

Arbeidsnotater

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

IO 76/22

8. juli 1976

UTVALGSREGISTRE OG RUTINER FOR TREKKING OG KLARGJØRING AV UTVALG VED UNDERAVDELINGEN FOR INTERVJUUNDERSØKELSER

av

Hans Viggo Sæbø

INNHold

	Side
1. Innledning	1
2. Byråets utvalgsregistre	1
2.1. Oppretting og ajourføring	1
2.2. Organisering av registrene	2
2.3. Oppretting av delregistre	5
2.4. Problemer ved å trekke utvalg fra et register	6
3. Rutine for trekking av utvalg	6
3.1. Beskrivelse av rutinen	6
3.2. Brukerinstruks	9
4. Rutiner for kontroll og klargjøring av utvalg	10
5. Spesielle rutiner for arbeidskraftundersøkelsene	13

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.

1. Innledning

I Byråets utvalgsplan foregår trekking av personer og husholdninger i to trinn. I første trinn trekkes 102 områder (primære utvalgseenheter), mens annet trinn består i å trekke familier eller personer fra disse områdene. De primære utvalgseenhetene holdes fast ved Byråets utvalgsundersøkelser, og et utvalgsregister omfatter derfor bare personer bosatt i disse områdene.

Byråets nye utvalgsplan ble tatt i bruk fra årsskiftet 1974/75. For en nærmere beskrivelse av utvalgsplanen henvises til Thomsen og Rideng (1974, arbeidsnotat IO 74/25). Tidligere hadde vi et register bestående av hullkort, med et kort for hver person bosatt i utvalgsområdene. Dette ble forandret ved overgangen til ny utvalgsplan, og alle registre ligger nå på magnetbånd (tape). Rutinene som brukes i forbindelse med trekking av utvalg er dermed også blitt forandret. I 1975 har en arbeidet med å forenkle og forbedre de nye rutinene. Særlig gjelder dette selve trekkerutinen. Da Oslo Folkeregisters magnetbånd er bygd opp litt forskjellig fra registrene ellers i landet har vi et eget utvalgsregister for Oslo. Dette ble ajourført og lagt over på bånd i slutten av 1975, slik at vi nå kan bruke de samme rutiner på dette som på registret for resten av landet.

Etter innkjøringsperioden ser nå alle rutinene ut til å virke tilfredsstillende. Hensikten med dette notatet er å beskrive utvalgsregistrene og rutinene som brukes ved trekking og klargjøring av utvalg.

I avsnitt 2 vil en ta for seg hvordan registrene dannes og ajourføres. Det blir også beskrevet hvordan de er oppbygd og sortert, og noen metodiske problemer ved å trekke utvalg fra et register blir behandlet. Avsnitt 3 inneholder en beskrivelse av og brukerinstruks for trekkerutinen, mens en i avsnitt 4 går gjennom andre maskinelle rutiner for kontroll og klargjøring av et trukket utvalg. Til sist går en gjennom Byråets arbeidskraftundersøkelser med de spesielle rutiner som inngår her. Når det gjelder beskrivelsen av registrene, vil en skille mellom register for Oslo og register for landet utenom Oslo.

2. Byråets utvalgsregistre

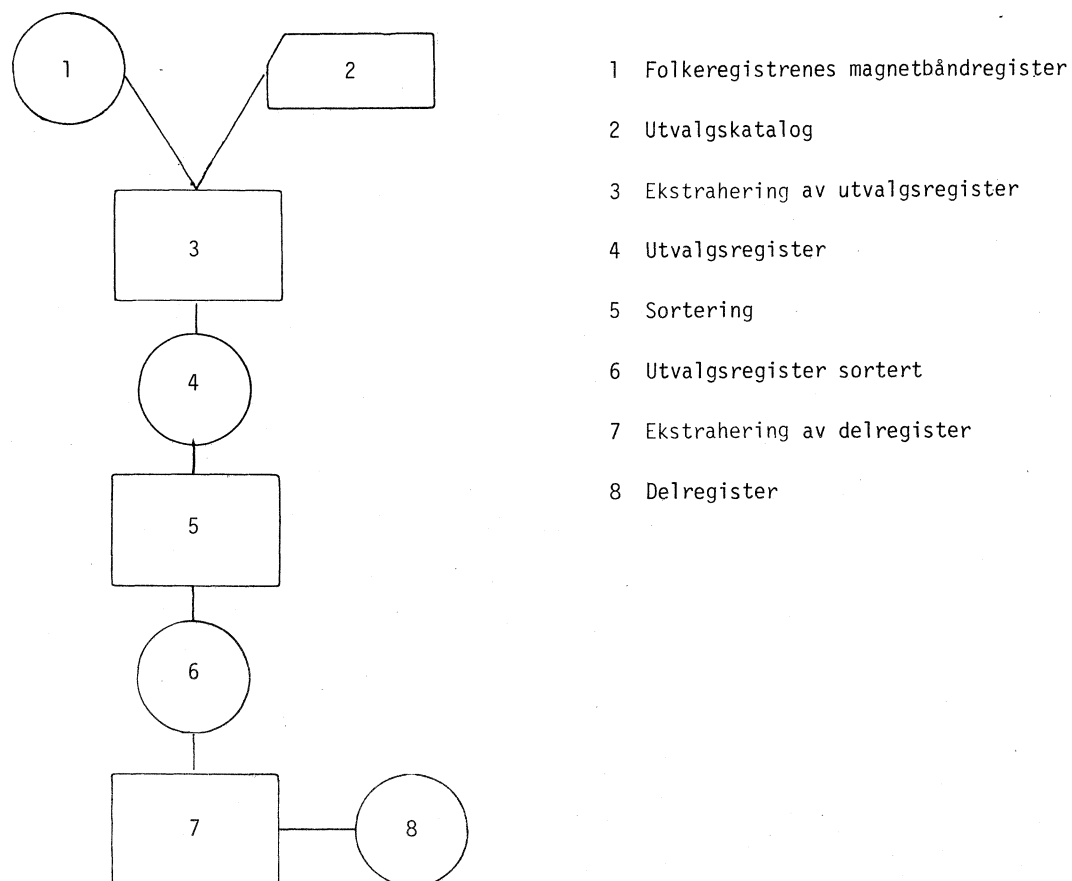
2.1. Oppretting og ajourføring

Grunnlaget for Byråets utvalgsregistre er folkeregistrenes magnetbåndregister. Dette inneholder alle familier og personer bosatt i Norge, og det blir ajourført etter oppgaver fra de lokale folkeregistre.

