

Arbeidsnotater

S T A T I S T I S K S E N T R A L B Y R Å

Dronningensgt. 16, Oslo-Dep., Oslo 1. Tlf. 41 38 20, 41 36 60

IO 72/19

22. desember 1972

METODEHFTE NR. 1

Notater om skogbrukstelling, MODIS IV, og aldersbegrepet i AKU

Innhold

	Side
Metodehefter i serien Arbeidsnotater	2
Ingolf Myklebust: "Forslag til utvalgstelling for skog- bruket" (IM, 25/8-71)	3
Jon D. Engebretsen: "Definisjoner og klassifikasjoner i MODIS IV" (JE/IVG, 5/4-72)	9
Jan M. Hoem og Olav Ljones: "Problemer knyttet til spesi- fikasjon av alderen for enkelte aldersgrupper (i arbeidskraftundersøkelsene)" (JMH/OLj/GH, 2/12-71)	16
Jan M. Hoem: "Aldersbegrepet i arbeidskraftundersøkelsene" (JMH/EH, 28/9-72)	26

Ikke for offentliggjøring. Dette notat er et arbeidsdokument og kan siteres eller refereres bare etter spesiell tillatelse i hvert enkelt tilfelle. Synspunkter og konklusjoner kan ikke uten videre tas som uttrykk for Statistisk Sentralbyrås oppfatning.

FORORD

Metodehefter i serien Arbeidsnotater

I tilknytning til mange prosjekter i Statistisk Sentralbyrå utarbeides det mindre, upretensiøse notater for avklaring av spørsmål av metodisk interesse. Det kan dreie seg om utvalgsteknikk, alternative spørsmålsformuleringer, presentasjonsmetoder, begrepsavklaringer, diskusjon av "funn" i data, system-idéer eller andre temaer. Selv om mange slike notater bare har begrenset interesse i ettertid, vil det blant dem være noen som kunne fortjene å bli mer alminnelig tilgjengelig enn de har vært hittil. Det kan også være nyttig å ha dem registrert sentralt slik at det blir lettere å få oversikt over det stoffet som foreligger, og å referere tilbake til det.

Byrået innfører nå en publiseringsordning for stoff av dette slaget. Etter forbilde av serien Technical Notes fra U.S. Bureau of the Census vil en leilighetsvis publisere et passende antall slike notater samlet i metodehefter i serien Arbeidsnotater. Inneværende hefte er det første av denne typen.

Forsker Jan M. Hoem er oppnevnt som redaktør av metodeheftene. Medarbeidere i Byrået som lager stoff som kan være aktuelt, bes sende dette til redaksjonen etter hvert som det blir ferdig.

Kontorlederne bes holde øynene åpne for denne nye publiseringsmuligheten.

Konsulent Bjørn L. Tønnesen er redaksjonssekretær.

IM 25/8-71

Forslag til utvalgstelling for skogbruket.

av Ingolf Myklebust.

Innhold

	Side
Innhenting av oppgaver	3
Trekking av utvalg	3
Distriktsinndeling	4
Størrelsesklasser og utvalgsprosenter	5
Oppgaveskjema og regler for utfyllingen	5
Bearbeiding av oppgavene	6
Den videre behandling av saken	6
Bilag I	7

Innhenting av oppgaver

Oppgavene innhentes direkte fra oppgavegiverne gjennom posten.

Trekking av utvalg

Utvalget trekkes fra eiendommer med oppgaver til Skogbrukstelingen 1967 etter at eiendommer som tilhører samme eier er slått sammen (virkelig eiendomsfordeling).

Utvalget trekkes bare fra eiergruppene 12 (Enkeltperson, innenbygdsboende med jordbruk), 13 (Enkeltperson, innenbygdsboende uten jordbruk), 14 (Enkeltperson, utenbygdsboende med jordbruk), 15 (Enkeltperson, utenbygdsboende uten jordbruk), 2 (Interessentskap), 3 (Aksjeselskap), 4 (Sameie), 5 (Husbrukskog, sams kog), 6 (Institusjon, stiftelser o.l.), 7 (Staten, ikke administrert av Direktoratet for statens skoger), 8 (Opplysningsvesenets Fond, ikke administrert av Direktoratet for statens skoger), 10 (Bygdeallmenning) og 11 (Kommune). Eiendommer i eiergruppene 7 (Staten, administrert av Direktoratet for statens skoger), 8 (Opplysningsvesenets Fond, administrert av Direktoratet

for statens skoger) og 9 (Statsallmenning) regner en med å kunne få de nødvendige opplysninger om i regnskapet fra Direktoratet for statens skoger.

Tellingen skal bare omfatte eiendommer med minst 100 dekar produktivt skogareal.

Eiendommene stratifiseres på størrelsesklasser og distrikter og utvalget trekkes særskilt for hvert stratum. Eiendommer med skog i flere kommuner plasseres før trekkingen i det distrikt de tilhører.

Jeg går ut fra at systemkontoret har standardprogrammer som kan brukes ved trekkingen av utvalget.

Distriktsinndeling

Både ved trekkingen av utvalget og ved bearbeidingen av oppgavene deles landet inn i 6 distrikter (distrikt I - VI). En vil da kunne nøye seg med et noe mindre utvalg enn om en skal gi tall for de enkelte fylker.

Ved inndelingen har en tatt sikte på å få distrikter med mest mulig ensartede driftsforhold.

Distrikt I. Østlandet søndre del (flatbygdene).

Østfold, alle kommuner.

Akershus og Oslo, alle kommuner.

Hedmark, kommunene Hamar, Kongsvinger, Ringsaker, Vang, Løten, Stange, Nord-Odal, Sør-Odal, Eidskog, Grue, Åsnes, Våler og Elverum.

Oppland, kommunene Lillehammer, Gjøvik, Østre Toten, Vestre Toten, Jevnaker, Lunner, Gran og Søndre Land.

Buskerud, kommunene Ringerike, Drammen, Kongsberg, Modum, Øvre Eiker, Nedre Eiker, Lier, Røyken og Hurum.

Vestfold, alle kommuner.

Distrikt II. Østlandet nordre del (dal- og fjellbygdene).

Hedmark, kommunene Trysil, Åmot, Stor-Elvdal, Rendalen, Engerdal, Tolga-Os, Tynset, Alvdal og Follidal.

Oppland, kommunene Dovre, Lesja, Sjøk, Lom, Vågå, Sel, Fron, Ringebu, Øyer, Gausdal, Nordre Land, Sør-Aurdal, Etnedal, Nord-Aurdal, Vestre Slidre, Øystre Slidre og Vang.

Buskerud, kommunene Flå, Nes, Gol, Hemsedal, Ål, Hol, Sigdal, Krødsherad, Flesberg, Rollag og Nore og Uvdal.

Distrikt III. Telemark-Agder.

Telemark

Aust-Agder alle kommuner.

Vest-Agder

Distrikt IV. Vestlandet.

Rogaland

Hordaland og Bergen

alle kommuner.

Sogn og Fjordane

Møre og Romsdal

Distrikt V. Trøndelag - Helgeland.

Sør-Trøndelag

alle kommuner

Nord-Trøndelag

Nordland, kommunene Bindal, Brønnøy, Vefsn, Grane, Hattfjelldal, Hemnes og Rana.

Distrikt VI. Nord-Norge, nordre del.

Nordland, kommunene Bodø, (Bodin), Narvik, Vega, Vevelstad, Herøy, Aistahaug, Leirfjord, Dønna, Nesna, Lurøy, Rødøy, Meløy, Gildeskål, Beiarn, Saltdal, Fauske, Skjerstad, Sørfold, Steigen, Hamarøy, Tysfjord, Lødingen, Tjeldsund, Evenes, Ballangen, Ankenes, Moskenes, Vestvågøy, Vågan, Hadsel, Bø, Øksnes, Sortland og Andøy.

Troms, alle kommuner.

Finnmark, statens og IO's skoger. Oppgavene innhentes direkte fra skogforvalterne og fylket holdes utenfor utvalgstillingen.

Størrelsesklasser og utvalgsprosjenter

Ved trekkingen av utvalget og ved bearbeidingen av oppgavene inndeles eiendommene i 4 klasser etter størrelsen av produktivt skogareal. Det brukes en egen utvalgsprosent for hver klasse.

Bilag I viser eiendommer i alt og eiendommer med i utvalget fordelt på størrelsesklasse og distrikter etter to forskjellige klasseinndelinger. Hvis alternativ 2 gir et tilstrekkelig stort utvalg, brukes dette.

Tallet på eiendommer i bilag I er tatt fra den kommunevise opptellingen. Ved virkelig eiendomsfordeling er tallet noe mindre og fordelingen på størrelsesklasser er noe forskjellig fra den kommunevise opptellingen. Forskjellen er imidlertid liten.

Oppgaveskjema og regler for utfyllingen

Postene på oppgaveskjemaet (bilag II) er stort sett i samsvar med forslaget fra Utvalg til å vurdere kostnadsstatistikken i skogbruket. Avsnitt 1

og 2 på side 2 avviker noe fra utvalgets forslag. Endringen er foretatt for å få opplysninger om utgifter pr. m³ når hele arbeidet er utført av leid arbeidskraft og med arbeidstakers utstyr. Dette er meget viktige opplysninger. Et spørsmål om hvor mye av utgiftene som er lønn til arbeidere og funksjonærer, er sløffet da oppgavene av forskjellige årsaker vil ha liten verdi.

