

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 73/3

24. januar 1973

METODEHEFTE NR. 4

Notater om forbruksundersøkelser

INNHold

	Side
Metodehefter i serien Arbeidsnotater	2
Det rådgivende utvalg for forbruksundersøkelsen: "Om utformingen av Statistisk Sentralbyrås forbruksundersøkelser i åra framover". (HiB/GH, 1/3-71)	3
Odd Skarstad: "Om bokføringsperiodens lengde ved forbruksundersøkelser". (OS/WA, 17/8-71)	24
Odd Skarstad: "Om prinsipper for registrering av forbruk ved forbruksundersøkelser". (OS/IH, 17/9-72)	30
Grete Dahl: "Om frafallsproblemet ved statistiske undersøkelser. En drøfting av problemet spesielt i forbindelse med forbruksundersøkelser". (GD/WA, 4/5-72)	43
Odd Skarstad: "Korreksjon for frafall ved forbruksundersøkelsen 1973. Spesielt om estimering av svars-sannsynligheter". (OS, 12/1-73)	59

ForordMetodehefter 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 har innført en publiseringsordning for stoff av dette slaget. En publiserer leilighetsvis et passende antall slike notater samlet i metodehefter i serien Arbeidsnotater. Inneværende hefte er det fjerde av denne typen. Det inneholder notater vedrørende Byråets forbruksundersøkelser.

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.

HiB/GH, 1/3-71

OM UTFORMINGEN AV
STATISTISK SENTRALBYRÅS FORBRUKSUNDERSØKELSER
I ÅRA FRAMOVER

Notat fra Det rådgivende utvalg for forbruksundersøkelsen

Innhold

	Side
1. Innledning	4
1.1 Oversikt	4
1.2 Definisjoner og avgrensinger	4
1.3 Litt historikk	5
1.4 Noen hovedsynspunkter	5
2. Formål, anvendelser, erfaringer	6
2.1 Vektgrunnlag for konsumprisindeksen	6
2.2 Støttemateriale for spesifikasjon av det private konsum i nasjonalregnskapet	7
2.3 Makro-økonomiske analyser	8
2.4 Markedsforskning og andre etterspørselsanalyser	9
2.5 Analyser av den personlige inntektsfordeling og av personskattleggingen	9
2.6 Levekårsundersøkelser, samordning med sosio-demografisk statistikk, forbruksforskning ellers	12
2.7 Sammenfatning	14
3. Drøfting av alternative utforminger av forbruksundersøkelser i framtida	14
3.1 Skisse av fire ulike opplegg	14
3.2 Drøfting av de skisserte oppleggene i relasjon til an- vendelsene av forbruksundersøkelser	16
3.3 Nærmere om løpende forbruksundersøkelser	18
3.4 Kostnadsoverslag	20
3.5 Tidsplan	23
4. Utvalgets anbefalinger	23

1. INNLEDNING

1.1 Oversikt

I dette notatet skal Det rådgivende utvalg for forbruksundersøkelser legge fram noen synspunkter på hvorledes Byråets forbruksundersøkelser bør legges opp i de nærmeste åra framover, og hvilke anvendelsesområder vi antar vil stå sentralt.

I avsnitt 2 blir det gjort rede for de erfaringer som er gjort med bearbeidingen av forbruksundersøkelsen 1967 i og utenfor Statistisk Sentralbyrå.

I avsnitt 3 skisserer vi noen alternative måter som forbruksundersøkelser kan utformes på, og drøfter oppleggene ut fra det vi vet om anvendelsene av materialet. Utvalget har også forsøkt å sette opp grove overslag over kostnadene ved de ulike oppleggene.

Avsnitt 4 inneholder utvalgets anbefalinger.

1.2 Definisjoner og avgrensinger

Forbruksundersøkelser tilhører den klassen av utvalgsundersøkelser som har husholdningen som statistisk enhet (husholdningsundersøkelser). I forbruksundersøkelser søker en ved regnskapsføring og/eller intervju å kartlegge forbruksutgiften og dens komponenter i en viss, avgrenset periode. Forbruksundersøkelser i denne mening gir en fullstendig oversikt over, og detaljert spesifisering av, den del av det private konsum som husholdningene selv betaler for eller produserer.

Denne definisjonen klarlegger bare minste-innholdet i en forbruksundersøkelse, og tjener til å avgrense forbruksundersøkelser mot andre typer av husholdningsundersøkelser. Målet med husholdningsundersøkelser vil være å skaffe til veie et mest mulig fullstendig bilde av private husholdningers atferd og levemåte. Siden det ikke er praktisk å samle inn alle opplysninger om disse forholdene i en eneste stor undersøkelse, utfører Byrådet en rekke undersøkelser som tar for seg ulike emner. Mange av disse vil berøre husholdningenes forbruk, og i en forbruksundersøkelse vil en trekke inn også andre kjennetegn enn de som knytter seg til forbruksutgift. Det er et særskilt problem, som vi skal komme tilbake til, hvordan de ulike husholdningsundersøkelser kan samordnes og tilpasses hverandre.

Vi mener med dette å ha gjort det klart at dette notatet tar for seg forbruksundersøkelser som minst dekker alle utgiftsposter som hører inn under husholdningenes private forbruk. Den præcise avgrensing mellom forbruksundersøkelser og andre husholdningsundersøkelser er ikke dermed låst fast.

1.3 Litt historikk

I Norge har vi hatt forbruksundersøkelser som bygger på oppgaver fra de enkelte husholdninger siden begynnelsen av dette århundret. Det primære formålet med de første granskningene var å kaste lys over levemåten til de fattige i samfunnet. Som regel var det familier fra arbeiderklassen som ble undersøkt. Helt fram til annen verdenskrig var utvalgene av husholdninger små, og de husholdningene som var med, ble ikke trukket ut etter statistiske prinsipper. Fra 1945 fram til 1958 ble det holdt en rekke større og mindre undersøkelser som dekket forskjellige sosiale grupper. Fra 1950-åra av tok en sikte på å kartlegge forbruksstrukturen i hele befolkningen. Begynnelsen var forbruksundersøkelsen blant fiskere og jordbrukere i 1954. Denne ble fulgt av undersøkelsen av alderstrygdede i 1955-56. Forbruksundersøkelsen i 1958 omfattet 3 500 husholdninger, og dekket alle sosialgrupper unntatt fiskere, jordbrukere og alderstrygdede. Fra 1958 av gikk en over til å bygge på sannsynlighetsutvalg. Byråets siste forbruksundersøkelse ble holdt i 1967, og dekket alle private husholdninger.

Opplysningene har som regel vært hentet inn i form av regnskapsoppgaver. Både i 1958 og i 1967 ble også intervjuetoden nyttet. I 1958 intervjuet man de husholdningene som av slo å føre regnskap. I 1967 ble alle husholdninger intervjuet, dels om visse bakgrunnskjenntegn, dels om utgifter til større engangsanskaffelser i løpet av året.

Både i 1958 og i 1967 førte husholdningene regnskap i en måned hver. Registreringsprinsippene har variert noe. Byrådet har også fulgt forskjellige opplegg når det gjelder hvilke opplysninger som ble samlet inn i tillegg til den rene forbruksutgift. I 1958 søkte en å få til husholdningsregnskap som gav en samlet oversikt over inntekt, forbruksutgift og sparing. I tillegg fikk en oppgaver over inntekt fra likningskontorene. I 1967 ble det ikke stilt spørsmål om inntekt og sparing, men husholdningene ble intervjuet om sin beholdning av varige konsumgoder. I tillegg ble det også i 1967 hentet inn oppgaver fra likningskontorene over husholdningenes inntekter.

1.4 Noen hovedsynspunkter

Erfaringene fra arbeidet med 1967-undersøkelsen viser at etterspørselen etter materiale fra forbruksundersøkelsen er stor, og kommer fra de forskjelligste hold. I løpet av 1960-åra er det kommet til flere nye forskningsinstitusjoner som arbeider med problemer hvor en trenger tallmateriale som belyser forbruk og forbrukeratferd. Også forbruksundersøkelsen

er blitt berørt av den alminnelige tendens til å legge større vekt på kvantitative metoder i samfunnsforskning og planlegging. Til de planleggingsmodellene som er utviklet innenfor og utenfor Statistisk Sentralbyrå trengs det detaljerte opplysninger om etterspørselsstrukturen.

Byråets planer for 1970-åra legger økt vekt på statistikk som belyser sosiale og menneskelige forhold, og statistikk som belyser levekår. Blant annet vil problemer som har med fordelingen av inntekt og andre goder å gjøre, stå sentralt i de kommende år, i samfunnsdebatten såvel som i Byråets statistikkproduksjon. Her må forbruksundersøkelser være en viktig statistisk kilde.

Utvalget mener på denne bakgrunn at tida nå er inne til å styrke Byråets beredskap på dette området. Ikke minst er det ønskelig å finne fram til opplegg som gjør at Byrådet kan levere forbruksstatistikk som er aktuell. Med aktuell sikter vi her både til det tidspunkt statistikken er samlet inn, og de problemstillinger den er egnet til å belyse.

Byrådet har allerede utviklet gode tekniske muligheter for raskt å planlegge, gjennomføre og bearbeide husholdningsundersøkelser. Fordi det er blitt enklere å trekke utvalg og gjennomføre store beregninger, behøver vi for eksempel ikke lenger føle oss bundet til å bygge på selvveiende utvalg.

Bredden i anvendelsene av materialet fra forbruksundersøkelsene øker, og Byrådet får stadig bedre muligheter for å legge opp forbruksundersøkelser som tilfredsstillende spesielle behov.

2. FORMÅL, ANVENDELSER, ERFARINGER

2.1 Vektgrunnlag for konsumprisindeksen

Tradisjonelt er fornying av vektene i konsumprisindeksen et av hovedformålene med å gjennomføre forbruksundersøkelser. Statistisk Sentralbyrå har tidligere tatt sikte på å revidere konsumprisindeksen om lag hvert tiende år. Det er ønskelig at vektene kan endres noe hyppigere når det er behov for det. Vektene i konsumprisindeksen bør skiftes når det er skjedd vesentlige endringer i sammensetningen av forbruket hos private husholdninger. For å følge disse endringene trengs det løpende statistikk over det private konsum. Den får vi nå fra nasjonalregnskapet. Uten hyppige forbruksundersøkelser vil det normalt ta lang tid fra behovet for å revidere vektene er registrert i nasjonalregnskapet til de nye vektene kan foreligge.

Beregninger som er foretatt på forbruksmaterialet fra 1967 viser at utvalg på 5 000 husholdninger, slik en hadde i 1957, er større enn det som trengs for å beregne vektorer til konsumprisindeksen med rimelig grad av presisjon. Anslag tyder på at 3 000 husholdninger vil være tilstrekkelig.

Av de to foregående punktene trekker vi den konklusjon at innenfor en gitt budsjetttramme blir hensynet til konsumprisindeksen best ivaretatt ved hyppigere og mindre omfangsrike forbruksundersøkelser enn de Byrået gjennomførte i 1958 og 1967.

2.2 Støttemateriale for spesifisering av det private konsum i nasjonalregnskapet

Slik forbruksundersøkelsene til nå har vært lagt opp, kan de ikke direkte gi tall til nasjonalregnskapet annet enn akkurat for det året forbruksundersøkelsen er holdt. Også for disse årene har man hittil i liten grad kunnet nytte materiale fra forbruksundersøkelsen i forbindelse med nasjonalregnskapet. I løpet av det siste året har det vært gjennomført en mer systematisk sammenlikning, og en del av uoverensstemmelsene mellom nasjonalregnskapet og forbruksundersøkelsen 1967 er oppklart. Dette betyr at forbruksundersøkelser i framtida kan bli et verdifullt materiale for konsumberegninger i nasjonalregnskapet. Definisjonen av hva som faller inn under privat konsum i nasjonalregnskapets forstand skiller seg riktignok noe fra de begreper en må legge til grunn i en forbruksundersøkelse som bygger på husholdningsregnskap. Men denne forskjellen er blitt betydelig mindre etter omleggingen til ny standard for nasjonalregnskapet (SNA).

Samtidig blir detaljomsetningsindeksen, som er en av hovedkildene for nasjonalregnskapets oppgaver over det private konsum, mindre anvendelig for spesifisering på underposter. Kvartalstallene for privat konsum bygger for varenes vedkommende i sin helhet på detaljomsetningsindeksen, og den er også viktig for årsregnskapet. Detaljomsetningsindeksen gir indekstall for totalomsetningen i detaljhandelen fordelt på bransje, eller type forretning. Ut fra bransjevise indekser beregner 10. kontor kvartalstall for hovedgrupper av varer. Dette går bra i den utstrekning det er entydige sammenhenger mellom omsetningen i en bransje og omsetningen av en bestemt varegruppe. Strukturendringene i varehandelen er imidlertid i ferd med å svekke grunnlaget for fordelingen av omsetningstall på varegrupper. En dagligvareforretning i dag omsetter varer fra nær sagt alle hovedgrupper av det private konsum, og spesialiserte bransjeforretninger svarer for en stadig synkende del av detaljomsetningen.

Nå er det klart at forbruksundersøkelsene må legges helt om til å bli kontinuerlige hvis nasjonalregnskapet skal bygge konsumberegningene sine på dem. Det vil bety at Byrået må holde forbruksundersøkelsen, ikke bare hvert år, men hvert kvartal.

I den grad nasjonalregnskapet skal gi tall for landet som helhet gjelder det samme som ble nevnt under avsnittet om konsumprisindeksen: En kan skjære betydelig ned på antall husholdninger i utvalget i forhold til undersøkelsen i 1967.

Imidlertid arbeider Byrået også med å utvikle regionale nasjonalregnskapstall. En har ingen erfaringer med å nytte forbruksundersøkelsen 1967 i denne sammenheng innen Byrået, men forbruksundersøkelsen 1958 er blitt brukt til noen beregninger. Transportøkonomisk Institutt har anslått konsumet i Finnmark fylke ved hjelp av tall fra 1967-undersøkelsen.

Forbruksundersøkelsen er den eneste direkte kilde Byrået har å bygge på for å beregne konsumtall til de regionale nasjonalregnskap. Vi anser det som utelukket å gjennomføre så store forbruksundersøkelser at de kan brukes til direkte å anslå konsumet og dets sammensetning i hvert fylke. Men selv om en bruker mer indirekte metoder, vil beregning av regionale konsumtall kreve større utvalg enn hensynet til konsumprisindeksen eller de årlige nasjonalregnskapene tilsier.

