

Interne notat

STATISTISK SENTRALBYRÅ

86/17

7. april 1986

TRE NOTAT OM REGISTER I SAMBAND MED AKU

av

Kjell Arne Brekke

Innhald:

1. OVERREPRESNTASJON AV ARBEIDSTAKARAR I BRUTTOUTVALGET TIL AKU
Eit tappert forsøk på å finne årsakene 1
2. KONSISTENS MELLOM ARBEIDSTAKER-REGISTERET OG
ARBEIDSKRAFTUNDERSØKELSEN 16
3. UFULLSTENDIGE FAMILIER I 10% REGISTERET 29

**OVERREPRESENTASJON AV ARBEIDSTAKARAR I
BRUTTOUTVALGET TIL AKU.
EIT TAPPERT FORSØK PÅ Å FINNE ÅRSAKENE.**

Innhald:

1. Innledning	2
2. Situasjonen etter innføring av ny estimeringsmetode	2
3. Nye eller gamle skjevheter	4
4. Forventa andel arbeidstakarar gitt første trinns trekking	5
5. Andre trinns trekking	5
6. Årsaker til overrepresentasjon.....	7
6.1 Klyngeeffekter	8
6.2 Tilgang og avgang.....	8
6.3 Merkerutiner	9
6.4 Kobling av AKU og A/A registeret	9
6.5 Trekkinga av utvalget.....	9
6.6 Definisjonar av populasjonen	10
7. Ledda mellom personregisteret og det endelige utvalget ...	11
7.1 Resultat	11
8. Konklusjon	12
8.1 Anbefaling	13
Appendiks	14

1. Innledning

Ved innføring av ny estimeringsmetode (registermetoden) for AKU har ein observert store skift i talla i forhold til tidligare estimeringar. Det viste seg at desse skifta kom av overrepresentasjon (OR) av arbeidstakarar i bruttoutvalget til AKU.

Vi ønska å finne ut meir om årsakene til denne overrepresentasjonen. AKU-utvalga blir trekkd ved ein to-trinns utvalgsplan. Første trinn er trekking av utvalgskommuner som skal representere eit utvalgsområde. Resultatet frå første trinn har blitt holdt fast sidan 1975. I andre trinn blir utvalget trekkd frå utvalgskommunene. Det blir trekkd eit husholdningsutvalg, ikkje eit personutvalg. Dei operative definisjonane av husholdning kan også ha effekt for overrepresentasjonen. Vi undersøkte derfor forventa andel arbeidstakarar (FAA) gitt 1.trinns trekking og gitt reglane for 2.trinns trekking.

FAA gitt første trinns trekking er tiligare beregna for FOB begrepet ansatt. Dette ansettings begrepet kan vi anta tilsvarar AKU's. Beregningane er beskrivne i byråets serie Rapportar 85/34 : "Statistiske egenskaper ved byråets standard utvalgsplan." av Tor Haldorsen.

Haldorsen beregna også enkelte parametrar som gir informasjon om FAA gitt definisjonane i andre trinns trekking. Ingen av desse ga overrepresentasjon. Det er vanskelig å tenke seg effektar som kunne føre til noko slikt .

Det endelige AKU bruttoutvalget er ikkje det same som det som blir trekkd i byrået. Under feltarbeidet blir enkelte personar strokne av utvalget, og andre blir lagt til. Vi har undersøkt kva dette betyr for andelen arbeidstakarar i utvalget.

Formelar og definisjonar for dei beregningane som blir presentert her er gjengitt i eit appendiks bakerst i notatet.

2. Situasjonen etter innføring av ny estimeringsmetode

Dei omtalte skifta etter overgang til ny estimeringsmetode er gjengitt etter yrkesstatus i tabell 1 nedanfor. Beregningane er utført ved 6.kontor. I gjennomsnitt blir estimert antall sjølvstendige 13 300 fleire ved registermetoden enn ved den gamle metoden. Dette er ein auke på 6.5%. Ansatte blir redusert med 41 700 som er ein reduksjon på 2.5% og total sysselsetting blir redusert med 25 600, ein reduksjon på

1.37.

Tabell 1: Skift i antall sysselsatte ved overgang til ny estimeringsmetode, etter yrkesstatus. 1000 personer.

Kvartal:	83 1	83 2	83 3	83 4	84 1	84 2	84 3
<u>Sjølvtendige:</u>							
Med reg.	207.6	207.4	210.8	211.6	219.9	202.9	193.9
Uten reg.	198.2	196.2	195.6	196.5	197.5	192.5	184.2
Differanse	9.4	11.2	15.2	15.1	22.4	10.4	9.7
<u>Ansatte</u>							
Med reg.	1648.9	1667.6	1651.0	1673.9	1673.7	1682.8	1658.9
Uten reg.	1670.6	1712.6	1702.2	1718.8	1729.1	1722.3	1691.3
Differanse	-21.7	-45.0	-51.2	-44.9	-55.4	-39.5	-34.4
<u>Familie medlemmer:</u>							
Med reg.	45.1	51.0	52.6	42.5	43.5	49.2	49.9
Uten reg.	44.2	50.5	49.0	41.7	43.1	48.2	48.1
Differanse	0.9	0.5	3.6	0.8	0.4	1.0	1.8
<u>Uoppgitt:</u>							
Med reg.	8.0	7.6	7.5	12.3	7.2	14.9	11.5
Uten reg.	7.9	7.8	7.6	12.0	7.4	15.5	12.0
Differanse	0.1	-0.2	-0.1	0.3	-0.2	-0.6	-0.5
<u>Totalt:</u>							
Med reg.	1909.7	1933.6	1922.0	1940.2	1944.2	1949.8	1914.2
Uten reg.	1921.0	1967.1	1954.3	1969.0	1977.1	1978.5	1935.7
Differanse	-11.3	-33.5	-22.3	-28.8	-32.9	-28.7	-21.5

For å finne kjeldene til disse skifta vart andel arbeidstakarar undersøkt. Vi såg då på andelen i heile populasjonen, i bruttoutvalget og i nettoutvalget. Også desse beregningane vart utført ved 6.kontor. Andelane er gjengitt i tabell 2 under:

Tabell 2. Andel registrerte arbeidstakarar. I heile populasjonen , i bruttoutvalget og i nettoutvalget.

år kvart.	1	2	3	4	5
	Totalt	Brutto Utvalget	Netto Utvalget	Differ 2 - 1	Differ 3 - 2
82 2	54.03	55.84	55.98	1.81	0.14
82 3	53.04	55.31	55.50	1.91	0.19
82 4	52.95	53.72	53.82	0.78	0.10
83 1	52.43	53.26	53.30	0.83	0.04
83 2	52.65	54.57	54.89	1.92	0.31
83 3	53.23	55.17	55.35	1.94	0.18
83 4	52.70	54.42	54.78	1.72	0.36
84 1	52.36	54.29	54.94	1.93	0.65
84 2	52.85	53.84	54.41	0.99	0.57
84 3	53.55	54.42	54.99	0.87	0.57
gj.snitt.	53.01	54.49	54.80	1.48	0.31

Differansen i kolonne 5 kjem av frafallet. Vi ser at denne er liten i forhold til diferansen i kolonne 4. Det er derfor rimelig å konkludere med at skiftet ved overgang til nye metoder kjem av overrepresentasjonen av arbeidstakarar i bruttoutvalget. I resten av dette notatet, vil vi prøve å finne årsakene til denne overrepresentasjonen.

3. Nye eller gamle skjevheter.

Eit naturleg spørsmål når vi observerer slike skift som i dette tilfellet er om det er gamle skjevheter som blir retta opp, eller om det er nye som blir introdusert.

Etter standard teori for etterstratifisering skulle eventuelle skjevheter bli redusert ved registermetoden. Vi forutset då at det ikkje er korrelasjon mellom mekanismene i registreringa til A/A registeret og trekkinga av våre utvalg. Det er vanskelig å sjå kva som skulle kunne skape ein slik korrelasjon.

Dersom det er skeive aldersfordelingar e.l. som ligg til grunn for dei observerte skjevhetene kan det være andre etterstratifiseringar som er meir effektive. Når det gjeld frafall og førstetrinnseffekt har vi derimot grunn til å tru at slike etterstratifiseringar er mindre effektive.