Rettledning for utfyllingen av skjemaet følger som bilag III og brev til oppgavegiverne som bilag IV.

Bearbeiding av oppgavene

Tabellforslag vil bli lagt fram senere.

Den videre behandling av saken

Forslaget til oppgaveskjema, rettledning for utfyllingen, brev til oppgavegiverne, distriktsinndeling og inndeling i størrelsesklasser, bør forelegges Skogdirektoratet, direktør Arne Solbraa i Norges Skogeierforbund og medlemmene av Utvalg til å vurdere kostnadsstatistikken i skogbruket (direktør Per Krog og forstkandidat Arnt Astrup, Skogbruksforeningen av 1950, sekretær Milian Ødegaard, Skogbrukets Arbeidsgiverforening og stipendiat Torstein Oppheim, Norges Landbrukshøgskole) til uttalelse før en tar endelig bestemmelse om dem.

Før en sender tellingsmaterialet ut til uttalelse må en få klarlagt hvilket av de to alternativer til inndeling i størrelsesklasser en skal velge.

Da tellingsmaterialet bør sendes oppgavegiverne omkring årsskiftet 1971-1972, må den videre behandling av saken skje forholdsvis raskt.

Utvalgstelling for skogbruket

Eiendommer etter størrelsen av produktivt skogareal. Statens og OF's skoger og statsallmenninger er fratrukket.

Størrelsesklasser etter alternativ 1. Kommunevis opptelling.

Distrikt		I alt	100 - 499 da 2% utvalg	500 - 1 999 da 10% utvalg	2 000 - 9 999 da 20% utvalg	10 000 dekar og mer 50% utvalg
I	I alt	5% 18 439	12 945	4 334	966	194
	Utvalg ca.	980	260	430	190	100
II	I alt	6% 10 786	6 896	3 011	778	101
	Utvalg ca.	660	140	300	170	50
III	I alt	6% 12 716	7 830	4 145	709	32
	Utvalg ca.	725	160	410	140	15
IV	I alt	4% 16 159	12 768	3 196	190	5
	Utvalg ca.	713	350	320	40	3
V	I alt	5% 12 302	8 020	3 674	559	49
	Utvalg ca.	665	160	370	110	25
VI	I alt	4% 11 260	8 786	2 196	271	7
	Utvalg ca.	444	170	220	50	4
Hele landet.						
	I alt	5% 81 662	57 245	20 556	3 473	388
	Utvalg ca.	4 187	1 240	2 050	700	197

Størrelsesklasser etter alternativ 2. Kommunevis opptelling.

Distrikt		I alt	100 - 999 da 2% utvalg	1 000 - 4 999 da 10% utvalg	5 000 - 19 999 da 20% utvalg	20 000 dekar og mer 50% utvalg
I	I alt	3,5% ^{18 439}	15 865	2 153	327	94
	Utvalg ca.	650	320	220	65	45
II	I alt	4% ^{10 786}	8 831	1 654	263	38
	Utvalg ca.	420	180	170	50	20
III	I alt	3,5% ^{12 716}	10 501	2 091	116	8
	Utvalg ca.	444	210	210	20	4
IV	I alt	2,5% ^{16 159}	15 157	985	16	1
	Utvalg ca.	404	300	100	3	1
V	I alt	3,5% ^{12 302}	10 488	1 700	80	34
	Utvalg ca.	410	210	170	15	15
VI	I alt	3% ^{11 260}	10 383	846	30	1
	Utvalg ca.	311	210	85	15	1
Hele landet.						
	I alt	3% ^{81 662}	71 225	9 429	832	176
	Utvalg ca.	2 539	1 330	955	168	86

JE/IVG, 5/4-72.

DEFINISJONER OG KLASSIFIKASJONER AV SKATTER I MODIS IV.

Av

Jon D. Engebretsen

Innhold

	Side
1. Innledning	9
2. Definisjon av skattebegreper	9
3. Sektorinndeling av den offentlige forvaltning	11
4. Nærmere om sektor for skatteinnkreving	12
5. Klassifikasjonsprinsipper	13

1. Innledning

I sammenheng med overgangen til SNA og ny versjon av MODIS er det behov for en endring i definisjoner og klassifikasjoner av skatter. I del 2 blir det foreslått å registrere skattene fra skatteyttersiden på det tidspunkt de påløper, og for den endelige mottaker å registrere skattene når de blir bokført. I del 3 foreslåes det å innføre en egen sektor for skatteinnkreving under den offentlige forvaltning. I del 4 blir de ulike poster under sektor for skatteinnkreving nærmere behandlet. I del 5 blir det foreslått at både indirekte og direkte skatter i prinsippet klassifiseres etter utlikningsgrunnlaget.

2. Definisjon av skattebegreper

Påløpne og bokførte skatter

Ulike tidspunkt kan legges til grunn ved registrering av skatter i

nasjonalregnskapet og MODIS. I SNA¹⁾ blir det anbefalt at skatter generelt skal registreres etter hvert som de blir innbetalt, men for analytiske formål i enkelte tilfeller suppleres med data over påløpne skatter. I dette notat foreslås det å nytte påløpne beløp som registreringsprinsipp fra skatteyter-siden, og etter kontantprinsippet (bokført beløp) for den endelige mottaker (statens bevilgningsregnskap, sosiale trygder, kommuneforvaltningen m.v.). Videre foreslås det å innføre en sektor for skatteinnkreving som representerer overgangen fra påløpne til bokførte verdier. Oppstillingen nedenfor gir en oversikt over betydningen og oppbyggingen av skattebegrepene.²⁾

Tabell 1.

Hovedbegrep	Underbegrep	Merknader til avgrensingen
Påløpne direkte skatter	Påløpne inntekts- og formuesskatter, og medlemsavgift til folketrygden	Direkte skatter som administreres av likningsvesenet. Disse skatter regnes påløpt det år inntekten er opptjent og formuesgrunnlaget registreres.
	Påløpen arbeidsgiveravgift til folketrygden	Arbeidsgiveravgiften regnes påløpt etter hvert som kontraktmessig lønn blir utbetalt.
	Andre påløpne direkte skatter	Registreringstidspunktet i offentlige regnskaper legges til grunn.
Påløpne avgifter og subsidier	Påløpne vareavgifter og varesubsidier	Avgiftene og subsidiene regnes som påløpt i det år beregningsgrunnlaget henfører seg til.
	Påløpne sektoravgifter og sektorsubsidier	
	Toll	Registreringstidspunktet i offentlige regnskaper legges til grunn.
Endring i skatteinnkreverbeholdninger	Endring i skatteinnkreverbeholdning, direkte skatter	Skatteinnkreverbeholdningene omfatter skatter som er påløpt, men som ikke er kommet inn i offentlige regnskaper.
	Endring i skatteinnkreverbeholdning, indirekte skatter	
Bokførte direkte skatter		Registreringstidspunktet i offentlige regnskaper legges til grunn.
Bokførte avgifter og subsidier		

1) Jfr. A System of National Accounts. New York 1968, side 128.

2) Utdrag av denne oppstilling er gitt i arbeidsnotatet IO.72/2.

Forslaget vil på enkelte punkter bety brudd med praksis i nasjonalregnskapet i dag. Omleggingen i registreringsprinsippene vil gi seg størst utslag for etterskottspliktige, da det i dag er innbetalte direkte skatter som blir registrert, og disse finner etter gjeldende skatteregler sted i gjennomsnitt 1,5 år etter at skattene påløper. For forskottspliktige vil forslaget ikke føre til vesentlige endringer i forhold til praksis i nasjonalregnskapet hittil, fordi restskatter og tilbakebetalinger erfaringsmessig ikke varierer mye fra år til år.

Avgifter og subsidier blir i nasjonalregnskapet generelt ført som inntekt for det offentlige det år de påløper¹⁾, slik at heller ikke her vil forslaget bryte vesentlig med innarbeidet praksis.

En tar videre sikte på å skille mellom påløpen og bokført arbeidsgiveravgift under folketrygden. Dette betyr ikke brudd med tidligere prinsipper, selv om en i nasjonalregnskapet hittil har latt påløpen og bokført arbeidsgiveravgift falle sammen. I dag er dette en lite tilfredsstillende løsning fordi arbeidsgiveravgiften under folketrygden i de senere år har økt betydelig, og forskjellen mellom påløpen og bokført arbeidsgiveravgift kan da lett bli stor, særlig i de år det skjer regelendringer.

3. Sektorinndeling av den offentlige forvaltning

Ved overgangen til SNA foreslås det følgende hovedsektorinndeling av den offentlige forvaltning:

1. Offentlig forvaltning
 11. Statsforvaltningen
 - 111 Statskassen
 - 112 Andre statsregnskap
 12. Trygdeforvaltningen
 13. Kommuneforvaltningen
 14. Sektor for skatteinnkreving

I forhold til tidligere standardgrupperinger (jfr. arbeidsnotatet IO 70/10) foreslås det i dette notat den endring at undersektoren 1115 Skatteoppkrevere får ny betegnelse og kalles sektor for skatteinnkreving. Denne undersektor vil bli mer omfattende enn sektor 1115 Skatteoppkrevere. Det er for å hindre misforståelser og forvekslinger med den innarbeidede betegnelsen "skatteoppkrevere" at det foreslås et annet navn på denne undersektor.

