

Arbeidsnotater

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

IO 68/22

Oslo, 26. september 1968

PERSONDATAREGISTRERING, -ARKIVERING OG -MODELLUTVIKLING

I STATISTISK SENTRALBYRÅ I NORGE^{x)}

Av Svein Nordbotten

I N N H O L D

1. Innledning
2. Persondataregistrering og -arkivering
3. Personmodeller
4. Henvisninger

x) Innlegg på Nordisk seminar om Arbeidskraft- og utdanningsprognoser, Saltsjöbaden, Stockholm, 11.-14. november 1968.

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 Statistisk Sentralbyrå i Norge har vi hittil lite å vise til når det gjelder utvikling og bruk av utdannings- og arbeidskraftmodeller. Når vi likevel er oppfordret til å bidra med et innlegg på dette seminaret, er det trolig fordi våre demografiske modeller vil være fundamentet å bygge utdannings- og arbeidskraftmodeller på og fordi våre dataarkiver etter hvert vil tillate oss å bygge slike modeller med aspekter som hittil ikke har vært mulig å ta hensyn til.

Den type av modeller som beskrives i dette innlegg, har vi kalt personmodeller. Med personmodell mener vi en matematisk formulert modell som tar sikte på å forklare sammensetning av og atferd og endring over tiden i en bestand av personer. Våre personmodeller er også karakterisert ved at de, ved tallfesting av parametrene, nyttiggjør individualdata fra våre persondataarkiver uten først å gå den tradisjonelle vei om en aggregering med tilhørende informasjonsreduksjon. Vanligvis vil våre personmodeller ta sikte på å forklare ikke-økonomiske forhold selv om økonomiske faktorer kan være viktige forklaringsvariable i modellen.

Vi har som mål å utvikle personmodeller som hver for seg kan bidra til å belyse forholdene på det spesielle felt de refererer seg til og til å stimulere videre metodeutvikling i forbindelse med modellanalyser på disse områder. Like viktig mener vi også at det er at hver modell kan betraktes som en integrert del av en større modellbygning som vi har kalt befolkningsmodellen. En gang i framtiden håper vi at denne vil kunne samkjøres med våre økonomiske modeller slik at vi i prognoseberegninger vil kunne ta vare på de viktigste samvirkningsforhold mellom økonomiske og ikke-økonomiske faktorer.

Når vi nå har tatt opp arbeidet med personmodeller i Byrået, har dette fire viktige årsaker. For det første har vi registrert et økende behov for å få belyst sannsynlige utviklingstendenser i befolkningen. For det andre har vår sentrale personregistrering ført til at vi nå etter hvert får kumulert verdifull informasjon i våre persondata-arkiver som vi ikke har råd til å la ligge unyttet. For det tredje har utviklingen av de moderne data-maskiner gitt oss det verktøy vi har behov for når en går over fra å bygge analysen på statistikk for grupper til direkte å utnytte de store masser av individualdata analytisk. For det fjerde har Byrået lyktes i å knytte til seg en mindre stab av kvalifiserte statistikere med interesse for å gjøre nybrottsarbeid på dette felt. Disse er organisert i en egen Gruppe for personmodeller.

2. Persondataregistrering og -arkivering

2.1 Fødselsnummer

Det som her kalles persondataregistrering omfatter bare individualdata identifisert med de norske fødselsnummer (1,16). Fødselsnummeret, som er et personlig identifikasjonsnummer, har følgende to viktige karakteristika:

1. Fødselsnummeret er en entydig personidentifikasjon på et hvilket som helst tidspunkt, dvs. det finnes i prinsippet aldri på noe tidspunkt to eller flere personer med samme fødselsnummer.
2. Fødselsnummeret er en permanent personidentifikasjon, dvs. fødselsnummeret følger personen uendret gjennom hele livet.

Statistiske systemer som systematisk knytter slike identifikasjonsnummer til data i registrering og arkiver med sikte på en høy fleksibilitet for å betjene bl.a. spesielle statistiske tabelleringer og analyser, er blitt kalt arkivstatistiske systemer (11, 12, 13). Etttersom en stor del av datagrunnet i personstatistikken oppstår som registreringer i administrative organer og prosesser, er det av avgjørende betydning at identifikasjonsnumrene også benyttes i disse administrative organer.