For å prøve å få eit visst empirisk grunnlag til å vurdere kva metode som gir minst skjevhet har vi estimert andel sysselsatte pr. år ifølge FOB, etter begge metodene. Estimator 1 er ein registermetode estimator, estimator 2 er uten register. Estimeringa er foretatt på nettoutvalget frå tilleggsundersøkelsen til AKU i 1980. Frafallet er her ekstra stort, og nettoutvalget var då også svært skjevt. Sysselsatte etter registeret var overrepresentert med nesten 5%.

Desverre var talla i utgangspunktet prosentuert med forskjellig grunnlag. Estimator 1 var prosentuert i forhold til heile befolkninga, medan estimator 2 var prosentuert i forhold til befolkninga 16-74 år. Vi har omregna ved å bruke tall frå FOB 1980 hefte II. Dersom aldersfordelinga på det tidspunkt TU vart gjennomført var forskjellig frå denne kan dette ha ført til skjevheter. Slike skjevheter er truleg likevel neglisjerbare.

Tabell 3: Andel sysselsatte estimert med (Est 1) og uten (Est 2) register, samanlikna med "fasit" frå FOB.

	Andel Syssels.	Diff Est - FOB
FOB	70.0	
Est 1	70.6	0.6
Est 2	73.0	3.0

Til samanlikning er standard-avviket 0.46 (Eg har då ikkje tatt hensyn til designeffekter og to-trinns trekking) Sjølv om netto-utvalget frå tilleggsundersøkelsen ikkje er representativ for AKU, tyder dette resultatet på at registermetoden reduserer skjevhetene svært effektivt.

4. Forventa andel arbeidstakarar gitt første trinns trekking

Vi har beregna andelen arbeidstakarar i dei kommunene som er trekk ut i første trinn. (Veid etter størrelsen på det strata dei er trekk frå) Dette kan vi så samanlikne med andel arbeidstakarar totalt. Resultatet er gjengitt i tabell 4.

Tabell 4. Andel arbeidstakarar i utvalgskommunene og totalt etter fylkespar og kjønn.

Ut.k - Utvalgskommunene	Tot. - Totalt i alle kommuner		Kvinner		Begge kjønn	
	Menn ut.k.	tot.	ut.k.	tot.	ut.k.	tot.
Akershus	67.95	67.83	55.61	55.47	61.78	61.65
Oslo	64.39	64.39	54.86	54.86	59.63	59.63
Øst- Vestfold	58.90	58.63	43.88	44.30	51.39	51.47
Hedem. - Oppl.	55.26	54.72	43.34	42.96	49.30	48.84
Buske. - Tele.	60.50	60.28	46.05	46.00	53.28	53.14
Agder - Rogal.	60.22	59.17	45.04	44.06	52.63	51.62
Hord. - So.&F.	59.10	57.72	47.90	45.87	53.50	51.80
Møre og Romsd.	52.73	52.42	38.98	41.35	45.86	46.89
Trøndelag	57.05	56.38	47.35	44.95	52.20	50.67
Nordland	51.12	49.95	42.62	41.37	46.87	45.66
Troms - Finnm.	50.77	49.23	46.37	45.94	48.57	47.59
Totlat					53.07	52.51

Vi ser at andel arbeidstakere i dei uttrukne kommunene ligg ca 0.567 over gjennomsnittet. Dette er lite samanlikna med den totale skjevhet i første trinns trekking. Det er likevel større enn den skjevhet som Tor Haldorsen har funne for FOB-begerpet ansatt. Den var på 0.35%. Forskjellen er likevel ikkje større enn at det ikkje gir grunn til å tru at registermetoden skal gi særleg skeive tall.

Det er likevel grunn til å merke seg at OR er ganske store for enkelte fylkespar. Dei kan bli enda større for einskildfylker. Dette betyr at

OR gitt første trinns trekking er av stor betydning når vi skal estimere antall ansatte i det enkelte fylket. Ved bruk av registermetoden vil vi likevel tru at det meste av OR blir retta opp.

5. Andre trinns trekking

Vi har prøvd å isolere den del av overrepresentasjonen som kjem av andre trinns trekking. Vi treng då å kjenne FAA gitt første trinn og samanlikne denne med faktisk andel i bruttoutvalget. Resultatet kan vi midle over tid.

FAA gitt første trinn finn vi ved å beregne andelen arbeidstakarar i dei komunene som er uttrukne frå kvart strata, og så gi dei vekt som tilsvarar stratastørrelsen. Då vi ikkje hadde utvalgsplanen tilgjengelig på maskinlesbar form har vi i staden gitt kvar utvalgt kommune vekt som tilsvarar størrelsen på bruttoutvalget frå komuna. Då utvalget er tilnerma sjølvveiande vil denne tilnerminga ikkje inføre nokon feil av betydning. Fullstendig beskriving av beregningsmetoden er gitt i appendiks.

Resulatet er gjengitt i tabell 5.

Vi har gjennomført beregningane berre for tre kvatal. På bakgrunn av tidligar beregningar av overrepresentasjonen kan vi likevel anslå gjennomsnittet til ca 0.9%. Vi har då forutsatt at overrepresentasjonen p.g.a første trinns trekking ikkje endrar seg vesetleg over tid.

Tabell 5. Forventa og faktisk andel arbeidstakarar i bruttoutvalget

<u>Kvartal</u>	<u>Faktisk</u>	<u>FAA / 1.tr.</u>	<u>Diff.</u>
1.kvart. 83	53.262	53.045	0.217
2.kvart. 83	54.575	53.292	1.283
3.kvart. 83	55.166	53.917	1.249

Vi har også sett på andelen arbeidstakarar i bruttoutvalget i komuner som er trekket med sannsyn 1. Denne har vi samanlikna med andelen i kommuna totalt. Resultatet er gjengitt i tabell 6. Vi ser at kommuner der bruttoutvalget overrepresenterer arbeidstakarar er i klar overvekt (19 - 5).

Tabell 6. Registerandel i komuner trekt i eit trinn. Bruttoutvalg og Heile kommunen. 2.kvartal 1983.

Kommune	Obs	Br.u.	tot.	Differanse
0101 Halden	62	56.1	52.1	+ 4.0
0103 Fredrikstad	77	46.3	49.1	- 2.8
0104 Moss	84	57.6	55.7	+ 1.9
0219 Bærum	233	63.4	62.8	+ 0.6
0220 Asker	114	63.7	63.4	+ 0.3
0231 Skedsmo	87	61.1	64.2	- 3.1
0301 Oslo	1307	59.7	59.1	+ 0.6
0412 Ringsaker	85	60.9	51.2	+ 9.7
0501 Lillehammer	57	61.1	56.2	+ 4.9
0602 Drammen	137	60.0	59.0	+ 1.0
0605 Ringerike	46	58.1	55.0	+ 3.1
0706 Sandefjord	47	58.6	50.1	+ 8.5
0805 Porsgrunn	103	54.5	53.9	+ 0.6
0806 Skien	131	61.2	54.4	+ 6.8
1101 Kristiansand	104	56.5	52.6	+ 3.9
1102 Sandnes	111	71.3	62.2	+ 9.1
1103 Stavanger	243	61.3	59.0	+ 2.3
1149 Karmøy	112	42.2	45.8	- 3.7
1201 Berg	518	55.5	57.5	- 2.0
1504 Ålesund	83	55.6	54.2	+ 1.4
1601 Trondheim	374	63.0	58.9	+ 4.1
1804 Bodø	96	53.9	61.5	- 7.6
1833 Rana	62	66.7	54.7	+12.0
1902 Tromsø	120	66.2	56.8	+ 9.4

Vi har gjort det samme for alle komuner som er trekkte i første trinn. Dersom vi antar at andre trinns trekking ikkje er skjev, kan vi sjå på forteiknet til differansen som realisasjon av ein stokastisk variabel, som er binomialfordelt med sannsyn 0.5. For alle kommuner uttrekkte i første trinn (inkl. dei med trekkesansyn 1.0) er differansen positiv i 64.6% av kommunene. Dette er 3 st.av. frå forventinga under antagelsen over. Vi kan difor konkludere med at andre trinns trekking fører til overrepresentasjon.