I den nye utvalgsplanen består hver primær utvalgseenhet av en kommune, eller i noen få tilfeller av to eller flere mindre kommuner. Hvert slikt område er delt opp i flere utvalgsområder. Hovedhensikten med dette er å kunne fordele intervjuobjektene (IO) på intervjuerne. I Stavanger, Kristiansand og Bærum er det 6 utvalgsområder, i Trondheim 9, i Bergen 18 og i Oslo 42, mens det i resten av de uttrukne kommunene er 3. Ved mindre undersøkelser bruker vi bare 1/3 av utvalgsområdene (ett sett), slik at vi i de fleste uttrukne kommuner unntatt de største byene bruker én intervjuer. Oppdelingen i utvalgsområder innen kommunene er foretatt etter valgkretsgrenser. I Oslo bruker vi sone- og rodegrensener ved denne oppdelingen.

Ekstrahering av et utvalgsregister for landet utenom Oslo fra folkeregistrenes magnetbånd foregår etter en katalog av parameterkort. Denne angir hvilke kommuner og valgkretser som hvert utvalgsområde består av. Figur 1 er et skjematisk oppsett over prosedyren ved opprettingen av dette registret. Den videre framgangsmåte med uttrekking av et delregister er også vist her. Denne prosedyren blir beskrevet i 2.3.

Figur 1. Rutiner for oppretting av utvalgsregister for landet utenom Oslo



Ajourføring foregår simpelthen ved at vi oppretter et nytt register med ajourførte data fra folkeregistrene som grunnlag. De uttrukne primære utvalgseenhetene holdes fast ved ajourføring av registret. Det kan imidlertid være aktuelt å endre den videre oppdeling av disse. Formålet med en justering av "tredelingen" er først og fremst å redusere feltkostnadene, men slike endringer kan også være nødvendige for å redusere utvalgsvariansen når vi trekker fra bare et sett. Vi må i alle tilfelle gå gjennom katalogen hver gang vi oppretter et nytt register, da valgkretsgrensene kan være endret. Nye valgkretser kan også ha kommet til. Byrådet får ajourførte bånd fra folkeregistrene i begynnelsen av hvert år. Det er derfor meningen å ajourføre registret hvert år i mars/april.

Oslo Folkeregisters bånd inneholder ikke familienummer. Det gjør imidlertid det sentrale personregister som finnes her i Byrådet. Dette har til gjengjeld ikke bydelskoder (koder for sone og rode), og vi må kombinere disse to registre for å få et familierregister med oppdeling i utvalgsområder innen Oslo. En tar da utgangspunkt i folkeregisteret. Første trinn består i å anvende en katalog som angir bydelskodene for hvert utvalgsområde. Deretter bruker vi en spørrerutine på det sentrale personregister for å få trukket ut familienumrene. Resultatet blir et utvalgsregister med alle nødvendige koder for familier og personer i Oslo. Den videre prosedyre med sortering av registret og ekstrahering av et delregister følger oppsettet i figur 1.

2.2. Organisering av registrene

Utvalgsregistrene inneholder en record for hver person bosatt i utvalgsområdene. Recordens innhold og redigering er angitt for landet og Oslo.

Recordbeskrivelse for register for landet utenom Oslo (pr. 1976).

Feltnr.	Posisjon	Feltbeskrivelse
1	1-3	Utvalgsmrådenr.
2	4-7	Kommunenr.
3	8-18	Fødselsnr.
4	19-41	Navn
5	42-62	Mellomadresse
6	63-66	Postnr.
7	67	Kode for matrikelnr.
8	68-72	Gatekode/gårdsnr.
9	73-76	Hus/bruksnr.
10	77	Bokstav til husnr.
11	78-80	Undernr.
12	81-82	Skolekrets
13	83-84	Valgkrets
14	85-95	Familienr.
15	96	Personkode
16	97	Ekteskapelig status
17	98	Statsborgerskap
18	99-100	Krets i rekkefølgenr.
19	101-116	Poststedets navn
20	117-118	Deltagelse i tidligere undersøkelser [blank]
21	119	Puljenr. [blank]
22	120	[Blank]

Recordbeskrivelse for register for Oslo

Feltnr.	Posisjon	Feltbeskrivelse
1	1-3	Utvalgsområdenr.
2	4-7	Kommunenr. [0301]
3	8-18	Fødselsnr.
4	19-43	Navn
5	44-73	Mellomadresse
6	74-77	Postnr.
7	78	Kode for matrikelnr.
8	79-83	Gatekode/gårdsnr.
9	84-87	Hus/bruksnr.
10	88	Bokstav til husnr.
11	89-91	Undernr. [000]
12	92-93	Sone-kode
13	94-95	Rode-kode
14	96-106	Familiennr.
15	107	Personkode
16	108	Ekteskapelig status
17	109	Statsborgerskap
18	110-111	[Blank]
19	112-127	Poststedets navn
20	128-129	Deltagelse i tidligere undersøkelser [blank]
21	130	Puljenr. [blank]
22	131-132	[Blank]

Recordene for Oslo avviker fra recordene for landet ellers ved at det er avsatt 2 ekstra posisjoner til navn og 9 ekstra posisjoner til mellomadresse. Dessuten erstatter bydelkodene i felt 12 og 13 kretskodene.

Redigeringen av recordene holdes fast ved alle sorteringer, ekstrahering av delregister og trekking av utvalg. De blanke feltene 20 og 21 blir benyttet under trekking til koder for deltagelse i tidligere undersøkelser og eventuelt puljenummer (se avsnitt 3).