1) Nasjonalregnskapets inntektshovedbok gir imidlertid bokførte avgifter og subsidier, med unntak av merverdiavgift, investeringsavgift m.v.

4. Nærmere om sektor for skatteinnkreving

Da sektor for skatteinnkreving vil få et annet innhold enn undersektoren "skatteoppkrevdere" i nasjonalregnskapet i dag, vil en nedenfor gi en oversikt over hvilke delsektorer sektor for skatteinnkreving består av, og hva slags endringer i skattebeholdningene de enkelte delsektorer representerer.

A. Skatteinnkrevdere, direkte skatter

- I. Personlige skatteyttere
 1. Endring i netto restskatt
 2. Endring i utskrevet, men ikke innbetalt forskottsskatt
- II. Selskaper
 1. Endring i selskapenes skattegjeld
- III. Arbeidsgivere
 1. Endring i arbeidsgiverbeholdning av forskottstrekk
 2. Endring i arbeidsgiverbeholdning av arbeidsgiveravgift
- IV. Kommunale skatteoppkrevdere
 1. Endring i margin
 2. Netto korreksjoner
 - (i) Statsskatt
 - (ii) Medlemspremie
 - (iii) Kommuneskatt

B. Skatteinnkrevdere, indirekte skatter

- V. Private skatteinnkrevdere
 1. Endring i skatteinnkreverbeholdning av merverdiavgift og investeringsavgift
 2. Endring i skatteinnkreverbeholdning av særavgift

En vil kort kommentere de enkelte poster i oppstillingen.

Post I.1. Det løpende forskottstrekk eller utskrevet forskottsskatt kan være større eller mindre enn faktisk påløpen (utliknet) skatt for et inntektsår. Det oppstår restskatter eller krav om tilbakebetaling (overskytende forskott). Oppgjør mellom publikum og de offentlige skatteoppkrevdere foregår stort sett året etter inntektsårets utgang. Ved årets utgang vil det på disse kontiene være innestående restskatter - overskytende forskott som er opparbeidet i løpet av året + restanser fra tidligere år som ikke er oppgjort. Det er bevegelsene på disse kontiene som registreres under post I.1. Etter forslaget vil kontiene i forhold til konsumentene virke som frivillige spare- eller gjeldskonti.

Post I.2. Ved årets begynnelse får selvstendig næringsdrivende en forskottsseddel som forfaller i fire terminer innenfor inntektsåret. I prak-

sis følges ikke betalingsterminene, og beløp blir utestående ved årets utgang.

Post II.1. For at denne kontoen skal virke etter sin hensikt må en tenke seg at selskapene ved årets utgang krediteres skattegjeldskonto med et beløp som tilsvarende påløpen skatt for vedkommende år (før beløpet er kjent), og debiteres kontoen for faktiske utbetalinger til det offentlige. Ved årets utgang vil selskapenes skattegjeld med det gjeldende system være påløpen skatt for vedkommende år + halvparten av påløpen (utliknet) skatt for foregående år. Det er endringene på denne kontoen som registreres under post II.1.

Post III.1. Arbeidsgivernes løpende forskottstrekk i lønnsinntekt forfaller til betaling annen hver måned. Ved utgangen av året består beholdningen stort sett av trekk fra november og desember.

Post III.2. Arbeidsgiveravgift som trekkes av lønnsinntekt, forfaller til betaling i samme terminen som forskottstrekket.

Post IV.1. Den kommunale skatteoppkrever holder tilbake en viss prosent av de løpende innbetalinger (margin). Dette beløp blir ikke fordelt mellom de ulike offentlige institusjoner samme år. Året etter blir beløpet enten brukt til finansiering av tilbakebetalinger til publikum eller fordelt mellom institusjonene.

Post IV.2. De løpende innbetalinger til den kommunale skatteoppkrever blir fordelt mellom institusjoner etter en foreløpig fordelingsnøkkel. Når den endelige fordelingsnøkkel er kjent, foretas et korreksjonsoppgjør mellom institusjonene. Nettokorreksjonen på tvers av institusjoner vil alltid være 0.

Post V.1. Etter gjeldende hovedregel for innbetaling av merverdiavgift er det en betalingsfrist på gjennomsnittlig ca. 3 måneder etter avgiften påløper (faseforskyvningen). For jordbruk, skogbruk og fiske er det spesielle regler, med en betalingsfrist på gjennomsnittlig ca. 10 måneder etter at avgiften påløper. Betalingsfristene for investeringsavgiften følger merverdiavgiftens bestemmelser. Det er kjøpere av investeringsvarene som skal betale inn avgiften til myndighetene.

Post V.2. For særavgifter er det ulike regler og fremgangsmåter ved innkrevningen. En har gjennom post V.2 åpnet muligheten for å registrere ulikheten mellom påløpne og bokførte særavgifter.

5. Klassifikasjonsprinsipper

Skatter kan grupperes på ulike måter avhengig av formålet med klassifiseringen og hvilke problemstillinger en mener det er relevant å belyse. Skatter kan for det første klassifiseres etter hvem som administrativt betaler

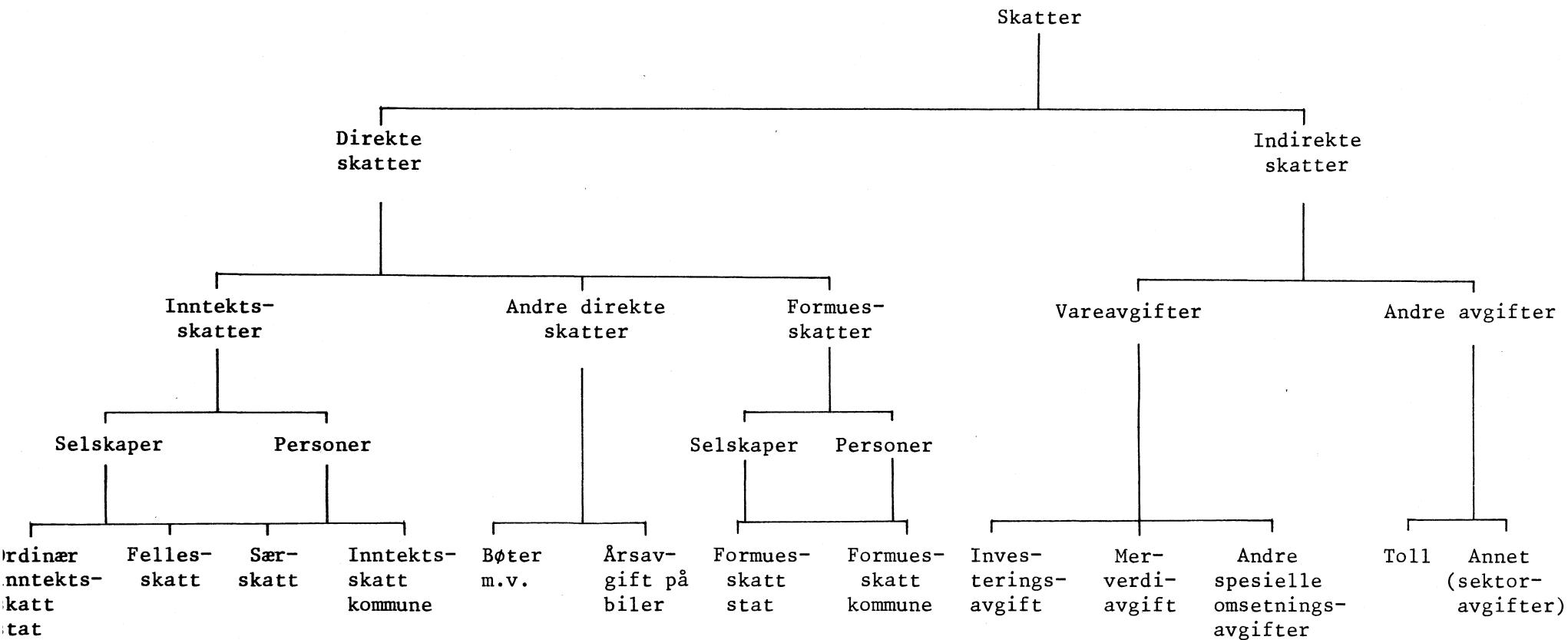
dem til offentlige forvaltningsorganer. En skiller mellom direkte skatter, som er skatter som blir betalt direkte av skattyterne til den offentlige forvaltning, og indirekte skatter, som blir betalt til den offentlige forvaltning av næringsdrivende, i forbindelse med omsetning, produksjon og kjøp av driftsmidler. Ut fra et økonomisk-teoretisk synspunkt er denne klassifikasjonen kanskje ikke så fruktbar. Viktigere er det å bestemme hvilket beregningsgrunnlag (utlikningsgrunnlag) de enkelte skatter er basert på. En foreslår derfor i prinsippet å klassifisere både indirekte og direkte skatter etter utlikningsgrunnlaget¹⁾. I fig. 1²⁾ er de viktigste norske skatter forsøkt klassifisert etter dette prinsipp. For direkte skatter har en videre skilt mellom personer og selskaper. Indirekte skatter er splittet opp i:

- (i) Vareavgifter og varesubsidier. Dette er avgifter og subsidier som er knyttet til varestrømmer som beregningsgrunnlag.
- (ii) Sektoravgifter og sektorsubsidier. Dette er avgifter og subsidier som er knyttet enten til varer eller til spesielle egenskaper ved institusjonelle sektorer.
- (iii) Toll. Dette er avgifter som er knyttet til varer som leveres fra importsektorene.