2.2 Eksisterende dataregistrering og -arkivering

Arbeidet med å tildele fødselsnummer og vedlikeholde et register over hele landets befolkning tok til i Byrået i 1963 med utgangspunkt i Folketellingen 1960 og den løpende folkeregistrering (17, 18). I dag registreres persondata med fødselsnummer i en rekke administrative prosesser. De fleste av disse data blir ved hjelp av fødselsnumrene organisert og lagret i våre persondata-arkiver.

Følgende tabell gir en oversikt over hvilke data vi nå har i vårt persondata-arkiv:

Innhold og registreringshyppighet i persondata-arkivet

Kjennemerke	Registreringshyppighet		
	løpende	årlig	10-årlig
1. Fødselsdata og -sted	x		
2. Bostedskommune	x		
3. Navn	x		
4. Adresse	x		
5. Ekteskapelig status	x		
6. Dødsdato og -sted	x		
7. Inn- og utvandringsdata	x		
8. Fødselsnummer for mor	x		
9. " " far	x		
10. " " ektefelle	x		
11. Inntektsdata		x (f.o.m. 67)	
12. Formuesdata		x (" ")	
13. Skattedata		x (" ")	
14. Examen artium		x (" 68)	
15. Familiedata			x (pr.1/11-60)
16. Utdanningsdata			x (" ")
17. Yrkesdata			x (" ")
18. Næringsdata			x (" ")
19. Boligdata			x (" ")

Det offisielle tidspunkt for oppstartning av denne registrering var 1. oktober 1964. Av den grunn mangler det enkelte kjennemerker for deler av befolkningen, f.eks. foreldrenes fødselsnummer for personer født før dette tidspunkt. Likeens vil vi heller ikke kjenne utdanning og yrke pr. 1. november 1960 for personer som har innvandret etter Folketellingen.

I tillegg til de ovenfor nevnte data, registreres det også andre, bl.a. i helsesektoren, som tilfredsstillende de ovenfor nevnte krav for å bli tatt med i persondata-arkivet, men som foreløpig holdes utenfor av praktiske grunner. Av de data som inngår, er vi lite tilfreds med at:

- 1) Utdanning blir registrert bare hvert 10-år.
- 2) Yrkesaktivitet og yrke blir registrert bare hvert 10-år.

Dessuten er vi ikke fornøyd med å få registrert familie-enheten bare ved hver folketelling. Data om den samarbeidende familie vil være av stor

betydning for en rekke analyser av f.eks. flyttinger og dermed etterspørselen etter utdanning på forskjellige steder og for arbeidskraft-tilbudet, særlig for kvinner. Sykeligheten innen de enkelte yrker kan være en annen faktor som kunne ha betydning blant annet for arbeidskraftprognoser.

2.3 Planlagt persondataregistrering

For oss vil en viktig målsetting ved Folketellingen 1970 være supplerings av den løpende persondataregistrering. Vi ønsker et bilde av utdannings- og yrkesstrukturen som kan la seg vedlikeholde ved løpende registreringer. Like ens håper vi at det i denne telling skal være mulig å etablere et familiebegrep som kan vedlikeholdes gjennom de registreringer vi allerede har på løpende basis.

Regelmessig registrering av individuelle utdanningsdata er en utvidelse som vi håper å kunne realisere i tilknytning til Folketellingen 1970 forutsatt at vi finner fram til tilfredsstillende registreringssystemer og får de nødvendige bevilgninger. Grunntanken er at det for hvert enkelt kurs/klasse skal samles inn elevlister med fødselsnummer-identifisering så vel ved kursets begynnelse som ved dets avslutning. Avbrudd o.l. forutsettes rapportert løpende.(10). Vi antar at en slik registrering ikke bare vil komme statistiske formål til gode, men også at den vil by på en rekke fordeler for utdanningssystemets administrative organer.

Et stort problem i denne sammenheng er hvordan en skal få gjort fødselsnumrene lett tilgjengelig i systemet. På lang sikt kan løsningen ligge i at numrene påføres alle fødselsattester og skolevitnemål. På kort sikt er problemene vanskeligere, og Byrådet overveier for tiden de alternativer som eventuelt kan tenkes.

Systemet praktiseres nå noe modifisert for studenter ved universiteter og høyskoler.(15). Likeledes registreres data med fødselsnummer for artianere.

