



Espen Dahl, Dr. Polit
Professor og
forskningsleder av
Helse, omsorg og
velferd (HOV),
Høgskolen i Oslo
espen.dahl@hio.no



Kjetil A. van der Wel,
Cand. Polit
Stipendiat ved Gruppe
for inkluderende velferd
(GIV), Høgskolen i Oslo
kjetil.wel@sam.hio.no

Å forske på marginaliserte grupper

Noen metodiske og forskningsetiske utfordringer

Behovet for å utforske sosiale forhold blant vanskeligstilte grupper har økt de siste årene. Samtidig møter slik forskning særlige metodiske utfordringer. Her presenterer vi en frafallsanalyse av spørreskjemaet i «Funksjonsevneundersøkelsen», gjennomført ved Høgskolen i Oslo i 2005. Undersøkelsen henvendte seg til nesten 1273 langtidsmottakere av sosialhjelp i 14 kommuner. Vi bruker registerdata fra FD-trygd for å studere mulige skjevheter mellom spørreskjemaet og registerutvalget. Analysene tyder på at spørreskjemaet ikke avviker vesentlig fra registerutvalget. Artikkelen drøfter også prinsipielt samtykkekravet ved bruk av registerdata, vurdert opp mot nytten av å analysere blant annet representativitet.

Forskning på vanskeligstilte og marginaliserte grupper har fått fornyet oppmerksomhet de siste årene i og med at fattigdom er blitt satt på den politiske dagsordenen. Interessen for fattigdom kom med Bondevik II-regjeringen i 2002. I en egen fattigdomsplan definerte den fattigdom i Norge som et alvorlig sosialpolitisk problem og foreslo å målrette stønader og skreddersynte tiltak for å bekjempe den (St.meld. nr 6 2002-2003, St.prp. nr. 1 2008-2009). Følgende grupper er overrepresentert blant inntektsfattige: unge, enslige forsørgere, innvandrere, langtidsmottakere av sosialhjelp, og mennesker med rusproblemer (SSBs årlige rapporter). Stoltenbergs rødgrønne regjering har fulgt opp denne satsingen ved å bevilge mer penger, jenen-

nomføre de reformene som Bondevik II tok initiativet til (NAV-reformen), og iverksette nye reformer, heri opptatt kvalifikasjonsprogrammet som skal bringe sosialhjelpsmottakere ut av «avhengighet», «passivitet» og fattigdom, og inn i inntektsgivende arbeid (St.meld. nr. 9 2006-2007). I lys av dette er forskningsbasert kunnskap om målgruppene, både når det gjelder levekår og livskvalitet og hvordan tiltak faktisk virker, blitt enda viktigere enn før. Behovet for pålitelig kunnskap om vanskeligstilte grupper er således voksende, men også metodisk sett svært utfordrende, spesielt fordi frafallet ved spørreskjemaundersøkelser kan bli både stort og skjevt. Det er også forskningsetisk krevende fordi behovet for forskningsbasert kunnskap av høy kvalitet kan komme på kant med kravet om de

undersøkt personvern, integritet og rett til medbestemmelse.

Utforskning av vanskeligstilte gruppers kår er en stolt tradisjon i norsk samfunnsforskning. Det som kjennetegnet «gullaldersosiologien» var fokuset på slike grupper, slik som fanger, psykiatriske pasienter og hjemløse (Mjøset 1991). Denne tradisjonen er fulgt opp av nye generasjoner levekårsforskere, som blant annet har tatt for seg fanger, arbeidsløse, sosialhjelpsmottakere og rusmisbrukere (Fyhn og Dahl 2000, Friestad og Hansen 2004).

Av politiske grunner er behovet for mer kunnskap om marginaliserte og vanskeligstilte grupper større enn på lenge. Den sosiologiske arven gir et imperativ til å rette søkelyset mot slike grupper, og mulighetene for å forske på disse gruppene er bedre enn noensinne på grunn av omfattende og nøyaktige offentlige registerdata om alle bosatte i Norge. Samtidig er personvernet for de «utforskede» blitt styrket betraktelig.

Formålet med denne artikkelen er å analysere skjevhetene i frafallet i en spørreskjemaundersøkelse – Funksjonsevneundersøkelsen blant langtidsmottakere av sosialhjelp – ved hjelp av registerdata fra FD-trygd, og å gi en mer prinsipiell drøfting av noen av de problemer som kravet om samtykke til bruk av registerdata skaper.

Bakgrunn

Forskning om langtidsmottakere av sosialhjelp er viktig fordi gruppa er en sentral målgruppe for tiltak mot

fattigdom og for arbeidsinkludering, samtidig som slik forskning kan kaste lys over en gruppe som ofte forblir mangelfullt beskrevet i befolkningsrepresentative spørreundersøkelser. Studier har funnet at langtidsmottakere av sosialhjelp kjennetegnes ved dårlige levekår på nær sagt alle områder. I 2006 hadde to av tre av alle sosialhjelpsmottakere ikke råd til å dekke ett eller flere grunnleggende behov, og på de fleste indikatorene var andelen rundt seks ganger høyere blant sosialhjelpsmottakerne enn i befolkningen (SSB 2008). Andelen som rapporterte svært dårlig helse var i den samme studien på 34 prosent, sammenlignet med ni prosent i hele befolkningen. Videre har sosialhjelpsmottakere lavere utdanning, mindre arbeidserfaring, færre mestringsressurser og sosiale ressurser enn andre (SSB 2008, van der Wel m.fl. 2006). Også andre studier har funnet dårligere helse, levekår og livskvalitet blant sosialhjelpsmottakere (Seim og Hjemdal 1992, Hove 1992, Brevik 1996, Underlid 2005, 2008). Rapporten fra spørreundersøkelsen vi her skal se nærmere på viste også at ulike barrierer mot inkludering i arbeidslivet hopet seg opp for en stor del av utvalget. En tredel hadde fire eller flere av syv barrierer mot arbeid, mens om lag en av fire hadde få problemer (van der Wel m.fl. 2006). Hva er grunnen til de store levekårsproblemene i denne gruppa? Enkelte studier peker på at behovet for sosialhjelp reproduseres fra én generasjon til en annen (Stenberg 2000, Lorentzen og Nielsen 2008). En beslektet teori fra sosial-epidemiologien framholder at sosiale problemer har en tendens til å akkumulere over livsløpet (Blane m.fl.

