord må utfallsvariabelen bli estimert ved hjelp av en annen person, fra sammenligningsgruppen med samme karakteristika som de i tiltaksgruppen bortsett fra tiltaksdimensjonen.

Randomisering sikrer at responsen  $(Y_0, Y_1)$  er uavhengig av tiltaksvariabelen  $D$ , slik at forventningen til  $Y_0$  er uavhengig av om personen er på tiltak eller ikke. Dvs.  $E(Y_0 | D=1) = E(Y_0 | D=0) = E(Y | D=0)$  der  $Y = DY_1 + (1-D)Y_0$  er det observerte utfallet.

Metoden matching forsøker å ligne den eksperimentelle eller randomiserte situasjonen ved å balansere fordelingen av høyresidevariable  $X$  for tiltaksdeltakere og ikke-deltakere slik at fordelingen ikke er systematisk forskjellig for de to gruppene. Direkte sammenligning gir da mening.

Følgende forutsetning trengs:

$$(A-1) (Y_0, Y_1) \perp\!\!\!\perp D | X$$

mao. betinget på høyresidevariablene  $X$  er  $(Y_0, Y_1)$  og  $D$  uavhengig ( $\perp\!\!\!\perp$  betegner uavhengighet).

Antagelsen (A-1) impliserer at  $P[D=1 | Y_0, Y_1, X] = P[D=1 | X]$  dvs. at den betingete sannsynligheten for  $D$  er uavhengig av potensielt utfall  $Y$ .

I tillegg må vi gjøre følgende forutsetning:

$$(A-2) 0 \leq P[D=1 | X] < 1 \equiv P(X) < 1 \text{ for alle } X$$

Antagelsen over sikrer at alle tiltaksdeltakere har en motspiller i sammenligningsgruppen og at alle i sammenligningsgruppen var potensielle tiltaksdeltakere.

Gitt antagelsene over, kan det manglende kontrafaktiske utfallet konstrueres av ikke-deltakere:

$$E(Y_0 | X, D=1) = E(Y_0 | X, D=0) = E(Y_0 | X).$$

Selve matchingen kan skje ved direkte kobling av tiltak- og sammenligningsgruppe på alle forklaringsvariablene,  $X$ . Ved mange forklaringsvariablene kan dette by på problemer i praksis. Balansert score,  $b(X)$ , kan brukes til å matche tiltak- og kontrollgruppen på en enklere måte, der  $b(X)$  er en funksjon av observerte variable slik at den betingete fordelingen til  $X$  gitt  $b(X)$  er lik både for de på tiltak og sammenligningsgruppen. Den betingete sannsynligheten for å havne på tiltak dvs. propensity scoren;  $P(X) = \Pr[D=1 | X]$  er en slik balansert score funksjon. Rosenbaum og Rubin (1983) beviste et resultat som er nyttig for å redusere antall dimensjoner i betingingsproblemet ved implementering av matchingen. De viste at for tilfeldig variable  $Y$  og  $X$ , en diskret variabel  $D$  med

verdi 0 eller 1 og funksjonen  $P(X) = \Pr(D=1|X)$  at  $E(D|Y, P(X)) = E(E(D|Y, X)|Y, P(X))$ , slik at  $E(D|Y, X) = E(D|X) = P(X)$  impliserer  $E(D|Y, P(X)) = E(D|P(X))$ .

Det impliserer at når  $Y_0$  utfallet er uavhengig av tiltaksdeltakelse betinget  $X$ , så er også  $Y_0$  utfallet er uavhengig av tiltaksdeltakelse betinget på propensity scoren,  $\Pr[D=1|X]$ . Matching på propensity score er derfor tilstrekkelig for å få forventningsrette estimater gitt at CIA holder. Dermed kan vi beregne effekten på forventet jobbsannsynlighet av tiltaksdeltakelse for de som har gjennomgått tiltak:

$E(Y_1 - Y_0|D=1, P(X)) = E(Y_1|D=1, P(X)) - E(Y_0|D=0, P(X))$  basert på gjennomsnitt for tiltaksdeltakere og ikke-deltakere som er matchet ved hjelp av estimerte propensity-score verdier.

Rent praktisk beregner vi for hvert individ propensity-score ut fra en logistisk regresjon der vi beregner sammenhengen mellom tiltaksindikatoren  $D$  og relevante forklaringsvariable  $X$ . Etter å ha fjernet enheter uten felles propensity-score overlapp, "common support", foretar vi en nærmeste nabo matching basert på propensity-scorene, ved å matche tiltaksdeltakere med ikke-deltakere som har likest mulig propensity-score verdi som tiltaksdeltaken.