2.3 Makro-økonomiske analyser

En sentral del av Byråets planleggingsmodeller i MODIS-serien består av et sett av etterspørselsrelasjoner som fordeler konsumet på ca. 50 varegrupper når samlet konsum og de relative prisene er kjent. I motsetning til i skatteinntektsanalyser inngår her bare totaltall for hele landet. Pris- og inntektselastisiteter til de makro-økonomiske modellene har delvis vært beregnet ut fra nasjonalregnskapets tall for privat konsum.

For MODIS IV, som nå er under arbeid, bygger en på etterspørselselastisiteter beregnet på grunnlag av forbruksundersøkelsen 1967. På lengre sikt er det meningen at MODIS IV skal ta hensyn til virkninger på forbrukets sammensetning av endringer i inntektsfordelingen. Tall for inntektsfordelingen regner man da med å ta fra inntekts- og skattestatistikken med forbruksundersøkelser som et verdifullt supplement. Men bare forbruksundersøkelser kan gi grunnlag for beregning av etterspørselselastisiteter spesifisert for ulike inntektsklasser.

Forbruksundersøkelser som refererer seg til et enkelt år, har en svakhet som rammer all økonomisk analyse av materialet. Prisvariasjoner

i løpet av året blir for små til at en kan beregne priselastisiteter ved direkte metoder.

Sosialøkonomisk institutt og planleggingsavdelingen i Finansdepartementet har nyttet analyser av forbruksundersøkelsen i den makro-økonomiske vekstmodellen MSG.

Det er stor etterspørsel etter forbruksmateriale i forbindelse med områdeplanlegging.

2.4 Markedsforskning og andre etterspørselsanalyser

Markedsforskning er tradisjonelt et viktig anvendelsesområde for forbruksundersøkelser. For slike formål blir det ofte stilt krav til detaljert spesifikasjon av forbruket som det er umulig å imøtekomme.

Proposisjonen om et merkantilt forskningsfond vil ventelig bli behandlet i Stortinget i løpet av året. Dette fondet skal blant annet kanalisere midler til markeds- og distribusjonsforskning, ved bevilgninger til eksisterende institusjoner eller ved å engasjere forskere til spesialoppdrag. Forbruksundersøkelsen har allerede levert materiale til et forprosjekt i forbindelse med opprettelsen av dette fondet. Prosjektet dreide seg om metoder for å anslå regionale forbrukstall.

Både 1958-undersøkelsen og 1967-undersøkelsen har vært nyttet til etterspørselsstudier og økonometriske analyser i og utenfor Byrådet. Det er et erklært formål med forbruksundersøkelser at de skal gi materiale til slik forskning. Likevel oppstår det ofte vansker. Det skyldes at de analytiske problemstillingene ikke har vært formulert tilstrekkelig klart da undersøkelsen ble planlagt. Det er også umulig å vente at planleggerne skal kunne ha klart for seg hvilke analyser som vil bli aktuelle i mange år framover. Det kan nevnes at materialet fra 1967 ikke egnert seg så godt som en hadde ventet til å analysere etterspørselen etter varige forbrugsgoder på grunn av uvesentlige detaljer ved registreringen av beholdninger.

2.5 Analyser av den personlige inntektsfordeling og personskattleggingen

Forbruksundersøkelsen 1967 arbeider Byrådet nå med å knytte sammen med inntektsundersøkelsen av samme år.

For studier av inntektsfordeling kommer forbruksundersøkelsen inn på flere måter.

For det første er det en rekke svakheter ved det inntektsbegrepet en er nødt til å bruke i statistikk som bygger på selvangivelser. Disse

svakhetene knytter seg dels til definisjonen av antatt inntekt. En rekke trygdeytelser, sjukepenger osv. er skattefrie og står derved overhodet ikke i selvangivelsen. Dels kan det være usikkerhet om påliteligheten av de opplysningene som står i selvangivelsen, og om hvor vidt begrepet antatt inntekt, om vi nå godtar det, dekker det samme i alle befolkningsgrupper, f.eks. både lønnstakere og sjølstendige. Disse svakhetene gjør at den inntekt som registreres via skattestatistikken kan være et tvilsomt mål på det vi kunne kalle konsumdisponibel inntekt. Derfor er det grunn til å registrere husholdningenes utgifter i tillegg til inntekten.

For det andre er sammensetningen av forbruket av betydning hvis en ønsker å studere virkningene på inntektsfordelingen av den indirekte beskatningen. Hvor stor skatt en husholdning betaler, vil blant annet være bestemt av den rolle avgiftsbelagte og subsidierte varer og tjenester spiller i husholdningens forbruk.

For det tredje gir sammensetningen av forbruket grunnlaget for analyser av husholdningenes inntekter i forholdet til de behov inntekten skal dekke. Ved skatteregler og stønader søker myndighetene å påvirke, ikke inntekten pr. husholdning eller inntekten pr. inntektstaker, men inntekten pr. forbruksenhet i en eller annen forstand. Det er enighet om at barnefamilier trenger større inntekter enn enslige, og det kan ytes ekstra klassefradrag ved sykdom eller uførhet. Å komme fram til brukbare forestillinger om hvor store barnetrygd og klassefradrag "bør" være, er en uhyre komplisert forskningsoppgave, men det er i alle fall klart at oppgaver over husholdningenes faktiske forbruk er en av de kildene slik forskning må støtte seg til.

Vi ser at forbruksundersøkelser i seg sjøl er en viktig kilde til å studere inntektsfordelingen i vårt samfunn. Dessuten er forbruksundersøkelser nyttige som supplement til inntektsstatistikk som en har fra annet hold.

Innen Byrådet er det særlig skatteforskningen som nytter forbruksundersøkelser til å analysere den personlige inntektsfordeling.

Skatteanalyser er i stor grad basert på skatteinsidensmodeller som belyser hvordan skatter og avgifter fordeler seg på de ulike grupper av befolkningen for alternative skatteregler og avgiftssatser. Til dette trenger man to typer opplysninger fra forbruksundersøkelsen: For det første forbrukets sammensetning hos de ulike grupper i en utgangssituasjon, for det andre en tallfesting av hvordan forbruket reagerer på endringer i priser og disponibel inntekt (pris- og inntektselastisiteter). I tillegg ville det blant annet være behov for oppgaver over inntekt etter inntekts-

kilde og sparing etter art. Det ble ikke stilt spørsmål om sparing i forbruksundersøkelsen 1967. Fra registre i Byrået har en overført tall for husholdningenes inntekt ifølge skattelikningen, men i tillegg til slike svakheter som ble drøftet ovenfor, kommer at inntekten ikke er delt opp etter art på en tilstrekkelig detaljert måte.

For skatteanalyser deler en befolkningen i grupper etter blant annet disse kjennetegnene: husholdningstype, inntektsklasse, yrkesstatus, og type boligstrøk. En ville også ønske å foreta en regional inndeling; det har vært spørsmål etter særskilte tall for Nord-Norge.

Forbruksundersøkelsen slik den ble lagt opp i 1967 manglet en del opplysninger sett fra skatteforskernes synspunkt. I tillegg til mangler i oppgavene over inntekt og sparing, gjaldt innvendingene definisjoner av visse kjennetegn. Inntektsgrupper og sosiale grupper står sentralt, og skatteforskningen rammes av det allmenne problem at mange interessante grupper blir for svakt representert i et selvveiende utvalg på 5 000 husholdninger. For disse gruppene ville det vært ønskelig med særskilte utvalg.

Det er stor pågang fra statistikkbrukere utenom Byrået etter materiale fra forbruksundersøkelser til å belyse inntektsfordelingen i Norge og virkningen på inntektsfordelingen av sosialpolitiske og andre tiltak. Komiteer for å utrede eldres levekår og barnefamilienes økonomiske situasjon har nyttet materiale fra forbruksundersøkelsen 1967. For den siste komiteen har Byrået arbeidet særskilt med beregninger av hva det koster å ha barn. Fra Arbeiderbevegelsens utredningskontor har det vært henvendelser om utnytting av forbruksundersøkelsen for å belyse både virkningen av å avløse trygdesystemet med en form for garantert minsteinntekt, og virkningene av å avskaffe ordningen med at renter på lån kan trekkes fra inntekten ved skattlegging. Ingen av disse problemstillingene lot seg analysere ut fra 1967-undersøkelsen, men det skyldtes detaljer i opplegget. Institutt for anvendt sosialvitenskapelig forskning har anvendt materiale fra forbruksundersøkelsen 1967 i en rekke sammenhenger, blant annet for å belyse den økonomiske situasjon hos nystiftede familier.

Ved slike problemorienterte anvendelser er det alltid detaljer ved registrering, koding og definisjoner som passer mindre godt for ett eller flere formål. Slike ting er uunngåelig ved statistikk som har så mange anvendelser som forbruksundersøkelser har. Når forbruksundersøkelsene holdes med flere års mellomrom, og en og samme undersøkelse derfor må nyttes over et lengre tidsrom, blir det ekstra vanskelig å sikre at materialet er egnet til å belyse aktuelle problemstillinger.

2.6 Levekårsundersøkelser, samordning med sosio-demografisk statistikk, forbruksforskning ellers

Som vi pekte på i innledningen, er forbruksundersøkelser bare ett av flere slag husholdningsundersøkelser som blir gjennomført i Statistisk Sentralbyrå. Til sammen er alle Byråets husholdningsundersøkelser med og gir et bilde av norske husholdningers atferd og levemåte. I arbeidet med å bygge ut sosio-demografisk statistikk og statistikk som belyser levekår er husholdningsundersøkelser et viktig ledd.

Med levekårsundersøkelser skal vi mene undersøkelser hvor en søker å belyse levestandard også ut fra ikke-pengemessige komponenter ved å trekke inn forbruk av offentlige goder og fellesgoder, tidsbruk, helse med videre. Forbruk av individuelle goder kjøpt i markedet er bare én komponent i et slikt utvidet levekår-begrep, men det er fortsatt en viktig komponent. Selv om et slikt synspunkt er omstridt, er det mange som vil finne det naturlig å betrakte levekårsundersøkelser som utvidede forbruksundersøkelser. Byrået har på sitt program at forbruksundersøkelsen skal utvides til å omfatte flere levekårskomponenter enn forbruk i snever forstand.

Nå kan en med god grunn hevde at så godt som alle intervjuundersøkelser Byrået utfører, dreier seg om levekår. På den annen side er det klart at det verken er mulig eller rasjonelt å trekke alle interessante faktorer som påvirker levekårene inn i én utvidet forbruksundersøkelse.

For å få et helhetsbilde bør forbruksundersøkelsene, uansett hva de omfatter, derfor i størst mulig grad samordnes med andre husholdningsundersøkelser, og med Byråets sosio-demografiske statistikk for øvrig. Med samordning mener vi:

1. To eller flere undersøkelser foretas for samme utvalg av husholdninger. Eksempel på dette er (i prinsippet) forbruksundersøkelsen 1967 og inntektsundersøkelsen 1967. Dette kan likevel ikke bli noen vanlig framgangsmåte av praktiske grunner.
2. Visse opplysninger, ikke standardkjennetegn, innhentes både i forbruksundersøkelsen og i andre undersøkelser, eller forbruksundersøkelsen tar med spørsmål som ligger utenom dens primære formål, men kommer sentralt inn i sosio-demografisk statistikk. Eksempel på dette vil være at en i forbruksundersøkelsen stiller visse spørsmål om tidsbruk, eller henter inn detaljerte opplysninger om utdanning.

3. Visse bakgrunnskjenntegn som en registrerer i alle fall både i forbruksundersøkelsen og i andre husholdningsundersøkelser defineres slik at de er strengt sammenliknbare fra undersøkelse til undersøkelse. Eksempler på slike kjennetegn er husholdningens samlede inntekt, husholdningstype, husholdningsfase, definisjonen av hva som er en husholdning.

Dette gjør at en kan kjøre ut tidsbruk og forbruksutgift etter de identiske samme forklaringsvariable, og på den måten se hvordan både tidsbruk og forbruk varierer med f.eks. inntekt.

Punkt 3 kan se ut som et selvfølgelig krav, men er det ikke. En rekke kjennetegn vil det være hensiktsmessig å definere ulikt for ulike formål. Byrået utfører husholdningsundersøkelser som skal dekke vidt forskjellige behov, og det kan være mange grunner til å operere med forskjellige husholdningsdefinisjoner i, la oss si, en boligundersøkelse og i en undersøkelse av behovet for sysselsetting blant gifte kvinner.

Det særskilte for forbruksundersøkelser er imidlertid at en vil ønske å se dem i sammenheng med praktisk talt alle andre husholdningsundersøkelser som utføres. Derfor er det et særskilt behov for å definere bakgrunnskjenntegn i forbruksundersøkelser slik at de passer inn i en rekke forskjellige sammenhenger, ikke bare som forklaringsfaktorer for sammensetningen av forbruket.

I den grad planleggerne av en forbruksundersøkelse vet hva undersøkelsen skal brukes til, og samordnes med, kan de legge undersøkelsen opp med sikte på dette. Men når én og samme undersøkelse skal være kilden for analyser av privat forbruk i en periode på 7 til 10 år, er det ikke til å unngå at det dukker opp behov som planleggerne ikke har kunnet forutse. Kanskje ønsker man å se en forbruksundersøkelse i sammenheng med undersøkelser hvor planleggingen ikke engang hadde tatt til da forbruksundersøkelsen ble gjennomført. Dette fører til urasjonell bruk av materialet. Forbruksundersøkelsen ligger der, men bagatellmessige forhold ved opplegget gjør at i mange sammenhenger hvor en ønsker å dra nytte av forbruksundersøkelsen, passer den likevel ikke inn.

Slike problemer gjør seg særlig gjeldende når vi skal levere statistikk til brukere utenom Byrået. Her må vi regne med at etterspørselen etter statistikk i forbindelse med studier av levekår og med forbruksforskning vil øke. Institutt for anvendt sosialvitenskapelig forskning er tidligere nevnt som avtaker av materiale fra forbruksundersøkelsen 1967. De har nyttet det også i forbindelse med arbeidet med sosiale indikatorer.

Statens Institutt for forbruksforskning og vareundersøkelser er opprettet, og Byrådet må vel regne med etterspørsel derfra. Det kan da bli aktuelt å registrere flere og andre kjennetegn ved husholdningene enn de Byrådet bruker ellers.