2007). Slike teorier står i kontrast til forestillinger om den «temporaliserte» og «demokratiserte» fattigdommen, som i korthet antar at fattigdom og økonomiske problemer oppstår akutt som følge av livshendelser, er av kortvarig natur, og kan ramme «hvem som helst» (Dahl og Lorentzen 2008). For både politiske og vitenskapelige formål er det således av stor betydning å drive empirisk forskning på en marginalisert og problembelastet gruppe som langtids sosialhjelpsmottakere, og den kunnskapen som frambringes må være pålitelig. Å kople spørreskjemadata med registerdata vil kunne øke påliteligheten.

FD-trygd, og andre offentlige, nasjonale registre som kan tilrettelegges for forskning, er en gullgrube for samfunnsforskere. FD står for Forløps Data, og databasen er, som navnet tilsier, en forløpsdatabase med opplysninger om en rekke typer statistiske data. Databasen inneholder individuelle opplysninger fra og med 1992 for hele den norske befolkningen og bygger på opplysninger i administrative registre og statistikkregistre (SSB.no 13.4.2009). Slike nasjonale databaser er spesielt verdifulle når det gjelder å frambringe ny kunnskap om marginaliserte grupper i befolkningen, det vil si grupper som befinner seg i randsonen av samfunnets sentrale institusjoner og arenaer som for eksempel lønnet arbeid (Halvorsen 2000). Disse teller ofte forholdsvis få mennesker, og er vanskelige å nå med standard spørreskjemaundersøkelser. Tradisjonelle samfunnsvitenskapelige datainnsamlingsmetoder bygger på forutsetninger som ofte ikke er til stede i slike grupper. Mange av deres problemer er nettopp av en karakter som gjør spørreskjemametoden mindre egnet. Eksempler er «skjemaskrekk», redusert kompetanse og kognitive evner, mangel på fast bopel, og at de er vanskelige å nå. Da kan det være fristende å basere seg kun på registerinformasjon om disse gruppene. Registerinformasjon på sin side har imidlertid en del mangler som spørreskjemaundersøkelser ikke har, nemlig informasjon om subjektive opplevelser, psykisk helse, holdninger, psykiske ressurser (opplevd mestringskontroll), generell livskvalitet, og vurdering av relasjonene til det

offentlige hjelpeapparatet. Denne typen subjektiv informasjon er viktig i seg selv, men særlig viktig er den for å evaluere hjelp og tiltak det offentlige tilbyr. Spørreskjemametodens akilleshæl – faren for stort og skjevt frafall – er registerdataenes styrke. Spørreskjemametodens styrke – subjektiv informasjon – er registerdataenes svakhet. Det «ideelle» kvantitative forskningsdesignet for utforskning av vanskeligstilte grupper er derfor å kombinere disse to metodene i en og samme undersøkelse.

Det er flere fordeler ved å kombinere data fra disse to forskjellige og uavhengige kildene. Ved å kople registerdata til spørreskjemadata kan vi spare respondentene for mange spørsmål, og få pålitelig informasjon om hendelser og forhold bakover, og ikke minst framover i tid. Funksjonsevneundersøkelsen er lagt opp for å utnytte de mulighetene FD-trygd gir her. Det fortrinnet vi skal utnytte i denne artikkelen er den muligheten FD-trygd – som omfatter hele den norske populasjonen – gir til å analysere og vurdere frafall i spørreskjemaundersøkelsen. I denne artikkelen presenterer vi en slik analyse av en vanskeligstilte gruppe som er sentral i politikken for å bekjempe fattigdom i Norge – langtidsmottakere av sosialhjelp. Konkret vil vi vurdere hvor representativ spørreskjemaundersøkelsen er ved hjelp av registerdata fra FD-trygd, og diskutere hvor nyttige registerdata kan være i en slik vurdering. Fordi bruk av registerdata til dette formålet også involverer noen forskningsetiske problemstillinger som berører samtykke, vil vi også drøfte dette.

Som hos andre marginale grupper er det et metodeproblem at sosialhjelpsmottakere er vanskelige å få i tale gjennom spørreskjemametoden. I mange undersøkelser er andelen som ikke svarer ofte stor, over 50 prosent (van der Wel 2006:6). Dersom bruttoutvalget er stort nok og gir tilstrekkelig statistisk styrke, er ikke det så alvorlig i seg selv. Alvorlig blir det først når forskjellene mellom netto- og bruttoutvalget blir stort og systematisk. I så fall, er det fare for at tall for forekomst av et problem, fattigdom for eksempel, blir undervurdert, siden det er blant de fattige vi kan forvente det høyeste frafallet. Sammenhenger

mellom variabler i materialet, estimert for eksempel med oddsratioer, kan imidlertid være mindre sensitive for slik seleksjonsbias (Søgaard m.fl. 2004, Kleinbaum m.fl. 1982). Generelt er det grunn til å tro at de som ikke svarer er vanskeligere stilt og mer marginalisert enn dem som svarer. Det betyr at det utvalget en sitter igjen med er systematisk positivt selektert i forhold til undersøkelsespopulasjonen. Undersøkelsen vil i så fall tegne et skjevt og for fordelaktig bilde av populasjonen. Dette kan føre til at myndighetene iverksetter feil politiske tiltak, noe som i verste fall kan bidra til å redusere gruppas velferd ytterligere. Hvor fortegnert bildet er, er det imidlertid umulig å vite uten å ha informasjon om hele populasjonen. Det er her registerdata kommer til nytte.