Registret (Oslo medregnet) sorteres etter utvalgsområdenummer, familienummer og personkode. Utvalgsområdene utenfor de største byene er nummerert fra 101-199, 201-299 og 301-399 slik at de 3 delene i hver primære utvalgsenhet har de to siste sifrene felles. I Oslo er utvalgsområdene nummerert fra 401-442. Bergen, Trondheim, Stavanger, Kristiansand og Bærum har nummer hvor første siffer er henholdsvis 5, 6, 7, 8 og 9. Den videre nummereringen av områdene her går fram av oppstillingen nedenfor.

I registret har vi 3 sett av utvalgsområder. Fordelingen av områdene på sett i landet utenom Oslo er som følger.

Sett I: 101-199, 511-516, 611-613, 711, 712, 811, 812, 911, 912.

Sett II: 201-299, 521-526, 621-623, 721, 722, 821, 822, 921, 922.

Sett III: 301-399, 531-536, 631-633, 731, 732, 831, 832, 931, 932.

Første siffer i områdenummeret angir altså sett dersom dette er 1, 2 eller 3. Dersom dette sifferet er 5, 6, 7, 8 eller 9 er det neste siffer som angir sett. I Oslo består sett 1 av områdene 401-414, sett 2 av 415-428 og sett 3 av 429-442. Ved mindre undersøkelser er det aktuelt å trekke utvalg fra et av settene.

Fødselsnummeret gir oss alder og kjønn. Familienummeret er fødselsnummeret til familiens hovedperson. Sorteringen foregår etter alderen på denne. Personkoden er 1 for hovedperson, 2 for ektefelle og 3 for barn. En familie i folkeregistret, (familieenhet) - består aldri av mer enn to generasjoner. Barn som flytter hjemmefra blir tildelt eget familienummer. Dette beholdes selv om de senere flytter tilbake til foreldrene. En husholdning kan ofte bestå av flere familier i registret, og dette er et problem ved husholdningsundersøkelser. I praksis løses dette ved at vi trekker familier, men gir intervjuerne instruks om å ta med alle husholdningsmedlemmer.

Koder for ekteskabelig status er som følger:

Kode 1 - ugift

" 2 - gift

" 3 - enke/enkemann

" 4 - skilt

" 5 - separert

Fødselsnummer gir oss mulighet til å trekke utvalg av personer innen en viss aldersgruppe og av samme kjønn. Kombinerer vi dette med kode for ekteskabelig status kan vi f.eks. trekke utvalg av gifte kvinner.

Kode for statsborgerskap er 1 for norske statsborgere og 2 for utlendinger. Kode 3 står for statsløse og kode 4 for norsk statsborger uten fulle rettigheter.

2.3. Oppretting av delregistre

Registret for landet utenom Oslo består av ca. 2 millioner records, mens vi i Oslo har med ca 450 000. For å trekke et utvalg må en gjennomløpe hele registret. Dette blir dyrt med så store registre, og vi har derfor opprettet et mindre delregister fra hvert register. Trekking av utvalg til undersøkelser foregår fra disse.

Vi sørger helst for at utvalg til en intervjuundersøkelse er selvveiende, slik at hver enhet får samme "oppblåsningsfaktor". I Byråets utvalgsplan er de primære utvalgsenhetene og dermed utvalgsområdene trukket med ulike sannsynligheter. For at det endelige utvalg skal bli selvveiende, må vi derfor trekke familier eller personer med ulik sannsynlighet i de forskjellige utvalgsområdene. Vi tar imidlertid hensyn til dette alt ved uttrekkingen av delregistrene, slik at disse består av et selvveiende utvalg av 10% av befolkningen. I områder som er trukket med sannsynlighet 1 i første trinn (gjelder f.eks. Oslo) trekker vi ut 10%, mens andelen fra andre områder er større. Delregistret vårt fra landet utenom Oslo inneholder fra 10% av befolkningen i de største byene til nær 100% i mindre

utvalgsområder. I alt består det av ca. 350 000 records, mens delregistret for Oslo har med 45 000. Vi må trekke ut nye delregistre ved hver ajourføring av hovedregistrene. I Oslo opprettes delregistret ved å trekke ut hver 10. familie systematisk, mens ekstraheringen ellers i landet styres av parameterkort hvor en angir hvor stor del av familiene som skal trekkes ut fra hvert utvalgsområde.

2.4. Problemer ved å trekke utvalg fra et register

Det er et problem at de enheter som registrene inneholder (person og familieenhet) ikke alltid faller sammen med de enheter vi ønsker å trekke utvalg av. I 2.2 har vi sett hvordan vi ved husholdningsundersøkelser må trekke utvalg av familieenheter, og så la intervjuerne supplere dette med eventuelle andre husholdningsmedlemmer. En slik framgangsmåte gir medlemmer av store husholdninger større trekkmulighet enn andre. Dersom mange husholdninger består av flere familieenheter fører dette til merarbeid både i felten og ved intervjukontoret ved at utvalget øker. Det er derfor også behov for en instruks som avgrensner husholdninger hvor alle skal intervjues fra husholdninger der en bare intervjuer uttrukket familieenhet.

Denne instruksjonen kan formuleres slik: En uttrukket familieenhet kan høyst trekke med seg 4 andre familieenheter til intervju. Er husholdningen større, skal bare den uttrukne familieenhet intervjues.

I praksis oppstår problemene med store husholdninger helst ved trekking av en pasient på pleie- eller sykehjem. Alle andre pasienter på vedkommende rom skal da intervjues hvis det totale antallet på rommet er 5 eller mindre. På rom med 6 eller flere pasienter skal bare den uttrukne intervjues.

Det vanligste problemet som oppstår ved alle undersøkelser er at uttrukne enheter er flyttet eller døde. Dette kan skje fordi registret aldri er helt ajour. Hovedregelen er at en i et slikt tilfelle intervjuer den person eller husholdning som er flyttet inn i samme leilighet/hus. Derved får også nyinnflyttede samme sannsynlighet for å delta i en undersøkelse som andre.

I undersøkelser med roterende utvalg som arbeidskraftundersøkelsene følges vanlige regler ved første gangs intervju. Folk som flytter i løpet av den perioden de er med i en slik undersøkelse skal imidlertid følges opp og intervjues på ny adresse.

For behandling av problemene med å trekke utvalg fra et ikke ajourført register henvises ellers til Thomsen (1971, arbeidsnotat IO 71/6).