2.7 Sammenfatning

Tall fra forbruksundersøkelser har et vidt anvendelsesområde. Dette henger sammen med at forbrukeratferd står i skjæringspunktet for interessene til en rekke fag: Medisin, økonomi, sosiologi, psykologi. Forbrukstall utgjør også et bindeledd mellom økonomisk statistikk på den ene side og sosio-demografisk statistikk på den annen side.

Etterspørselen etter data fra forbruksundersøkelser, og bredden i anvendelsesområdet, må antas å bli større i framtida, både innenfor og utenfor Byrådet.

De forskjellige brukere av forbruksundersøkelsene stiller krav som er vanskelige å forene. Det kreves også et vidt register av bakgrunns-kjennetegn, og samordning med mange typer av annen statistikk.

Forbruksundersøkelser med 7-10 års mellomrom krever et stort apparat, og er dårlig egnet for mange anvendelsesområder. En undersøkelse planlegges, gjennomføres og lagres. Når det så dukker opp anvendelsesområder og behov som ingen hadde mulighet til å forutse på det tidspunkt undersøkelsen ble planlagt, vil tallene ofte være foreldet, ikke bare rent tidsmessig, men også fordi de ikke passer til de problemer en nå ønsker å belyse.

Det gjelder all statistikk at en ikke kan lage den slik at alle behov blir tilgodesett på samme tid. Men for forbruksundersøkelser gjelder dette i utpreget grad fordi de brukes i så mange ulike sammenhenger.

3. DRØFTING AV ALTERNATIVE UTFORMINGER AV FORBRUKSUNDERSØKELSER I FRAMTIDA

3.1 Skisse av fire ulike opplegg

Det er selvsagt ingen grense for hvor mange varianter av forbruksundersøkelser en kunne finne fram til. Nedenfor skal vi skissere fire hovedprinsipper for forbruksundersøkelser, hvor de trekkene ved opplegget som varierer er hyppigheten av undersøkelsene og utvalgets størrelse og sammensetning.

V a r i a n t (i)

En større undersøkelse hvert 5.-7. år. Denne må da dekke alle anvendelsesområder. Periodene mellom hver telling kan nyttes til analyse av den foregående undersøkelse og planlegging av den kommende.

V a r i a n t (ii)

Hyppige mindre tellinger, men ikke nødvendigvis hvert år. Hver forbruksundersøkelse konsentrerer seg om en delpopulasjon, f.eks. funksjonshemmede, trygdede, bønder, barnefamilier. En kan også tenke seg forbruksundersøkelser som tar sikte på å dekke en spesiell problemstilling, f.eks. etterspørsel etter varige forbruksgoder. Det ville være rasjonelt å legge slike undersøkelser opp mest mulig enhetlig, og gjerne nær hverandre i tid, slik at de kunne slås sammen og behandles som deler av én stor undersøkelse.

V a r i a n t (iii)

Løpende forbruksundersøkelser med husholdningsregnskaper fra hver måned, i alle år. De regnskaper som samles inn i løpet av et kvartal eller et år, skal komme fra et generelt utvalg av alle private husholdninger.

V a r i a n t (iv)

Dette er en kombinasjon av variantene (ii) og (iii). Løpende forbruksundersøkelser for et generelt utvalg kombineres med særundersøkelser av delpopulasjoner.

For å forklare nærmere: Vi tenker oss et generelt utvalg på f.eks. 1 200 husholdninger i året. Disse fører regnskap en måned eller kortere tid hver, og slik at regnskapene dekker hele året, som nå. Derimot er det ingen sterke grunner til å holde fast ved at hvert husholdningsregnskap skal dekke en kalendermåned. For å registrere kvartalsvise endringer kan det være hensiktsmessig å be hver husholdning om å føre regnskap en uke i hvert av 2 eller flere kvartaler.

I tillegg til regnskapene bør det være intervju med hver husholdning, i tilknytning til regnskapsførselen og/eller ved årets utgang. Intervjuene faller logisk i to deler: En fast kjerne av spørsmål om utgifter til større engangsanskaffelser (som i årsintervjuet 1967); dessuten tilleggsspørsmål som skal dekke behovet for å ta opp spesielle problemstillinger.

I tillegg til det generelle utvalg på 1 200 husholdninger trekkes særutvalg etter behov. Husholdningene i disse særutvalgene fører regnskap og blir intervjuet på samme måten som de øvrige.

3.2 Drøfting av de skisserte oppleggene i relasjon til anvendelsene av forbruksundersøkelser

I det følgende skal vi drøfte de fire variantene fra punkt 3.1 i forhold til anvendelsene av forbruksundersøkelsen. For å lette drøftingen, skal vi nå ordne de nåværende og framtidige anvendelser som ble drøftet i avsnitt 2, i fire typer:

- A. Konsumprisindeksen (punkt 2.1)
- B. Kvartalsvis og årlig nasjonalregnskap (punkt 2.2)
- C. Statistikk som skal belyse velstandsfordeling og levekår (punktene 2.5, 2.6 og 2.2 (regionalt nasjonalregnskap))
- D. Beregning av etterspørsels-elastisiteter og annen analyse av forbrukeratferd (punktene 2.3, 2.4, 2.6)

Vi har prøvd å dele inn slik at de krav som stilles til forbruksundersøkelsen blir mest mulig rendyrket. Under punktene A og B dreier det seg om makro-tall og om tidsrekker. Under punktene C og D kommer de mer mikro-betonte anvendelser.

A. Konsumprisindeksen

Innenfor ganske vide tidsgrenser er det av liten betydning for beregning av vekter til konsumprisindeksen hvor langt tidsrom innsamlingsperioden for beregningsgrunnlaget dekker. For å beregne vekter til konsumprisindeksen kan en derfor tenke seg både større engangstillinger og løpende forbruksundersøkelser, hvor beregningsgrunnlaget strekker seg over en to-tre års periode. En sammenslåing av spesialundersøkelser er også mulig, og alle framgangsmåter er like gode.

For konsumprisindeksen har løpende forbruksundersøkelser to fordeler framfor variantene (i) og (ii). For det første kan de gi en løpende registrering av endringer i den gjennomsnittlige forbruksstruktur. Denne "overvåkings-tjenesten" gjør det lettere å vurdere når det er på tide å skifte vektgrunnlag. Det er ikke nødvendig å binde seg til å skifte vekter på bestemte tidspunkt.

Den største fordel er at med løpende forbruksundersøkelser kan vekter skiftes på kort varsel uten at det går lang tid fra den dag en har besluttet å endre vektgrunnlaget og til den dag en ny forbruksundersøkelse er planlagt, gjennomført og bearbeidet.

B. Kvartalsvis og årlig nasjonalregnskap

Løpende forbruksundersøkelser er - selvsagt - det eneste mulige opplegget hvis vi ønsker å registrere løpende endringer over tid i forbruket.

Helst burde en ha et fast utvalg hvis en vil måle endringer. Dette er imidlertid ikke praktisk mulig. Det vi kan tenke oss, er en delvis utskifting av utvalget hvert kvartal. Vi trenger ikke nødvendigvis binde oss til at alle hushold skal føre regnskap i en måned av gangen. En mulighet er for eksempel at hver husholdning i utvalget fører regnskap en uke av gangen i to eller flere kvartaler. Kombinert med delvis utskifting av husholdningene hvert kvartal, fører dette til at i hvert fall en del av utvalget er fast fra kvartal til kvartal og fra år til år.

Utvalget har ikke drøftet i detalj hvordan dette skulle gjennomføres. Men vi kan slå fast at med en litt fantasifull utvalgsplan kan vi registrere kvartalsvise endringer på løpende basis med god presisjon.

C. Levekårs- og fordelingsstatistikk

Her står vi overfor det problem at små grupper blir for dårlig representert i et selvveiende utvalg. Disse gruppene må overrepresenteres i utvalget hvis vi skal kunne få noenlunde presise tall for dem. Derfor er variant (i) dårlig. Det er praktisk umulig å ta hensyn til mer enn en eller høyst to grupper av gangen når vi skal legge opp en stor engangstelling. Det er heller ikke den beste løsningen å studere hver interessant gruppe for seg. Vi mister informasjon hvis undersøkelsene ikke er lagt opp så vi har en referansegruppe å sammenlikne med. Det beste er derfor å ta undersøkelser som dekker et tverrsnitt av befolkningen, og med én og én særgruppe overrepresentert, med andre ord det opplegget som er skissert under variant (iv).

Over en lengre periode vil et slikt opplegg gi svært god fordelingsstatistikk. På grunn av kontinuiteten i undersøkelsen kan vi slå undersøkelser fra mange år sammen og betrakte dem som én stor undersøkelse over en lengre periode. Legg merke til at selv om vi tar for oss en særskilt delpopulasjon hvert år, vil alle befolkningsgrupper være representert med regnskaper fra hvert år fordi grunnstammen hele tida er et generelt utvalg. Ved å betrakte undersøkelser fra flere år samtidig kan vi bygge opp utvalg nær sagt så store vi vil. Dette kan ha særlig betydning for regional statistikk.

D. Etterspørsels-elastisiteter og annen analyse av forbrukeratferd

Heller ikke for dette formålet er variant (i) noen helt god løsning. Dette skyldes ikke egenskaper ved utvalget. Det henger sammen med at når en undersøkelse skal analyseres ut fra en bestemt problemstilling, bør undersøkelsen helst legges opp med dette for øyet. I annet fall vil det nesten alltid være detaljer i koding, registrering, definisjoner og spørsmålsformuleringer som svikter.

Formodentlig kan vi legge opp engangsundersøkelser som er godt egnet til å beregne de etterspørsels-elastisiteter som trengs i Byråets egne modeller. Når det gjelder forbruksforskere utenfra, viser det seg praktisk umulig å forutse de problemstillinger som kan bli aktuelle.

Det ideelle er undersøkelser skreddersydd for hver enkelt av de problemstillinger vi mener det er verd å bruke ressurser på. Så langt kan vi ikke gå, men et opplegg med noe større innebygd fleksibilitet enn det vi har nå ville være en god hjelp. Av de framgangsmåtene som er skissert, anser vi variantene (iii) og (iv) som best egnet for forbruksforskning.

Løpende forbruksundersøkelser vil gjøre det mulig å beregne pris-elastisiteter direkte.

3.3 Nærmere om løpende forbruksundersøkelser

Det rådgivende utvalg mener at løpende forbruksundersøkelser slik som skissert i variant (iv) er det opplegget som best vil dekke behovet for forbruksstatistikk. Hvis Byrådet først skal endre vesentlig på opplegget for forbruksundersøkelsene, må det være løpende forbruksundersøkelser som blir aktuelt. I det følgende skal vi derfor gå noe nærmere inn på hvordan løpende forbruksundersøkelser kunne legges opp, og vurdere kostnadene ved løpende forbruksundersøkelser sammenliknet med periodiske.

Det rådgivende utvalg har ikke drøftet detaljene i hvordan løpende forbruksundersøkelser skal gjennomføres i praksis. Vi er kommet fram til noen hovedtrekk vi vil anbefale, og som vi skal gjøre rede for nedenfor.

Forbruksundersøkelsen av 1958 bygde på at husholdninger førte fullstendige regnskap over inntekter og utgifter i en måned. De hushold som nektet regnskapsførsel, ble intervjuet. I 1967 var også månedsregnskap hovedkilden for opplysninger om utgift. Det ble ikke stilt spørsmål om husholdningenes inntekt, med unntak av naturalinntekt og visse trygdeinntekter. I tillegg til månedsregnskapene ble det holdt et intervju med

hver husholdning. Disse årsintervjuene ble gjennomført i januar 1968, og gjaldt utgifter til større engangsanskaffelser i løpet av hele året 1967.

Vi har tenkt oss å fortsette med en kombinasjon av regnskap og intervju for hver husholdning. Hvis forbruksoppgavene skal bearbeides kontinuerlig, er det ulemper ved å gjennomføre årsintervju på fast tidspunkt, som i 1967. Dersom vi skal fortsette å intervju husholdningene om utgifter til større engangsanskaffelser, bør dette intervjuet skje i forbindelse med innsamlingen av regnskapsbøker. Hvilken periode intervjuet skal dekke, er det ikke nødvendig å ta standpunkt til nå. Det kan dreie seg om et kvartal, et halvår, et år eller en kombinasjon av disse.

Ved løpende undersøkelser bør en overveie å skjære regnskapsperioden ned til 14 dager. Det er en rekke argumenter for og mot kortere og lengre regnskapsperioder. Det viktigste argument mot å innskrenke perioden ser ut til å være at utgift til forbruksvarer fordeler seg svært ujevnt over måneden, særlig for husholdninger med månedslønn, men også for andre (husleie, månedskort). Denne ulempen kan imidlertid i hvert fall delvis oppveies ved at regnskapene suppleres med intervju, og ved at regnskapsføringen spres over ukene i hver måned slik at det føres regnskap fra hver uke i hver måned. Av argumenter for å skjære ned regnskapsperioden kan nevnes at kortere regnskapsføring kan ha betydning for oppslutningen om undersøkelsen. Forbruksundersøkelser med regnskapsføring har vanligvis høyere frafall enn andre husholdningsundersøkelser. I og med løpende forbruksundersøkelser kan antallet husholdninger i utvalget innskrenkes. Med et lite utvalg må vi legge større vekt på frafallsproblemet, og det er grunn til å tro, i hvert fall håpe, at en kortere regnskapsperiode vil minske frafallet. Et annet argument er at kostnadene ved revisjon, koding og punching øker nærmest proporsjonalt med regnskapsperiodens lengde. Men vi vet lite om hva regnskapsperioden betyr for frafall og for godheten av utgiftsanslagene. Tendensen internasjonalt går mot regnskapsperioder på en uke eller 14 dager. Det burde være mulig å bearbeide forbruksundersøkelsen 1967 særskilt for å belyse hva virkningene av kortere regnskapsperiode ville blitt, siden husholdningenes utgifter er ført inn i regnskapsboka med dato.

Foreløpig anbefaler vi at en tar sikte på 14-dagers regnskapsperiode for de løpende forbruksundersøkelser, i hvert fall som et eksperiment. Det bør være et intervju med husholdningene før, og et umiddelbart etter, regnskapsføringen. Innledningsintervjuet dreier seg om bakgrunnskjennetegn for husholdningene, avslutningsintervjuet skal kartlegge

utgiftsposter som må antas å bli dårlig dekket ved regnskapsføringen, varige konsumgoder og eventuelle andre opplysninger en måtte ønske (inntekt, sparing).