6. Årsaker til overrepresentasjonen

Vi skal her nevne dei moglege grunnane til OR som vi har undersøkt. Dette er inga liste over alle moglege årsaker til OR, men heller ein oversikt over dei hypotesene vi har funne det naturleg å undersøke for å finne fram til årsakene til den aktuelle OR i dette prosjektet.

6.1 Klyngeeffekter

Ei mogleg årsak til den OR vi finn i andre trinns trekking er at personane har ulik trekkesannsyn då det blir trekkte husholdningar og ikkje enkeltpersonar. Tor Haldorsen har beregna effekten av dette på FOB-data. Han fant at denne effekten tenderer til at bruttoutvalget skulle underrepresentere dei ansatte. Vi har liten grunn til å tru at det motsatte skulle vere tilfelle for registrerte arbeidstakarar.

6.2 Tilgang og avgang

Det endelige bruttoutvalget er ikkje det samme som det utvalget som blir trekkte i byrået. Ein del personar går ut av utvalget (avgang) og ein del blir lagt til (tilgang).

Tilgangen kjem av reglar intervjuarane har for feltarbeid. Dersom det bur fleire personar i ei husholdning enn den uttrekte familien, skal desse inkluderast. Også når nokon har flytta inn på ei uttrekt families adresse, skal desse inkluderast. Tilgangen var 1. kvartal 1984 vel 10% av bruttoutvalget.

Dersom ein intervjuar kjem til eit hus som er tomt, nedbrent e.l. eller dersom det er flytta inn andre på adressa. Blir familien ein vente å finne der stoken av utvalget. Vidare er opplysingane vi har om dei trekte personane er ikkje heilt oppdaterte. Difor kan det vise seg at intervju-objektet har dødd eller emigrert nyleg. Alt dette vil føre til avgangar. Også avgangen var 1. kvartal 1984 på vel 10% av bruttoutvalget

Reglane som bestemmer kven som blir avgang og tilgang er blitt endra frå 3. kvartal 1985. Etter dei nye reglane vil både tilgang og avgang bli av langt mindre omfang.

Vi har også ein type avgang som vi kan kalle falsk avgang. Falsk avgang har vi når personar blir tatt ut av utvalget p.g.a feil selektering av aldersgrupper i byrået. AKU-utvalget skal omfatte berre personar som er 16-74 år pr. 31/12 intervju-året. Vi må difor ta med alle trekte familier med minst ein person i aldersgruppa 16-74 år. Teknisk er det enklare å ta med berre familier der hovudpersonen er i denne aldersgruppa. Normalt er det ca 1.7% av utvalget som er i familier der hovudpersonen er eldre. For pulja som er trekkte 4. kvartal 1983 er det derimot ingen slike. Det kan difor tyde på at den teknisk enklare løysinga er valgt i dette kvartalet. Denne falske avgangen er likevel ikkje stor nok til å forklare ein OR på rundt 1%. Vi finn heller ingen falsk avgang i andre puljer.

6.3 Merkerutiner

Alle personar som blir trekkt til ein undersøkelse får merke i 10% registeret. Desse merka skal stå i to år. 10% registeret blir ajourført 2 gangar årleg. Under ajourføringa blir først alle merka personar overført til det nye registeret, før det blir trekkt nye personar for å oppfylle registeret til å omfatte 10%.

Det er to typer merker. Alle personar som blir trekkt til ein undersøkelse får person-merke. I tillegg blir hovedpersonen i ei familie merka i eigen posisjon dersom familien blir trekkt til eit husholdnings-utvalg. Heile familien får samstundes personmerke. Ein familie kan bli trekkt til eit husholdningsutvalg, sjølv om ein eller fleire medlemmer tidligare har deltatt i ei personundersøking.

Vi vil kome tilbake til effekten av merkerutinene i avsnitt 7.

6.4 Kopling av AKU og A/A registeret

Det omtalte skiftet i talla ved overgang til registermetoden kunne kanskje forklarast ved 1.trinns-effekt og frafall. Når vi likevel ikkje kan godta denne forklaringa som fullstendig, kjem det av at når vi koblar AKU-utvalga med A/A-registeret finn vi gjenomgåande ein OR registrerte arbeidstakarar som er større enn det vi kan forklare på denne måten. Ein mogleg forklaring på dette kunne være at det har oppstått feil under koblinga vi måtte foreta for å beregne størrelsen på OR. For personar både i AKU-utvalget og i A/A registeret har vi opplysingar frå både AKU og registeret. For å teste om koblinga var korrekt har vi jamført desse opplysingane. Vi fant ingen uoverensstemmelse i fødselsnummeret frå dei to kildene. Konklusjonen må bli at det ikkje er grunn til å tru at det er gjort feil ved denne koblinga.

6.5 Trekkinga av utvalget

Trekkinga av bruttoutvalget blir gjort ved at ein trekker personar frå eit register som skal innehalde 10% av alle personar 16-74år i kvart strata. Trekkinga foregår ved at ein trekker familiekjerner med eit visst intervall frå registeret som er sortert etter utvalgsområde og alder på hovedpersonen i familiekjernene. Alle familier som har minst eit medlem med alder 16-74år kan bli trekkt. Andelen som blir trekkt til registeret frå kvar kommune er besemt etter folketallet i 1975 når rutinene vart utarbeidde. Dette har ført til at det i dag berre

inneheld 9.9% av befolkninga. Enkelte kommuner har blitt større sidan den gang og utgjer derfor kanskje ein større andel av strataet. Trekkessannsynet er likevel det samme som i 1975. Det er truleg at vi har hatt størst vekst i kommuner som har mange arbeidsplassar mens kommuner som har ferre arbeidsplassar har gått tilbake. Dette kan føre til at 10% utvalget inneheld for mange arbeidstakarar.

Den beregninga av første trinns skjevhet som er gjort i dette notatet er gjort ved at vi har tenkt oss at det blir trekkd like mange frå kvar kommune som det faktisk vart gjort i 1.kvartal 1983, men at andelen arbeidstakarar tilsvarte den faktiske andelen for heile kommunen. Den skjevheten vi har funne skulle difor inkludere effekten av faste andre trinns trekkesannsyn. Vi fant 5.6 promille mens Haldoresen fant 3.5 promille for ein rein førstetrinnseffekt. Det kan difor vere nærliggande å anslå effekten av faste andre-trinns trekkesannsyn til ein OR på 2 promille.

6.6 Definisjonar av populasjonen

Den definisjonen av AKU-populasjonene som ligg til grunn for beregningar av andel arbeidstakarar i heile populasjonen er: Alle personar som er 16-74 år pr. 31.12 samme år, som har gyldig kommunenr. og som er registrert i personregisteret med regtype 1.

Eit spørsmål er no om AKU's bruttoutvalg blir trekkd frå den samme gruppa. Aldersrestriksjonen er ikkje noko problem, dersom vi har korrekte fødselsnummer. Sjølv om opplysingane er henta frå forskjellige register vil vi anta at fødselsnummer samsvarar. Ved etableringa av 10% registeret blir det testa på om kommunenummer er gyldig. Dette skaper altså heller ingen problem. Når det gjeld regtype derimot så har vi ikkje dette tilgjengelig i 10% registeret. Vi kan derfor ikkje utelukke personar med regtype 2-9 under trekking. I A/A registeret er det i alt ca 4800 personar med regtype 2-9 (3.kv. 1984). Av desse stod berre 37 personar i 10% registeret. Dette tyder på at heller ikkje manglande opplysingar om regtype er av særlig betydning.

Når vi opprettar 10% registeret avgrensar vi oss til ei undermengde av dei som står i skattedirektørens band. Vi tar ikkje med personar som er i familier med familienummer 0 eller blank. 3 kvartal 1984 utgjorde denne gruppa ca 0.01% av massen. Dette kan ikkje skape endringar av FAA av den størrelsesorden vi her er ute etter.

Alt i alt tyder dette på at konklusjonen må bli at den definisjonsforskjellen som er mellom trekkegrunnlaget for AKU og definisjonen av populasjonen er så liten at den vanskelig kan føre til nevneverdige skjevheter. Vi har då gjort antagelsar om samsvar i fødselsnr og komunenr mellom personregisteret og skattedirektørens register. Desse antagelsane kan berre skjekkast med å koble 10% mot personregisteret.