3. Rutine for trekking av utvalg

3.1. Beskrivelse av rutinen

Rutinene for trekking og videre behandling av utvalgene er i prinsippet like for utvalg fra Oslo og for utvalg fra landet forøvrig. Den eneste forskjell i registrene er at feltene ikke er helt i overensstemmelse. I det følgende vil en derfor bare omtale registret for landet utenom Oslo.

Utgangspunktet for trekkerutinen er delregistret. Da dette er selvsveiende skal alle enheter ha samme sannsynlighet for å bli trukket ut. Dette kan oppnås ved systematisk trekking, og rutinen virker slik at det fra et vilkårlig startpunkt trekkes ut 10-er med faste mellomrom i registret. For å regne ut trekkesannsynligheten må vi imidlertid kjenne antall aktuelle trekkenheter i delregistret. Første gang en skal trekke et utvalg fra et nytt register må en derfor gjennomløpe dette og telle opp innholdet. Første trinn i trekkerutinen består av en slik opptelling. Familiene grupperes etter sett og personene etter sett, alder, kjønn og ekteskapeleg status. Resultatet av opptellingen skrives ut i en tabell over såkalte basistall. Denne består egentlig av 4 tabeller, en for hvert sett og en for hele delregistret.

Figur 2 er et eksempel på en tabell over basistall. Sett 4 står for hele delregistret. I forspalten står 5-årige aldersgrupper slik at gruppe 18 svarer til personer over 85 år. For alle grupper er det angitt et bruttotall og et nettotall. Bruttotallet er antallet i delregistret, mens nettotallet er antallet som ikke har vært med i undersøkelser før. Disse tallene er like første gang et register tas i bruk.

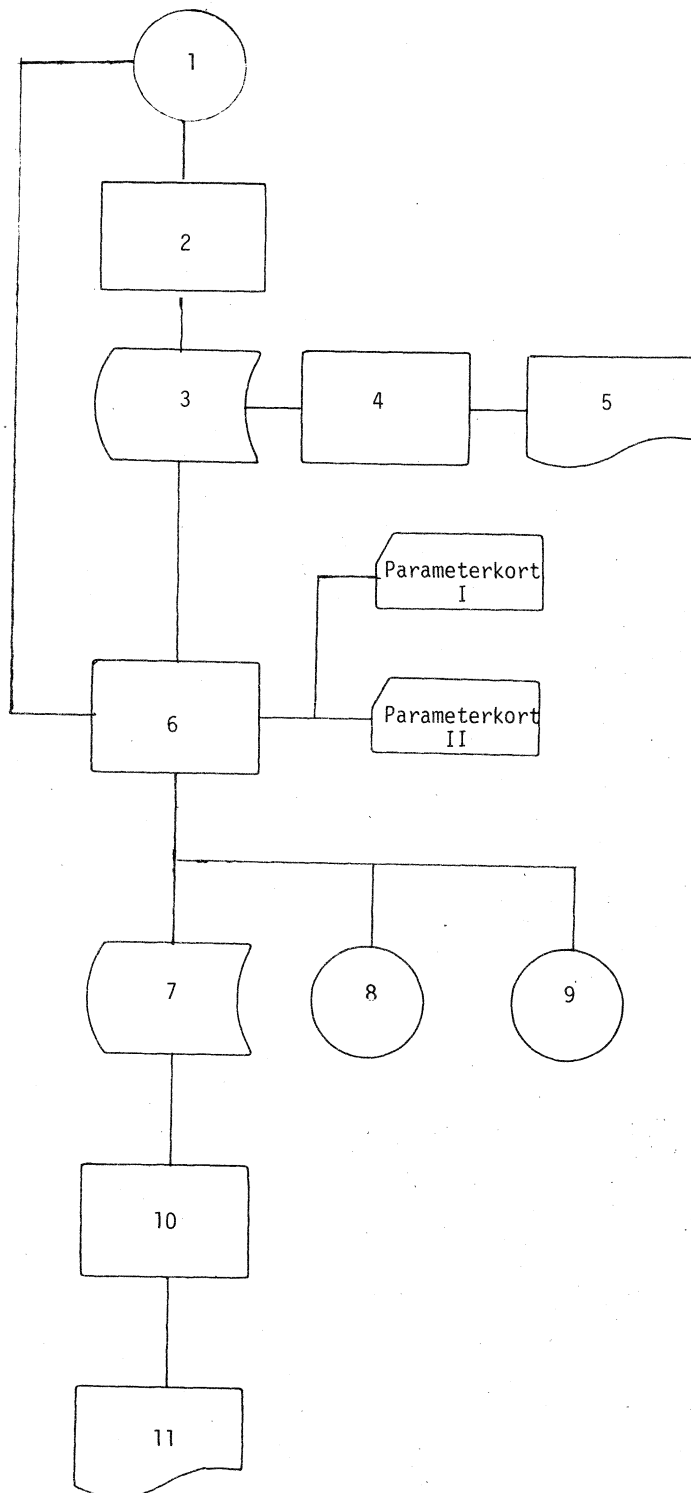
Med hjelp av denne tabellen kan vi regne ut trekkesannsynligheten enten trekkenheten er familie, person, gifte menn mellom 30 og 40 år etc.. Dette gjelder også hvis vi trekker fra bare ett av de tre settene. Når et utvalg trekkes ajourføres disse tabellene automatisk (nettotallene forandres) og skrives ut. Derved slipper en å gjennomløpe registret en ekstra gang før hver trekking.

Figur 3 er et skjematisk oppsett over trekkerutinen med danning av basistall, trekking av utvalg og utlistering av nye basistall. Oppsettet av parameterkortene blir gjennomgått i neste avsnitt.