Selve regnskapsføringen kan foregå på en av to måter. For det første sammenhengende regnskapsføring i 14 dager, for det andre rotering av utvalget slik at hver husholdning fører regnskap en uke i hvert av to kvartaler. Dette opplegget skal vi kalle delt regnskapsføring.

Delt regnskapsføring gir oss et utvalg som i hvert fall delvis er fast fra kvartal til kvartal, og som derfor særlig er egnet til å anslå kvartalsvise endringer i forbruket.

Ved delt regnskapsføring må husholdningene intervjues om visse utgiftsposter ved avslutningen av hvert ukeregnskap. Nøyaktig hvorledes intervjuingen skal fordeles på de to regnskapsperiodene har utvalget ikke drøftet. For praktiske formål regner vi foreløpig med en fordobling av intervjuetida pr. husholdning ved delt regnskapsførsel.

Hovedtrekkene i løpende forbruksundersøkelser blir etter dette: Hver husholdning fører regnskap i 14 dager, delt eller sammenhengende. I tillegg gjennomføres innledningsintervju og avslutningsintervju ved begynnelsen og slutten av hver regnskapsperiode.

Koding og revidering skjer fortløpende, etter hvert som regnskapene kommer inn. Det blir også nødvendig å legge opp databehandlingen slik at resultatene kan kjøres ut og eventuelt offentliggjøres fortløpende (hvert kvartal).

3.4 Kostnadsoverslag

I kostnadsoverslaget nedenfor har vi anslått de gjennomsnittlige kostnader pr. år ved periodiske forbruksundersøkelser, og kostnadene pr. år ved løpende forbruksundersøkelser med og uten delt regnskapsføring.

Vi har antatt at uansett hvilken utforming forbruksundersøkelsene får i framtida, vil Byrådet ha en viss kontinuerlig drift av dem. Derfor har vi regnet at faste kostnader til planleggere og til systemarbeid vil bli lite berørt av valget mellom løpende og periodiske forbruksundersøkelser.

De kostnadene vi skal betrakte, er følgelig slike som henger direkte sammen med innsamlingen og den første bearbeiding av materialet. Her vil de årlige kostnader stort sett være proporsjonale med det gjennomsnittlige antall husholdninger som intervjues/fører regnskap. Kostnader pr. husholdning vil avhenge av hvilken godtgjøring husholdningene får for sin regnskapsføring, av intervjuetidas lengde og av lengden på regnskaps-

perioden. Antall intervjutimer pr. husholdning påvirker kostnadene på to måter: Intervjuerne har timebetaling, og jo lengre intervjuet er, jo mer tid går det til revisjon, koding m.v. Lengden av regnskapsperioden bestemmer tallet på linjer i regnskapsboka, og følgelig hvor mye tid som går med til revidering, koding og punching av den.

Periodiske forbruksundersøkelser svarer til variant (i) som er beskrevet i punkt 3.1. Vi antar at det ikke i noe fall blir aktuelt å gjennomføre periodiske forbruksundersøkelser med så lange mellomrom som 10 år, og har i kostnadsoverslaget forutsatt 5-årige undersøkelser. Det gir 1 000 regnskapsbøker/intervju i gjennomsnitt pr. år. Videre kan det ikke komme på tale å skjære ned på utvalget hvis hver enkelt forbruksundersøkelse skal gi tall til alle anvendelser, så vi antar at utvalget blir på 5 000 husholdninger hver gang. En del av argumentasjonen for å skjære ned på regnskapsperioden faller bort i forbindelse med utvalg av en slik størrelsesorden. Det er forutsatt en regnskapsperiode på en måned. Endelig har vi regnet med innledningsintervju og avslutningsintervju på til sammen 4 timer pr. husholdning.

Løpende forbruksundersøkelser antar vi skal omfatte 2 000 husholdninger i året, et generelt utvalg på 1 200 husholdninger og et særutvalg med 800 husholdninger. Regnskapsperioden er her forutsatt å være 14 dager. I forhold til periodiske undersøkelser kan intervjuene rasjonaliseres betraktelig ved at de dekker et avgrenset problemfelt i hver periode. Ved sammenhengende regnskapsføring er intervjutida anslått til to timer pr. husholdning, ved delt regnskapsføring til fire timer.

Selve beregningen av kostnadene er gjort slik:

Godtgjøring til regnskapsførere. Til forbruksundersøkelsen 1973 er det foreløpig budsjettert med 90 kroner pr. husholdning. Ved 14 dagers regnskapsføring har vi regnet med 50 kroner pr. husholdning.

Lønn til intervjuere er kr. 19,- pr. time inklusive trygdeutgifter.

Arbeidskraft i Byrået. For koding osv. av regnskapsbøker kan vi halvere den arbeidstida som trengs når regnskapsperioden halveres. I tillegg kommer arbeid med intervjustjemaer, og innsamlingsarbeid. Alt i alt har vi regnet 4 årsverk pr. 1 000 husholdninger ved periodiske undersøkelser, 3 ved løpende, og noe mer hvis regnskapsføringen er delt. Årslønn pr. assistent/fullmektig er satt til 24 000 kroner.

Under disse forutsetningene vil variable kostnader til løpende forbruksundersøkelser (2 000 husholdninger i året) bli på snutt 230 000 kroner i året ved sammenhengende regnskapsføring, og på snutt 420 000 kroner ved delt regnskapsføring. Til sammenlikning vil 5-årslige forbruksundersøkelser (1 000 husholdninger pr. år) tilsvare om lag 270 000 kroner i gjennomsnitt pr. år. Kostnadsoverslaget er nærmere spesifisert i tabell 1.

Tabell 1. Årlige variable kostnader ved periodiske (5-årslige) og ved løpende forbruksundersøkelser

	Periodiske forbruksundersøkelser	Løpende forbruksundersøkelser	
		Sammenhengende regnskapsføring	Delt regnskapsføring
I. Husholdninger pr. år..	1 000	2 000	2 000
II. Årsverk i Byrået (innsamling, revisjon, koding, punching m.v.)	4	6	6½
III. Intervjutid i timer ..	4 000	4 000	8 000
IV. Kostnader i penger:			
Lønn i Byrået	96 000	144 000	156 000
Lønn til intervjuere..	76 000	76 000	152 000
Reisekostnader	4 000	8 000	8 000
Godtgjøring til regnskapsførere	90 000	100 000	100 000
I a l t	266 000	328 000	416 000

I tillegg til disse variable kostnadene er det for alle varianter forutsatt en funksjonærstab i Statistisk Sentralbyrå på to konsulenter og en sekretær ved fagkontoret, og en sekretær ved Systemkontoret.

3.5 Tidsplan

Det er vedtatt at Byrådet skal gjennomføre en større forbruksundersøkelse i 1973. Utvalget vil anbefale at denne legges opp med 14 dagers regnskapsperiode, og avslutningsintervju i tilknytning til regnskapsføringen. Systemarbeidet bør også planlegges slik at bearbeidingen i prinsippet kan skje fortløpende.

På den måten vil forbruksundersøkelsen 1973 bli tilpasset overgangen til løpende forbruksundersøkelser, og løpende forbruksundersøkelser fra 1/1 1974 vil bli en naturlig fortsettelse av den undersøkelse som allerede er i gang. Det eneste prinsipielt nye som kommer til, vil bli en eventuell delt regnskapsføring. Til gjengjeld vil delt regnskapsføring kreve en del arbeid med prøveundersøkelser, løsning av praktiske problemer i forbindelse med flytting og frafall m.v. Her kan man eventuelt ha løpende forbruksundersøkelser med sammenhengende regnskapsføring i en overgangperiode.

4. UTVALGETS ANBEFALINGER

Det rådgivende utvalg for forbruksundersøkelser anbefaler at Statistisk Sentralbyrå går over til løpende forbruksundersøkelser fra januar 1974.

Argumentasjonen for løpende forbruksundersøkelser kan summeres opp i følgende punkter:

1. Løpende forbruksundersøkelser vil gi Byrådet høy beredskap når det gjelder forbruksstatistikk, vekter til konsumprisindeksen og statistikk som vedrører private husholdninger.
2. Løpende forbruksundersøkelser vil gi muligheter til systematiske studier av levekår hos spesielle grupper i befolkningen.
3. Med løpende forbruksundersøkelser er det mulig innenfor en moderat utvidelse av kostnadsrammen å dekke behovene både til de brukere som trenger ferske tall og til de brukere som krever at forbrukstall bygges på store utvalg.
4. Hvis de løpende forbruksundersøkelser legges opp med delt regnskapsføring, vil dette gi prinsipielt ny informasjon om kvartalsvise endringer i forbruket.

OM BOKFØRINGSPERIODENS LENGDE VED FORBRUKSUNDERSØKELSER

av Odd Skarstad

Et ofte diskutert problem ved forbruksundersøkelser er bokføringsperiodens lengde. Det viser seg at de registrerte utgifter pr. husholdning pr. tidsenhet er påvirket av registreringsperiodens lengde. Kennsley og Turner har foretatt studier av dette forhold på engelske forbruksdata[1]. For forbruksdata fra 1953-54 fant Kennsley f.eks. ukentlige forbruksutgifter for samlet utgifter for de tre uker:

1. uke	2. uke	3. uke
226,04 shilling	209,31 shilling	211,04 shilling

Han beregnet også tall for de enkelte goder. Han viser at 1. ukes utgifter er signifikant høyere enn de to andres. Derimot er det ikke signifikant forskjell på uke nr. 2 og 3. Han fant tilsvarende resultater når det gjaldt antall registrerte utgiftsposter. Kennsley mener årsaker til forskjellene kan være:

- a) helt eller delvis at det ved føringen av 1. ukes utgifter også er tatt med utgiftsposter fra perioden umiddelbart før bokføringen skulle begynne,
- b) at husholdningene er blitt mer likeglad med føringen av bøkene etter hvert slik at stadig flere poster er blitt utelatt,
- c) at husholdningene delvis har endret konsumatferd som følge av føringen.

Han mener også det vil være vanskelig å avgjøre årsakene bare ved å studere tallmaterialet.

Turner har gjennomført tilsvarende undersøkelser fra en senere forbruksundersøkelse med to-ukers bokføringsperiode. Særlig har han konsentrert seg om å studere utgifter til te. Han finner bl.a. signifikant fall i utgiftsbeløp fra første til annen uke, blant de husholdninger som har utgifter til te i begge uker. Videre finner han at det er flere som har utgifter til te bare i første uke enn bare i annen (signifikant forskjell).

Disse engelske undersøkelser tyder ikke på at de funne resultater avhang av årstid, geografiske områder eller spesielle husholdningstyper.

Det kunne være av interesse å få utført lignende undersøkelser på norske forbruksdata. Både ved forbruksundersøkelsen i 1958 og 1967 ble det brukt en måneds bokføringsperiode. Uheldigvis er ikke datoene for utgiftsbeløpene blitt kjørt inn på bånd. Det kunne kanskje være av interesse å få punchet opp en del bøker.

Som det foreløpig forsøk har jeg gått gjennom en del bøker fra 1967 for hånd (godt og vel hundre).

En måneds periode kan for en del goder betraktes som en hel "syklus", mange betaler husleie, trikkekort, bladpenger osv. en gang i måneden. En må av denne grunn vente å finne systematiske variasjoner i faktiske utgifter over en måned, særlig fordi de nevnte utgifter ofte er av betydelig størrelse. En kunne selvsagt utelukke disse utgifter og bare studere andre utgifter, men utgifter som husleie o.l. vil antakelig - via budsjettbetingelsen - påvirke også andre utgiftsbeløp.

Jeg har valgt å se på antall registrerte utgiftsposter (inkludert bruk av varer fra egen produksjon) og studere hvordan dette varierer over de forskjellige uker innen måneden. Jeg har delt inn måneden i fire perioder à 7 dager, altså f.o.m. 1.-t.o.m. 7., f.o.m. 8.-t.o.m. 14., f.o.m. 15.-t.o.m. 21. og f.o.m. 22.-t.o.m. 28. i måneden. Her vil alle ukedager bli representert akkurat én gang i hver delperiode. Et forhold som fordelingen av lønningsdager over måneden må antas å spille en rolle for fordelingen av antall utgiftsposter (og utgiftsbeløp) for lønnstakere. For selvstendige påvirker kanskje fordelingen av inntektsutbetalingen over måneden i mindre grad utgiftsfordelingen. Jeg har sondret mellom gruppene ansatte, selvstendige og øvrige (trygdede og pensjonister). For å unngå utslag p.g.a. de store høytider har jeg utelatt månedene desember og mars (påsken falt i mars i 1967). De geografiske områder som er tatt med her er: område 3 i Oslo (Sandaker, Torshov) (37 hush.), Asker (25 hush.), Rennebu (34 hush.) og Harstad (22 hush.).

Fordeling av antall registrerte vareposter

Konsumentgruppe	Delperiode i måneden				Antall observasjoner
	1.-7.	8.-14.	15.-21.	22.-28.	
Ansatte	3 663	3 534	3 326	3 417	66
Selvstendige ..	1 344	1 231	1 201	1 194	32
Øvrige	790	538	543	555	20
Alle grupper ..	5 797	5 303	5 070	5 166	118

En ser at antall utgiftsposter for 1. uke er større enn for de andre ukene for alle konsumentgruppene. Det er av interesse om forskjellene er signifikante. For å avgjøre dette måtte en først vite noe om hvilken teoretisk fordeling antall utgiftsposter pr. uke hadde. Jeg har regnet med en normalfordeling. Dette er antakelig urealistisk, bl.a. må regne med at fordelingen av antall utgiftsposter er skjev (en "langhale" mot høyre). Standardavvik for antall ukentlige utgiftsposter pr. husholdning (for en vilkårlig uke) er ca. 20. Med dobbelsidig test med signifikansnivå 0,05 viser det seg at antall utgiftsposter 1. uke er signifikant forskjellig fra de andre ukene. Mellom ukene 2, 3 og 4 er det ikke signifikante forskjeller. Dette gjelder for alle konsumentgrupper samlet.

For å lette sammenligningen beregnes den prosentvise fordelingen av antall utgiftsposter på de fire ukene for konsumentgruppene.