Registrering av yrkesaktivitet og yrke i forbindelse med hver folketelling gir et svakt datagrunnlag for estimering av strukturer i modeller som skal forklare yrkesvalg og tilbud av forskjellige typer av arbeidskraft. Flere alternativer for en hyppigere registrering eksisterer. I dag registreres yrke eller stilling i forbindelse med de årlige selvangivelser til likningsmyndighetene, men disse opplysningene blir ikke kodet og punched. Andre data fra selvangivelsene punches med fødselsnummer og gir grunnlaget for den årlige registrering av inntektsdata og vil også kunne gi grunnlag for vurdering av yrkesaktivitet. Koding og punching av yrke for de personer som på selvangivelsen oppgir annet yrke enn tidligere skulle ikke være noen

uoverkommelig oppgave. Selv om en neppe vil kunne regne med en detaljert yrkesspesifikasjon, vil disse data representere et viktig tillegg til de nåværende detaljerte 10-årslige yrkesdata.

En annen mulighet er yrkesregistrering i forbindelse med medlemskap etc. i syketrygden. Her er forholdene på grunn av omleggingen foreløpig mer uklare.

2.4 Noen generelle trekk ved persondata

Fra modellbyggerens synsvinkel er det visse generelle trekk ved persondata-arkivenes innhold som kan strekes under:

- a. Tallet på observasjonssett vil være stort sammenliknet med de relativt få statistiske standardgrupper en tidligere hadde. Statistisk sett gir dette grunnlag for en mer presis estimering.
- b. Tallet på kjennemerker som direkte og indirekte kan knyttes til den enkelte person vil i praksis være langt større enn tallet på kjennemerker en vil kunne karakterisere en gruppe med. En vil derfor ha større frihet for å velge ut variable og formulere mer realistiske teorier enn tidligere.
- c. En rekke begivenheter vil registreres på løpende basis med tidsangivelse for når de fant sted. Dette gir oss empirisk datamateriale som berettiger modellbyggeren til å arbeide med modelltyper som betrakter begivenheter som realisasjoner av en tidskontinuerlig prosess hvor ikke to begivenheter kan skje samtidig. Denne modelltype er etter all sannsynlighet i mange situasjoner en bedre framstilling av virkeligheten enn den tradisjonelle, og gir også høve til å arbeide med et mer hensiktsmessig begrepsapparat.
- d. Vi har nå oppgaver for den atferdsmessige enhet, personen, og unngår derved å måtte gjøre så mange forutsetninger om homogenitet i persongrupper på grunn av foreliggende statistiske data.

3. Personmodeller

3.1 Modelltyper

Vi har ikke låst anvendelsen av persondataene til noen bestemt modelltype. Det foregår utnytting av materialet på flere fronter. Arbeidet i Gruppen for personmodeller er imidlertid karakterisert ved at vi tenker oss at de skal utgjøre en samling av konsistente, matematisk formulerte relasjoner i en stor befolkningsmodell.

Modellens relasjoner vil være av de to vanlige typer:

1. Definisjonsrelasjoner som ivaretar de nødvendige definisjonsmessige restriksjoner mellom alle de variable som innføres i modellene.
2. Strukturrelasjoner som forklarer biologisk, medisinsk, psykologisk og institusjonelt betingede forhold mellom de variable.

Hensikten med å ha en slik omfattende modell i tankene er i størst mulig utstrekning å lede innsatsen mot et felles mål. Persondata-arkivenes særlige struktur har også ledet til at en i stor utstrekning arbeider med tidskontinuerlige Markov-kjedemodeller med parametre definert som overgangsintensiteter av forskjellig slag.

Når en arbeider med aggregerte data for grupper, står en overfor et valg mellom følgende to alternativer: En kan forutsette at det er den samlede gruppes atferd en er interessert i. De aggregerte data representerer da et observasjonsmateriale som ikke legger noen restriksjoner på funksjonsformer for de relasjoner som utgjør den hypotese som det arbeides med. På den annen side gir en avkall på å studere sentrale problemer som tilgang, fordeling og avgang av enheter som begivenheter som er knyttet til de enkelte individer. Alternativt kan vi velge den mer realistiske problemstilling og forutsette at det er individenes atferd vi vil studere. Med statistiske aggregater som eneste observasjonsmateriale, vil imidlertid den allerede foretatte aggregering da ofte legge restriksjoner på utformingen av relasjonene. Med adgang til individuelle persondata vil en imidlertid kunne se bort fra disse problemer under problemformuleringen.

Den nye modelltype og de detaljrike arkiver utfordrer også til utvikling av estimeringsmetodikk som nyttiggjør seg informasjonen på en mest mulig effektiv måte. Selv om vi gjør bruk av individualdata, er vi ikke opptatt av modeller som forklarer det enkelte navngitte menneskes historie og sannsynlige utvikling. Det er den typiske person som er av interesse, og våre modeller kan derfor kanskje karakteriseres som makromodeller som nyttiggjør seg mikrodata under estimeringen.