Figur 2. Tabell over basistall (delregister for landet utenom Oslo), sett 4

Aldersgruppe	Familier		Personer total		Menn total		Menn gifte		Kvinner total		Kvinner gifte		Kvinner ugifte			
	Brutto	Netto	Brutto	Netto	Brutto	Netto	Brutto	Netto	Brutto	Netto	Brutto	Netto	Brutto	Netto		
01	49	41	18 092	15 094	9 291	7 758	5	4	9 286	7 754	8 801	7 336	3	2	8 798	7 334
02	130	104	28 753	23 836	14 733	12 219			14 733	12 219	14 020	11 617	5	5	14 015	11 612
03	207	175	27 217	22 343	13 977	11 512	2	1	13 975	11 511	13 240	10 831	5	4	13 235	10 827
04	637	504	26 979	20 661	13 692	10 482	6	6	13 686	10 476	13 287	10 179	137	118	13 150	10 061
05	5 321	4 070	25 970	19 802	13 440	10 209	1 688	1 310	11 752	8 899	12 530	9 593	4 195	3 255	8 335	6 338
06	11 202	8 607	25 856	19 823	13 362	10 241	7 635	5 881	5 727	4 360	12 494	9 582	9 597	7 389	2 897	2 193
07	11 051	8 411	21 776	16 610	11 241	8 580	8 830	6 728	2 411	1 852	10 535	8 030	9 168	6 984	1 367	1 046
08	8 959	6 838	17 335	13 245	8 721	6 639	7 307	5 557	1 414	1 082	8 614	6 606	7 589	5 792	1 025	814
09	8 646	6 537	16 627	12 681	8 414	6 415	7 106	5 394	1 308	1 021	8 213	6 266	7 362	5 645	851	621
10	9 951	7 555	18 382	13 973	9 334	7 084	7 852	5 934	1 482	1 150	9 048	6 889	7 826	5 939	1 222	950
11	11 576	8 805	20 610	15 668	10 360	7 905	8 668	6 585	1 692	1 320	10 250	7 763	8 507	6 434	1 743	1 329
12	12 011	9 139	20 351	15 608	10 070	7 651	8 358	6 339	1 712	1 312	10 281	7 957	8 010	6 204	2 271	1 573
13	11 831	8 991	18 514	14 119	9 103	6 936	7 451	5 664	1 652	1 272	9 411	7 183	6 564	5 013	2 847	2 170
14	11 134	8 516	16 096	12 369	7 608	5 843	6 018	4 607	1 590	1 236	8 488	6 526	4 925	3 827	3 563	2 699
15	9 805	7 564	13 113	10 215	5 816	4 452	4 307	3 325	1 509	1 127	7 297	5 763	3 307	2 648	3 990	3 115
16	8 167	6 679	9 950	8 132	4 310	3 547	2 826	2 349	1 484	1 198	5 640	4 585	1 807	1 474	3 833	3 111
17	5 369	4 404	6 115	5 024	2 511	2 080	1 385	1 155	1 126	925	3 604	2 944	758	626	2 846	2 318
18	3 867	3 106	4 100	3 293	1 554	1 264	547	440	1 007	824	2 546	2 029	241	191	2 305	1 838
Total	129 913	100 046	335 836	262 496	167 537	130 817	79 991	61 279	87 546	69 538	168 299	131 679	80 006	61 550	88 293	70 129

Figur 3. Rutine for trekking av utvalg



1. Selvveiende delregister
2. Program for danning av basistall
3. Basistall
4. Program for utlisting av basistall
5. Tabell over basistall
6. Trekkeprogram
7. Nye basistall
8. Nytt delregister
9. Trukket utvalg
10. Program for utlisting av nye basistall
11. Tabell over nye basistall

Prosedyren 2-5 er bare nødvendig ved første undersøkelse fra et nytt delregister

3.2. Brukerinstruks

Parameterkort I inneholder 10 vilkårlige siffer. Disse brukes i rutinen for å danne et vilkårlig startpunkt for trekkingen. Dette kortet fylles ut på driftskontoret.

Parameterkort II inneholder data som spesifiserer hvilke enheter som har trekkmulighet og trekkesansynligheten for disse enhetene. Oppsettet er vist i figur 4.

Figur 4. Parameterkort II

Kolonne	Lengde	Feltbeskrivelse
1	1	Sett
2- 5	4	Fødselsår fom/tom (alderseleksjon)
6-10	5	Ekt. status
11	1	Kjønn
12	1	Husholdning/person
13-14	2	Undersøkelsesnr.
15	1	Puljenr.
16	1	Merking/ikke merking
17-19	3	Trekketetthet

Koden for sett er 1, 2, 3 eller blank dersom utvalget skal trekkes fra alle tre sett.

Trekkenhetenes alder angis med 4 siffer slik at de to første gir fødselsår for de eldste og de to siste fødselsår for de yngste. Ønsker vi f.eks. personer mellom 16 og 74 år som i arbeidskraftundersøkelsene, blir de to første siffer 02 og de to siste 60 (undersøkelse i 1976).

Vi kan trekke personer med bestemt ekteskapielig status ved å fylle ut en eller flere av kolonnene som er avsatt til dette. Ønsker vi f.eks. å trekke blant gifte og separerte setter vi kode 2 i det andre feltet og 5 i det siste (se kodeliste i 2.2). Dersom vi ikke skal selekttere på ekteskapielig status trenger vi ikke å punche noen av de 5 kolonnene for dette.

Kjønn angis med 1 for menn, 2 for kvinner eller blank for menn og kvinner.

Gyldige koder for husholdning/person er henholdsvis H og P. Står det H i dette feltet har bare familiene representert ved hovedpersonen (personkode 1) trekkmulighet. Dette feltet skal alltid fylles ut. Det samme gjelder feltet for undersøkelsesnummer.

I noen undersøkelser ønsker vi å dele opp et trukket utvalg i puljer, slik at utvalget også sorteres etter pulje innen hvert utvalgsområde. Dette oppnås ved å angi antall puljer i parameterkortet.

Det er uheldig å trekke ut samme person eller familie til flere undersøkelser, og vi merker derfor uttrukne IO-er i registret ved å sette kode 1 på parameterkortet. Det er bare umerkede enheter som har trekkesansynlighet, og det er antallet slike som angis som nettotall i tabellen over basistall. Dersom en undersøkelse bare omfatter en bestemt gruppe personer kan det komme på tale å ikke merke de uttrukne. Hvis merking foretas vil denne gruppa bli underrepresentert blant enheter med trekkmulighet i registret, og dermed også i senere utvalg. Ungdomsundersøkelsen 1975 er et godt eksempel på en undersøkelse hvor vi ikke merket de uttrukne. Det ble her trukket ca. 9 000 personer i alderen 17-19 år, eller halvparten av alle personer i denne aldersgruppa i delregistret.