Kanskje også koble personregisteret mot skattedirektørens register. Det vi har funne til no gir ingen grunn til å gjennomføre slike omfattande koblingar.

7. Ledda mellom personregisteret og det endelige utvalg

Dei utvalga vi trekker blir ikkje trekkte direkte frå personregisteret. Av ulike årsaker er det fleire ledd mellom personregisteret og det endelige utvalg. For å finne ut meir om kva som skaper skjeheter kan det lønne seg å gå alle desse ledda nermare etter i saumane.

Det er skattedirektørens band som ligg til grunn for trekkinga. Personregisteret blir ikkje nytta fordi den manglar opplysingar som kan nyttast til finare geografiske inndelingar.

Frå skattedirektørens band blir det trekkte eit utvalg av familier som tilsvarar ein tidel av befolkninga i kvart utvalgsområde. Dette utvalget kallar vi for 10% registeret. Det er dette registeret utvalga blir trekkte frå.

Alle personar som blir trekkte til ei undersøking blir merka. Utvalga blir trekkte blandt dei som ikkje har merke. Feil på merkerutinene kan føre til skjevheter. Det er altså naturleg å teste den umerka delen av 10% registeret.

Frå 10% registeret blir det trekkte eit utvalg av personar i alle aldersgrupper. Før dette utvalget blir sendt til intervjuarane må ein skilje ut alle familier uten medlemmer i alderen 16-74 år.

Det trekkte utvalget blir forandra under feltarbeidet. Dette kjem av tilgangar og avgangar som tidligare omtalt. Det endelige bruttoutvalget inkluderer tilgangen og ekskluderer avgangen.

Alle desse ledda kan gi opphav til skjevheter. Det er difor naturleg å sjekke andel arbeidstakrar i alle ledda i denne lista:

1. Personregisteret
2. Skattedirektørens band
3. 10% registeret
4. Den umerka delen av 10% registeret.
5. Det trekkte utvalget.
6. Utvalget etter feltarbeid.

Vi har i denne sammenheng ikkje sjekka punkt 2. på lista. Grunnen til dette er at eventuelle OR på dette punkt skulle bli oppdaga under punkt 3.

7.1 Resultat

Vi har her følgt andelen arbeidstakarar i dei ledda som er beskrivne over. Som tidligare nevnt har vi ikkje sett på skattedirektørens band.

Tabell 7.: Andelar på ulike ledd fram til endelig utvalg.

Ledd	Kvartal:	84 k1	84 k3
Personregisteret		52.36	53.55
10% registeret		52.69	54.08
Den umerka del av 10% reg		---- ¹	53.89
Det trekke utvalg		53.80	54.38 ²
Utvalg etter feltarbeid		54.29	54.54 ²

¹ Den umerka delen vart ikkje undesøkt dette kvartalet då resultatet frå 3 kvartal 1984 tyda på at det ikkje var merkerutinene som skapte skjevheter.

² På grunn av tekniske problem er 3.gangs IO ikkje inkludert i beregningane.

Vi ser tabellen over at skilnaden mellom 10% registeret og personregisteret er av samme størrelsesorden som 1.trinns-effekten. Det er ingenting i dette som tyder på at det er noko galt med 10% registeret. Vidare ser vi at andelen arbeidstakarar i den umerka delen er lavare enn i heile 10% registeret. Dette tyder på at den merka delen har ein høg andel arbeidstakarar. Forklaringa må være at dette er personar som er trekk ut i tidligare utvalg, og dei var altså skjeve. Vi ser at det skjer eit skift på ca 0.35% under feltarbeidet, og det visate seg at dette kom av avgangen. Det største skiftet er likevel mellom 10% registeret og det trekkte utvalg. Det kunne difor være naturleg å tru at OR kjem av uheldige sider ved trekkeprogrammet.

På 10% registeret frå 1.kvartal 1984 vart trekkeprogrammet som trekker utvalga prøvd. Vi trekkte ca 170 000 personar frå 10% registeret og sammanlikna andelen arbeidstakarar blandt dei som var trekk med andelen i heile 10% registeret. Det var ingen forskjell. Andelen var 52.69% som i 10% registeret.

8. Konklusjon

Vi har funne at fast første trinns trekking fører til ein overrepresentasjon på ca 0.55%.

Vidare fører feltarbeidet til eit skift i andel arbeidstakarar på ca 0.35%. Då dette kjem av avgangar kan årsakene til dette ligge i 10%

registeret. Det er personar som står i 10% registeret men som likevel er døde eller emigrerte. Desse høyrer sjølvsagt ikkje til i AKU-populasjonen, og vil bli avgang under feltarbeidet. Dersom det er slik avgang som gir skift, er skilnaden mellom 10% registeret og personregisteret større en det ser ut til i tabell 7. Vi treng altså meir informasjon om grunnen til avgangen for å kunne slutte noko om årsakene til dette skiftet.

Dessverre gjenstår det ein overrepresentasjon på ca 0.6% som vi ikkje har funne noka forklaring på. Denne er stor nok til å være signifikant.

8.1 Anbefaling

Slik avgang og tilgang blir behandla i dag er det svært vanskelig å ha nokon kontroll med dei. Når vi no kan koble med A/A registeret gir dette oss eit svært godt høve til å få oversikt over effektane av desse. Vi vil difor foreslå at avgangen blir med på den fila som blir kobla mot A/A registeret. Vi treng då også opplysingar om årsaken til avgangen. Likeeins bør opplysingar om tilgang og tilgangsårsak være med i denne filen. I dag blir desse opplysingane redigert bort. Det er viktig at vi tar med avgang og tilgang for alle dei fire puljene som er med i undersøkinga, og ikkje berre for første gangs IO.

På denne måten vil vi lett kunne få vite arbeidstakar-andel blandt både tilgang og avgang etter ulike årsaker. Viare vil vi også kunne beregne dei effektane det vil ha for andelane i heile utvalget. Vi vil då fort oppdage eventuelle uheldige sider ved enkelte av undersøking-rutinene.

Appendiks

Definisjonar:

N_i - antall i kommune i

N_s - antall i strata s

Vi har då:

$$N_s = \sum_{i \in S} N_i$$

n_i - antall i bruttoutvalget i kommune i , 1.kv. 1983

n_s - antall i bruttoutvalget i strata s , 1.kv. 1983

p_i - andel arbeidstakarar i kommune i

p_s - andel arbeidstakarar i dei trekkte kommunene i strata s

p_i^* - andel arbeidstakarar i bruttoutvalget i kommune i , 1.kv. 1983

L - antall lommuner i heile landet.

T - indeksmengde for dei kommunene som er trekkte i første trinn.

Andel arbeidstakarar totalt blir no

$$p = \frac{\sum_{i=1}^L N_i p_i}{\sum_{i=1}^L N_i}$$

Betinga forventning gitt første trinn.

$$E(p^*) = \frac{\sum_{s=1}^{102} N_s p_s}{\sum_{s=1}^{102} N_s}$$

For å kunne beregne denne størrelsen må vi kjenne korleis landet er inndelt i strata og kva for kommuner som er trekkte ut frå kvart strata. Dette hadde vi ikkje tilgjengelig, så vi brukte difor ein tilnerming.