1) Forslaget er gi samsvar med tilrådinger gitt i OECD's Committee on Fiscal Affairs (CFA (71)1).

2) Jfr. også Charles M. Allan "The Theory of Taxation". Manchester 1971, side 29.

Fig. 1. KLASSIFIKASJON AV SKATTER



JMH/OLj/GH, 2/12-71.

Problemer knyttet til spesifikasjon
av alderen for enkelte aldersgrupper
[i arbeidskraftundersøkelsene]

av

Jan M. Hoem og Olav Ljones.

Innhold

	Side
1. Innledning	17
2. Beregningsteknikk for visse rater	18
3. Antall personer i utvalget med gitt beregnet alder	19
4. Antall personer i utvalget med gitt oppnådd alder	22
Vedlegg 1: Utdrag fra notat JMH/OLj/NTD, 3/6-71	24
Vedlegg 2: Rotasjonsplan 4 i håndskrevet notat fra Hoem 22/11-71	25

1. Innledning

I et tidligere notat med samme tittel (JMH/OLj/WTD, 3/6-71)*) påpekte vi at man vil få et varierende antall personer med i utvalget i enkelte aldre med den utvalgsplan vi da baserte oss på. Det skyldes at et kriterium for at et potensielt intervjuobjekt skal bli med i utvalget, er at det er mellom 16 og 74 år (beregnet alder, dvs. alder pr. 31/12) første gang den husholdning det tilhører, kommer med i en intervju-pulje. Dette kunne få betydning for ulike rater spesifisert for større aldersklasser, slik som klassen 16-19 år.

Vi påviste at disse problemene ville falle bort dersom en endret de nominelle aldersgrensene fra 16-74 år til 15-75 år. I de to prøveundersøkelsene i 1971 har man imidlertid brukt grensene 16-74 år.

Det er nå vedtatt at vi skal bruke en annen rotasjonsplan enn den notatet av 3/6-71 var basert på, nemlig plan 4 i Hoems håndskrevne notat om rotasjonsplanen, datert 22/11-71. I kapittel 3 og 4 nedenfor skal vi påvise at en får problemer analoge med de ovennevnte også med den nye rotasjonsplanen. Disse problemene vil til og med utvides til å gjelde for 17-åringene også, mens disse ikke var problematiske i forrige notat.

En kan igjen løse disse problemene ved å utvide de nominelle aldersgrensene til 14-75 år. En slik løsning vil imidlertid innføre problemer av en annen art, idet vi da nemlig kommer til å samle inn opplysninger f.eks. fra en del 14- og 15-åringer uten å ha til hensikt å bruke disse opplysningene til noe.

Det er imidlertid mulig å løse problemene med det varierende antall personer i enkelte aldersgrupper på en annen måte. I kapittel 2 påviser vi at en annen oppblåsningsteknikk enn den nåværende automatisk vil kompensere for de varierende persontall. Vi foreslår derfor at man snarest mulig går over til denne nye oppblåsningsteknikken, som ikke er mer komplisert enn den vi har brukt hittil.

I hele dette notatet gjør vi samme slags forenkling forutsetninger som i notatet av 3/6-71.

*) Utdrag gjengis i vedlegg 1.

2. Beregningsteknikk for visse rater

For å konkretisere diskusjonen i dette kapitlet skal vi tenke oss at man ønsker å estimere yrkesfrekvensen for kvinner i alderen 16-19 år i befolkningen på grunnlag av én enkelt utvalgsundersøkelse. La Y_x være antall sysselsatte kvinner i alder x i befolkningen, og la y_x være det antall av disse som kommer med i utvalget. La N_x være samlet antall x -årige kvinner i befolkningen, og la n_x være det tilsvarende antall i utvalget. (Her vil både n_x og y_x være stokastiske variable slik våre utvalg trekkes, og N_x vil være prognostisert, ikke kjent, men vi skal se bort fra slike momenter nå.) Den yrkesfrekvensen vi ønsker å estimere, er da

$$f = \frac{\sum_{x=16}^{19} Y_x}{\sum_{x=16}^{19} N_x}.$$

Det finnes minst to måter å estimere denne på med de data vi har for hånden:

(i) Vi kan basere oss utelukkende på dataene i utvalget, og estimere f med

$$f^* = \frac{\sum_{x=16}^{19} y_x}{\sum_{x=16}^{19} n_x}.$$

Dette tilsvarer forøvrig en metodikk der vi først estimerer $\sum Y_x$ ved å "blåse opp" tallene for hele aldersklassen 16-19 under ett, og så estimerer f ved divisjon med $\sum N_x$. I så fall ville nemlig vår estimator for $\sum Y_x$ være

$$\sum_{x=16}^{19} y_x \left(\frac{\sum_{x=16}^{19} N_x}{\sum_{x=16}^{19} n_x} \right).$$

Divisjon med $\sum N_x$ gir naturligvis f^* .

(ii) Alternativt kan vi "blåse opp" tallene for hver enkelt aldersklasse for seg før vi estimerer f , slik at vi får estimatoren

$$\hat{f} = \frac{\sum_{x=16}^{19} y_x \frac{N_x}{n_x}}{\sum_{x=16}^{19} N_x}.$$

Mens estimatoren f^* åpenbart er influert av variasjoner f.eks. i antall 16-åringer, er dette ikke tilfelle med \hat{f} . (Vi antar her at Y_{16}/N_{16} avviker fra Y_x/N_x for høyere x .) Hvis n_{16} fluktuerer fra kvartal til kvartal, vil y_{16} fluktuerer i takt med n_{16} (forutsatt konstant Y_{16}/N_{16}) slik at y_{16}/n_{16} holder seg i ro. Man bør altså bruke \hat{f} , ikke f^* .

En kan merke seg at den betingede variansen på \hat{f} er større enn det den ville vært om vi hadde hatt full representasjon av 16-åringer.

I dag brukes en beregningsmetode som ved nærmere ettersyn viser seg å gi f^* . Det har ikke vært mulig å finne ut hvorfor det er tilfelle - det ser ut for at det bare "er blitt slik av seg selv". Dette tyder på at saken ikke ble behørig diskutert da bestemmelsen ble tatt.

Vi anbefaler overgang til \hat{f} , dvs. omlegging av oppblåsnings- teknikken slik at hver ettårig aldersklasse blåses opp for seg selv.

3. Antall personer i utvalget med gitt beregnet alder

Vi skal her innføre en formelapparat som gjør det mulig å beregne antall personer med gitt beregnet alder i utvalget for enhver aktuell utvalgsplan. Vi skal så gi noen eksempler på anvendelse av formlene på plan 4 i Hoems notat av 22/11-71.

La t betegne tellingsår og k tellingskvartal, $t = 1, 2, \dots$ og $k = 1, 2, 3, 4$. Utvalgspuljene nummereres fra 1 og oppover. Pulje nr. i er med i undersøkelsene for første gang i kvartal k_i i år t_i . Denne puljen inneholder da $m_i(x, k_i, t_i)$ personer med beregnet alder x . Disse vil være født i år $t_i - x$. Vi ser bort fra dødsfall, flyttinger, skifte av utvalgsområde, osv.

Rotasjonsplanen er definert gjennom spesifikasjon av indeksvariable $I_i(k, t)$, der

$$I_i(k,t) = \begin{cases} 1 & \text{hvis pulje nr. } i \text{ er med i utvalget} \\ & \text{som intervjues i kvartal } k \text{ i år } t, \\ 0 & \text{ellers.} \end{cases}$$

Antall personer med beregnet alder x i utvalget i kvartal k i år t , er da lik

$$(1) \quad b_x(k,t) = \sum_{r=0}^{t-1} \sum_{\{i:t_i=t-r\}} I_i(k,t) m_i(x-r, k_i, t-r).$$

La oss nå sette

$$(2) \quad m_i(x,k,t) = \begin{cases} m & \text{for } 16 \leq x \leq 74, \\ 0 & \text{ellers.} \end{cases}$$

Hvis

$$\delta(x) = \begin{cases} 1 & \text{for } 16 \leq x \leq 74, \\ 0 & \text{ellers,} \end{cases}$$

blir altså $m_i(x,k,t) = m \delta(x)$. Da blir

$$(3) \quad b_x(k,t) = m \sum_{r=0}^{t-1} \sum_{\{i:t_i=t-r\}} I_i(k,t) \delta(x-r).$$

Formel (1) og (3) gjelder for alle utvalgsplaner. I "plan 4" er $I_i(k,t) = 0$ hvis $t-t_i > 2$, så vi trenger bare summere fra $r=0$ til $r=2$. Vi kan da skrive (2) på formen

$$\begin{aligned} \frac{1}{4m} b_x(k,t) &= \frac{1}{4} \sum_{\{i:t_i=t\}} I_i(k,t) \delta(x) \\ (4) \quad &+ \frac{1}{4} \sum_{\{i:t_i=t-1\}} I_i(k,t) \delta(x-1) \\ &+ \frac{1}{4} \sum_{\{i:t_i=t-2\}} I_i(k,t) \delta(x-2). \end{aligned}$$

For $18 \leq x \leq 74$ blir $\delta(x) = \delta(x-1) = \delta(x-2) = 1$, og (4) reduserer seg til

$$\frac{1}{4m} b_x(k,t) = 1 \quad \text{for } 18 \leq x \leq 74.$$

For $x = 17$ blir siste sum i (3) lik 0, og for $x = 16$ blir begge de siste summene i (3) lik 0. Nærmere beregninger gir de verdiene for

$$\frac{100}{4m} b_x(k,t) \quad \text{for } x = 16; 17,$$

som finnes i tabell 1.