3.2. Personmodeller under utvikling

En minner om at dette innlegget er begrenset til arbeid med modeller som gjør bruk av persondata-arkivet. Tidligere er det i Byrået utført en rekke demografiske og andre analyser av befolkningen som bare unntaksvis er nevnt her.

Vår demografiske prognosemodell omfatter flere hundre tusen definisjonsrelasjoner, dødelighetsrelasjoner, fødselsrelasjoner og flytterela-

sjoner (4). Den har som siktemål å gi en forklaring på befolkningsutviklingen i ett-årige aldersgrupper for hvert kjønn i hver av de vel 450 kommuner vi har i Norge. På det nåværende utviklingsstrinn er strukturrelasjonene av tradisjonell og primitiv type.

Modellen er operativ på en datamaskin gjennom to sett av programmer. Det første programsettet spesifiserer estimeringen av strukturparametrene på grunnlag av data fra persondata-arkivene. Det andre omfatter rutine for anvendelse av modellen for beregning av befolkningsprognoser. Det operative systemet er bygd opp modulært slik at når relasjoner i modellen endres eller nye føyes til, vil dette kunne ivaretas ved utskifting av avgrensede rutiner.

Modellen er ikke formelt vesensforskjellig fra tidligere modeller. Det som er nytt ved denne er de detaljerte resultater modellen gir, den mulighet den representerer for bedre utnytting av eksisterende data og dens hensiktsmessige operative utforming som gir en sterk beredskapstilstand med hensyn til utkjøring av prognoser.

På et tidligere stadium ble virkningene av spesifikasjonsfeil i modellen studert (2, 3). Som en kan tenke seg er særlig prognoser følsomme for endringer i de relasjoner som skal forklare flyttinger. Relasjonene som brukes nå, tar hensyn til personlige kjennemerker som alder og kjønn hos flytterne. De faktorer som representerer miljø i ut- og innflyttingskommuner er implisitte og forutsatt uforanderlige over tiden.

Vi har tatt opp arbeid med å forbedre flytterelasjonene både ved å trekke inn flere personlige faktorer og ved å innføre eksplisitt spesifiserte miljøfaktorer som kan forandres i samsvar med kommunens struktur (7).

Blant de personlige faktorer som er sentrale i forklaring av flyttingene regner vi familieforhold, personlig inntekt, utdanning, yrke etc. Vi må her anta at det er en vekselvirkning mellom de forskjellige faktorer og at det kan være hensiktsmessig å utvikle modeller som forklarer flere faktorer simultant. Fødselsdata og løpende registrering gjør det også mulig å legge opp analyser som en kontinuerlig generasjonsanalyse.

Viktige miljøfaktorer som kan påvirke flyttetilbøyeligheten er arbeidsforhold, boligsituasjon, utdanningssituasjon, muligheten for fritidsysler og økonomiske forhold i innflyttingskommune, utflyttingskommune og i de resterende kommuner. Det praktiske problem vil i første rekke være å utvikle et hensiktsmessig system for klassifisering av kommunene etter type.

Ekteskapelig status er et viktig kjennemerke både på grunn av sin sosiologiske betydning og som forklarende faktor for blant annet variasjoner

i fruktbarheten (5, 6, 8). Ekteskapelig status studeres både isolert og simultant med fruktbarheten i Markovkjedemodeller. I utforming av modellene kan en med det datagrunnlag en nå har til disposisjon ta omsyn til kvinnens alder og bosted når en begivenhet observeres og hennes fødselstidspunkt. Mulighetene vil også være til stede for å trekke inntekt, familiestørrelse og sammensetning inn i disse modeller på et senere tidspunkt.

Dødelighetsanalyser har hittil ikke vært tatt opp som selvstendig studieprosjekt i personmodellgruppen.

3.3 Videre modellutbygging

Den videre bygging av personmodeller vil være bestemt av tilgang på kvalifisert arbeidskraft og av at Byrået får de nødvendige bevilgninger til videreføring av persondataregistring og arbeid med personmodeller. Dersom disse betingelser oppfylles, gir dataarkivene grunnlag for omfattende modellbygging som uten tvil vil kunne komme samfunnsplanleggingen til gode.