Prosentvis fordeling av antall utgiftsposter over ukene:

Konsumentgruppe	Delperiode i måneden				Sum 1.-29.
	1.-7.	8.-14.	15.-21.	22.-28.	
Ansatte	26,3	25,4	23,9	24,5	100,0
Selvstendige ..	27,0	24,8	24,2	24,0	100,0
Øvrige	32,6	22,2	22,4	22,9	100,0
Alle grupper ..	27,2	24,9	23,8	24,2	100,0

Som sagt ligger alle grupper høyest i alle ukene, men at dette er mest markert for gruppen "Øvrige". Her er det imidlertid bare 20 observasjoner, så en bør vel være forsiktig med å legge for mye i tallene.

Videre synes det ikke å være noen markert forskjell mellom fordelingen på ukene for ansatte og selvstendige.

Materialet viser at det er en del husholdninger med ekstremt skjev fordeling av antall utgiftsposter over ukene (muligens dårlig tilnærming til normalfordelingen). Det kan derfor være av interesse å se på fordelingen mellom ukene når det gjelder h.h.v. størst og minst antall utgiftsposter (ved likhet mellom to eller flere uker er alle tatt med). En finner:

	Uke 1	Uke 2	Uke 3	Uke 4	Total
Tilfelle med flest poster	53	32	18	22	125
" " færrest "	22	34	40	28	124

Av tabellen fremgår bl.a. at det var 1. uke som hadde flest observasjoner med størst antall og færrest med minst antall utgiftsposter. Omvendt hadde tredje uke flest observasjoner med minst antall og færrest med størst antall utgiftsposter. Dette funn passer godt med det en fant foran angående fordelingen av antall utgiftsposter på ukene.

Det finnes flere muligheter når det gjelder hvilke typer poster som er årsak til ulikheter mellom ukene. Det ble foran antydnet at det kunne tenkes det var tatt med en del poster første uke som ikke skulle vært med. En kan kanskje anta at dette særlig er store poster (varige konsumgoder). Det kunne derfor være av interesse å se om det er ulikheter mellom ukene m.h.t. sammensetningen av konsumet. Dette vil i så fall være meget uheldig fordi tallmaterialet da trolig kan inneholde systematiske feil i den relative fordeling mellom godegrupper. For å undersøke dette måtte en undersøke hvordan ukenes utgifter, evt. antall utgiftsposter for de forskjellige varegrupper, fordelte seg på ukene. En kunne her ha en antakelse om at ved evt. forskjeller ville det trolig være 1. uke som var "feil", mens de andre var "riktige". Jeg har ikke for øyeblikket anledning til å undersøke hypotesen på noe tallmateriale for Norge. Det kan imidlertid være av interesse å se nærmere på de tall som Kennsley fant. Han beregnet ukentlige utgifter for en 3-ukers periode for de forskjellige varegrupper. Med utgangspunkt i tabell 1 i hans artikkel har jeg beregnet prosentvis fordeling for forskjellige utgiftsgrupper for hver av de tre uker.

Prosentvis fordeling av ukentlige utgifter

Utgiftsgruppe	1. uke	2. uke	3. uke	Alle tre uker
Matvarer	36,0	37,4	37,3	36,9
Drikkevarer og tobakk	10,6	11,3	11,1	11,0
Klær og skotøy	14,8	12,9	14,7	14,2
Husholdningsartikler .	7,7	8,1	6,7	7,5
Andre varer	7,7	7,9	7,8	7,8
Reiser	6,8	6,4	6,1	6,4
Tjenester	9,3	9,3	9,6	9,4
Andre utbetalinger ...	4,1	4,0	3,9	4,0
Hus og brensel	3,0	2,7	2,7	2,8
Alle grupper	100,0	100,0	100,0	100,0

Hvis hypotesen om "fasadekjøp" i første uke er riktig, skulle en vente relativt høye utgifter til husholdningsartikler, klær og skotøy og lave andre utgifter, f.eks. matvarer. Matvareutgiftene er relativt lavest i første uke. For klær og skotøy er 1. uke høyest, men ikke for husholdningsartikler. Utslagene synes så lite markerte at en neppe bør hevde at tallene underbygger hypotesen.

En annen mulighet angående årsaker til forskjell mellom ukene kan være at regnskapsførernes glemsomhet øker jo lenger de har drevet med bokføringen. En må videre vente at det er særlig de mindre utgiftsposter som blir glemt. En kunne kanskje tenke seg å sammenligne gjennomsnittsutgifter pr. post for de forskjellige ukene, men som nevnt foran vil utgifter som husleie o.l. som ofte faller innen en bestemt del av måneden vanskeliggjøre tolkningen av resultatene.

En kunne alternativt studere forskjeller mellom ukene i måneden m.h.t. størrelsesfordelingen på postene, f.eks. utgiftsposter på 0-3 kr. pr. post, 3-5, 5-10 osv. En skulle da kunne komme på spor etter eventuell utelatelse av småposter eller sammenslåing av småposter hvis dette har foregått i større utstrekning mot slutten enn tidligere i regnskapsmåneden. Som nevnt foran er det et problem at datoene for postene ikke er punchet inn på bånd. Jeg har gått gjennom en del bøker for hånd (48 bøker; 26 for Asker og 22 for Harstad). Månedene mars og desember er utelatt. Etter litt prøveregning bestemte jeg meg for å gruppere postene i tre grupper: 0-2 kroner, 2-6 kroner og 6 kroner eller mer pr. post. Summering av alle bøker gav følgende tall:

Antall utgiftsposter gruppert etter utgiftsstørrelse pr. post

Uke	0-2 kroner	2-6 kroner	6 kroner eller større	Alle poster
1. uke (1.-7.)	1 022	1 229	585	2 836
2. " (8.-14.)	882	1 138	524	2 544
3. " (15.-21.)	897	1 131	506	2 534
4. " (22.-28.)	876	1 014	557	2 447
Alle uker (1.-28.)	3 677	4 512	2 172	10 361

Samme tabell: Prosentvis fordeling

Uke	0-2 kr.	2-6 kr.	Over 6 kr.	Alle poster
1. uke	36,0	43,3	20,6	100,0
2. uke	34,7	44,7	20,6	100,0
3. uke	35,4	44,6	20,0	100,0
4. uke	35,8	41,4	22,8	100,0
Alle uker	35,5	43,5	21,0	100,0

En bør antakelig være forsiktig med å trekke konklusjoner når ikke observasjonsmaterialet er større enn her. Materialet tyder imidlertid ikke på at omfanget av svikt i registreringen av småposter eller sammenslåing av disse til større poster relativt tiltar mot slutten av måneden.

En hypotese som ble nevnt foran går ut på at en p.g.a. "fasadementalitet" kan få registrert for stort kjøp av varige konsumgoder i første uke. Dette skulle en her vente ville gi seg utslag i relativt mange store utgiftsposter i 1. uke. Materialet tyder ikke på at så er tilfelle. (Materialet kan selvsagt ikke bevise noe om dette, bl.a. kan det av en eller annen grunn være få av andre store utgiftsposter i første uke.)

Det fremgår av tabellene at prosent antall poster på seks kroner eller mer er størst i 4. uke (f.o.m. den 22. t.o.m. den 28.). Dette mener jeg skyldes at månedlig skattetrekk, trygdetrekk, pensjonstrekk o.l. blir utgiftsført særlig i denne uken.

Den konklusjon jeg kommer til må bare bli at jeg ikke kan forklare årsaken til fallet i antall registrerte utgiftsposter utover i registreringsperioden.

Det norske tallmaterialet har en vesentlig ulempe når det gjelder bruk av det til å belyse variasjon over måneden i registreringen; nemlig fordi bokføringsperioden er én måned. En måned er på flere måter en "økonomisk" periode, en har månedslønn, månedlige skattetrekk, trygdetrekk, husleieutbetaling osv. Dette vil vanskeliggjøre en tolkning av observerte variasjoner i utgifter og antall utgiftsposter over bokføringsperiodene. Hvis en hadde tid og anledning, kunne det være av interesse å skaffe seg adgang til tallmateriale fra andre land. Momentet bør kanskje også tillegges betydning ved planlegging av senere forbruksundersøkelser i Norge, særlig ved eventuelle prøveundersøkelser.

OS/IH, 17/9-71

OM PRINSIPPER FOR REGISTRERING AV FORBRUK VED FORBRUKSUNDERSØKELSER

Av Odd Skarstad

Innhold

	Side
I Generelt om registreringsmåter for privat forbruk	31
a. Om registreringsprinsipper for goder som omsettes	31
b. Om registreringsprinsipper for egenproduserte goder	32
II Spesielle registreringsproblemer	33
a.1 Kjøpte matvarer	33
a.2 Egenproduserte matvarer	35
b. Tobakk og drikkevarer	37
c. Klær og skotøy	37
d. Bolig, lys og brensel	38
e. Møbler og husholdningsartikler	39
f. Transportmidler	39
g. Leid hjelp i hjemmet	40
h. Utgifter til ferie	40
i. Oppsummering	40
I II Om revisjon av regnskapsbøker i "felten"	41

I. GENERELT OM REGISTRERINGSPROBLEMER

a. Om registreringsprinsipper for goder som omsettes

Det kan tenkes flere mulige prinsipper for registrering av forbruk ved forbruksundersøkelser. En tenker seg ofte følgende muligheter:

- a) Kjøpsavtaleprinsippet
- b) Betalingsprinsippet
- c) Leveranseprinsippet
- d) Forbruksprinsippet

Kjøpsavtaleprinsippet betyr som en forstår at det er inngåtte kjøpsavtaler som registreres. Hvis det f.eks. kjøpes en bil, skal dette registreres når kjøpsavtalen inngås, og uavhengig av når bilen leveres eller betales. En slik registrering vil antakelig virke forvirrende og være vanskelig gjennomførbar ved en undersøkelse der opplysningene delvis innhentes ved at husholdningene selv fører regnskap.

Betalingsprinsippet består i å registrere godtgjørelse (motytelser) for leverte goder. I en økonomi med pengehushold registreres dette som uttelling i form av penger eller andre betalingsmidler. Ved tidligere forbruksundersøkelser i Norge er det dette prinsipp som har vært brukt. En grunn til dette er at en har ansett uttellingene lette å huske og dermed lette å registrere for regnskapsførerne.

Ved leveranseprinsippet registreres de varer og tjenester husholdningen mottar i regnskapsperioden. Leveranse- og betalingsprinsippet vil ofte være sammenfallende hvis alle leveringer foregår mot kontant oppgjør. For en del goder, særlig tjenester (f.eks. boligjenester) vil det aldri være samtidighet mellom betaling og leveranser.

Ved forbruksprinsippet er det altså selve forbruket av godene en søker å registrere. Prinsippet synes ikke ugjennomførbart for en del goder som på en direkte måte forbrukes i regnskapsperioden (matvarer f.eks.) men de praktiske vansker en får med avskrivninger ved varer som er av en viss varighet byr på store problemer.

Det er naturlig å vurdere registreringsprinsippene i relasjon til formålene med forbruksundersøkelsene.

To viktige formål med tallmaterialet fra forbruksundersøkelsen er at det skal danne grunnlag for vektene i konsumprisindeksen og gi nasjonalregnskapstall for fordelingen av det private konsum på de enkelte goder og godegrupper. En er her interessert i total- og gjennomsnittstall for hele befolkningen. I den utstrekning de totale lagerendringer av varer i registreringsperioden er små, vil leveranse- og forbruksprinsippet gi samme resultater. Hvis heller ikke kredittmønsteret endrer seg i perioden, vil dessuten betalingsprinsippet gi samme resultater som de to nevnte. Det er neppe trolig at lagerendringer og endringer i kredittmønsteret er relativt særlig store for de fleste forbruksvarer i løpet av registreringsperioden. Den praktiske gjennomførbarhet synes derfor å bli avgjørende for valget mellom registreringsprinsippene. Som regel velger en derfor enten betalings- eller leveranseprinsippet (se ellers senere).

b. Om registreringsprinsipper for egenproduserte goder

Forbruk av egenproduserte goder kan registreres ved

- a) produksjonen eller
- b) forbruket av godene.

Mengdeberegningene vil by på spesielle problemer ved egenproduserte goder, fordi beregningene må foretas av husholdningene selv, uten å være en nødvendig følge av en omsetningsprosess. Registrering av mengder kan foregå enten ved bokført forbruk fra dag til dag eller for lengre perioder (f.eks. pr. uke eller måned). Tilsvarende kan en registrere forbruk ved intervju (daglig, ukentlig, månedlig, årlig).

Prisfastsettingen på forbruk av egenproduserte goder må foretas her i Byrået. I forbindelse med prisfastsettingen ved egenproduserte goder må en ta stilling til hvilke priser en skal benytte: produsert- eller forbrukerpriser, evt. om en bør ta sikte på å skaffe begge deler. Det er som sagt meningen å benytte oppgaver fra forbruksundersøkelser til å beregne fordelingen på de forskjellige varer og tjenester for privat konsum i nasjonalregnskapet. I nasjonalregnskapet er forbruk av egenproduserte varer vurdert til produsentpriser. Dette tilsier at en bør registrere forbruk av egenproduserte varer til produsentpriser også ved forbruksundersøkelsene.

En annen anvendelse av forbruksdata er som sagt til beregning av vektene i konsumprisindeksen. Her er det rimelig å benytte prissett som viser hva f.eks. reduksjon i forbruket av en vilkårlig vare med én enhet vil bety i innsparing på budsjettet. Dette tilsier at en bør registrere

forbruk av egenproduserte varer til produsentpriser (innsparinger på budsjettet avhenger jo av salgsprisen). (I denne sammenheng er å nevne at en idéelt også burde legge til grunn produsentpriser på egenproduserte varer ved utregning av konsumprisindeksen for "hele landet" siden hensikten med indeksen er å belyse endringer i gjennomsnittlige leveomkostninger over tiden.)

Konklusjonen blir at det synes mest påkrevet å få beregnet verdi av forbruk av egenproduksjon til produsentpris. Hvor vidt en også skal beregne til detaljpris, bør vel vurderes nærmere.