Tabell 1. Prosentvis representasjon av personer med beregnet alder 16 og 17 i en del intervjurunder,

		$\frac{100}{4m} b_x(k,t)$.	
Intervjurunde		16-åringer	17-åringer
1971,	3. kvartal	100	100
	4. "	100	100
1972,	1. "	50	100
	2. "	50	100
	3. "	50	100
	4. "	75	100
1973,	1. "	25	100
	2. "	50	100
	3. "	50	100
	4. "	50	100
1974,	1. "	25	75
	2. "	50	100
	3. "	50	100
	4. "	50	100
1975,	1. "	25	75
	2. "	50	100
	3. "	50	100
	4. "	50	100

4. Antall personer i utvalget med gitt oppnådd alder

Av de $m_i(x, k_i, t_i)$ personene vi innførte i begynnelsen av forrige kapittel, vil $2/12$ ha fødselsdag før undersøkelsesuken i første kvartal, $5/12$ før undersøkelsesuken i annet kvartal, og de tilsvarende brøkene for tredje og fjerde kvartal blir $8/12$ og $11/12$. I k -te kvartal er brøken $(3k-1)/12$. Resten, dvs. $1 - (3k-1)/12 = (13-3k)/12$, har fødselsdag etter undersøkelsesuken i k -te kvartal.

Av de $m_i(x-r, k_i, t-r)$ personene fra pulje i med beregnet alder x i år t , har altså andelen $(3k-1)/12$ oppnådd alder x , mens resten har oppnådd alder $x-1$. I utvalget i kvartal k i år t er antall personer med oppnådd alder x derfor lik

$$(5) \quad c_x(k, t) = \frac{3k-1}{12} \sum_{r=0}^{t-1} \sum_{\{i:t_i=t-r\}} I_i(k, t) m_i(x-r, k_i, t-r) \\ + \frac{13-3k}{12} \sum_{r=0}^{t-1} \sum_{\{i:t_i=t-r\}} I_i(k, t) m_i(x+1-r, k_i, t-r).$$

Under (2) reduserer dette seg til

$$(6) \quad c_x(k, t) = \frac{3k-1}{12} m \sum_{r=0}^{t-1} \sum_{\{i:t_i=t-r\}} I_i(k, t) \delta(x-r) \\ + \frac{13-3k}{12} m \sum_{r=0}^{t-1} \sum_{\{i:t_i=t-r\}} I_i(k, t) \delta(x+1-r).$$

I "plan 4" kan vi spesielt skrive dette som

$$(7) \quad \frac{1}{4m} c_x(k, t) = \frac{1}{48} \sum_{r=0}^2 \sum_{\{i:t_i=t-r\}} I_i(k, t) \left[(3k-1) \delta(x-r) + (13-3k) \delta(x+1-r) \right]$$

Vi får

$$\frac{1}{4m} c_x(k, t) = 1 \quad \text{for } 18 \leq x \leq 73.$$

Verdier for $\frac{1}{4m} c_x(k,t)$ for $x = 16$ og 17 er gitt i tabell 2. Verdier for $x = 74$ kan regnes ut på tilsvarende måte, idet vi for denne x får

$$\begin{aligned} \frac{1}{4m} c_{74}(k,t) &= \frac{3k-1}{48} \sum_{\{i:t_i=t\}} I_i(k,t) \\ &+ \frac{1}{4} \sum_{\{i:t_i=t-1\}} I_i(k,t) + \frac{1}{4} \sum_{\{i:t_i=t-2\}} I_i(k,t). \end{aligned}$$

Også denne vil variere med (k,t) . Eksempelvis blir $c_{74}(k,t)/4m$

i 3. kvartal 1971 lik $\frac{2}{3}$, og

i 4. kvartal 1971 lik $\frac{11}{12}$.

Det er verd å merke seg at brøken altså ikke blir lik 1, slik den blir for alle x der $16 \leq x < 74$ i disse to kvartalene.

Tabell 2. Representasjon av personer med oppnådd alder 16 og 17 i en del intervjurunder, $\frac{1}{4m} c_x(k,t)$

Intervjurunde		16-åringer		17-åringer	
		Andel	I prosent	Andel	I prosent
1971,	3. kvartal	1	100	1	100
	4. "	1	100	1	100
1972,	1. "	11/12	91,7	1	100
	2. "	19/24	79,2	1	100
	3. "	2/3	66,7	1	100
	4. "	37/48	77,1	1	100
1973,	1. "	7/8	87,5	1	100
	2. "	19/24	79,2	1	100
	3. "	2/3	66,7	1	100
	4. "	13/24	54,2	1	100
1974,	1. "	2/3	66,7	23/24	95,8
	2. "	19/24	79,2	1	100
	3. "	2/3	66,7	1	100
	4. "	13/24	54,2	1	100
1975,	1. "	2/3	66,7	23/24	95,8
	2. "	19/24	79,2	1	100
	3. "	2/3	66,7	1	100
	4. "	13/24	54,2	1	100

Utdrag fra notat JMH/OLj/WTD, 3/6-71.

I utvalgsundersøkelsene har vi nominelt en nedre aldersgrense på 16 år og en øvre på 74 år. Dette tenkes praktisert slik:

Når en husholdning kommer med i utvalget for første gang, tar man med i undersøkelsen samtlige husholdningsmedlemmer som på fødselsdagen i det aktuelle kalenderår er mellom 16 og 74 år, begge aldre inklusive. Dette korresponderer med at man ved inntreden i utvalget bruker en beregnet alder lik alderen pr. 31/12. Denne beregnede alderen er naturligvis også det samme som differensen mellom nummeret på det aktuelle kalenderår og fødselsåret.

Med denne fremgangsmåten kan en person godt være 15 eller 75 år gammel (regnet i antall fylte år) på et intervju tidspunkt.

Personer som var for unge til å bli opptatt ved første intervju, tas ikke med i noen av de senere intervjurundene selv om de skulle ha blitt gamle nok i mellomtiden. Analogt utelates ingen av dem som blir 75 år før intervju syklusen er omme, hvis de var med i første intervjurunde.

Dette vil bety at vi får et varierende antall 15- og 16-åringer i utvalget, selv "i gjennomsnitt", når alderen regnes i fylte år. Analogt for 74- og 75-åringer. (Dette gjelder ikke for andre aldre.)

Hvis utvalgsmedlemmene ordnes etter beregnet alder (istedenfor etter alder i fylte år), vil det naturligvis aldri kunne bli noen 15-åringer i utvalget. Vi vil imidlertid fortsatt få et varierende antall 16-åringer. Antall 74-åringer vil ikke vise systematiske fluktuasjoner i dette tilfellet, men antall 75-åringer vil gjøre det.

Under beregningen av ulike rater spesifisert for aldersklasser som omfatter f.eks. 16-åringene, slik som rater for aldersgruppen 16-19 år, er det viktig å være oppmerksom på dette. 16-åringene vil jo da inngå med varierende grad av underrepresentasjon.

Vi skal se litt nærmere på disse spørsmålene. For å rendyrke de effekter vi her vil studere, skal vi se bort fra dødelighet i intervjuåret, flyttinger, variasjoner i størrelsen av fødselskullene, utvalgsvariasjon, sesongvariasjoner i fødslene (over året), osv.

Rotasjonsplan 4 i håndskrevet notat fra Hoem 22/11-71

Arbeidskraftundersøkelsene benytter et roterende utvalg av intervjuobjekter. Rotasjonsplanen er slik at ingen IO skal intervjues mer enn fire ganger. Derimot kan man intervju noen færre enn fire ganger dersom det passer slik.

Hver fjerdedel av dem som intervjues i et gitt kvartal, utgjør en IO-pulje. I 1971-72 brukte man en utvalgsplan som i håndskrevet notat fra Hoem 22/11-71 er kalt plan 4. Planen kan fremstilles som i nedenstående figur. Hver linje i figuren representerer en IO-pulje og hver kolonne et intervjukvartal. Kryss i en rute indikerer at IO-puljen blir intervjuet.

	1971	1972	1973	1974
—	X X X			
—	X X X X			
—	X X	X X		
—	X X		X X	
—		X	X X	
—		X X	X X	
—		X X	X X	
—			X X	X X
—			X X	X X
—				X X
—			X X	X X
—				X X

ALDERSBEGREPET I ARBEIDSKRAFTUNDERSØKELSENE

av

Jan M. Hoem

I n n h o l d

	Side
1. Ulike definisjoner av aldersbegrepet	26
2. Genuine alderseffekter	28
3. Fiktive alderseffekter	30
4. Mer om utvalgsramme og oppblåsningsmetode ...	33
5. Oppsummering av diskusjonstemaer reist i dette notatet, og sluttkommentar	35
Referanser	36
Tabell 1	37

1. Ulike definisjoner av aldersbegrepet

1A. I de beregninger som hittil har vært utført i tilknytning til arbeidskraftundersøkelsene på utvalgsbasis, har en forsøksvis operert med to aldersbegreper, nemlig oppnådd alder (alder i fylte år) og alder (i fylte år) ved utgangen av kalenderåret (begge spesifisert i hele år). Beregninger på grunnlag av dataene fra arbeidskraftundersøkelsene viser at det i praksis ikke er likegyldig hvilken aldersdefinisjon som benyttes. I tabell 1 har jeg sakset fra tabellvedlegget (SH/IH, 24/8-72) til sakspapirene til utvalgsmøtet 4/9-72. Eksempelene viser tilfeller der forskjellene i aldersdefinisjon har slått særlig kraftig ut. Det er noen flere slike tilfeller i det fremlagte materialet.