Dersom utvalget hadde vore sjølvveiande kunne vi ersatatte N_s med n_s :

$$E(p^*) = \frac{\sum_{s=1}^{102} n_s p_s}{\sum_{s=1}^{102} n_s} \approx \frac{\sum_{i \in T} n_i p_i}{\sum_{i \in T} n_i}$$

Den siste tilnerminga ser vi av:

$$\sum_{s=1}^{102} n_s p_s = \sum_{s=1}^{102} n_s \left(\frac{\sum_{i \in S} N_i p_i}{\sum_{i \in S} N_i} \right) \approx \sum_{s=1}^{102} n_s \left(\frac{\sum_{i \in S} n_i p_i}{\sum_{i \in S} n_i} \right) = \sum_{i \in T} n_i p_i$$

Andelen i bruttoutvalget er

$$p^* = \frac{\sum_{i \in T} n_i p_i^*}{\sum_{i \in T} n_i}$$

Vi har no splitta den totale skjevhet $p - p^*$ i to ledd:

Første trinn:

$$p - E(p^*) \approx \frac{\sum_{i=1}^L N_i p_i}{\sum_{i=1}^L N_i} - \frac{\sum_{i \in T} n_i p_i}{\sum_{i \in T} n_i}$$

Andre trinn:

$$E(p^*) - p^* \approx \frac{\sum_{i \in T} n_i (p_i - p_i^*)}{\sum_{i \in T} n_i}$$

**KONSISTENS MELLOM
ARBEIDSTAKER-REGISTERET OG
ARBEIDSKRAFTUNDERSØKELSEN**

Innhold:

1.	Innledning	17
1.1	Kilder til totale sysselsettingstall	17
1.2	Kvifor konsistensvurdering	18
1.3	Hovudresultat	18
1.4	Årsaker til manglande konsistens	18
1.5	Mål for samsvar mellom A/A registeret og AKU	19
2.	Konsistens innan fylka	21
3.	Konsistens innan aldersgruppene	24
3.1	Tid for registrering gruppene	26
4.	Konsistens for andre strata	27
5.	Konklusjon	27

1. Innledning

Vi skal i dette notatet bruke koblingar mellom Arbeidsgiver-/Arbeidstaker (A/A) registeret og Arbeidskraftundersøkelsen (AKU) til å vurdere kvaliteten av registeret. A/A-registeret er eit administrativt register. Då registeret omfattar heile befolkningen kan det være interessant å publisere statistikk frå det. Særlig interessant er det at vi kan få svært detaljert statistikk. T.d. med fine geografiske inndelingar. Registeret kan også brukast til å styrke den vanlige AKU.

I dette arbeidet har vi berre brukt opplysingar frå koblingar mellom A/A registeret og AKU. Vi har ikkje gjort forsøk på å følge opp med direkte observasjonar om meldingsgang e.l. På denne måten kan vi heller ikkje skilje mellom manglar ved registeret og definisjonsforskjellar mellom registeret og AKU. Kobling vil likevel kunne gi intrykk av størrelsen av slike feil/manglar og den vil dessuten kunne gi nyttig informasjon ved ein eventuell empirisk oppfølging. Dei kvalitetsmåla vi her brukar er dessuten dei mest interessante for å vurdere registeret sin verdi som støtte til AKU.

1.1 Kilder til totale sysselsettings tall

Vi skal her sjå på to kilder til totale sysselsettingstall. For det første A/A registeret som blir administrert av rikstrygdeverket og Arbeidskraftundersøkelsen (AKU) som er ein kvartalvis intervju-undersøkelse som blir utført ved Statistisk Sentralbyrå.

Alle lønnstakarar (untatt sjøfolk) i Norge skal innmeldast til A/A registeret dersom arbeidsforholdet er venta å vare i meir enn 6 dagar og dersom gjenomsnittlig arbeidstid per veke overstig 3 timar. Då ikkje alle arbeidsgivarar melder alle arbeidsforhod som oppfyller krava i definisjonen vil mange arbeidsforhold ikkje finnast i registeret. På grunn av sein inn og utmelding vil likevel personar stå i registeret som har slutta i den jobben vedkomande er registrert under. Det er likeleis mange som er lønsmottakarar men ikkje har kome inn i registeret.

Den klassifiseringa i AKU som ligg mest opp til det som er kriteriet for å stå i A/A registeret er lønnstakarar. Dette omfattar både lønnstakarar i inntektsgivande arbeid og lønnstakarar som er midlertidig fraverande. Vi vil her sjå nærare på samsvaret mellom A/A registeret og denne klassifiseringa. Spesielt er vi interessert i å skille ut grupper der samsvaret er speselt godt eller dårleg.

1.2 Kvifor konsistensvurdering

Bakgrunnen for denne interessen er at vi i A/A registeret har opplysingar om alle i alderen 16-74 som er registrert bosatt i Norge, dvs. vi veit om dei står i registeret eller ikkje. I ein intervju-undersøkelse som AKU har vi berre opplysingar om ein liten del av populasjonen. Sjølv om AT aleine ikkje skulle være god nok til å være grunnlag for statistikk kan desse opplysingane om dei som ikkje er med i undersøkelsen stykre resultatet i AKU svært mykje. Ein slik metode for å nytte A/A registeret til å styrke AKU vil truleg bli satt i produksjon i 1986. Beregningar i denne samheng tyder på at gevinsten ved å ta med registeropplysingar svarar til gevinsten vi ville ha fått om vi dobla utvalget i AKU.

Opplysingar om samheng mellom klassifiseringar i AKU og A/A registeret kan også avdekke spesielle grupper det er spesielt vanskelig å halde gode register over. Dette kan være av interesse for vedlikehaldet av A/A registeret. Vi ser då på klassifiseringa i AKU som "fasit" og måler feilen i A/A registeret i forhold til denne.

Vi kan også bruke desse samanstillingane av A/A registeret og AKU til å vurdere effekten av frafallet i AKU. Det kan då sjå ut til at AKU aleine vil føre til ei overestimering på rundt 1% p.g.a. frafallet.

1.3 Hovudresultat

Hovedresultatet i dette notatet er at det ikkje er mogleg å fastslå nokon forskjell i samsvar mellom A/A registeret og AKU for dei ulike fylka. Dei forskjellane vi kan observere kan forklarast med utvalgsvarians sjølv om vi antar at det ikkje eksiterer nokon forskjell. Derimot kan vi finne store forskjellar i samsvaret for dei ulike aldersgruppene. Det er her hos dei yngste og dei elste gruppen ein finn dårligast samsvar. Særlig er det mange i alderen 20-24 år som ikkje står i registeret men som likevel er lønnstakar i følge AKU i tredje kvartal. Samsvaret ser også ut til å være betre for dei som har stått lenge i registeret enn hos dei som nett har blitt registrert.

1.4 Årsaker til manglande konsistens.

Det er to typer manglande konsistens mellom AKU og registeret. For det første er det ein del som ikkje er lønnsotakarar etter AKU, men som likevel vil stå i registeret. Den andre typen manglande samsvar er personar som etter AKU definisjonene er lønsmottakarar, men som likevel ikkje står i registeret.

Det er sjølvsagt ikkje heilt uprobelmatisk å bertakte AKU som "fasit". Ein person skal være innmeldt til A/A registeret dersom han har eit arbeidsforhold som er antatt å vare i meir enn 6 dagar, og innebære mer enn gjenomsnittlig tre timars arbeidstid per veke. Sjøfolk skal ikkje registrerast. Då kravet til å bli klassifisert som sysselsatt i AKU berre er at ein har hatt inntektsgivande arbeid med over ein times varighet er det klart at det er ein del som vil bli klassifisert som lønnstakarar i AKU som likevel ikkje skal innmeldast til AT. Vidare er frafallet i AKU (dei som ikkje har svart) rundt 10%.

Tabell 1. Arbeidstid ifølge AKU for ulike AKU- A/A register klassifiseringar. Totaltall og prosent av kolonnesum. 1.kv. 1983.

Lønnstakar:	ja	ja	nei	nei				
Registrert:	ja	nei	ja	nei				
Arb.tid:								
0 timar	377	7.6	97	14.1	57	14.1	205	5.1
1-9 timar	182	3.7	121	17.6	-	-	27	0.7
10-19 timar	292	5.9	91	13.2	11	2.7	79	2.0
20-29 timar	519	10.5	61	8.9	12	3.0	95	2.3
30-39 timar	1090	22.0	79	11.5	22	5.4	90	2.2
Over 40 t.	2454	49.6	205	29.8	53	13.1	419	10.5
Uoppgitt	33	0.7	34	4.9	-	-	10	0.2
Anna	-	-	-	-	250	61.7	3077	76.9
Totalt	4947	100.0	688	100.0	405	100.0	4002	99.9

Frå tabell 1 ser vi at 31.7% av lønnstakarane som ikkje står i registeret har arbeidstid 0-9 timar, medan det berre er 11.3% av dei som står i registre som har slik arbeidstid. Truleg kjem dette av at enkelte av desse arbeidsforholda etter definisjonen ikkje skulle meldast til A/A-registeret, eller av at slike arbeidsforhold blir sett på som meir lause, og difor ikkje blir meldt til registeret sjølv om dei skulle oppfylle definisjonen. Dersom vi antar at heile forsjellen kjem av ulike definisjonar, kan vi beregne at det resterande antall lønnstakarar som ikkje står i registeret er 518. Dette kan tyde på at denne gruppa er den viktigaste type feilklassifisering. Gruppa omfattar både arbeidsforhold som aldri blir meldt til registeret, og forhold som ikkje står i registeret p.g.a. sein meldingsgang. Det er difor rimelig at den er størst, då den andre (registrerte som ikkje er lønnstakarar) berre kjem av sein meldingsgang.