Vår interesse for skjevhet i frafallet er i første rekke knyttet til grad av marginalisering, sosialt og i forhold til arbeidsmarkedet. Registerinformasjonen i FD-trygd er uten tvil bedre for sistnevnte enn for førstnevnte. For førstnevnte må vi basere oss på grovere indikatorer som ekteskapelig status, statsborgerskap og innvandrersstatus. For sistnevnte, arbeidsmarkedstilknytning, har vi indikatorer som måler arbeidsinntekt, yrkesaktivitet, mottak av diverse trygdeytelser, mottak av sosialstønning, og antall måneder som sosialhjelpsmottaker. Vi har innhentet og koplet på registerinformasjon seks års yrkes-, trygde- og sosialhjelps-karrierer, det vil si for perioden 2000-2005. Grad av marginalisering kan måles på mange måter. En målestokk er arbeid på den ene siden, og sosialhjelp på den annen. Trygden befinner seg mellom disse «ytterposisjonene». Det er derfor av interesse å studere ikke bare indikatorer på tidligere «avhengighet» av sosialhjelpen og integrasjon i arbeidslivet, men også relasjonen til trygdesystemet. Det skal vi her gjøre i form av tidligere overføringer fra trygden generelt og mottak av dagpenger i forbindelse med arbeidsledighet spesielt. Også helse og funksjonsevne er sentrale anliggender i Funksjonsevneundersøkelsen, men siden informasjon om helse er nærmest ikke-eksisterende i FD-trygd, er det ikke mulig å studere frafall i relasjon til helse. (Helseatferd, som for eksempel sykefravær,

er dekket, men er problematisk å bruke som helseindikator i denne sammenhengen fordi yrkestilknytning er et vilkår for å motta sykepenger).

Data og metode

Funksjonsevneundersøkelsen var en undersøkelse av nesten 1273 langtids sosialhjelpsmottakere i 14 norske kommuner. Disse kommunene er identiske med de kommunene/sosialkontorene som deltok i det nasjonale «Tiltaksforsøket» og som vi hadde et utstrakt samarbeid med (Lødemel og Johannessen 2005). Formålet med undersøkelsen var dels å foreta en øyeblikkskartlegging av mottakernes ressurser, problem og levekår, og dels å legge grunnlaget for en prospektiv oppfølgingsstudie med utgangspunkt i registerdata for å finne ut hvordan det går med dem og belyse hvorfor noen klarer seg godt mens andre ikke lykkes så vel. Studien ble gjennomført som et tverrfaglig samarbeid ved Høgskolen i Oslo og involverte forskere med samfunnsvitenskapelig og helsefaglig kompetanse. Den ble finansiert av daværende Sosial- og helsedirektoratet. Detaljer om selve undersøkelsen, herunder om spørreskjemaet, samt sentrale funn finnes i van der Wel m.fl. (2006).

I Funksjonsevneundersøkelsen anvendte vi flere metoder for å øke svarprosenten. For det første mobiliserte vi sosialkontorene i de kommunene som langtidsmottakerne ble rekruttert fra til aktivt å finne og følge opp de sosialhjelpsmottakerne som inngikk i utvalget. For det andre brukte vi usedvanlig lang tid på datainnsamlingen, ni måneder (fra februar til november). For det tredje søkte vi å motivere respondentene til å svare ved å utlove flax-lodd til dem som besvarte og returnerte spørreskjemaene. For det fjerde gjennomførte vi fire purrerunder. Til tross for disse tiltakene og de ressurser som ble satt inn, oppnådde vi ikke en svarprosent på mer enn ca. 50. Vurdert ut fra faren for lav svarprosent og mulig utvalgsskjevhet, planla vi å benytte registerdata fra starten av. Vi ba derfor hver respondent om lov til å bruke registerdata om vedkommende, og hver enkelt ble gitt mulighet til å reservere seg skriftlig og muntlig mot dette. Prosjek-

tet ble vurdert av Norsk etisk komité for samfunnsvitenskap og humaniora (NESH) og ble godkjent av alle relevante instanser.

Avgrensning av respondenter ble gjort etter nærmere angitte kriterier. Respondentene skulle:

- Pr. 24. januar 2005 ha mottatt sosialhjelp i minst seks måneder i løpet av de siste tolv månedene.
- Være mellom 18 og 60 år.
- Ha sosialhjelp som hovedinntektskilde.

I tillegg inkluderte vi sosialhjelpsmottakere som pr. 24. januar 2005 deltok på tiltak i «Tiltaksforsøket» (eller en ekvivalent), eller som deltok på andre arbeidsrettede tiltak i kommunal eller statlig regi. Alle som fylte inklusjonskriteriene i de 14 kommunene ble innlemmet i bruttoutvalget. Kommunene er ikke tilfeldig valgt fra populasjonen av kommuner i Norge. Vi kan derfor ikke hevde at utvalget av «våre» langtidsmottakere er representativt for alle langtidsmottakere av sosialhjelp i Norge. De valgte kommunene har imidlertid en god geografisk spredning og omfatter både landkommuner og bykommuner. Denne artikkelen studerer dermed ikke representativiteten til et nasjonalt sannsynlighetsutvalg, men samsvarert mellom populasjonen av langtidsmottakere av sosialhjelp i de 14 kommunene, som alle var invitert, og de som svarte på spørreskjemaet og som tillot kopling til registerdata.