De matchede parene er utgangspunkt for beregning av andelen arbeidstakere på evalueringstidspunktet både for tiltaksdeltakere og for sammenligningsgruppen.



## De sist utgitte publikasjonene i serien Notater

- 2004/88 G. Daugstad og B. Lie: Kvalitativ forstudie til leveårsundersøkelse blant ikkevestlige innvandrere. 138s.
- 2004/89 S. Lien og Ø. Sivertstøl: Langtidsmottakere av sosialhjelp 1997-1999. 64s. ISSN 0806-3745
- 2005/1 S. Hansen og T. Skoglund: Sysselsetting og lønn i historisk nasjonalregnskap. Beregninger for 1949-1969. 36s.
- 2004/2 FoU og innovasjonstatistikk 2001 og 2002-dokumentasjon. 82s.
- 2005/3 M. Steinnes, J. Monsrud, E. Engelién og V.V. Holst Bloch: Samferdsel og miljø. Utvikling av et norsk indikatorsett tilpasset et felles europeisk sammenligningsgrunnlag. 80s.
- 2005/4 E. Falnes-Dalheim og A. Falnes-Dalheim: Dokumentasjon av FoB2001. Spesifikasjoner, bearbeiding, flytdiagram for spørreskjemadelen av tellingen. Del I. 117s.
- 2005/5 E. Falnes-Dalheim, A. Falnes-Dalheim: J. Sjørbotten og B. Østvedt: Dokumentasjon av FoB2001. Spesifikasjoner, bearbeiding, flytdiagram for spørreskjemadelen av tellingen. Del II Vedlegg. 146s.
- 2005/6 E. Falnes-Dalheim: Bearbeiding av prøvetellingen i Stange 2000. Folke- og bolig tellingen 2001. 126s.
- 2005/7 S. Kwesi Baateng og S. Ferstad: Dokumentasjonsnotat for FylkesKOSTRA vidregående opplæring. Publisering av 2003-tallene. 221s.
- 2005/8 Ø. Linnestad og O.K. Lien: SM08 Prisindekser. Fraktindeks på utenriks sjøfart. 56s.
- 2005/9 E. Cometa Rauan og R. Johannessen: Forventningsindikator - konsumprisene. November 2004 - mai 2005. 18s.
- 2005/10 A.S. Abrahamsen: Analyse av revisjon - Feilkoder og endringer i utenrikshandelstatistikken. 71s.
- 2005/11 A-K. Mevik: Usikkerhet i ordrestatistikken. 22s.
- 2005/12 A. Akselsen, S. Lien, Ø. Sivertstøl: FD - Trygd. Variabelliste. 56.
- 2005/13 T. Seland Forgaard: Monitor for sekundærflytting. En deskriptiv analyse om sekundærflyttinger blant flyktninger som ble bosatt i Norge i perioden 1994-2003. 48s.
- 2005/14 O. Villund: Kvalitet på yrke i registertbasert statistikk. Resultater og utfordringer. 48s.
- 2005/15 E. Engelién, M. Steinnes og V.V. Holst Bloch: Tilgang til friluftsområder. Metode og resultater 2004. 38s.
- 2005/16 G. Dahl: Uførepensjonisters bakgrunn. 56s.
- 2005/18 A. Rolland: KOSTRA, tjenestekvalitet og kompetansefordeling i supermarkedstaten. 45s.
- 2005/19 H. Tønnseth. Årsrapport 2004. Kontaktutvalget for helse- og sosialstatistikk 10s.
- 2005/20 N.K. Buskoven: Vertskommunekompensasjon - kartlegging av kommunenes utgifter til asylmottak. 49s.
- 2005/21 H.C. Hougen: Omnibusundersøkelsen oktober/november 2004. Dokumentasjonsrapport. 52s.
- 2005/22 D. Sve, L. Solheim og G. Haraldsen: Eldres kvalitet. Dokumentasjon av datafangsten. 64s.
- 2005/23 E. Rauan: Undersøking om foreldrebetaling i barnehagar, januar 2005. 45s.
- 2005/24 L. Østby: Bruk av velferdsordninger blant nyankomne innvandrere fra de nye EØS-medlemslandene. 36s.
- 2005/25 A. Fagereng: Reestimering av faktoretterterspørselen i KVARTS. 72S.