Arbeidet i tiden framover vil stort sett kunne tenkes konsentrert om:

- a. Forbedring av de nåværende relasjoner i prognosemodellen på grunnlag av de pågående spesialstudier.
- b. Utviding av modellen til å omfatte relasjoner som forklarer ekteskapelig status.
- c. Utviding av modellen til å omfatte relasjoner som forklarer etterspørsel etter utdanning.
- d. Utviding av modellen til å omfatte relasjoner som forklarer tilbudet av arbeidskraft.

Når det gjelder de to første punkter er forarbeidene allerede omtalt, og vi skal her nøye oss med å slå fast at det er av stor betydning ikke minst for forklaring av tilbudet av arbeidskraft å få spesifisert ekteskapelig status. Vi nytter betegnelsene "etterspørsel etter utdanning" og "tilbudet av arbeidskraft" for å understreke at vi her vil studere dette fra elevens eller arbeidstakerens synsvinkel. Mer presist ville det imidlertid være å si at det er forklaringer på "oppnådd utdanning" og "sysselsatt arbeidskraft" relasjonene vil gi.

Første steg i utviding av modellen med sikte på å forklare etterspørselen etter utdanning vil bli å innføre enkle tradisjonelle utdanningsrelasjoner som bygger på koeffisienter for utdanningstilbøyelighet og som svarer til spesifikasjonene i befolkningsmodellen, dvs. relasjoner som uttrykker tilbøyeligheten for å søke utdanning av forskjellig slag etter alder,

kjønn og tidligere utdanningsnivå for de enkelte geografiske områder. Slike koeffisienter har vært tallfestet tidligere, men ikke med den spesifisering det her er tale om. Vi regner med at Folketellingen her vil gi det nødvendige datagrunnlag.

Med utgangspunkt i en løpende registrering av utdanning, vil vi kunne ta opp utdanningsanalyser på et modellgrunnlag som stort sett svarer til det som tidligere er antydnet i forbindelse med demografiske problemstillinger. Etterspørsel etter utdanning kan tenkes forklart ved individuelle faktorer som alder, kjønn, ekteskapelig status, tidligere utdanning, fødselsdato, ved familiemiljøfaktorer som familiestørrelse og sammensetning, familieinntekt, foreldres yrke og utdanning og ved kommunemiljøfaktorer karakterisert ved arbeidskraftsituasjonen, økonomi og utdanningstilbud i bostedskommunen sammenholdt med landet forøvrig.

Spesialanalyser etter disse linjer antar vi vil gi grunnlag for å erstatte de primitive utdanningskoeffisientene med mer meningsfylte og stabile utdanningsrelasjoner i befolkningsmodellen.

For tilbudet av arbeidskraft vil framstøtene bli tilsvarende. I første omgang vil vi søke å etablere enkle relasjoner med tilbudstilbøyeligheter for de forskjellige grupper av årskull, kjønn, utdanning og ekteskapelig status. Sammenliknet med det som tidligere er gjort av Byrået (9), vil persondataarkivet så fremt en oppnår å få en hyppigere registrering av yrkesendringer, gi grunnlag for en differensiering i flere grupper.

På samme måte som for utdanning er målsettingen å søke og erstatte disse enklere relasjoner med andre utviklet gjennom spesialstudier. Vi håper at disse vil gi bedre forklaringer på arbeidskraftens variasjoner ved hjelp av de personlige kjennemerker og de miljøfaktorer som kan knyttes til den enkelte arbeidstaker. Det vil her være viktig å utvikle relasjoner som ikke bare gir en forklaring på hvilke faktorer som er avgjørende for om en person tilbys arbeidskraft eller ikke, men også gir uttrykk for hvor meget som tilbys under forskjellige forhold.

Både utdannings- og arbeidskraftrelasjonene planlegges bygd inn i det operasjonelle systemet slik at prognoser for utdanningsbehov og arbeidskrafttilbud kan lages på tilsvarende måte som de demografiske prognosene. Etter hvert som eksogene miljøfaktorer trekkkes inn i relasjonene, vil modellen gi resultater som har et preg av å være prognoser som er betinget av de eksplisitte forutsetninger som gjøres om utviklingen av forhold som ikke omfattes av modellen, f.eks. den framtidige utvikling på arbeidsmarkedet, de økonomiske konjunkturer og utbygging av utdanningssystemet. Dette vil kreve at befolkningsmodellen må være kompatibel med blant annet våre økonomiske

modeller slik at vi skal kunne gi konsistente vekselvirkende prognoser på de forskjellige områder.