Ved hjelp av tabellen over netto basistall er det lett å finne antall enheter med trekkmulighet i delregistret. Trekketettheten som skal angis til slutt på parameterkortet finnes ved å dividere dette antallet på ønsket utvalgsstørrelse. Anta at vi f.eks. skal trekke et utvalg på 6 000 familier blant 100 000 med hovedperson 16-74 år i delregistret. Trekketettheten blir $100\ 000/6\ 000 = 16.7$, og det er disse tre sifrene som punches på parameterkortet. Trekkerutinen virker slik at gjennomsnittlig hver 16.7-ende familie med trekkmulighet blir trukket ut.

4. Rutiner for kontroll og klargjøring av utvalg

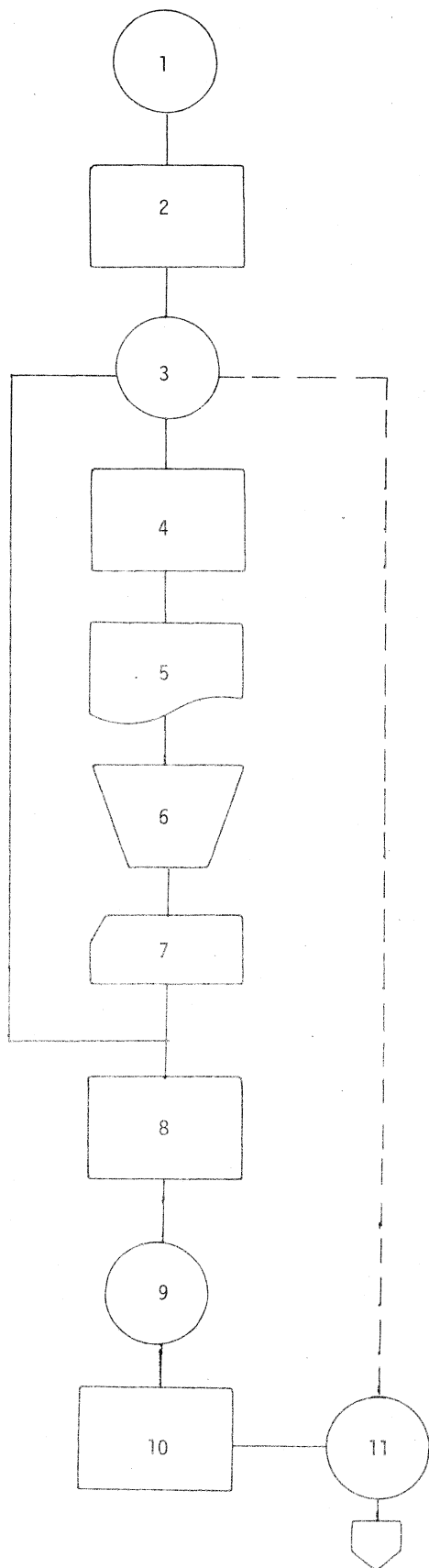
I behandlingen av et trukket utvalg inngår diverse sorteringer, en korreksjonsrutine og utskrivning av slipper og lister til intervjuerne. Dersom det er foretatt puljeinndeling under trekkingen, sorteres først utvalget etter pulje innen hvert område. Sorteringen etter familienummer og personkode holdes fast innen hver pulje.

Første trinn i korreksjonsrutinen er utskrivning av en opprettingsliste. Denne inneholder alle familiene eller personene som er trukket med alle personlige data fra registret. Lista bearbeides på fagkontoret (som regel gruppe for metoder), og eventuelle korreksjoner blir punchet. Oppretting på utvalgsbåndet foregår i et korreksjonsprogram (KOKS-KORREK). Korreksjonsrutinen vil bli nærmere beskrevet i avsnittet om arbeidskraftundersøkelsene.

I de fleste undersøkelsene er det ikke nødvendig med korreksjoner til et trukket utvalg, og en kan kjøre ut lister og slipper direkte. Figur 5 er et skjematisk oppsett over hele prosedyren med klargjøring og eventuell kontroll av et utvalg. Rutinene er i prinsippet like for person- og husholdningsundersøkelser med unntak av at vi kan sette inn et parameterkort i utlistningsprogrammet for familier. Dette benyttes dersom vi ønsker at bare hovedpersonene (personkode 1) skal listes ut. I posisjon 1-18 på kortet skrives det i så fall "BARE HOVEDPERSONER". Kortet utelates hvis alle familiemedlemmer skal listes.

Sorteringen på listene er vanligvis som sorteringen i registret og på utvalgsbåndet. Vi får en egen liste for hvert utvalgsområde (ev. pulje). Familiene eller personene er her sortert etter familienummer og personkode. I familienummeret sorteres det først etter hovedpersonens [fødselsår x måned x dag], slik at de "eldste" familiene kommer først. I enkelte undersøkelser ønsker vi å foreta en manuell "puljeinndeling" på listene. Et eksempel er forbruksundersøkelsene hvor vi fordeler husholdningene på regnskapsuker. For å lette denne fordelingen kan vi bruke et program for alternativ sortering, slik at rekkefølgen mellom husholdningene på listene blir tilfeldig. Dette programmet sorterer utvalget etter familienummerets [fødselsdag x måned x år], og bestilles av metodegruppa etter behov.