I de fleste andre av det store antallet sammenligninger som er foretatt, er allikevel utslagene små.

Det ser ut for at betydningen av valget mellom disse to aldersdefinisjonene er størst for personer i særlig lav og særlig høy alder. Vi har også tidligere oppfattet aldersbegrepet som særlig problematisk for slike personer, og har behandlet det i flere notater. (Se referanselisten.)

1B. Under diskusjonen av aldersbegrepet bør en ikke begrense seg til de to definisjonene nevnt ovenfor. Andre definisjoner kan også komme på tale. Eksempelvis antyder Møglestue i et nylig notat (IM/IH, 25/8-72) alder i fylte år ved utgangen av kvartalet som "en typisk kompromissløsning". I dette notatet skal jeg lansere alder i fylte år ved nærmeste årsskifte som et seriøst alternativ. (Alder ved utgangen av kalenderåret er naturligvis det samme som alder ved neste årsskifte.)

Det har ikke ennå vært foretatt beregninger der en har brukt disse to siste aldersdefinisjonene. Alder i fylte år ved nærmeste årsskifte vil imidlertid bli brukt under en presentasjon av resultater fra arbeidskraftsundersøkelsene for et interdepartementalt arbeidsutvalg for yrkesprognoser i tilknytning til Langtidsprogrammet 1974-1977.

1C. Alderen, regnet i antall hele år, er naturligvis et kjennetegn som vi tillegger den enkelte, og som endrer seg med tiden ved at den får et tillegg på 1 hvert år. Den enkelte aldersdefinisjon fremkommer da ved at vi spesifiserer individets alder ved fødselen og det tidspunkt i året da alderen skal øke med 1. I det vi ser bort fra avvikende aldersdefinisjoner i andre kulturer, som her bare har kuriositetsinteresse, vil jeg bemerke at det ikke er opplagt at alderen skal settes til 0 år ved fødselen, selv om dét er det vanlige. I tilknytning til befolkningsprognosene settes alderen til -1 ved fødselene, og den får et tillegg på 1 ved hvert årsskifte. Dette gir alder i fylte år ved begynnelsen av kalenderåret. Jeg nevner dette vesentlig for å fremheve at vi bør kunne stille oss rimelig fritt til gjeldende konvensjoner når vi fastlegger en aldersdefinisjon for et gitt formål.

Aldersdefinisjonene nevnt i underkapitlene 1A og 1B fremkommer formelt slik:

Ved oppnådd alder settes alderen til 0 ved fødselen og den stiger med 1 hver fødselsdag.

Ved alder i fylte år ved utgangen av året settes alderen lik 0 ved fødselen, og den stiger med 1 ved hvert årsskifte.

Ved alder i fylte år ved utgangen av kvartalet settes alderen lik -1 ved fødselen og stiger med 1 ved utgangen av det kvartal fødselsdagen faller i.

Ved alder i fylte år ved nærmeste årsskifte stiger alderen med 1 pr. 1/7 hvert år. Alderen ved fødselen settes lik -1 hvis fødselsdagen faller i første halvår, og den settes lik 0 hvis den faller i annet halvår.

I det følgende skal jeg ikke komme mere inn på alder i fylte år ved utgangen av kvartalet. Den synes ikke egentlig å ha så stor interesse her.

2. Genuine alderseffekter

2A. En rekke av de kjennetegn vi observerer på individbasis i arbeidskraftundersøkelsene, påvirkes betydelig av individets alder fordi så mye av en persons handlingsmønster avhenger av hvor gammel han er. Når vi studerer hvordan et kjennetegn avhenger av alderen, regnet etter en gitt definisjon, postulerer vi på sett og vis at de som har samme alder, regnet etter den gitte definisjonen, stort sett har et homogent handlingsmønster, mens handlingsmønsteret til dem som har ulik alder, kan avvike fra hverandre. Hvis det er grunn til å regne med et slikt mønster av homogenitet og heterogenitet ved anvendelse av den gitte aldersdefinisjon, vil vi si at det foreligger en genuin alderseffekt.

La meg gi noen eksempler.

2B. For unge mennesker vil deres opptreden på ekteskapsmarkedet antakelig avhenge nokså mye av oppnådd alder. Lovreglene fastsetter aldersgrenser for krav om forskjellige slags samtykke til giftermål, og disse grensene er uttrykt ved alder i fylte år. Det finnes imidlertid særlig populære årstider å gifte seg på, og hvis folk må fullføre skolegang o.l. før de gifter seg, kan skoleåret få en viss inflytelse. Det finnes derfor også momenter som virker i retning av å bringe inn andre faktorer enn oppnådd alder.

2C. De unges skolegang, og følgelig deres opptreden på arbeidsmarkedet, vil særlig være påvirket av alderen i forhold til skoleåret. Slik det er organisert her i landet, vil de fleste begynne i første klasse i barneskolen om høsten det året de fyller syv år, de fortsetter i andre klasse neste høst, osv. De fleste vil derfor gå ut av ungdomsskolens niende klasse om våren det året de fyller 16 år. Det betyr at de går i første klasse i det skoleåret da deres alder i fylte år pr. nærmeste årsskifte er syv, i annen klasse det året da deres alder pr. nærmeste årsskifte er åtte, osv. De fleste går gjennom niende klasse i det skoleåret da de er 15 år gamle pr. nærmeste årsskifte.

2D. Dette bildet gjøres mindre enkelt ved at det finnes enkelte skoleslag der undervisningen avsluttes ved slutten av kalenderåret, ikke ved slutten av det vanlige skoleåret. Ved universiteter og høyskoler kan man jo avslutte sin utdanning etter hvert semester, og det finnes en del utdanningsveier, særlig halvårlige skoler som husmorskoler og handelskoler, som har to eksamenskull i året. Andre eksempler finnes sikkert også.

Møglestue mener allikevel at omfanget av utdanningsavslutning ved årsskiftet teller lite i det store bildet. Selv om man ikke bør glemme at det er en del som går ut av skoleverket på den tiden av året også, vil hovedvirkningen på arbeidskraftundersøkelsene sannsynligvis komme fra det ordinære skoleåret. For sikkerhets skyld skal han imidlertid se nærmere på dette til neste utvalgsmøte.

2E. De fleste som oppnår alderspensjon, slutter vel å arbeide omtrent på den fødselsdag da de når pensjonsalderen. Pensjonering er derfor trolig nært forbundet med oppnådd alder. Dette motvirkes muligens noe av at enkelte "står i stillingen" ut kalenderåret, ut skoleåret, o.l. Det bør være mulig å finne ut hvor mye dette momentet eventuelt kan telle. Hoffmann nevner at undersøkelser ved Gerontologisk institutt kanskje kan være til hjelp her. Jeg antar at I. kontor vil ta seg av dette hvis det finnes nødvendig.

2F. En genuin alderseffekt kommer naturligvis klarest frem når dataene blir analysert etter en aldersdefinisjon som passer godt med det fenomen som har forårsaket effekten. Slike alderseffekter kan allikevel være så sterke at de slår igjennom selv om en bruker en aldersdefinisjon som ikke egentlig passer så bra. Ovenfor har jeg hevdet at alderseffekten av skolegang kommer klarest frem i arbeidskraftundersøkelsene når alder defineres i fylte år ved nærmeste årsskifte. En annen aldersdefinisjon vil kunne gjenspeile denne effekten på ulike vis, men i så fall må denne andre definisjonen bare betraktes som et substitutt for den vi "egentlig" burde bruke.

Noen alternative aldersdefinisjoner er bedre substitutter enn andre. Eksempelvis er alder ved utgangen av året en nærliggende erstatning for alder ved nærmeste årsskifte. De to definisjonene faller imidlertid ikke sammen når de anvendes i arbeidskraftundersøkelsene, slik en kanskje kunne tro umiddelbart. Jeg skal begrunne dette nærmere i underkapittel 3D nedenfor.

2G. En inadekvat aldersdefinisjon vil naturligvis gi et noe fordreiet bilde av en genuin alderseffekt. Noen ganger er fordreilingen så stor at den får betydning for tolkningen av de bearbejdede observasjonsresultatene. Enkelte inadekvate aldersdefinisjoner har f.eks. vist seg å skape fluktuasjoner i de tabellerte observasjonsverdiene, fluktuasjoner som vanskelig kan ha sin bakgrunn i den virkelighet som skal belyses. Dette

vil jeg kalle fiktive alderseffekter. Vi skal nå se nærmere på hvordan slike kan oppstå.

3. Fiktive alderseffekter

3A. I arbeidskraftsundersøkelsene skal vi velge aldersdefinisjon i to sammenhenger, nemlig

- (i) i forbindelse med trekkingen av utvalget, og
- (ii) i tilknytning til bearbeidingen av resultatene.