1.5 Mål for samsvar mellom A/A registeret og AKU

Det finst fleire mål for konsistens mellom to kilder. Ofte prøver ein å finne fram til ein enkelt parameter for konsistensen. Vi har likevel valgt å bruke to parameterar. Dette fordi desse to parametrane gir

uttrykk for ulike årsaker til manglande konsistens.

Dei to parametrane vi har valgt, tar utgangspunkt i dei to typene manglande konsistens som vi har omtalt over.

$$P_{11} = \Pr(\text{registrert} | \text{lønnstaker})$$

$$P_{01} = \Pr(\text{registrert} | \text{ikkje lønnstaker})$$

Pr står her for sannsynet. T.d. $\Pr(\text{registrert} | \text{lønnstaker})$ er sannsynet for å være registrert gitt at ein er lønnstaker.

Resultatet av undersøkelsen kan sammenfattast i følgende tabell.

	registrert	ikkje registrert	
lønnstakar	n11	n10	n1.
ikkje lønnst.	n01	n00	n0.
	n.1	n.0	n

Her er n antall personer i nettoutvalget. Punktum betyr at vi har summert over alle verdisar av vedkomande variabel. F.eks. er n.1 antall registrerte personar i utvalget, uansett klassifisering i AKU.

Vi vil no bruke som estimat for P_{11} og P_{01} :

$$P_{11}^* = n11/n1.$$

$$P_{01}^* = n01/n0.$$

I tillegg til å vurdere konsistensen for heile bestanden, skal vi vurdere konsistensen innan undergrupper av bestanden mot konsistensen på landsnivå.

For å hevde at ei undergruppe har konsistens som skil seg frå resten av landet er det naturleg å kreve at P_{11}^* og P_{01}^* begge er gode, eller begge er dårlige.

P_{11} er andelen av faktiske lønnstakarar som står i registera, og denne bør difor vere så høg som mogleg. Tilsvarande er P_{01} andelen som står i registeret uten å være lønsmottakarar, og denne bør vere lav.

Også summen av dei kan være interessant. Dersom ingen blir meldt inn til registeret vil $P_{11} + P_{01} = 0$. Tilsvarande om alle blir meldt til

registeret uansett yrkesstatus vil $P_{11} + P_{01} = 2$. Dersom summen altså er spesielt lav for ein delpopulasjon kan dette tyde på at ein mindre andel av lønnstakarane blir innmeldt i denne populasjonen.

Nivået på desse størrelsane seier ikkje så mykje om kvaliteten på registeret, då vi ikkje kan isolere effekten av definisjonsforskjellar. Derimot vil variasjonen i størrelsane gi informasjon om variasjon i registerkvalitet for ulike grupper. Dersom P_{11} er lav og samtidig P_{01} høg så tyder dette på at dårleg meldingsgang gir seg større utsalg i denne gruppa enn blandt gjennomsnittet. Dersom $P_{11} + P_{01}$ er lav tyder det på at ein lav andel av arbeidstakarane i denne gruppa er innmeldt.

2. Konsistens innan fylka.

Det er naturleg å undersøke om det finst geografiske variasjonar i konsistens mellom AKU og AT. Med den utvalgsstørrelsen vi har er det ikkje mulig å gå stort lenger enn til å undersøke om det er variasjon mellom fylker. Sjølv for enkelte fylker er utvalgsstørrelsen så liten at talla blir av liten verdi. I Aust-Agder td. er nettoutvalget rundt 35 personar.

Dei aktuelle andelane er estimert for fire kvartal i 1983. Resultatet er oppgit nedanfor.

Tabell 2 : Estimater for P₁₁ . 1983.

Kvartal:	Menn				Kvinner			
	1.	2.	3.	4.	1.	2.	3.	4.
Fylke								
Østfold	88.7	92.6	89.1	89.6	88.9	88.6	87.9	88.7
Akershus	94.0	90.0	90.1	91.2	88.8	86.2	87.2	89.1
Oslo	89.7	89.0	90.8	92.6	87.3	85.0	82.8	86.6
Hedemark	92.8	89.3	90.6	92.1	87.4	85.3	88.2	88.2
Oppland	92.4	92.2	90.9	93.2	83.7	87.5	90.6	86.8
Buskerud	97.5	92.9	91.2	91.0	84.7	84.8	85.2	87.1
Vestfold	93.3	88.4	91.4	90.2	81.0	86.4	85.4	80.9
Telemark	90.1	88.9	87.4	86.2	83.3	84.6	91.2	86.7
Aust-Agd	85.7	77.8	77.8	87.0	82.4	94.1	83.3	88.2
Vest-agd	79.8	81.7	85.8	88.7	65.3	79.2	82.3	86.9
Rogaland	91.4	92.7	93.7	92.4	85.1	83.4	89.9	88.3
Hordala.	87.4	86.9	84.2	89.4	84.4	81.5	82.5	81.5
Sogn Fj.	93.2	88.9	91.7	83.9	84.6	86.6	91.7	86.0
Møre Ro.	81.4	82.0	79.9	76.9	86.0	85.9	81.1	87.3
S-Trønd.	89.8	90.2	86.9	90.3	89.0	89.9	92.6	88.4
N-Trønd.	90.5	88.2	92.8	94.4	88.7	90.7	89.8	79.0
Nordland	86.7	81.7	84.0	84.7	81.8	84.9	81.6	85.8
Troms	89.3	86.8	87.1	86.6	86.8	79.9	86.5	78.3
Finnmark	77.6	84.2	82.1	93.1	88.6	85.4	86.7	94.2
Totalt	89.8	88.5	88.4	89.5	85.3	85.3	86.3	86.4

Tabell 3 : Estimater for P_{01} . 1983.

Kvartal:	Menn				Kvinner			
	1.	2.	3.	4.	1.	2.	3.	4.
Fylke								
Østfold	9.6	12.4	12.1	8.6	8.7	5.3	7.9	4.8
Akershus	14.5	12.8	17.3	12.1	6.4	8.4	12.4	5.8
Oslo	12.9	11.8	12.5	14.9	5.0	9.1	9.9	6.2
Hedemark	7.4	16.0	13.9	11.2	12.1	12.0	6.6	9.0
Oppland	14.0	8.8	9.1	9.3	7.1	9.7	14.5	10.8
Buskerud	14.5	21.7	14.3	5.5	8.0	10.6	13.9	7.7
Vestfold	13.8	10.3	10.3	7.6	2.2	6.1	6.8	2.8
Telemark	9.8	19.0	13.3	9.6	8.9	10.1	12.6	7.7
Aust-Agd	14.3	5.6	27.8	5.6	-	13.0	5.3	4.3
Vest-agd	10.7	12.5	13.6	5.9	7.0	7.3	9.6	3.2
Rogaland	14.2	9.3	14.4	11.0	3.6	6.6	6.6	8.9
Hordala.	12.2	11.7	17.7	12.4	9.4	7.7	10.1	9.6
Sogn Fj.	8.0	14.1	15.7	9.7	15.6	10.0	11.1	9.7
Møre Ro.	8.4	15.8	16.5	12.4	5.3	7.1	7.4	5.3
S-Trønd.	8.5	19.1	18.3	8.3	8.4	12.5	17.5	10.6
N-Trønd.	11.1	11.1	15.1	8.8	14.3	11.4	17.2	9.2
Nordland	11.3	7.6	17.1	8.7	3.7	5.8	6.9	8.7
Troms	15.9	17.2	20.6	15.0	9.5	10.9	12.9	11.3
Finnmark	6.7	12.5	18.2	9.5	18.8	15.7	16.7	18.9
Totalt	11.6	13.0	15.1	10.3	7.4	8.7	10.6	7.8

Tabell 4 : Estimat for $P_{01} + P_{11}$. 1983.