Vi sendte spørreskjemaer til sosialkontorene, som formidlet dem videre gjennom posten til den enkelte respondent i kommunen. Respondentene sendte utfylte skjemaer direkte til forskerne i vedlagt frankert konvolutt. I en del tilfelle drev også sosialkontoret aktiv innsats for å lete opp respondenter og oppmuntre dem til å svare. Sosialkontorene sendte oss informasjon om personnumrene til «sine» sosialhjelpsmottakere. Ved hjelp av personnumrene foretok Statistisk sentralbyrå sammenkoplingen av registerdata og spørreskjemadata og sendte oss de endelige filene i aidentifisert form.

I frafallsanalysen opererer vi dels med gjennomsnitt og dels med prosentandeler. Rundt disse estima-

tene har vi beregnet 95 prosent konfidensintervaller for hhv. registerutvalget og spørreskjemaet (Lane 13.4.2009). Et 95 prosent konfidensintervall angir at i 95 av 100 utvalg vil den sanne variabelverdien for populasjonen ligge mellom øvre og nedre grense for konfidensintervallet. Det innebærer at dersom konfidensintervallene rundt en variabelverdi (punkt-estimatet) for hvert av de to utvalgene overlapper, er det lite sannsynlig at de er statistisk forskjellige fra hverandre.

Resultater

Tabell 1 viser den noe komplekse datastrukturen i Funksjonsevneundersøkelsen. Sosialhjelpspopulasjonen i de 14 kommunene som sosialkontorene identifiserte i januar 2005 talte i alt 1273 individer som vist i cellen nederst i høyre hjørne i tabell 1. Registerutvalget omfatter 1129 individer, og spørreskjemaet 551.

Tabell 1

Datastrukturen i Funksjonsevneundersøkelsen. Antall respondenter i registerutvalget, spørreskjemaet og totalt.

Tabell 1 viser videre at vi mangler registerdata for 144 individer, eller 11,7 prosent av populasjonen. 123 av disse reserverte seg mot bruk av registerdata. 21 personer hadde ugyldige personnumre. Vi ser av tabellen at 75 av de 144 uten registerdata likevel har besvart spørreskjemaet.

For spørreskjemaet mangler vi kopling til registerdata for 75 personer, eller 14 prosent, i første rekke fordi respondentene reserverte seg. Endelig framgår det av Tabell 1 at de to utvalgene som skal sammenliknes har henholdsvis 1129 observasjoner (registerutvalget) og 476 observasjoner (spørreskjemaet). Andelen av populasjonen på 1273 personer som vi har både spørreskjemaopplysninger og registerdata om utgjør 37,4 prosent. Sosialkontorene oppga at 225 personer var ute av stand til å svare på spørreskjemaet, grunnet for eksempel rus, sykdom

og norskproblemer (van der Wel m.fl. 2006). Det kan diskuteres om disse med rette skal inngå i den populasjonen spørreundersøkelsen er ment å uttale seg om. Om disse trekkes fra, blir svarprosenten med opplysninger fra begge kilder 45,4 prosent. Rapporten fra Funksjonsevneundersøkelsen opererer med noe høyere bruttotall (van der Wel m.fl. 2006:9-10). Det skyldes at vi har sanert datasettene fordi registerinformasjonen tydet på at noen respondenter ikke tilfredsstilte inklusjonskriteriene.

SPØRRESKJEMA	JA	NEI	SUM
JA	476	75	551
NEI	653	69	722
SUM	1129	144	1273

Vi vil sammenligne svarresponsen langs 13 variabler hentet fra FD-trygd i hhv. registerutvalget og spørreskjemaautvalget (se Tabell 2). Tabell 2 inneholder punkttestimatene for disse variablene samt konfidensintervallene inne i parentesene.

Tabell 2 (se neste side) viser at det er merkbare prosentforskjeller mellom de to utvalgene for kjønn og en av aldersgruppene, 41-50 år. Forskjellen etter kjønn er drøyt 3 prosentpoeng (62-58,8=3,2); og for aldersgruppa 41-50 år er den på 3,3 (23,1-19,8).

Forskjellene mellom utvalgene for variablene 3-6 er gjennomgående små, rundt 1-2 prosentpoeng. Prosentandelen ugifte er imidlertid 4,7 prosentpoeng høyere i registerutvalget, mens prosentandelen skilte er noe høyere i spørreskjemaautvalget (3,1 prosentpoeng).

Variablene 7-12 måler aspekter ved arbeidsmarkedsmarginalisering. Vi ser at forskjellene mellom de to utvalgene varierer en del. For eksempel er gjennomsnittlig yrkesinntekt i spørreskjemaautvalget en del høyere enn i registerutvalget, mens andel uten overføringer fra folketrygden er omtrent den samme i begge utvalg.

For variabel 13, bostedskommune, er det også varierende, men jevnt over moderate forskjeller mellom

utvalgene. For eksempel er andelene i de to utvalgene omtrent like store i Moss, mens forskjellen er tydelig, og signifikant i Narvik. (Konfidensintervallene overlapper ikke).

Med unntak av forskjellen for Narvik, ser vi at konfidensintervallene for punkttestimatene for alle variablene i Tabell 2 overlapper, og det som regel med god margin. Det betyr at ingen av forskjellene i Tabell 2, med unntak av den for Narvik, er statistisk signifikante. Forskjellene skyldes sannsynligvis tilfældigheter.