Figur 5. Rutiner for kontroll og klargjøring av utvalg

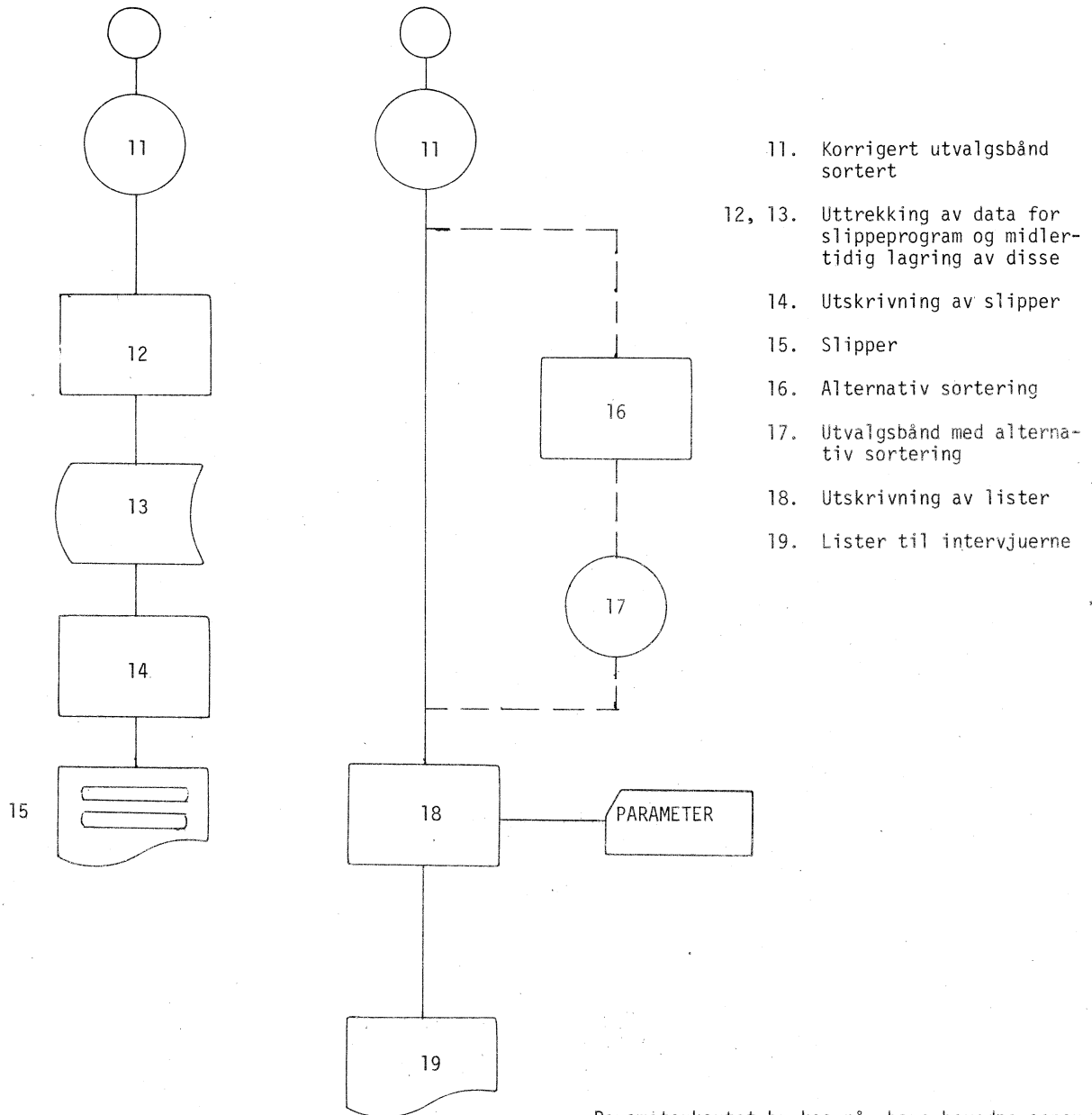


1. Trukket utvalg
2. Sortering etter pulje
3. Sortert utvalg
4. Program som redigerer utvalget med hensyn på oppretting
5. Opprettingsliste
6. Bearbeiding på fagkontor (metodegruppa)
7. Korreksjonskort
8. Korreksjonsprogrammet
KOKS KORREK
9. Korrigert bånd
10. Sortering
11. Korrigert utvalgsbånd sortert

Når vi ikke kjører korreksjonsrunde går rutinen rett fra 3 til 11.

Neste side

Forts. fra forrige side



Parameterkortet brukes når bare hovedpersoner skal listes ut i en husholdningsundersøkelse.

5. Spesielle rutiner for arbeidskraftundersøkelsene

Byråets arbeidskraftundersøkelser (AKU) foregår hvert kvartal med et utvalg på ca. 6 000 husholdninger. Utvalget består av 4 puljer som skiftes ut hvert kvartal i følge en rotasjonsplan. Denne virker slik at 2 puljer er felles i to etterfølgende undersøkelser, mens en pulje alltid er ny. Hver pulje deltar i alt i 4 undersøkelser, og det går 5 kvartaler mellom første og siste gang en pulje er med.

Vi trekker 4 nye AKU-puljer hvert år (7 første år). En av disse skal med som ny pulje i hvert kvartal. Rutinen for trekking og utskrivning av opprettingsliste ("recordnummerliste") og slipper er som for en vanlig husholdningsundersøkelse. Eneste unntak er at vi kan foreta omkodning av pulje-numrene. Dette må gjøres for å unngå at to puljer som skal være med i utvalget samme kvartal har samme nummer. Som eksempel kan en nevne at det i 1975 var med puljer nummerert fra 1 til 7. Av disse er puljene 3-7 med en eller flere ganger i 1976. De 4 nye puljene som ble trukket i 1976 ble derfor omkodet til nummer 1, 2, 8 og 9. Deretter ble pulje 1 og 2 på AKU-båndet for 1975 fjernet, og dette båndet ble koblet sammen med det nye utvalgsbåndet. Utvalgsbåndet for AKU 1976 inneholder slik alle puljer som skal være med dette året.

Første gang en pulje er med i AKU blir utvalget korrigert av intervjuerne. Der hvor uttrukket familie er flyttet, blir de som er flyttet inn i samme leilighet/hus intervjuet isteden. I tillegg blir alle personer/familier som bor sammen med uttrukne familier tatt med, slik at det korrigerede utvalg består av komplette husholdninger. Når listene kommer tilbake til Byrådet og innsjekkingen er ferdig, blir utvalget også rettet opp maskinelt. Dette skjer ved at alle rettelsene blir punchet og kjørt inn på utvalgsbåndet i korreksjonsprogrammet KOKS-KORREK (se avsnitt 4). En opererer her med tre typer opprettinger: Tilganger, avganger og korreksjoner (parametertype T, A og K). Det må punches to tilgangskort for hver ny person for å få plass til alle feltene. Hvert avgangskort inneholder bare nummeret i følge opprettingslista på den som skal tas ut, mens hvert korreksjonskort angir recordnummer, startposisjon og ny verdi på de felter en ønsker å korrigere.