II. SPESIELLE REGISTRERINGSPROBLEMER

a.1 Kjøpte matvarer (dagligvarer)

Som nevnt foran har en benyttet betalingsprinsippet ved regnskapsføringen ved tidligere forbruksundersøkelser i Norge. En registrerer altså forbruket på det tidspunkt utbetalingene foregår. Hvis en husholdning f.eks. skaffet seg en halv gris, så ble grisen regnet som konsumert når betalingen foregikk, uansett når grisen ble levert og uansett når den ble spist. Det synes opplagt at det for den enkelte vare og den enkelte husholdning kan bli betydelige avvik mellom utbetalinger og forbruk, selvsagt avhengig av lagerendringer og endringer i størrelse og sammensetning av husholdningens matvaregjeld. Argumenter til forsvar for betalingsprinsippet er at lagerendringene vil være relativt små for den enkelte vare for alle husholdningene i utvalget (de individuelle avvik vil jevne seg ut i et stort materiale), og at kredittmønstret innen kolonialhandelen endres lite i løpet av det tidsrom forbruksundersøkelsen pågår. Hvis en mener kredittmønstret kan endres vensetlig i løpet av registreringsperioden, vil det være et argument for å velge leveranseprinsippet fremfor betalingsprinsippet. Men slike endringer over tiden er neppe store for denne varegruppe.

I den norske forbruksundersøkelsen i 1967 tok man sikte på å registrere både utbetalinger og leveranser ved avbetalingskjøp av varige goder. Hvis en husholdning f.eks. kjøpte en bil til 15 000 kroner og betalte 5 000 kroner kontant, så skulle en registrere 5 000 kroner i kontantrubrikken og 15 000 kroner i kreditttribrikken. Nå viste det seg at registrert kredittkjøp utgjorde forholdsvis lite totalt. Dette førte til at man lot være å bearbeide særskilt oppgaver over bokført kredittkjøp. Å ha egen rubrikk i regnskapsbøkene til registrering av avbetalingshandel kan synes å være misbruk av plass i bøkene. F.eks. kan det tenkes plassen heller bør nyttes til registrering av mengdeoppgaver (for matvarer). I 1967-undersøkelsen hadde

Ved 1967-undersøkelsen viste betalingsprinsippet seg å ha en vesentlig svakhet ved registreringen for husholdninger som handler "på bok". Varene som ble anskaffet skulle ikke registreres før de ble betalt. Hvis en husholdning betalte f.eks. 1 000 kroner til kjøpmannen for varer levert tidligere (over et lengre tidsrom), var det meningen dette skulle spesifiseres i regnskapet. Dette var det en del husholdninger som unnlot å gjøre. Videre kunne det tenkes de 1 000 kronene var bare avdrag på gjeld. I så fall var det umulig å vite hvilke varer utgiftene refererte seg til. Tilsvarende forhold ville en få hvis en husholdning handlet for f.eks. 100 kroner, men betalte bare 50 kroner straks. Hvilke varer skulle en si de 50 kroner gjaldt? Nå viser det seg, når en studerer regnskapsbøkene fra 1967, at en del husholdninger har misforstått hensikten med rubrikken for kredittkjøp i regnskapsbøkene. De har der åpenbart registrert leverte varer kjøpt på bok. I kontantrubrikken har de så registrert uspesifiserte større beløp "betalt til kjøpmannen". Ved bearbeidingen av materialet er alle spesifiserte beløp registrert som kontantkjøp, mens de store uspesifiserte beløpene i slike tilfelle er strøket. Det er dermed åpenbart at i en del av de tilfellene hvor utbetalinger og leveranser avviker fra hverandre i tid, er det leveranser som er blitt registrert (leveranseprinsippet blitt fulgt). I den grad regnskapsførerne har spesifisert utbetalingene kan det også ha forekommet dobbeltregistreringer når en bearbeidet på denne måten.

Det synes klart at kredittkjøp av kolonialvarer og de vanskeligheter det medfører, er forhold en må ta hensyn til ved opplegg av senere forbruksundersøkelser. I en del geografiske områder (særlig på landsbygda) og for en del yrkesgrupper (fiskere f.eks.) er antakelig kjøp på bok hos kolonialkjøpmannen nokså vanlig. Ved dårlige oppgaver fra disse områder og yrkesgrupper kan en få systematiske skjevheter i materialet.

For å få en registreringsmetode som er bedre enn den tidligere ved kjøp av dagligvarer på bok, finnes åpenbart to muligheter:

- 1) En kan gjennomføre en total omlegging til leveranseprinsippet
- 2) En bruker betalingsprinsippet som før for de fleste husholdninger, men for dem som kjøper dagligvarer på bok, benyttes leveranseprinsippet.

Leveranseprinsippet brukes i Danmark, Sverige og Finland. I Finland er det gjort unntak for husleie, elektrisitets- og telefonregninger, sykehusutgifter, utgifter til seriebilletter og avisabonnement hvor betalingsprinsippet nyttes. En måtte antakelig gjøre liknende unntak ved en eventuell overgang her i landet også, idet leveranseprinsippet er problematisk ved tidsutstrakte leveranser (en del tjenester) og når

utgiftsbeløpet er ukjent på leveringstidspunktet (telefonregninger f.eks.). Tilhengere av betalingsprinsippet fremhever dets styrke, d.v.s. at det er lett å forstå og at utbetalinger er lette å huske. Derfor vil de frarå en generell omlegging til leveranseprinsippet.

Alternativ 2) innebærer bruk av leveranseprinsippet bare for de - husholdninger som handler på bok. En må her ved intervju eller på annen måte skaffe rede på hvilke husholdninger i utvalget som handler på bok. Når det gjelder den praktiske siden ved gjennomføringen, f.eks. om en må operere med to typer regnskapsbøker, så må dette vurderes nærmere i forbindelse med utarbeidelsen av bøkene.

Hvis en finner at det ikke byr på for store praktiske vanskeligheter å operere med forskjellige prinsipper, vil jeg foreslå alternativ 2).

a.2 Egenproduserte matvarer

Registrering av forbruk av egenproduserte goder er som nevnt ofte vanskelig. Dette gjelder også matvarer. Hvilke registreringsprinsipper en bør velge vil avhenge av varens art (d.v.s. når den produseres, konsumeres, lagerhold, husholdningens evne og vilje til å gi riktige oppgaver m.m.). Jeg tar nå for meg noen av de vanligste varer.

i) Kjøtt, kjøttvarer og fleisk

I forbruksundersøkelsen i 1967 ble forbruk av egenprodusert kjøtt og fleisk registrert på det tidspunkt det ble konsumert (forbruksprinsippet). Hvis f.eks. en bonde slaktet en sau til eget bruk ble dette ikke registrert. En registrerte det som ble spist fra dag til dag. Til sammenlikning ble et innkjøpt saueslakt altså registrert ved anskaffelsen (betalingstidspunktet). En kan stille spørsmål om ønskeligheten av å ha forskjellige prinsipper her, m.a.o. om en ikke burde registrere levert kjøtt og fleisk fra gårdsdriften til husholdningen på leveringstidspunktet. En slik registreringsmåte ville være lettvintere for husholdningen, og en burde kunne få vel så nøyaktige oppgaver som når de må huske på hver dag å veie det de spiser. Opplysningene kunne innhentes ved intervju. Jeg tror de fleste gårdbrukere vet noenlunde hvor mye oksekjøtt, fleisk osv. de har produsert for eget bruk i løpet av ett år. Det er imidlertid et spørsmål om alle gårdbrukere ville være villige til å oppgi hvor mye de har slaktet til eget forbruk i løpet av et år. Produksjon av kjøtt og fleisk skal det jo forøvrig gis oppgaver over i selvangivelsene, men det kan være tvil om hvor pålitelige disse oppgaver er.

Min konklusjon er at vi bør holde fast ved det tidligere prinsippet her (forbruksprinsippet) til tross for at det virker noe tungvint for husholdningene under føringen av regnskapene. Husholdningene kjenner neppe ukentlig eller månedlig forbruk, og en må vel ta sikte på å registrere forbruk fra dag til dag ved bokføring som tidligere gjort.

ii) Fisk og fiskevarer

I 1967 ble forbruk av fisk fra egen produksjon registrert på samme måte som kjøtt. En bør vel holde fast ved det tidligere prinsipp også her. Når det gjelder fiske til eget bruk, er det vel også vanskelig for husholdningen å huske omfanget av dette over et lengre tidsrom, hvorved intervju blir en dårlig metode her. Det vil neppe være lettere å få registrert det som blir fisket til eget bruk hver dag enn det som blir brukt.

iii) Melk og egg

Melk og egg produseres i vesentlig grad for eget bruk. Disse varer spises både som de er og særlig ofte som ingredienser i annen mat. Dette gjør det vanskelig å finne gode registreringsprinsipper, i hvert fall hvis en skal benytte en eller annen utgave av forbruksprinsippet.

Jeg tror de fleste gårdbrukerfamilier vet noenlunde hvor mye melk av egen produksjon de bruker pr. dag, pr. uke eller pr. måned. Jeg foreslår at en ved siden av bokføring av det daglige forbruk også registrere f.eks. ukeforbruk ved intervju.

Forbruk av egg vil en vel få de sikreste oppgaver over ved å få bokført forbruk fra dag til dag. Men jeg vil foreslå at det dessuten spørres om ukentlig forbruk ved intervju, i hvertfall i kommende prøveundersøkelse.

iv) Poteter og grønnsaker

Også her er det tidligere brukt forbruksprinsippet, d.v.s. registrert forbruk fra dag til dag. Det bør en antakelig gjøre for fremtiden også. Jeg tviler på om produsentene kjenner f.eks. årsproduksjonen til eget bruk her.

v) Bær og frukt

Tidligere er produksjonsprinsippet blitt brukt. En registrerte via regnskap hva som ble syltet i syltetida. Dette kunne en sikkert registrere ved intervju også. De fleste husmødre som sylter selv vet trolig hvor meget de syltet siste høst av de forskjellige bærslag. For disse varere er det antakelig vanskelig å registrere forbruk fra dag til dag.

b. Tobakk og drikkevarer

En har gjort den erfaring tidligere ved forbruksundersøkelser (også her i landet) at en har fått dårlige oppgaver over tobakk- og alkoholforbruket (for lave tall jamt over). Delvis kan dette skyldes at det husholdningsmedlem som fører hovedheftet ikke kjenner til hele husholdningens forbruk. Hvis f.eks. husmoren fører regnskapet, vet hun ofte ikke hva ektemannen kjøper av disse varer. Det er mulig at en mer utstrakt bruk av egne regnskapshefter for husholdningens andre medlemmer kunne hjelpe på dette (de såkalte Ø-heftene). Videre hersker det antakelig ofte uvilje mot å oppgi slikt forbruk.

I den finske forbruksundersøkelsen 1966 ble det stilt kontrollspørsmål om tobakk- og alkoholforbruket i forbindelse med innsamlingen av regnskapsheftene. Når det gjelder tobakkforbruket, synes jeg man kunne gjøre det samme her i landet. Svært mange har trolig noenlunde greie på sitt forbruk av sigaretter, røyketobakk osv. pr. uke, samtidig som de fleste neppe har større motvilje mot å gi pålitelige svar. Jeg foreslår at utgifter til tobakk registreres både via regnskap og intervju. Når det gjelder kontrollspørsmål om alkoholforbruket, er det mer usikkert om man ville få pålitelige svar i vårt land, selv om man kanskje fikk det i Finland. Innstillingen til å gi slike opplysninger er muligens noe spesiell i Norge.

c. Klær og skotøy

I forbruksundersøkelsen 1967 ble kjøpte klær registrert ved én måneds bokføring på samme måte som dagligvarer. Kjøp av en del større klesartikler som frakker, kapper, dresser, kåper og drakter ble dessuten registrert ved årsintervjuet. Ved forbruksundersøkelsen i Sverige 1969 hadde man én måneds bokføringsperiode, men bokførte dessuten i ytterligere to måneder utgiftsposter på over 25 kroner (ikke matvarer). Kjøp av klær og skotøy utgjorde antakelig en vesentlig del av disse postene. Etter min oppfatning er tre måneder en noe lang bokføringsperiode. En kan lett risikere at husholdningene blir lei bokføringen.

I Finland gjorde man i 1966 forsøk på å sammenlikne resultatene av bokføring og intervju som registreringsmåte for utgifter til klær og skotøy. Det ble foretatt intervju for januar og februar. Setter en bokførte utgifter til 100 får en følgende tall:

Konsumvare:	Intervju (bokføring=100)
Ytterklær	91
Sko	88
Diverse ekviperingsartikler (luer, hansker, slips o.l.)	74
Underklær	65
Tøyer og garn	60

Tallene antyder den tendens at intervjumetoden gir lavere tall enn bokføringsmetoden og at dette er mer utpreget desto mindre kjøp det dreier seg om.

Analyser av Petra Vestbye på tallmaterialet i den norske forbruksundersøkelsen 1967 viste også uoverensstemmelser mellom tall for årsintervju og månedsregnskap over utgifter til klær. Her gikk imidlertid uoverensstemmelsene stort sett den andre veien, i det årsintervjuet som regel ga høyere utgifter enn månedsregnskapet (dette gjaldt alt sammen større bekledningsartikler). Det ville vært av stor interesse å forsøke å analysere årsakene til uoverensstemmelsene. Inntil videre bør en vel ta sikte på som tidligere å registrere utgifter til mindre klesartikler bare ved regnskap og de større dessuten ved årsintervju.

Egenproduksjon av klær er antakelig vanskelig å registrere, ikke minst er prisfastsettingen problematisk. Det en har registrert er innkjøpt garn, tøyer o.l. Produksjon av garn av egen ull foregår antakelig i bare liten utstrekning og skulle vel være neglisjerbar.

d. Bolig, lys og brensel

I 1967-undersøkelsen ble husleie registrert ved månedsbokføringen. Ved årsintervjuet ble eierforhold, skattelikning, driftsutgifter, byggearbeid i forbindelse med boligen, registrert. En del husholdninger glemte å ta med husleie i månedsregnskapet; noen husholdninger betaler heller ikke månedlig husleie (visse former for framleie e.l.). Husleie bør vel også kunne registreres ved intervju. Hvis en her skulle bruke leveranseprinsippet, må en beregne husleien for registreringsperioden. Hvis perioden er en annen enn én måned, ville dette prinsippet registreringsmessig sett være mer komplisert enn betalingsprinsippet.

Lys og brensel ble i 1967 registrert ved månedsbokføring, men dessuten ved årsintervjuet, idet en spurte om årsutgiftene til elektrisitet, kull, koks osv. Dette synes å være en brukbar framgangsmåte.