Hovedinntrykket av Tabell 2 er at de som har svart på spørreskjemaet er noe mindre marginalisert enn personene i registerutvalget, slik vi kunne forvente. De har noe sterkere arbeidslivstilknytning, noe høyere yrkesinntekter, noe mindre «avhengighet» av sosialhjelpen målt med antall måneder de har mottatt sosialhjelp og utbetalte beløp over perioden registerdataene dekker. Det er videre en viss, men ingen sterk tendens til at spørreskjemaespondentene har lavere mottak av folketrygdytelse og dagpenger under arbeidsløshet. Det er imidlertid viktig å understreke at forskjellene mellom de to utvalgene i disse henseender jevnt over ikke er dramatiske. Forskjellene mellom de to utvalgene er heller ikke statistisk signifikante i og med at konfidensintervallene overlapper.

Dette er ikke stedet for en substansiell beskrivelse av langtids sosialhjelpsmottakeres levekår. Det er likevel fristende å peke på at registerdataene underbygger det bilde som ble tegnet av sosialhjelpsmottakerne i spørreskjemaundersøkelsen (van der Wel m.fl. 2006); de har få ressurser, opplever mange og store problemer og er relativt svakt integrert i det norske samfunnet, særlig i arbeidslivet.

Diskusjon

I denne analysen har vi ved hjelp av registerdata fra FD-trygd vurdert hvor skjevt frafallet er i Funksjonsevneundersøkelsen, en survey av langtidsmottakere av sosialhjelp. Vi har sett på forskjeller i andeler og gjennomsnittsverdier langs 13 variabler i et registerutvalg og et spørreskjemaautvalg basert på samme sosialhjelpspopulasjon. Variablene måler sosiodemografiske

Tabell 2. Fordelingen på utvalgte variable i registerutvalget og spørreskjemaet. Data fra FD-trygd. 95 prosent konfidensintervaller i parentes.

	Registerutvalget N=1129	Spørreskjemaet N=476
1. Kjønn		
Mann	62,0 (59,1 – 64,8)	58,8 (54,4-63,2)
Kvinne	38,0 (35,2 – 40,9)	41,2 (36,9-45,7)
2. Alder		
19-30	46,0 (43,1-48,9)	44,3 (39,9-48,8)
31-40	26,4 (23,9-29,1)	24,2 (20,5-28,2)
41-50	19,8 (17,5-22,2)	23,1 (19,6-27,1)
51-60	7,9 (6,5-9,6)	8,4 (6,2-11,4)
3. Innvandrersstatus		
Norsk	78,3 (75,8-80,6)	79,6 (75,7-83,0)
Ikke norsk	21,7 (19,4-24,2)	20,4 (17,0-24,3)
4. Statsborgerskap		
Norsk	89,2 (87,2-90,9)	88,3 (85,1-91,0)
Ikke norsk	10,8 (9,2-12,8)	11,7 (9,1-14,9)
5. Utdanning		
Ingen utd/ikke fullført grunnskole	3,2 (2,2-4,2) 21,0 (18,6-23,4)	4,8 (1,3-5,1) 19,7 (17,4-24,6)
Ungdomsskole	52,5 (49,6-55,4)	51,7 (48,0-57,0)
Videregående, grunnutdanning	13,2 (11,2-15,2)	14,5 (10,0-16,4)
Vgs, avsluttende + påbygging	3,8 (2,7-4,9)	3,4 (2,2-5,4)
Universitets- og høyskolenivå	6,3 (4,9-7,7)	5,9 (4,2-8,4)
Uoppgitt		
6. Ekteskapelig status 2005		
Ugift	71,7 (69,1-74,4)	67,0 (62,7-71,3)
Gift	10,5 (8,7-12,3)	11,2 (8,3-14,1)
Enke/enkemann	1,2 (0,5-1,8)	1,1 (0,1-2,0)
Skilt	12,6 (10,6-14,5)	15,7 (12,4-19,0)
Separert	3,5 (2,5-4,6)	5,0 (3,0-6,9)
Skilt partner	0,1	
7. Andel uten yrkesaktivitet i hele observasjonsperioden (0 i yrkesinntekt)	25,4 (22,9-28,0)	23,2 (19,6-27,3)
8. Yrkesinntekt (kroner akkumulert over seks år, gj.snitt)	134 311 (+-12497)	147 437 (+-20715)
9. Sosial stønad (kroner akkumulert over seks år, gjennomsnitt/andel < 100 000)	300658 (+-12551)/ 17,4 (15,3-19,7)	278342 (+-19569)/ 21,3 (18,1-25,5)
10. Måneder på sosialhjelp (akkumulert over seks år, gj.snitt/andel under 24 mnd)	38,4 (+1,82) / 28,8 (26,2-31,5)	36,3 (+1,80) / 31,7 (27,5-36,2)
11. Overføringer fra folketrygden (akkumulert over seks år, gj.snitt/andel uten)	80913 (+-8024) / 57,5 (54,7-60,4)	92303 (+-13988) / 57,2 (52,6-61,7)
12. Dagpenger under arbeidsløshet (akkumulert over seks år, gj.snitt/andel uten)	29373 (+-4826) / 78,2 (75,7-80,5)	31813 (+-7372) / 75,1 (70,9-78,9)

13. Bostedskommune		
Bergen/Åsane	21,7 (19,3-24,1)	17,9 (14,4-21,3)
Gjøvik	8,2 (6,6-9,8)	9,7 (7,0-12,3)
Harstad	5,5 (4,2-6,8)	5,9 (3,8-8,0)
Kautokeino	0,9 (0,3-1,4)	1,3 (0,3-2,3)
Kristiansund	6,9 (5,4-8,4)	9,0 (6,5-11,6)
Meldal	0,9 (0,3-1,4)	1,7 (0,5-2,8)
Molde	7,2 (5,7-8,7)	6,1 (3,9-8,2)
Moss	21,3 (19,0-23,7)	21,2 (17,5-24,9)
Narvik	8,2 (6,6-9,8)	4,6 (2,7-6,5)*
Nedre Eiker	7,7 (6,2-9,3)	8,6 (6,1-11,1)
Oppegård	4,3 (3,2-5,5)	2,7 (1,3-4,2)
Orkdal	3,3 (2,2-4,3)	5,9 (3,8-8,0)
Randaberg	1,2 (0,6-1,9)	1,7 (0,5-2,8)
Skaun	2,6 (1,6-3,5)	3,8 (2,1-5,5)