Kvartal:	Menn				Kvinner			
	1.	2.	3.	4.	1.	2.	3.	4.
Fylke								
Østfold	98.3	105.0	101.2	98.2	97.6	93.9	95.8	93.5
Akershus	108.5	102.8	107.4	103.3	95.2	94.6	99.6	94.9
Oslo	102.6	100.8	103.3	107.5	92.3	94.1	92.7	92.8
Hedemark	100.2	105.3	104.5	103.3	99.5	97.0	94.8	97.2
Oppland	106.4	101.0	100.0	102.5	90.8	97.2	105.4	97.6
Buskerud	112.0	114.6	105.5	96.5	92.7	95.4	99.1	94.8
Vestfold	107.1	98.7	101.4	97.8	83.2	92.5	92.2	83.7
Telemark	99.9	107.9	100.7	97.8	92.2	94.7	103.8	94.4
Aust-Agd	100.0	83.4	105.6	92.6	-	107.1	88.6	92.5
Vest-agd	90.5	94.2	99.4	94.6	72.3	86.5	91.9	90.1
Rogaland	105.6	100.0	108.1	103.4	88.7	90.0	96.5	97.1
Hordala.	99.6	98.6	101.9	101.3	93.8	88.2	92.6	91.1
Sogn Fj.	101.2	103.0	106.8	93.6	100.2	96.6	102.8	95.7
Møre Ro.	88.5	97.8	96.4	89.3	91.3	93.0	88.5	92.6
S-Trønd.	98.3	109.3	105.2	98.6	97.4	102.4	110.1	99.0
N-Trønd.	101.6	99.3	107.9	103.2	103.0	102.3	107.0	88.2
Nordland	97.0	89.3	101.1	93.4	85.5	90.7	88.5	94.5
Troms	105.6	104.0	107.7	101.6	96.3	90.8	99.4	99.6
Finnmark	84.3	96.7	100.3	102.6	107.4	101.1	103.4	103.1
Totalt	101.4	101.5	103.5	99.8	92.7	94.0	96.9	94.2

Hovudkonklusjonen her må bli at det er liten forskjell i samsvaret mellom AT og AKU for dei ulike fylka. Sjølv om enkelte fylke skil seg litt ut er forskjellane så små at dei godt kan tilskrivast utvalgsvariansen i AKU.

Derimot er det klart ein forskjell mellom kjønna. Denne forskjellen er mest tydelig i tabell 4. Dette kan tyde på at det eir fleire arbeidsforhold som ikkje blir meldt til registeret blandt kvinner enn blandt menn. Ei mogleg forklaring på dette kan vel være at kvinner oftare har arbeidsforhold med kort arbeidstid pr. veke, enkelte av desse skal ikkje meldast til registeret. Viare kan det tenkast at desse forholda lavare status og blir tatt midre seriøst av arbeidsgivaren.

3. Konsistens innan aldersgruppene.

Det er grunn til å vente at dei ulike aldersgruppene ikkje er like stabile i sine arbeidsforhold. Vi kunne vente oss tre hovudklasser. Ungdom som f.eks. startar på utdanning etter eit kortare arbeidsforhold. Eldre som sluttar når dei går av med pensjon. Og ei spesielt stabil midtgruppe i alderen rundt 30-60 år. Då det truleger

vanskeligare å halde gode register over ustabile grupper kan vi også vente store forskjellar i samsvaret for dei ulike aldersgruppene. Dei empiriske resultatene støtter desse antagelsane.

Vi har brukt den standard for aldersinndeling som er brukt i AKU.

Tabell 5 : Estimat for P_{11} . 1983.

Kvartal:	Menn				Kvinner			
	1.	2.	3.	4.	1.	2.	3.	4.
Alder								
16-19 år	69.6	59.7	58.6	67.3	53.3	47.3	52.6	60.4
20-24 år	82.8	83.0	81.3	83.1	78.4	82.5	82.1	83.3
25-29 år	88.6	87.2	87.3	88.9	85.8	85.1	85.2	86.1
30-39 år	92.6	91.2	90.9	91.9	88.6	88.5	89.1	89.1
40-59 år	92.5	93.8	93.5	95.1	90.6	91.7	92.0	91.9
60-64 år	91.4	93.6	95.9	94.2	91.6	90.1	92.1	91.0
65-66 år	90.7	89.8	94.9	98.0	87.5	78.0	88.6	75.0
67-69 år	87.8	81.8	83.9	61.3	88.0	90.5	88.2	93.8
70-74 år	75.0	76.5	66.7	47.1	25.0	40.0	40.0	57.1
Totalt	89.8	88.5	88.4	89.5	85.3	85.3	86.3	86.4

Tabell 6 : Estimat for P_{01} . 1983

Kvartal:	Menn				Kvinner			
	1.	2.	3.	4.	1.	2.	3.	4.
Alder								
16-19 år	9.8	8.8	11.1	8.9	8.1	10.3	11.0	11.6
20-24 år	19.8	24.2	32.0	21.6	11.3	13.3	31.0	17.8
25-29 år	22.0	19.6	30.6	21.9	8.2	14.9	14.3	9.5
30-39 år	12.7	14.6	18.1	10.0	9.0	9.0	10.1	8.5
40-59 år	13.8	13.0	13.7	9.6	9.1	10.1	10.9	6.9
60-64 år	10.3	15.4	15.3	11.1	6.2	9.4	8.3	4.9
65-66 år	10.8	14.0	12.6	6.6	2.7	7.0	6.1	6.0
67-69 år	3.9	9.1	6.6	5.4	4.1	5.2	9.4	5.7
70-74 år	4.8	3.4	4.7	4.7	3.5	1.1	1.3	0.7
Totalt	11.6	13.0	15.1	10.3	7.4	8.7	10.6	7.8

Tabell 7 : Estimert for $P_{01} + P_{11}$. 1983

Kvartal:	Menn				Kvinner			
	1.	2.	3.	4.	1.	2.	3.	4.
Alder								
16-19 år	79.4	68.7	69.7	76.2	61.4	57.6	63.6	72.0
20-24 år	102.6	105.2	113.3	104.7	89.7	95.8	113.1	101.1
25-29 år	110.6	106.8	117.9	110.8	94.0	100.0	99.5	95.6
30-39 år	105.3	105.8	109.0	101.9	96.6	97.5	99.2	97.6
40-59 år	106.3	106.8	107.2	104.7	99.7	101.8	102.9	98.8
60-64 år	101.7	110.0	111.2	105.3	97.8	99.5	100.4	95.9
65-66 år	101.5	103.8	107.5	104.6	90.2	85.0	94.7	81.0
67-69 år	91.7	90.9	90.5	66.7	92.1	95.7	97.6	99.5
70-74 år	79.8	79.9	71.4	51.8	28.5	41.1	41.3	57.8
Totalt	101.4	101.5	103.5	99.8	92.7	94.0	96.9	94.2

Vi ser av tabell 4 og 5 at det er klare forskjellar i samsvaret for dei ulike aldersklassene. Det kan sjå ut til at ein del av forskjellane kjem av forskjellar i det totale sysselsettingsnivået for dei ulike aldersklassene. T.d. er størstedelen av menn i aldersklasse 4 lønntakarar, samtidig som denne gruppa har størst verdiar for P_{11} og også store verdiar for P_{01} .

Ein del av forskjellane kjem truleg likevel av ulik mobilitet. T.d. er det mange i aldersgruppa 20-24 år som ikkje er lønsmottakarar men som likevel står i registeret i kvartal 3. Ei forklaring på dette kan være at det er eit etterheng av personar som har hatt sommarjobb men som ikkje er utmeldt av registeret. AKU vart utført 22-28/8 for dette kvartalet. Dei fleste i sommarjobb har då truleg slutta i jobben.

Som vi ser av summen i tabell 7 blir arbeidsforhold blandt yngre og eldre sjeldnare meldt til registeret enn arbeidsforhold i andre grupper.