Vi trenger ikke bruke samme aldersdefinisjon for de to formålene. Ulike hensyn kan tvert om tilsi at det kan være hensiktsmessig å bruke forskjellige definisjoner. Imidlertid vil det naturligvis være et visst samspill mellom dem, og det må en ta med i betraktningen når en velger definisjonene.

3B. Valget av aldersdefinisjon får sin betydning fordi det fastlegger et prinsipp for utvelgelse eller organisering av dataene. Den bestand (eller populasjon) som skal studeres, er primært avgrenset etter alder. Både ved utvalgstreking og analyse begrenser vi oss til personer i aldre mellom 16 og 74, regnet etter den aldersdefinisjon som legges til grunn i det enkelte tilfelle. Dersom vi bruker forskjellig aldersdefinisjon ved utvalgstreking og analyse, og mye taler for at vi bør gjøre det, vil utvalgspopulasjon og analysepopulasjon avvike fra hverandre. Vi må da passe på at utvalgspopulasjonen er størst, og at vi har adekvate data for alle deler av analysepopulasjonen.

Vi skal diskutere disse momentene og distinksjonene nærmere i nedenstående underkapitler.

3C. Aldersdefinisjonen kommer som kjent inn i utvalgstrekingen på følgende måte:

Når en ny pulje av husholdninger tas inn i utvalget, vil vi bare intervju de husholdningsmedlemmene som har en alder ved utgangen av kalenderåret på minst 16 og høyst 74. I senere intervjurunder beholdes disse personene i utvalget, mens f.eks. husholdningsmedlemmer som får alder 16 (regnet etter samme definisjon) etter første intervjurunde, holdes utenfor.

Motivet for å bruke alder ved utgangen av året for å avgjøre om en person skal tas med i første intervjurunde, var vel først og fremst at

dette ville lette utvalgsarbeidet for intervjueren. Kanskje har forestillingen om at tabelleringene skulle skje etter samme aldersdefinisjon også spilt en rolle. Nedenfor skal jeg argumentere for å beholde dette utvelgelseskriteriet også ut fra det hensyn at det vil gjøre det mulig for oss å observere avgangen fra ungdomsskolen.

Vi bestemte oss for å holde personer som ikke var "utvalgbare" i første intervjurunde, helt utenfor undersøkelsen også i senere intervjurunder primært fordi vi ville unngå forskjellig intervjunummereffekt blant medlemmer i samme husholdning. Dette bygget på amerikanske og kanadiske erfaringer med intervjunummereffekter. Svenskene har ikke funnet tilsvarende resultater i sine undersøkelser, så jeg er nå blitt usikker på hvor tungtveiende dette argumentet er. Det er nok grunn til at vi diskuterer momentet på ny i arbeidsutvalget. Jeg skal derfor gå litt nærmere inn på det i underavsnitt 4B. nedenfor.

Den fremgangsmåten vi nå bruker på dette punkt i arbeidskraftsundersøkelsene, vil vel i alle fall lette intervjuerens arbeid.

3D. La oss så ta for oss aldersdefinisjonen i analysearbeidet. Ovenfor fremholdt jeg at den betydning skolegangen har for arbeidskraftundersøkelsene, best belyses hvis en analyserer dataene etter alder i fylte år ved nærmeste årsskifte. Alder i fylte år ved neste årsskifte (dvs. ved utgangen av kalenderåret) er ingen god erstatning for dette.

En kan kanskje innvende at dette først og fremst er en diskusjon av formaliteter. Ved undersøkelser utført i annet (kalender-) halvår vil jo alder ved neste årsskifte og alder ved nærmeste årsskifte være den samme. Ved undersøkelser utført i første halvår vil de to aldersbegrepene atskille seg med akkurat ett aldersår for alle personer. (En som har alder 16 ved neste årsskifte vil ha alder 15 ved nærmeste årsskifte ved en undersøkelse utført i januar - juni.) Dette er for så vidt riktig, og det gjør det lett å benytte beregninger utført og tabellert for ettårige aldersklasser etter den ene aldersdefinisjonen, når en vil gå over til å bruke den andre definisjonen.

Hvilken definisjon en bruker, får imidlertid betydning på flere måter:

(i) La oss si at en skal sammenligne yrkesaktiviteten til 16-åringer høsten 1971 og våren 1972. Hvis tabelleringen er foretatt etter alder ved utgangen av året, må en huske på at 16-åringene høsten 1971 er å finne blant 17-åringene våren 1972. De som er tabellert som 16-åringer våren 1972, var blant 15-åringene høsten før.

(ii) Tabelleringene skjer imidlertid ikke etter enkelte aldersår, men etter klasser på flere aldersår. Det blir vanskelig å sammenligne tabellresultater f.eks. for aldersklassen 16-19 på de to tidspunktene nevnt ovenfor. En reell sammenligning bør foretas mellom aldersklassen 16-19 om høsten og 17-20 om våren (etter alder ved utgangen av året), men den siste er ikke tabellert.

(iii) Den samlede analysebestanden om våren atskiller seg fra den tilsvarende bestanden høsten før når en bruker alder ved neste årsskifte som kriterium, og forskjellen viser seg å være ganske vesentlig for enkelte av våre tabelleringer. Om våren får en med i bestanden bl.a. en god del personer som var 15 år ved siste årsskifte, og som derfor ikke var med i bestanden året før. Disse vil i stor utstrekning være elever i ungdomsskolen, så en må vente et sprang i estimert antall skoleelever og studenter i første kvartal i forhold til det foregående fjerde kvartal, slik det også fremgår klart av første linje i tabell 1. Dette er en fiktiv alderseffekt som det bør være vår oppgave å eliminere. Tilsvarende fiktive effekter finnes i andre av tallseriene.

3E. Noen få kommentarer om oppnådd alder som analysebasis kan også være på sin plass.

Forskjellene mellom korresponderende linjer merket U og E skyldes naturligvis forskjeller i analysepopulasjon. Vi ser at disse kan gå begge veier, noen steder er elementet i linjen merket U størst og noen ganger er det minst. Eksempelvis er det estimerte antall arbeidsaktive ugifte kvinner større i linjen merket U enn i den merket E, noe som vel skyldes at det i første tilfelle er en del kvinner med oppnådd alder 15 år i bestanden, i motsetning til det andre tilfellet.

I linjen merket E under "skoleelever, studenter" har tallene en merkelig utvikling. I store trekk er det vel de samme personene som er skoleelever og studenter gjennom de fire undersøkelseskvartalene. En bør derfor vente en jevn vekst fra kvartal til kvartal over de tre første kvartalene i tallene i denne linjen. Det blir jo stadig flere skoleelever som har passert sin 16. fødselsdag etter som en går fra ett kvartal til det neste. Et årskull er på ca. 60 tusen personer, og hvert kvartal har vel omtrent fjerdeparten av disse fødselsdag. En burde vel da vente en kvartalsvis økning på noe under 15 tusen i linjen. (Vi får vel ikke fullt 15 tusen, for det er ikke alle i kullet som går på skoler.) I observasjonstallene er økningen på 12 tusen fra tredje til fjerde kvartal i 1971, og

dét er pent og rimelig. Fra siste kvartal i 1971 til første kvartal i 1972 er imidlertid økningen på hele 21 tusen, og for meg ser dette ut for å være et avvik (fra under 15 tusen) som langt overstiger tilfeldige fluktuasjoner.

Mellom de to kvartalene i 1972 går tallet ned med 26 tusen. Det er vanskelig å vurdere rimeligheten av størrelsesordenen her, men retningen er som ventet. Undersøkelsesuken i annet kvartal av 1972 falt nemlig i slutten av mai, og da må mange skoleelever og studenter ha vært i jobb.

4. Mer om utvalgsramme og oppblåsningsmetode

4A. I de to underkapitlene i dette siste hovedkapitlet skal jeg ta opp et par spørsmål som har noe perifer tilknytning til bestemmelsen av aldersdefinisjonen, som er vårt hovedtema, men som er kommet frem i tilknytning til dette notatet.

4B. Det er klart at yrkesaktiviteten til de svært unge på mange måter er særlig interessant. Det kan derfor være grunn til å velge en utvalgsplan som gir oss en særlig god representasjon av disse aldersklassene. Den planen vi nå bruker, gjør dessverre det stikk motsatte. Som vi har påvist i Notat JMH/OLj/GH, 2/12-71, blir 16-åringer (og tildels også 17-åringer) dårligere representert enn andre aldersklasser, både når en bruker alder ved utgangen av året og oppnådd alder. En vil få tilsvarende resultat om en bruker alder ved nærmeste årsskifte.

Dette skyldes at vi ikke tar med i utvalget personer som blir 16 år etter første intervjununde. Vi ville ikke hatt problemet dersom vi intervjuet f.eks. alle husholdningsmedlemmer med alder minst 16 ved utgangen av året. Vi bør diskutere om de momentene jeg nevnte på slutten av avsnitt 3C (intervjunummereffekten og intervjuerarbeidet) er så tungtveiende at vi ikke bør endre vår fremgangsmåte, f.eks. samtidig med omleggingen av skjemaet. Mitt foreløpige standpunkt er at vi bør endre utvalgsreglene slik som nettopp antydnet.