* Signifikant forskjellig andel

kjennetegn, arbeidsmarkedsressurser og aspekter ved integrasjon og marginalisering i forhold til arbeid og stønadssystem. Det er en tendens til at respondene av spørreskjemaet er noe bedre stilt, noe bedre sosialt integrert og står noe nærmere arbeidsmarkedet enn registerutvalget. Gjennomgående er disse forskjellene moderate, og ingen er statistisk signifikante vurdert ut fra konfidensintervallene (bortsett fra i én av 14 kommuner). Det synes således forholdsvis trygt å generalisere fra spørreskjemaet til registerutvalget (og trolig også til populasjonen, se nedenfor), i hvert fall i forhold til de 13 variablene vi har vurdert i denne analysen. At forskjellene mellom spørreskjemaet og registerutvalget ikke var større med et frafall på nesten 60 prosent, er overraskende. Én mulig forklaring kan være at indikatorene brukt her, som var tilgjengelige i registerdata, ikke fanger opp relevante skjevheter i frafallet. Kanskje er det slik at tilgjengelige «objektive» registervariabler ikke er like relevante indikatorer i undersøkelser av marginale grupper som i befolkningen generelt, men at andre faktorer som helse, rusbruk, mestring og generell tillit er mer avgjørende? Datainnsamlingen foregikk som nevnt i fire bølger, og en tidligere analyse av responstid viste at innvandrere og respondenter med lese- og skrivevansker var overrepresentert ved de senere innsamlingsrundene, mens det ikke var signifikante for-

skjeller i responstid for helse, kjønn, alder, utdanning, rus- og alkoholmisbruk (van der Wel m.fl 2006). Gitt at de som svarer sent, ligner på dem som ikke svarer (Filion 1975), gir dette langt fra full støtte til en hypotese om at det har foregått systematisk og konsistent frafall på variabler som ikke er observert gjennom registerdata. Vi ser det som svært nyttig å ha tilgang til et såpass rikt tilfang av landsomfattende registerdata som vi har i Norge slik at vi kan foreta denne typen analyser. En survey som Funksjonsevneundersøkelsen blir langt mer pålitelig og verdifull både for vitenskapelige og sosialpolitiske formål.

Et komplett befolkningsregister som FD-trygd har flere sterke sider. Frafallet er lite selv om vi opplevde at en del fødselsnumre ikke lot seg kople. Det er en styrke å ha data som dekker mer enn ett år. På den måten spiller årlige tilfeldige variasjoner mindre rolle, og funnene blir mer pålitelige. Svakheten ligger ikke først og fremst i registeret, men gjelder informasjon som ikke finnes i registeret, opplysninger om helse blant annet. Antakelig vil den muligheten vi etter hvert har fått til å kople på informasjon fra Reseptregisteret avhjelpe denne situasjonen et stykke på vei (Reseptregisteret 13.4.2009).

Som nevnt har kravet om informert samtykke til bruk av registerdata ført til at vi ikke har registerdata for hele sosialhjelpspopulasjonen på 1273 individer.

Populasjonen er blitt redusert med 123 personer eller 10 prosent fordi de reserverte seg. Her er en kilde til en mulig skjevhet som vi har begrensede muligheter til å kontrollere. Likevel skal det mye til at frafallet som skyldes reservasjon skal kullkaste resultatene fra denne analysen fullstendig. For at frafallet skal gjøre det, må det være ekstremt skjevt, i og med at antallet er forholdsvis beskjedent. Frafallsanalysen utført i denne artikkelen tillater oss imidlertid kun å slutte at spørreskjemaet er representativt for registerutvalget, og strengt tatt ikke for hele sosialhjelpspopulasjonen. Dette er selvsagt en begrensning, en begrensning som skyldes kravet om samtykke til bruk av registerdata. Spørreskjemaet ble av samme grunn redusert med 75 personer eller 14 prosent. Disse besvarte altså spørreskjemaet men samtykket ikke til at vi kunne innhente registerdata om dem. Dette innebærer at vi ikke har hatt mulighet til å vurdere frafallet i hele spørreskjemaet ved hjelp av registerdata, kun for 476 av totalt 551. Siden vi har spørreskjemaopplysninger om alle 551 er mulighetene likevel gode til å vurdere akkurat dette frafallet. Det er likevel et problem som gjenstår. Som nevnt ble Funksjonsevneundersøkelsen planlagt som en prospektiv oppfølgingsstudie der vi ville basere oss på årlige registerdata fra FD-trygd for å analysere sosialhjelpsmottakernes videre bevegelser i arbeidsmarkedet og i stønadssystemet i årene etter 2005. Opplysninger fra spørreskjemaet vil være vesentlige for å kunne analysere årsakene til at noen klarer seg bra, mens andre strir. I spørreskjemaet har vi en lang rekke opplysninger som ikke finnes i norske registerdata, og som derfor vil kaste helt nytt lys over sosialhjelpsmottakernes «livsløp». For eksempel har vi en rekke opplysninger om helse i spørreskjemaet som vi antar vil være av vesentlig betydning for å forutsi hvordan det går med sosialhjelpsmottakerne over tid. Helseopplysninger er mangelvare i registerdataene. Det er denne kombinasjonen av spørreskjemaopplysninger og årlige, prospektive forløpsdata fra FD-trygd som vi tror vil gjøre Funksjonsevneundersøkelsen spesielt verdifull. Imidlertid vil et frafall på 14 prosent (av totalt 551 respondenter) svekke den statistiske styr-

ken ved undersøkelsen.