Når Byrået, til tross for det utmerkede arbeid som utføres av andre institusjoner i Norge med hensyn til analyser av utdannings- og arbeidskraftproblemer, tar sikte på å ta opp utdannings- og arbeidskraftmodeller på sitt program, er det fordi vi er overbevist om at den informasjonsmasse som bygges opp i persondata-arkivene er uten sidestykke. På grunn av de restriksjoner som konfidensiell behandling av persondata pålegger Byrået, kan en effektiv utnyttning av denne informasjonskapital bare foregå i Byrået.

4. Henvisninger

Nedenfor er det gitt en del henvisninger til arbeider utført i Byrået med tilknytning til persondata-arkiver og personmodeller. En stor del av de opplistede arbeider er arbeidsnotater som ikke er alminnelig tilgjengelig. Dersom seminardeltakerne er interessert i enkelte av disse notater, vil de i de fleste tilfelle kunne skaffes for personlig bruk ved henvendelse til Statistisk Sentralbyrå, Oslo-Dep., Oslo 1.

1. B. Bendiksen: Valg av personnummersystem, Arbeidsnotater IO 63/2, Statistisk Sentralbyrå, Oslo 1963.
2. K. Bratland: Fødselshyppigheter. En undersøkelse av regionale variasjoner og endringer over tiden. Arbeidsnotater IB 67/1, Statistisk Sentralbyrå, Oslo 1967.
3. E. Gilje: Befolkningsprognoser. En undersøkelse av de virkninger feil i forutsetningene har for resultatene. Arbeidsnotater IB 66/3, Statistisk Sentralbyrå, Oslo 1966.
4. E. Gilje: Modell for befolkningsprosjeksjoner for norske regioner, Arbeidsnotater IO 68/15, Statistisk Sentralbyrå, Oslo 1968.
5. J. M. Hoem: A Probabilistic Model for Primary Marital Fertility, Arbeidsnotater IO 68/12, Statistisk Sentralbyrå, Oslo 1968.
6. J. M. Hoem: Fertility Rates and Reproduction Rates in a Probabilistic Setting. Arbeidsnotater IO 68/6, Statistisk Sentralbyrå, Oslo 1968.
7. J. M. Hoem: Befolkningsprognosemodellens flyttingsrelasjoner. I. Arbeidsnotater IO 68/11, Statistisk Sentralbyrå, Oslo 1968.
8. J. M. Hoem: Four Demographic Papers. Arbeidsnotater IO 68/18, Statistisk Sentralbyrå, Oslo 1968.
9. G. S. Lettenstrøm og G. Skancke: De yrkesaktive i Norge 1875 og 1960 og prognoser for utviklingen frem til 1970. ARTIKLER nr. 10, Statistisk Sentralbyrå, Oslo 1964.

10. I. Møglestue: Utbygging av elev- og eksamensstatistikk. Arbeidsnotater IO 68/19, Statistisk Sentralbyrå, Oslo 1968.
11. S. Nordbotten: Elektronmaskinene og statistikkens utforming i årene framover. De Nordiska Statistikermøter i Helsingfors 1960, Helsingfors 1961.
12. S. Nordbotten: Om arkivstatistiske systemer. Arbeidsnotater IR 66/2, Statistisk Sentralbyrå, Oslo 1966.
13. S. Nordbotten: Purposes, Problems and Ideas Related to Statistical File Systems. Arbeidsnotater IR 67/2, Statistisk Sentralbyrå, Oslo 1967.
14. Rådgivende utvalg for undervisningsstatistikk: Artianerstatistikk og artiumsmatrikkel. Arbeidsnotater IO 66/9, Statistisk Sentralbyrå, Oslo 1966.
15. Rådgivende utvalg for undervisningsstatistikk: Studentstatistikk. Arbeidsnotater IR 68/2, Statistisk Sentralbyrå, Oslo 1968.
16. E. S. Selmer: Registration Numbers in Norway: Some Applied Number Theory and Psychology. The Journal of the Royal Statistical Society, Series A, Vol.130, Part 2, pp.225-231, 1967.
17. H. Skaug: The Norwegian Population Register System. Arbeidsnotater IO 67/2, Statistisk Sentralbyrå, Oslo 1967.
18. H. Skaug og K. Karlsen: Statistisk Sentralbyrås sentrale registre. ARTIKLER nr. 22, Statistisk Sentralbyrå, Oslo 1968.