3.1 Tid for registrering

Ein kan ha grunn til å tru at personar som har vore ansatt i samme jobben over eit lengre tidsrom er meir stablile enn andre. Vi har difor forsøkt å dele registrerte personar i grupper etter når dei har blitt ansatt i den jobben dei er registrert under, og prøve å finne om registeropplysingane er betre for enkelte av gruppene.

Vi kan ikkje bruke det same målet for samsvar i dette tilfellet, då vi er interresert i forskjellig oppførsel berre blandt dei som står i registeret. Vi avgrensar oss altså til dei som står i registeret, og ser om andelen av faktiske lønntakarar er større i enkelte grupper enn i andre.

Resultatet i tabell 6 støttar vår hypotese om at dei med lengst ansettelse ar mest stabile. Ein finn at andelen faktiske lønnstakarar er klart større for dei som har vore ansatt sidan 1980 eller før, i forhold til dei ansatt i 1982-83. Dei som er ansatte i 1981 utgjer ei mellomgruppe som er skildt ut for seg sjølv.

Tabell 8: Andelen faktiske lønnstakarar for tre grupper registrerte.

Kvartal	Menn				Kvinner			
	1.	2.	3.	4.	1.	2.	3.	4.
Ansatt 1980 eller før	94.7	93.4	93.4	95.9	93.2	93.0	92.4	94.9
Ansatt 1981	87.9	90.9	88.4	92.5	91.5	92.5	90.4	92.5
Ans. 1982-83	89.7	90.1	86.1	90.3	90.5	88.1	84.9	89.1

4. Kvalitet for andre strata

Vi er også interessert i samsvaret for andre strata for eventuelt å kunne skille ut grupper der samsvaret er spesielt godt eller spesielt dårlig. Dette for å kunne predikere antall lønnsmotakarar betre. Dei opplysingane vi no har frå AT ser ikkje spesielt lovande ut for ei slik vidare inndeling. Derimot kan det vere ein del å vinne på å innhente opplysingar frå andre kjelder. Ei interessant opplysing i denne sammenheng er inntekt.

Dei opplysingane i AT som kunne vore interessante er opplysingar om forventa sluttdato, forventa varighet av arbeidsforholdet, og forventa arbeidstid. Desse opplysingane ser ut til å vere brukt svært lite, og dei som fanst ga liten informasjon.

5. Konklusjonar.

Vi har her sett på korleis konsistensen varierer med fylke kjønn og aldersgruppe. Dei tabellane vi har funne tyder på at alder og kjønn er dei viktigaste kjennemerka. Det ser ut til å vere små forskjellar i konsistens frå fylke til fylke.

Vi har og sett at arbeidsforhold bland yngre og eldre blir sjeldnare meldt til registeret enn kva som gjeld for andre grupper. Yngre har

også ofte kortvarige jobbar i sommarmånadane, etter sommaren står det då mange yngre i registeret som etter AKU ikkje er lønnstakarar.

Det ser også ut til at konsistensen er større når vi avgrensar oss til dei personane som har stått lenge i registeret.

Vi har ikkje funne andre kjennemerke som kan skille ut grupper som har særleg konsitens, men det er mogleg at det kan gjerast dersom ein kan få fleire opplysingar om personane frå andre register. Ein parameter som er aktuell her er inntekt.

UFULLSTENDIGE FAMILIER I 10%REGISTERET.

1. Innledning

Frå og med tredje kvartal 1985 vart utvalga til arbeidskraftundersøklesen endra frå husholdningsutvalg til familieutvalg. Ved denne endringa vart enkelte problem med 10%-registeret oppdaga.

Hovedregelen i det nye opplegget er at intervjuarane berre skal intervju dei personane dei har fått oppgitt frå byrået. Tidligere skulle alle som budde i samme husholdning også intervjuast. Intervjuarane kunne no melde at dei i enkelte tilfelle hadde fått oppgitt ektemenn, men ikkje resten av familien. Dette til tross for at det skulle være eit familieutvalg og at vedkomande hadde vore gift i mange år. Ein fant så at vedkomande i mange tilfeller hadde familie i personregisteret, men ikkje i 10%registeret. Då 10%-registeret er eit familieregister, tyder dette på at det har oppstått feil i registeret.

I dette notatet skal vi beskrive årsaka til denne feilen. Den har blitt retta i seinare 10%-register, og skulle være borte frå første kvartal 1986.

2. Årsak til feilen

Feilen hadde oppstått under ajourføringa av 10%registeret. Ajourføringa skal skje to gangar årlig. Under køyringa her vart eit program erstatta med ein forelda versjon av samme programmet. Difor vart tidligare merka personar handsama på ein uheldig måte.

2.1 Merkerutinene

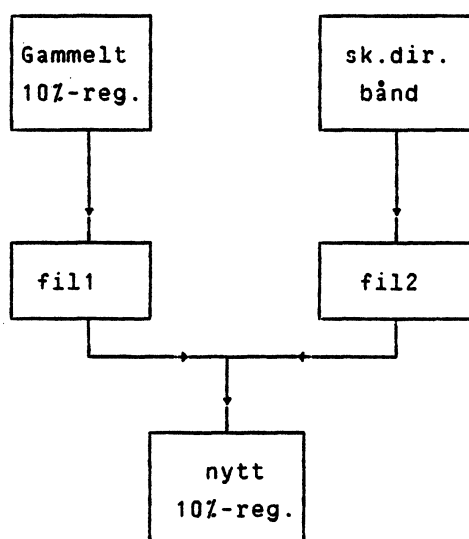
For å unngå at samme IO blir trekk ut til fleire ulike undersøkingar med korte mellomrom blir alle trekkte personar merka. Vi har to typer merker. Alle personar som blir trekk til ein undersøking får eit mekre som vi kan kalle personmerke. Når det blir trekk husholdning/familie utvalg får hovudpersonen i familien i tillegg eit husholdningsmekre i ein eigen posisjon. Regelen er no at dersom hovedpersonen ikkje tidlegare er trekk til eit husholdnings/familie utvalg kan familien bli trekk til eit slikt utvalg sjølv om ein eller fleire personar i familien har personmerke.

Alle personar med merke skal overførast til det nye 10%-registeret når dette blir ajourført. Då enkelte er trekkte til personutvalg vil det ofte vere berre ein person i ein familie som har merke. Det blir då berre ein person i denne familien som blir overført til det nye 10%-registeret. For at denne enkeltpersonen ikkje skal kunne bli trekkte til eit husholdnings/familie utvalg, skal vedkomande ha husholdningsmerke. Dei som ikkje har dette frå før blir gitt merket 00.

2.2 Ajourføring av 10%-Registeret

Vi skal her sjå på hovudpunkta under ajourføringa av 10%-registeret.

Under ajourføringa blir først alle som har merke i det gamle 10%-registeret tatt vare på i ei eiga fil som vi her kallar fil1. Vi beregnar så kor mange til vi må trekke i kvart utvalgsområde for at 10%-registeret skal innehalde 10% av befolninga i kvart utvalgsområde. Det beregna antall personar blir så trekkte frå skattedirektørens bånd. Dei trekkte personane blir lagt på ei eiga fil som vi kallar fil2. Tilslutt blir desse to filene sortert sammen til det nye 10%-registeret. Skjematisk kan vi framstille dette slik.



Under sammenkoblige av fil1 og fil2 til nytt 10%-register må ein sørge for at personar frå fil1 som ikkje har med familien får husholdningsmerke. Dette blir gjort ved at alle hovudpersonar i fil1 som ikkje har husholdningsmerke får husholdningsmerke 00 dersom dei ikkje også står i fil2. Grunnen til dette unntaket er at fil2 berre inneheld heile familier frå skattedirektørens band. Dersom ein hovudperon i fil1 også står i fil2 vil han dermed få med heile familien frå fil2. Han skal i dette tilfellet ha samme merke som han hadde før ajourføringa.

Det var rutinen som skulle utføre dette som vart brukt i forelda utgave ved ajourføringa 1. og 2. halvår 1985. Dette førte til at ingen fekk merke 00. Personar som berre hadde merke frå personaundersøkingar og der resten av familien ikkje har blitt trekk inn att kunne difor likevel bli trekk til ein husholdningsundersøkelse.