De tekniske, forskningsmessige problemene vi nå har pekt på gir anledning til på et prinsipielt grunnlag å drøfte sambandet mellom kravet til individuelt samtykke og samfunnets behov for relevant og pålitelig kunnskap om vanskeligstilte og marginaliserte grupper.

I de forskningsetiske retningslinjene til Norsk etisk komité for samfunnsvitenskap og humaniora (NESH) (NESH 13.4.2009) punktene 5-9 framheves at forskningen skal rette seg etter krav om respekt for menneskeverdet, integritet, frihet og medbestemmelse. Dette inkluderer et krav om at respondentene har rett til å gi fritt og informert samtykke. NESHs diskusjon av dette temaet dreier seg om innhenting av nye data. Forskeres bruk av allerede innsamlede, eksisterende data som FD-trygd drøftes ikke eksplisitt. Spørsmålet er om kravet til samtykke også skal gjelde uavkortet for bruk av registerdata. Det erkjennes under punkt 9 at kravet om frivillig og informert samtykke ikke kan gjelde absolutt. Her står det:

«Selv om informert samtykke er den klare hovedregelen også i prosjekter hvor deltakerne ikke selv deltar aktivt, kan unntak fra kravet om informert samtykke i visse tilfeller være aktuelt i situasjoner der forskningen ikke innebærer fysisk kontakt med forsøkspersonene, opplysningene som behandles er lite sensitive, og der forskningen samtidig har en nytteverdi som klart overstiger ulemper som kan påføres den inkluderte». (NESH 13.4.2009:14)

Dette utsagnet er høyst relevant for forskeres bruk av registerdata som allerede er samlet inn. De registerdataene vi har innhentet innfrir to av de tre kravene i sitatet, ikke fysisk kontakt og stor nytteverdi. Det tredje kravet, «lite sensitive» vil også være delvis oppfylt i det opplysninger om arbeid og yrkesdeltakelse neppe kan karakteriseres som sensitive, mens mottak av sosialhjelp nok vil.

Kravet til samtykke gjelder altså ikke absolutt. Vi har allerede pekt på, og forhåpentlig godtgjort, verdien ved å bruke registerdata til å vurdere hvor representativt spørreskjemaet er, og til å foreta

prospektive oppfølgingsundersøkelser av materialet. Samtykkekravet har redusert verdien av registerdataene, men i vårt tilfelle er verdiforringelsen neppe dramatisk. Men dette spørsmålet har mer prinsipiell rekkevidde. I framtida kan en risikere at en langt høyere prosentandel reserverer seg mot forskeres bruk av registerdata. I så fall blir verdien av en frafallsanalyse som gjør bruk av registerdata kraftig svekket. Usikkerheten knyttet til hvor skjevt et spørreskjema utvalg av en vanskeligstilt og marginalisert gruppe er, vil vedvare. I sin tur kan uoppdaget skjevhet i utvalget føre til at det utvikles en sosialpolitikk for marginaliserte grupper som er feilaktig innrettet og galt dimensjonert. I verste fall kan dermed en streng håndhevelse av personvern og samtykkekrav føre til at vanskeligstilte grupper får det enda vanskeligere, eller at de ikke får en forbedring de ellers kunne ha fått. Sekretariatsleder i Den nasjonale forskningsetiske komité for medisin (NEM), Knut Ryuter, har framhevet dette poenget i et intervju med bladet *Forskningsetikk* (2005): «En sårbar gruppe blir ikke mindre sårbar ved ikke å få tilført ny kunnskap. Derfor er denne forskningen ekstra viktig. Den kan føre til et gode for individet og for gruppen som man ellers ikke ville finne ut av». Å la være å forske på marginaliserte grupper kan også ha skadevirkninger for gruppa som sådan og for enkeltindivider i gruppa. Det samme kan selvsagt upålitelig

forskning. På den annen side kan forskning som gjør bruk av registerdata også ha skadevirkninger og føre til belastninger for dem det forskes på. Et strengt krav om samtykke til bruk av registerdata vil vektlegge hensynet til de mulige skadevirkninger forskning kan ha for de aktuelle gruppene tyngre enn de fordeler forskningen kan ha for de samme gruppene. I en slik avveining er det viktig at de fordelene bruk av registerdata gir, og som vi har pekt på i denne artikkelen, kommer behørig fram.

Konklusjonen er at registerdata som er så rike som i Norge, og som omfatter praktisk talt hele populasjonen, er nyttige til å vurdere skjevhet i frafall når vi utforsker marginaliserte grupper i befolkningen ved hjelp av spørreskjemaundersøkelser. Vi finner at spørreskjema utvalget i Funksjonsevneundersøkelsen ikke avviker vesentlig fra registerutvalget (som utgjør 88 prosent av populasjonen). Det er derfor liten fare forbundet med å bruke spørreskjema utvalget til å trekke slutninger om populasjonen av langtidsmottakere av sosialhjelp i de 14 kommunene som inngår i undersøkelsen. Krav om samtykke til å bruke registerdata representerer en fare for at fordelene ved å bruke registerdata vil svekkes hvis mange reserverer seg, noe som i sin tur vil gi mindre pålitelig kunnskap og dårligere grunnlag for å utvikle en effektiv sosialpolitikk for marginaliserte grupper.