Egenprodusert ved er imidlertid problematisk. En bør vel kanskje prøve å få registrert ved intervju antall favn hjemkjørt i løpet av året, og som det ikke er betalt for. Forøvrig tror jeg bruk av egen ved som oppvarming har hatt sterkt avtagende betydning de senere år.

e. Møbler og husholdningsartikler

Ved 1967-undersøkelsen ble det spurt om utgifter til transportmidler, radio o.l., møbler, husholdningsmaskiner, tepper og malerier i årsintervjuet. En registrerte utbetalinger i 1967 og dessuten varens hele kostnad (bare for varer kjøpt på avbetaling i 1967). Dette innebærer at hvis f.eks. en husholdning kjøpte en vaskemaskin på avbetaling mot slutten av 1966, så var det bare avdragene i 1967 som skulle registreres. I visse tilfelle var det antakelig vanskelig for husholdningene å huske hvor mange av avdragene som falt i 1967. En bør vel derfor ta sikte på å registrere leveranser i årsintervjuet. I regnskapsbøkene blir bare utbetalingene registrert.

f. Transportmidler

I 1967 ble det registrert kjøp av bil, både ny og brukt bil, der utbetalingene og leveranse ble registrert. I innledningsintervjuet ble folk spurt om bilen var kjøpt ny. Det kunne vel være av interesse å utvide undersøkelsen litt på dette punkt ved også å registrere bilens alder og nybilpris, og kanskje også bilmerke. Disse opplysninger kunne gi grunnlag for beregninger av avskrivninger (noe som kanskje også kunne gjøres maskinelt). Forøvrig finnes det prislister over de forskjellige typer og årganger, og disse kunne også brukes til slike beregninger av årlige avskrivninger.

Registrering av nypris og alder på bilene kan være av betydning også for skatteforskerne, i det det f.eks. kan gi grunnlag for å beregne inntektsfordelingseffekten av alternative måter for utforming av bilbeskatningen. I årsintervjuet bør vel bilkjøp registreres ved leveranser.

g. Leid hjelp i hjemmet

Leid hjelp i hjemmet ble registrert ved årsintervju i 1967. Her vil det være korrekt å ta med lønn til hushjelp som utgift selv om hushjelpen er medlem av husholdningen. En må da også regne med hushjelpens inntekt i husholdningens inntekt. Registreringen kan vel gjøres ved årsintervjuet som tidligere.

h. Utgifter til ferie

Feriereiser utenlands og innenlands bør registreres ved årsintervju. Dette er antakelig nødvendig for å få tilstrekkelig mange med utgifter til ferie i utvalget.

i. Oppsummering av konklusjoner (foreløpige forslag)

Jeg foreslår at forbruk blir registrert ved

- 1) Innledningsintervju
- 2) Regnskapsføring for en kortere periode (2-4 uker)
- 3) Årsintervju

Regnskapsføring er særlig nødvendig av hensyn til registreringen av dagligvarer, mens årsintervjuet er nødvendig av hensyn til registrering av større artikler som anskaffes relativt sjelden. Hvis det er aktuelt å skjære ned bokføringsperiodens lengde fra en måned (som den var i 1958 og 1967) til f.eks. 2 uker, vil dette gjøre nødvendigheten av å ha årsintervju større. Det er ellers vanskelig å si noe om hva som er den "beste" metoden for registrering av forbruk av de forskjellige goder: regnskap eller intervju, så lenge en ikke kjenner årsakene til forskjellene de medfører i resultatene.

I regnskapsbøkene har jeg foreslått som hovedregel å bruke betalingsprinsippet, dvs. at en bare registrerer utbetalinger. For noen husholdninger bør en av hensyn til dagligvarekjøp på bok bruke leveranseprinsippet (med unntak for noen goder der betalingsprinsippet brukes). I regnskapsbøkene registreres alle utbetalinger (h.h.v. leveranser) til såvel varige som ikke-varige forbruksgoder.

I årsintervjuet registreres leveranser av større artikler. I 1967 gjaldt dette utgifter til utbedring og vedlikehold av bolig, utbetalinger til lys, kjøpt brensel, større bekledningsartikler, husholdningsartikler og møbler, kjøp av transportmidler og driftsutgifter til disse, leid hjelp i hjemmet og ferieutgifter.

Ukentlig forbruk av tobakksvarer registreres ved intervju av husholdningens medlemmer i forbindelse med kommende prøveundersøkelse. Utgiftene bokføres også i regnskapet.

Forbruk av egenprodusert kjøtt, fisk, melk, egg, poteter og grønnsaker registreres ved bokført forbruk fra dag til dag.

Årsforbruk av egenprodusert brensel registreres ved årsintervjuet.

Produksjon av syltetøy registreres ved regnskap når syltingen foretas (som tidligere). Dessuten intervjues om årlig produksjon av syltetøy i årsintervjuet.

III. OM REVISJON AV REGNSKAPSBØKER I "FELTEN"

Problemet er om revisjonsarbeidet med husholdningsregnskapene skal foregå bare i Byrået, eller skal intervjuerne pålegges å gå gjennom regnskapene "på stedet" for å rette opp eventuelle feil og ufullstendigheter før bøkene sendes hit. Revisjonsarbeid av intervjuerne må i tilfelle foregå etter en bestemt instruks gitt av oss.

Her i landet har praksis til nå vært at intervjuerne ikke skal se bokføringsbøkene etter at bokføringen var begynt. Begrunnelsen for dette har vel vært at en fryktet at intervjuernes "kontroll" kunne innvirke på bokføringen, på den måten at de som førte bøkene ville prøve å vise et "bedre" forbruksmønster enn deres egentlige. F.eks. kunne det tenkes at forbrukstall for alkohol- og tobakksvarer kunne bli systematisk for lave.

I Sverige og Finland har man latt lokalombudet gå gjennom regnskapsbøkene. I Sverige f.eks. ble bøkene gått igjennom ellefte dagen og etter bokføringsperiodens slutt. Ved bokføringsperiodens slutt ble bøkene gransket etter en bestemt plan og kompletterende opplysninger innhentet.

Svenskene lot videre alle personer over 12 år føre en egen bok for personlige utgifter. På denne måten skulle det kanskje være mulig bl.a. å få registrert bedre enn ellers f.eks. husfarens små utgifter (tobakk, aviser o.l.). (Her i landet har bruken av de såkalte Ø-heftene vært mer begrenset bl.a. pga. budsjettmessige hensyn.) Så vidt jeg skjønner lot man i Sverige lokalombudet gå gjennom disse regnskapsheftene på samme måte som hovedheftene.

Jeg vil foreslå at det i den kommende prøveundersøkelsen blir eksperimentert med å la intervjuerne kontrollere en del av bøkene.

GD/WA, 4/5-72

OM FRAFALLSPROBLEMET VED STATISTISKE UNDERSØKELSER.

EN DRØFTING AV PROBLEMET SPESIELT I FORBINDELSE MED FORBRUKSUNDERSØKELSER

Av Grete Dahl

Innhold	Side
I. Generelle betraktninger	44
II. Begrep og definisjoner	44
III. Hva menes med frafall	45
IV. Frafallsårsaker og metoder til å redusere fra- fallet	47
V. Virkninger av frafall - teoretiske betraktninger .	50
VI. Korrigering for frafall	52
VII. Om frafallets størrelse ved en del norske under- søkelser	53
VIII. Tabellvedlegg	55

I. GENERELLE BETRAKTNINGER

Ved en utvalgsundersøkelse skiller en mellom tilfeldige og ikke-tilfeldige feil. De tilfeldige feil skyldes at en tar et utvalg, dvs. undersøker bare en del av telleenhetene i populasjonen istedenfor å foreta en fullstendig telling. Når utvalget tas som et sannsynlighetsutvalg kan en ved teoretisk-statistiske metoder gi anslag for størrelsesorden av disse tilfeldige feil.

De ikke-tilfeldige feil kan oppstå ved en fullstendig telling så vel som ved en utvalgsundersøkelse. Etter sin art deles de inn i følgende undergrupper

- i) frafallsfeil
- ii) målefeil
- iii) feil i forbindelse med revidering, koding, bearbeiding m.v.

Frafallsfeil forekommer ved praktisk talt alle statistiske undersøkelser. En god plan for gjennomføringen av en statistisk undersøkelse bør derfor inneholde forslag til løsning av de potensielle frafallsproblemer.

Dette notatet er et forsøk på å beskrive hvordan frafall kan oppstå ved en statistisk undersøkelse, samt å gjøre rede for de problemer frafallet kan forårsake.

En oversikt over det registrerte frafall ved norske forbruksundersøkelser etter 1945 og ved en del andre utvalgsundersøkelser foretatt ved Kontoret for intervjuundersøkelser etter 1967, finner en i tabellvedlegget, avsnitt VIII, tabellene A og B.

II. BEGREP OG DEFINISJONER

- | | |
|--------------------|--|
| Enhet (telleenhet) | - den elementære enhet ved undersøkelsen |
| Populasjon | - det antall - N - enheter som en vil undersøke |
| Utvalg | - de enheter - n - som trekkes ut for undersøkelsen |
| Svarsenheter | - de enheter - n_1 - i utvalget som en får opplysninger fra |
| Frafallsenheter | - de enheter - n_2 - i utvalget som en ikke får opplysninger fra (som faller fra). ($n_1 + n_2 = n$) |

III. HVA MENES MED FRAFALL

I en utvalgsundersøkelse på n enheter med n_1 svarsenheter og n_2 frafallsenheter, defineres den totale frafallsprosent ($100 \cdot k$) ved

$$k = \frac{n - n_1}{n} = \frac{n_2}{n}$$

De variable en observerer for utvalgseenhetene kan deles inn i grupper etter størrelsen på de variable, og en kan gi en detaljert beskrivelse av frafallet ved å angi størrelsen av det i disse gruppene eller for kombinasjoner av grupper av variable - tabellcellefracfall.

Tabellcellefracfallet kan belyses som i oppstillingene (a - c) nedenfor, hvor a_{lij} betegner estimatet for en parameter i populasjonen og n_{lij} er det antall svarsenheter estimatet er basert på.

a)

Geografisk inndeling \ Alder	Område 1		Område 2		Sum	
Gruppe 1	a_{111}	n_{111}	a_{112}	n_{112}	$a_{11.}$	$n_{11.}$
Gruppe 2	a_{121}	n_{121}	a_{122}	n_{122}	$a_{12.}$	$n_{12.}$
Sum	$a_{1.1}$	$n_{1.1}$	$a_{1.2}$	$n_{1.2}$	$a_{1..}$	$n_{1..}$

Første fotskrift refererer til typen av enheter - svarsenheter (1)/ frafallsenheter (2). Annen fotskrift refererer til variabelen "alder". Tredje fotskrift refererer til variabelen "geografisk inndeling".

Dersom en også kjenner fordelingen i de tilsvarende grupper (for variablene "alder" og "geografisk inndeling") av alle enheter i utvalget - $n_{.ij} = n_{lij} + n_{2ij}$ - kan en stille opp følgende tabell:

b)

Geografisk inndeling \ Alder	Område 1	Område 2	Sum
Gruppe 1	$n_{.11}$	$n_{.12}$	$n_{.1.}$
Gruppe 2	$n_{.21}$	$n_{.22}$	$n_{.2.}$
Sum	$n_{..1}$	$n_{..2}$	$n_{...}$

n_{2ij} er antall frafallsenheter i tabellcelle ij .

Frafallsprosenten ($100 \cdot k_{ij}$) i tabellcelle ij defineres ved

$$k_{ij} = \frac{n_{.ij} - n_{lij}}{n_{.ij}} = \frac{n_{2ij}}{n_{.ij}}$$

En fullstendig dokumentasjon av tabellcellefrafallet foreligger om samtlige tabellcellers frafallsprosent beregnes i overensstemmelse med formelen ovenfor. Disse frafallsprosentene kan i så fall stilles sammen som i nedenstående tabell:

c)

Geografisk inndeling	Område 1	Område 2	Sum
Alder			
Gruppe 1	k_{11}	k_{12}	$k_{1.}$
Gruppe 2	k_{21}	k_{22}	$k_{2.}$
Sum	$k_{.1}$	$k_{.2}$	$k_{..}$

$k_{..}$ er den totale frafallsprosent.

Frafallet i de forskjellige tabellceller (grupper av variable) kan variere sterkt fra celle til celle (gruppe), og også avvike meget fra den totale frafallsprosent. Det er helt urealistisk å tro at en f.eks. i en forbruksundersøkelse hvor en undersøger mange variable og kombinasjoner av variable, vil kunne anslå eller i beste fall korrigere for frafallet i alle tabellceller. En vil i praksis måtte velge ut et fåtall variable og eventuelt korrigere for frafallet med hensyn på disse (se ellers avsnitt VI).

Tabellene nedenfor viser tabellcellefrafall slik det ble registrert ved de norske forbruksundersøkelser av 1958¹⁾ og 1967²⁾.

Tabell 1. Fra forbruksundersøkelsen 1958 ($n = 4\ 800$ husholdninger - selvsveiende utvalg). Tabellcellefrafall (i %) etter sosialgrupper og geografiske områder

	Bygder	Byer	Etter handelsfelt					Hele landet
			Østre	Søndre	Vestre	Midtre	Nordre	
Selvst. næringsdrivende	26,5	32,0	29,9	15,7	20,6	33,0	48,4	29,0
Lønnstakere	25,7	28,4	25,4	32,1	22,7	27,7	41,9	26,9
Komb. næringsdrivende/lønns- takere	26,9	35,3	16,1	42,9	22,6	19,2	45,3	27,7
Pensjonister/ ikke yrkesvirk- somme	48,1	45,9	45,0	43,6	42,2	50,8	66,7	47,0
Alle sosialgr. .	28,0	31,4	28,8	31,2	24,3	30,5	46,2	29,5

1) Forbruksundersøkelsen 1958 - Hefte I. 2) Forbruksundersøkelsen 1967 - Hefte I. Vesthve. P.: Metodisk vurdering av forbruksundersøkelsen 1967 (IO 70/6).

I tabellen er innkomne regnskap, som senere ble kassert på grunn av ufullstendighet, også regnet som frafall.

Ved forbruksundersøkelsen 1958 var den totale frafallsprosent (basert på innkomne regnskap) = 26.