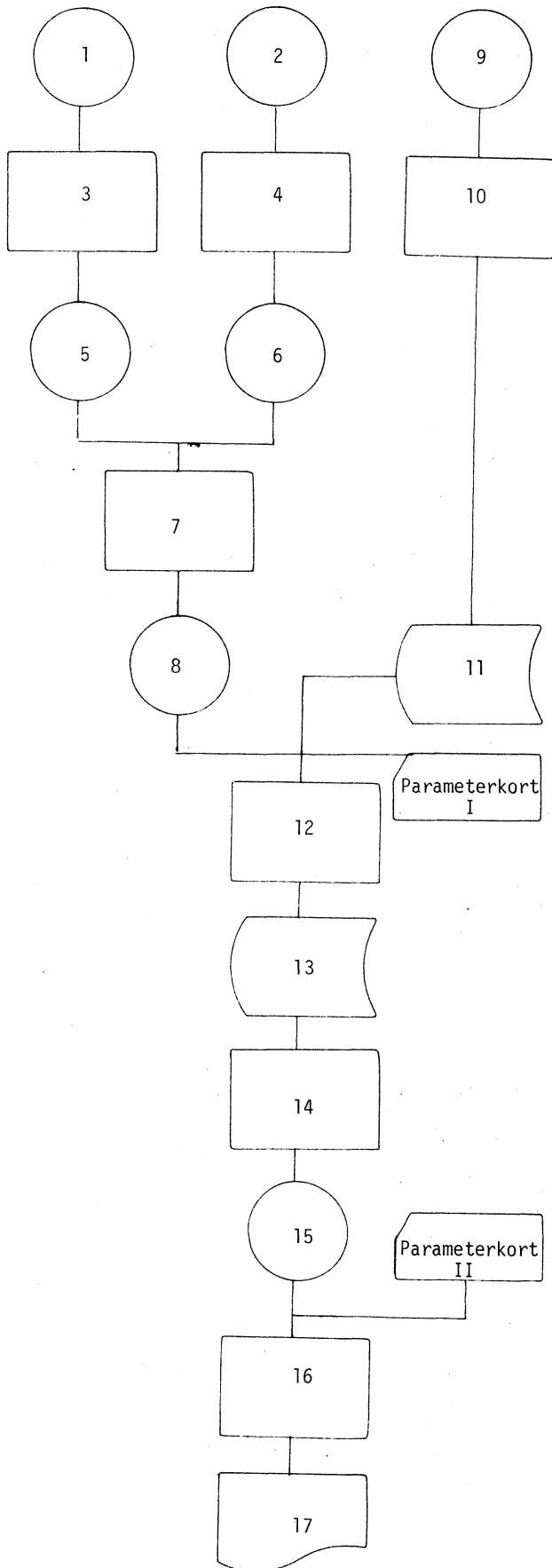
Før det kan skrives ut lister til intervjuerne i et kvartal, må oppretting av utvalget fra forrige kvartal være avsluttet. Den videre rutinen skiller seg litt fra rutinen som brukes ved en vanlig husholdningsundersøkelse. Dette skyldes at vi ønsker å angi på listene hvem som ikke har svart på spørsmål om utdanning tidligere. Dette gjelder foruten alle i den nye puljen også personer som har vært midlertidig fraværende ved tidligere intervju. Dessuten skal alle under 25 år spørres om utdanning hver gang.

Rutiner fra et korrigeret utvalgsbånd til utskrivning av lister for et kvartal er vist skjematisk på figur 6. I utvalget for et kvartal N er det med 2 puljer som også var med i kvartal (N-1) og en pulje som var med i kvartal (N-3). Ved å kjøre bearbeidingsbåndene fra kvartal (N-1) og kvartal (N-3) mot utvalgsbåndet får vi derfor sjekket hvem som ikke har svart på spørsmål om utdanning tidligere. Rutinen virker slik at alle disse samt alle i den nye puljen får koden UTD påført til høyre for persondataene på listene. Det samme gjelder alle under 25 år, men dette må angis på parameterkort I ved å skrive P i posisjon 1 og årstallet i posisjon 2 og 3. I 1976 skriver vi f.eks. P 76 her.

Det korrigerede utvalgsbåndet er som vanlig sortert etter [område x pulje x familienummer x personkode]. Dette må imidlertid omsorteres etter fødselsnummer ([fødselsdato x øvrige siffer]) for å kunne sjekke det mot bearbeidingsbåndene. Etter at utdanningskodene er påført sorterer vi utvalget på vanlig måte igjen, slik at vi får en liste for hver pulje innen hvert område.

På parameterkort II angis i posisjon 1-2 kjøreår, i posisjon 3 kvartal nr. og i posisjon 4-7 hvilke puljer som skal være med dette kvartalet (4 siffer å et nummer). I posisjon 8-11 skal det stå for hvilken gang de respektive puljer er med, slik at nummeret i posisjon 8 korresponderer med pulje-nummeret i posisjon 4 osv. Alle opplysninger på dette parameterkortet blir skrevet ut på listene.

Figur 6. Rutiner for påføring av utdanningskode og utlisting av AKU-utvalg



1. Bearbeidingsbånd for kvartal (N-1)
2. Bearbeidingsbånd for kvartal (N-3)
- 3, 4, 10. Sortering etter fødselsnummer
- 5, 6. Bearbeidingsbånd for kvartal (N-1) og (N-3) sortert
7. Program for å sette sammen bånd 5 og 6. ("merging")
8. Bearbeidingsbånd for kvartal (N-1) og (N-3) satt sammen
9. Korrigert utvalgsbånd sortert på [område x pulje x familienummer x personkode]
11. Utvalget sortert etter fødselsnummer
12. Sjekking av utvalget mot bearbeidingsbåndet
13. Utvalg med utdanningskoder
14. Sortering etter [område x pulje x familienummer x personkode]
15. Sortert utvalg
16. Utlistingsprogram
17. Lister til intervjuerne