SUMMARY

Recent years have seen an increased demand for scientific knowledge about disadvantaged and marginalised groups. At the same time, this kind of research faces certain methodological challenges. Here we will present an analysis of selection bias and attrition in the survey "Social and physical functioning", conducted by Oslo University College in 2005. The survey targeted a population of 1273 long term social assistance recipients living in 14 municipalities. We use linked register data on the target population to investigate possible bias in the survey sample. The findings indicate no significant bias in the survey sample compared to the recorded target population. The paper discusses the ethical claim for free and informed consent in the collection of register data, in relation to the advantage of using such data to account for attrition bias in social research surveys.

LITTERATUR

- Akselsen, Anders, Sandra Lien og Øyvind Sivertstøl** (2007) *FD-Trygd variabelliste*. SSB notat 2007/17. Oslo: Statistisk sentralbyrå.
- Breivik, Ivar** (1996) *Sosialhjelpsklienters arbeid, helse og oppvekst*. Oslo: Norsk institutt for by- og regionforskning.
- Friestad, Christine og Skog Hansen, Inger Lise** (2004) *Levekår blant innsatte*. Fafo-rapport 429. Oslo: Fafo.
- Fyhn, Anders Brede og Dahl, Espen** (2000) *Mot tredje generasjons levekårsforskning*. Fafo-notat 2000:1. Oslo: Fafo.
- Blane, David, Netuveli, Gopalakrishnan og Juliet Stone** (2007) The development of life course epidemiology. *Revue d'épidémiologie et de sante publique*, Feb. 55(1), side 31-8.
- Filion, F.L.** (1975) Estimating bias due to nonresponse in mail surveys. *The Public Opinion Quarterly*, Vol. 39, No. 4 (Winter, 1975-1976), pp. 482-492
- Halvorsen, Knut** (2000) Sosial eksklusjon som problem. *Tidsskrift for velferdsforskning*, nr. 3/2000, side 157-171.
- Hove, Olav** (1992) *En ulykke kommer sjelden alene. Langtidsmottakere av sosialhjelp – levekår og livssituasjon*. Oslo: NotaBene.
- SSB.no** <http://www.ssb.no/emner/03/fd-trygd/> (Hentet 13.4.2009).
- Kleinbaum, David G., Kupper, Lawrence L. og Morgenstern, Hal** (1982) «Selection Bias» I Kleinbaum, David G., Lawrence L. Kupper og Hal Morgenstern (red) *Epidemiologic Research*. New York: Van Nostrand Reinhold Company Inc. side 194-219.
- Lane, David** http://davidmlane.com/hyperstat/confidence_intervals.html. (Hentet 13.4.09).
- Lorentzen, T., Nielsen, R.** (2008) *Går fattigdom i arv? Langtidseffekter av å vokse opp i familier som mottar sosialhjelp*. Fafo-rapport 2008:14.
- Lødemel, I. og Johannessen, A.** (2005) *Tiltaksforsøket: mot en inkluderende arbeidslinje?* HIO-rapport 2005 nr. 1. Oslo: Høgskolen i Oslo.
- Mjøset, L.** (1991) *Kontroverser i norsk sosiologi*. Oslo: Universitetsforlaget.
- NESH** <http://www.etikkom.no/retningslinjer/NESHretningslinjer/NESHretningslinjer/o6> (Hentet 13.4.2009).
- Reseptregisteret.** <http://www.reseptregisteret.no/> (Hentet 13.4.2009)
- Ryuter, Knut** (2005) Redusert samtykkekompetanse – og hva så? Intervju med Knut Ryuter i Forskningsetikk nr. 3 (5) 4-7.
- Seim, Sissel og Hjemdal, Ole K.** (1992) *Den som intet har – livskvalitet og levekår blant klienter med langvarig kontakt med sosialkontoret*. NKHS-rapport 6. Oslo: Norges kommunal- og sosialhøgskole.
- SSB** (2008) *Økonomi og levekår for ulike grupper, 2007*. Oslo: Statistisk sentralbyrå.
- Stenberg, Sten-Åke** (2000) Inheritance of Welfare Reciprocity: An Intergenerational Study of Social Assistance Reciprocity in Postwar Sweden. *Journal of Marriage and Family*, 62(1), side 228-239.
- Stortingsmelding nr. 6** (2002-2003) *Tiltaksplan mot fattigdom*. Oslo: Arbeids- og inkluderingsdepartementet.
- Stortingsmelding nr. 9** (2006-2007) *Arbeid, velferd og inkludering*. Oslo: Arbeids- og inkluderingsdepartementet.
- Stortingsproposisjon nr. 1** (2008-2009) *Handlingsplan mot fattigdom*. Særskilt vedlegg. Oslo: Arbeids- og inkluderingsdepartementet.
- Søgaard, Anne Johanne, Selmer, Randi, Bjertness, Espen og Thelle, Dag** (2004) The Oslo Study: The impact of self-selection in a large, population-based survey. *International Journal for Equity in Health*, 3,3.
- Underlid, Kjell.** (2005) *Fattigdommens psykologi. Oppleving av fattigdom i det moderne Norge*. Oslo: Det Norske Samlaget.
- Underlid, Kjell.** (2008): *Fattigdommens lukt og smak – Funn frå ein studie om korleis fattigdom vert opplevd. Tidsskrift for psykisk helsearbeid*, 3: 213-222.
- van der Wel, Kjetil A., Dahl, Espen, Lødemel, Ivar, Løyland, Borghild, Ohrem Naper, Sille og Slagsvold, Marit** (2006). *Funksjonsevne blant langtidsmottakere av sosialhjelp*. HIO rapport 2006:29. Oslo: Høgskolen i Oslo. (<http://www.hio.no/content/view/full/18968>).