Jeg tror imidlertid at vi bør beholde alder ved utgangen av året som kriterium ved utvelgelsen av intervjuobjektene. Om våren får vi da med en del ungdommer hvis alder ved nærmeste årsskifte er 15, og vi kan observere avgangen fra ungdomsskolens niende klasse.

Et annet alternativ, som vi har drøftet og avvist tidligere, er å intervjuer også et par yngre aldersklasser, f.eks. 14- og 15-åringer (alder ved utgangen av året). Det er vel ikke grunn til å drøfte denne muligheten om igjen.

40. Det andre spørsmålet vi skal se på her, er knyttet til systemet for "oppblåsning" i estimeringsformlene. En bruker nå ettårige aldersklasser ved oppblåsningen. Hensikten med dette er å eliminere den virkning utvalgsplanen har på størrelsen av de enkelte aldersklasser i utvalget. (Sml. Notat JMH/OLj/GH, 2/12-71, s. 2-3.) Alle de resultatene som tabell 1 er et utdrag av, er regnet på denne måten.

Prinsipielt skal oppblåsningsmetoden ikke påvirke beregningsresultatene i uheldig retning, tvert imot. Derimot er det mulig at den praktiske utforming metoden har fått, kan ha hatt noen uheldige konsekvenser, uten at jeg har noen sikker oppfatning om det siste.

Oppblåsningsmetodens praktiske utforming har tidligere vært behandlet bl.a. i Notat SØP/SH/IH, 14/3-72 og i avsnitt 6 av Referat SH/SØP/WA, 27/3-72. La meg først minne om hva den går ut på ved å gjengi følgende eksempel.

La det være gitt en viss aldersdefinisjon, la Y_x være antall sysselsatte kvinner i alder x (etter denne aldersdefinisjonen) i befolkningen, og la y_x være det antall av disse som kommer med i utvalget. La N_x være samlet antall x -årige kvinner i befolkningen, og la n_x være det tilsvarende antall i utvalget. Da estimeres Y_x ved

$$\hat{Y}_x = \frac{N_x}{n_x} y_x.$$

Legg merke til at N_x i prinsippet skal være samlet antall x -årige kvinner i befolkningen i undersøkelsesuken. Selv om alderen er definert som alder ved utgangen av året, skal altså N_x ikke være f.eks. antall x -årige kvinner i befolkningen ved utgangen av året.

En annen sak er at vi ikke kjenner N_x , og derfor må anslå den på beste og enkleste måte. I dag bruker vi prognostisert antall x -årige kvinner pr. 31/12-71 for undersøkelsene i 1971, og prognostisert antall pr. 31/12-72 for undersøkelsene i 1972. Dette gjelder tabelleringer etter alder ved utgangen av året. (Vår praktiske utforming av metoden strider altså for så vidt mot vårt prinsipielle standpunkt.) Ved tabelleringer etter oppnådd alder har vi beregnet verdien av N_x ved begynnelsen av undersøkelsesuken. Vi har da tatt utgangspunkt i størrelsen av N_x ved siste årsskifte (personregisterets tall pr. 31/12-70 for undersøkelsene i 1971, og prognostiserte tall pr. 31/12-71 for undersøkelsene i 1972), antatt at personenes fødselsdager er jevnt fordelt over året og at dødsfall inntreffer i henhold til dødelighetstabellene for 1966-70.

Når vi har diskutert disse spørsmålene ved tidligere anledninger, har vi tatt det standpunkt at det må være av underordnet betydning hvilken av de aktuelle metoder en brukte til å anslå N_x . (Se f.eks. Referat SH/SØP/WA, 27/3-72, side 3, siste avsnitt.) Jeg er ikke lenger så sikker på at dét er riktig. Vi bør nok diskutere om en ikke bør utføre noen prøveregninger for å vurdere effekten av metoden til anslag for N_x .

5. Oppsummering av diskusjonstemaer reist i dette notatet og sluttkommentar

5A. Oppsummering

(i) Vi bør få foretatt tabelleringer analoge til dem vi alt har fått, men nå etter alder ved nærmeste årsskifte.

(ii) Utvalgskriteriene bør diskuteres, særlig m.h.t. om personer som blir 16 år skal tas inn i utvalget i intervjurunder etter den første.

(iii) Den praktiske utforming av oppblåsingsmetoden bør diskuteres.

5B. Sluttkommentar

Det er ikke sagt at alle nærmere analyser av dataene fra arbeidskraftundersøkelsene bør bruke samme aldersdefinisjon. Sannsynligvis bør aldersdefinisjonen trolig tilpasses det enkelte tilfelle. I standardtabellene fra undersøkelsene bør vi derimot ta én definisjon som brukes over alt.

Ulike brukerformål kan da gi argumenter for ulike aldersdefinisjoner for tabellsettet. Det kan derfor være at man må søke å forene motstridende interesser. Så vidt jeg kan se, er imidlertid de unges yrkes- og utdannelsesaktivitet særlig problematisk i denne sammenheng, for det er for dem at vi får inn fiktive alderseffekter ved inadekvate aldersdefinisjoner. Det blir vel da særlig viktig å ta hensyn til disse.

Vi har ennå ikke sett tabellresultater der dataene er organisert etter alder ved nærmeste årsskifte. En endelig oppfatning om hvor adekvat denne aldersdefinisjonen er, kan en derfor ikke ha ennå. Overlegningene i dette notatet får imidlertid meg til å trekke den tentative konklusjon at denne definisjonen ventelig er den beste av dem som nå er foreslått.

Ved valg av aldersdefinisjon i standardtabellene må en også ta hensyn til sammenlignbarheten med andre data. For meg ser det ut for at dette hensynet vil være ivaretatt om en bruker alder ved nærmeste årsskifte, og samtidig angir undersøkelsestidspunktet. Imidlertid håper jeg at Møglestue kan se nærmere på dette til neste møte i det rådgivende utvalg.

Referanser

- [1] Hoem, J.M. og Ljones, O.: "Problemer knyttet til spesifikasjon av alderen for enkelte aldersgrupper." Notat JMH/OLj/WTD, 3/6-71.
- [2] Hoem, J.M. og Ljones, O.: "Problemer knyttet til spesifikasjon av alderen for enkelte aldersgrupper." Notat JMH/OLj/GH, 2/12-71.
- [3] Hoem, J.M.: "Arbeidskraftundersøkelsene. Store og små problemer som trenger avklaring, eller som har fått det." Notat JMH/GH, 27/1-72. Avsnitt 4 og 5.
- [4] Hovind, S.: "Arbeidsgruppens merknader til notater fra Sosiodemografisk forskningsgruppe, JMH/GH, 27/1-72, OLj/San, 28/1-72." Notat SH/IH, 15/3-72. Avsnitt 3.
- [5] Hovind, S.: "Vurdering av resultatene fra arbeidskraftundersøkelsene." Notat SH/IH, 24/8-72.
- [6] Hovind, S. og Østerlund Petersen, S.: "Sammenlikning av resultater fra arbeidskraftundersøkelsene med annen statistikk." Notat SH/SØP/IH, 4/5-72.
- [7] Hovind, S. og Østerlund Petersen, S.: "Drøfting av hvilket aldersbegrep som skal legges til grunn." Notat SH/SØP/IH, 16/6-72.
- [8] Møglestue, I.: "Aldersbegrepet i arbeidskraftundersøkelsene." Notat IM/IH, 25/8-72.
- [9] Østerlund Petersen, S. og Hovind, S.: "Oppblåsing av tallene." Notat SØP/SH/IH, 14/3-72.
- [10] Østerlund Petersen, S.: "Hvordan skal 75-åringer behandles i tabellverket?" Notat SØP/IH, 2/5-72.
- [11] Referater fra møter i Det rådgivende utvalg for arbeidskraft- og sysselsettingsstatistikk.

Møtedato	Referatidentifikasjon	Referatavsnitt
25/6-71	SH/IH, 1/7-71	7
22/3-72	SH/SØP/WA, 27/3-72	3, 6
10/5-72	PO/SH/IH, 16/5-72	2
21/6-72	PO/SH/IH, 29/6-72	1, 2
7/9-72	PO/IH, 7/9-72	1

Tabell 1.

Utdrag fra tabellvedlegget (SH/IH, 24/8-72) til sakspapirene til møte 4/9-72 i Det rådgivende utvalg for arbeidskraft- og sysselsettingsstatistikk.

Estimerte antall 1 000 personer

	*)	Arbeidskraftsundersøkelse				Kilde
		III	IV	I	II	
		1971	1971	1972	1972	
Skoleelever, studenter	U	177	182	249	204	Tabell 1
	E	166	178	199	173	
Personer i arbeidsstyrken, 16-19 år	U	118	111	78	104	Tabell 2
	E	122	112	102	120	
Sysselsatte, 16-19 år	U	110	106	73	92	Tabell 6
	E	114	107	97	108	
Personer i arbeidsstyrken, 70-74 år	U	24	21	29	25	Tabell 2
	E	23	20	23	20	
Arbeidsaktive ugifte kvinner	U			300	297	Tabell 1B
	E			275	280	
Ikke arbeidsaktive menn	U			143	142	Tabell 1B
	E			155	151	
Arbeidsaktive menn ikke i arbeidsstyrken	U			167	146	Tabell 1B
	E			141	129	

*) U = alder ved utgangen av kalenderåret.
E = oppnådd alder.