Tabell 2. Fra forbruksundersøkelsen 1967 (n = 5 008 husholdninger - selveiende utvalg). Tabellcellefracfall¹⁾ (i %) etter husholdningenes størrelse og inntekt, og etter husholdningenes størrelse og hovedpersonens alder

	Personer pr. husholdning						Alle husholdninger
	1	2	3	4	5	6 eller flere	
Alle husholdninger	42,9	25,8	15,9	13,5	11,7	8,7	21,8
Husholdningens inntekt:							
Under kr. 10 000	44,6	22,6	16,5	15,8	15,6	20,0	31,8
Kr. 10 000 - 14 900	40,9	21,1	13,5	9,6	13,8	0,0	22,1
" 15 000 - 19 900	45,8	25,8	11,3	7,7	8,3	5,9	20,5
" 20 000 - 24 900	37,2	27,5	17,3	11,0	5,3	3,4	18,4
" 25 000 - 34 900	35,9	29,4	12,8	9,8	11,2	7,5	16,8
" 35 000 - 49 900	46,2	26,4	16,9	17,6	7,2	9,3	17,8
" 50 000 og over	50,0	31,6	28,0	23,3	5,7	19,0	26,2
Hovedpersonens alder:							
Under 30 år	25,7	22,7	10,9	8,7	4,8	0,0	15,2
30 - 39 år	43,9	22,6	17,0	11,7	9,5	7,9	14,7
40 - 49 "	52,9	29,4	12,0	11,3	9,2	9,7	15,8
50 - 59 "	46,1	24,2	17,7	20,1	16,2	9,3	22,6
60 år og over	45,1	26,4	19,0	15,9	23,1	7,1	29,9

1) Som frafall er regnet husholdninger som ikke har levert både intervjueskjema og regnskapsbøker. Husholdninger hvis skjema/regnskapsbøker ble kassert på grunn av ufullstendighet, er også regnet som frafall.

Ved forbruksundersøkelsen 1967 var den totale frafallsprosent (basert på innkomne regnskap) = 21.

IV. FRAFALLSÅRSAKER OG METODER TIL Å REDUSERE FRAFALLET

De faktorer som forårsaker frafall kan ha sammenheng med

- i) intervjuenheten (-objektet)
- ii) intervjueren
- iii) undersøkelsens art
- iv) økonomiske og organisatoriske forhold

i) Intervjuenheten (-objektet)

Uten at dette er noe forsøk på å systematisere de frafallsårsaker som skyldes intervjuenheten (IO), består grupperingen nedenfor av de viktigste faktorer til forklaring av frafallet:

- a) IO er å treffe, men IO
 - 1) nekter
 - 2) er syk
- b) IO er ikke å treffe på grunn av
 - 1) midlertidig fravær
 - 2) flytting
 - 3) dødsfall

Punktene b 2) og b 3) skyldes egentlig feil ved det registeret utvalget er trukket fra, og er ikke frafallsfeil i egentlig forstand. Ved norske forbruksundersøkelser blir ikke punktene b 2) og b 3) registrert som frafallsårsaker. Registeret som nyttes ved trekking av utvalget er et bostedsregister slik at en trekker direkte på adresser og ikke på husholdninger. En husholdning som er trukket ut til å delta i undersøkelsen, men som er fraflyttet den oppgitte adresse, regnes således ikke som frafall. Intervjuobjektet blir den "nye" husholdningen som er flyttet inn på den uttrukne adresse.

Ved forbruksundersøkelsen 1973 vil en konfrontere IO-listene med tall fra folkeregistrene for på den måten å kunne avdekke feil ved registeret før undersøkelsen kommer i gang.

Frafall på grunn av midlertidig fravær (pkt. b 1) ovenfor), vil i de fleste tilfelle kunne reduseres noe gjennom en rutine med flere besøk hos IO når denne er fraværende ved første besøk. Hvor mange kontaktforsøk intervjueren skal foreta med IO avhenger av den tid en har til rådighet for gjennomføring av undersøkelsen og den arbeidsmengde intervjueren på forhånd er pålagt.

Sykdom under pkt. a 2) ovenfor, er vanligvis årsak til større frafall ved personundersøkelser enn ved husholdningsundersøkelser. Selv om ett medlem av familien er syk, så vil et intervju med hele husholdningen som intervjuobjekt, i de fleste tilfelle likevel komme i stand.

Et forhold som forårsaker stort frafall ved de fleste statistiske undersøkelser, er at IO nekter å delta i undersøkelsen, (pkt. a 1) ovenfor). Årsakene til nektingen er som oftest av en slik art at en kan redusere dette frafallet betraktelig ved å motivere IO til å delta i undersøkelsen, f.eks. ved å gjøre undersøkelsen bedre kjent blant folk gjennom presse og kringkasting, anvende dyktige intervjuere, presentere undersøkelsen for IO i et introduksjonsbrev (IO-brev) m.v.

Et godt og velskrevet IO-brev er antagelig det beste middel en har til å motivere IO. IO-brevet bør ikke holdes i altfor generelle vendinger, men ha et lite islett av å virke personlig slik at en kan få vekket (eventuelt bevart) IO's tillit til undersøkelsen. Brevet bør videre være saklig informerende med hensyn til den tid det vil ta og det arbeid det vil kreve av IO å delta i undersøkelsen.

ii) Intervjueren

Intervjuerens fremtreden og atferd kan være sterkt avgjørende for om IO skal si seg villig til å delta i undersøkelsen eller ikke. En bedre utdanning av intervjuerne med henblikk på deres spesielle arbeid vil kunne redusere frafallet noe.

I Norge har det vært vanlig med spesialkurs for intervjuerne før de store forbruksundersøkelsene blir satt i gang. Ved disse kurs blir rutinen ved gjennomføringen av undersøkelsen klarlagt, og intervjuerne videre gjort oppmerksomme på spesielt vanskelige ting som kan forventes å oppstå i forbindelse med innsamlingen av materialet.

iii) Undersøkelsens art

Frafallet ved statistiske undersøkelser varierer i størrelse med den type undersøkelse en gjennomfører. En tenker da på forskjellen mellom fullstendig telling og utvalgsundersøkelse, undersøkelsens emneområde, utvalgets geografiske fordeling, observasjonsmetoden og hvor krevende metoden er for IO, undersøkelsesperiodens lengde, observasjonstidspunktet m.v.

Det forholdsvis høye frafallet ved forbruksundersøkelser (se avsnitt VIII, tabell A) har nok sammenheng med at regnskapsføringen (observasjonsmetoden) er krevende for IO. Ved forbruksundersøkelsene av 1958 og 1967 var regnskapsperioden én måned. For 1973-undersøkelsen er det gått inn for en 14 dagers regnskapsperiode og en håper med dette å få redusert frafallet. I tidligere forbruksundersøkelser (før 1958) var regnskapsperioden ett år. Regnskapsperiodens lengde må for disse undersøkelser antas å være hovedårsak til det store frafallet, tatt i betraktning at husholdningene som deltok, på forhånd var blitt spurt om de ville delta eller ikke (ikke-tilfeldig utvalg).

iv) Økonomiske og organisatoriske forhold

Problemet med å forhindre og/eller redusere frafallet er i første rekke et omkostningsspørsmål. Spesielt synes det viktig med en ikke for stram kostnadsramme for feltarbeidet. I Norge har det ved forbruksundersøkelsene

vært vanlig å gi et lite pengebeløp som godtgjørelse til de husholdninger som deltar i undersøkelsene.

Av organisatoriske forhold som har betydning for frafallet kan nevnes tilgangen på intervjuere, feltarbeidsledere og andre funksjonærer som arbeider i feltgruppen. Ved SSB's intervjuorganisasjon er det 300 faste intervjuere. Disse vil ved de fleste undersøkelser til en viss grad kunne avløse hverandre f.eks. ved sykdom, og på den måten forhindre at frafall oppstår.

V. VIRKNINGER AV FRAFALL - TEORETISKE BETRAKTNINGER

Effekten av frafallet på de estimerte størrelser i populasjonen kan deles opp i flere komponenter:

- a) Større varians på estimatorene på grunn av at utvalget blir redusert med frafallsenheter.
- b) Skjevhet ved beregning av estimatene, siden det er lite sannsynlig at fordelingen av telleenheter etter forskjellige kjennetegn i den masse som er falt fra i alle henseende er lik de i den observerte masse.

En kan lett se at "usikkerheten" ved beregning av estimatene er større jo større frafallet er, gitt samme utvalgsstørrelse, og at en i en utvangsundersøkelse uten frafall av telleenheter kan oppnå samme "sikkerhet" for estimatene selv med et atskillig mindre utvalg.

Eksempel¹⁾

Som en forenkling tenker en seg populasjonen (N) delt inn i to grupper, gruppe 1 og gruppe 2 med henholdsvis N_1 og N_2 enheter ($N = N_1 + N_2$). N_1 er de antall enheter i populasjonen som en vil få opplysninger fra dersom de blir trukket ut, og N_2 er de enheter i populasjonen som en ikke vil få opplysninger fra ("nonresponse"), selv om de blir trukket ut.

$$W_1 = \frac{N_1}{N}, \quad W_2 = \frac{N_2}{N}$$

En trekker et tilfeldig utvalg på n enheter, hvorav n_1 er svarsenheter - alle fra gruppe 1. Hos de n_1 enhetene observeres verdien av en variabel y_1 , og en vil bruke disse observasjoner til å si noe om den tilsvarende variabel Y i hele populasjonen. Dersom en nytter \bar{y}_1 som en estimator for \bar{Y} får en følgende uttrykk for forventningsskjevheten på utvalgsgjennomsnittet:

$$1) \quad E(\bar{y}_1) - \bar{Y} = \bar{Y}_1 - \bar{Y} = \bar{Y}_1 - (W_1\bar{Y}_1 + W_2\bar{Y}_2) = W_2(\bar{Y}_1 - \bar{Y}_2)$$

1) Cochran, W.G.: Sampling Techniques.

Forventningsskjevheten er lik produktet av den relative andel av enheter i populasjonen som faller fra hvis de blir trukket ut, og differansen mellom gjennomsnittene i de to grupper. Siden utvalget ikke gir opplysninger om \bar{Y}_2 , er størrelsen på skjevheten ukjent.

Dersom Y er en binær variabel, vil \bar{Y}_2 ligge mellom 0 og 1, og en vil kunne finne et konfidensintervall for \bar{Y}_1 (og \bar{Y}) gitt et bestemt sannsynlighetsnivå (hvis n er stor) - intervallestimering.

Siden Y nå antas å være en binær variabel, innføres i overensstemmelse med vanlig terminologi

P	istedenfor	\bar{Y}	
P_1	"	\bar{Y}_1	
P_2	"	\bar{Y}_2	
P_1	"	\bar{y}_1	og $q_1 = 1 - p_1$

Hvis n_1 er stor så er et 95 % konfidensintervall for P_1 gitt ved

$$2) \quad P_1 \pm \sqrt{2 p_1 q_1 / n_1} \quad 1)$$

Når en på tilsvarende måte vil finne konfidensgrenser (Cochran) for P er en på trygg grunn ved å anta at $P_2 = 0$ når en finner nedre konfidensgrense (\hat{P}_{nedre}), og $P_2 = 1$ for øvre konfidensgrense ($\hat{P}_{\text{øvre}}$).

Et sannsynlighetsnivå på 95 % gir

$$3) \quad \hat{P}_{\text{nedre}} = W_1(p_1 - 2\sqrt{p_1 q_1 / n_1}) + W_2(0)$$

$$4) \quad \hat{P}_{\text{øvre}} = W_1(p_1 + 2\sqrt{p_1 q_1 / n_1}) + W_2(1)$$

Oppstillingen nedenfor gir verdier av \hat{P}_{nedre} og $\hat{P}_{\text{øvre}}$ for forskjellige størrelser av frafallet, W_2 , og av variabelen p_1 . Utvalget (n) er satt lik 1 000 og $n_1 = nW_1$. Sannsynlighetsnivå på 95 %.

Tabell 3. 95 % "konfidensintervall" for P (%).

Frafall (%) ($100 \cdot W_2$)	Verdi av undersøkelsesvariabel, ($100 \cdot p_1$)							
	5		10		20		50	
0	(3,6	6,4)	(8,1	11,9)	(17,5	22,5)	(46,7	53,2)
5	(3,4	11,1)	(7,6	16,3)	(16,5	26,5)	(44,4	55,6)
10	(3,2	15,8)	(7,2	20,8)	(15,6	30,4)	(42,0	58,0)
15	(3,0	20,5)	(6,8	25,2)	(14,7	34,3)	(39,6	60,4)
20	(2,8	25,2)	(6,3	29,7)	(13,7	38,3)	(37,2	62,8)

1) En har her regnet med et variansuttrykk som om populasjonen var uendelig.

En kan se at "konfidensintervallet" blir større når frafallet øker.

Det kan være av interesse å se hvilken størrelse på utvalget en trenger for samme størrelse på "konfidensintervallet" dersom $W_2 = 0$ (dvs. ikke noe frafall).

For f.eks. $100 \cdot p_1 = 50$ og $100 \cdot W_2 = 5$ er halve vidden av "konfidensintervallet" lik 5,6 (se tabellen ovenfor). Utvalgsstørrelsen (n_e) forutsatt ikke noe frafall, finner en da fra ligning 2) ovenfor:

$$5) \quad 5,6 = 2 \sqrt{(50)(50)/n_e} \quad , \text{ som gir}$$

$$\underline{n_e = 320}$$

Noe upresist kan dette formuleres slik: I en undersøkelse uten frafall, sammenlignet med en med 5 % frafall, vil en oppnå samme "sikkerhet" på estimatene selv om utvalgsstørrelsen reduseres til 1/3.

For $100 \cdot W_2 = 10, 15$ og 20 er verdiene av n_e henholdsvis 155, 90 og 60. En ser således at det er hensiktsmessig å bestrebe seg på å redusere frafallet, da en i en undersøkelse med lite frafall vil kunne ha større tiltro til resultatene.

VI. KORRIGERING FOR FRAFALL

Den skjevhet en får på estimatene på grunn av frafall, er det vanskelig å si noe om både med hensyn til dens absolutte størrelse og i hvilken retning skjevheten går (over-/underestimering). Med korrigering for frafall forstår en i en eller annen forstand en "oppretting" av de feil på estimatene som frafallet forårsaker. Mulighetene for å korrigere for frafallet er større jo flere opplysninger en har om frafallsenheter. Det er således viktig ved fremtidige forbruksundersøkelser å skaffe så rik informasjon som mulig om frafallsenheter. Slik informasjon vil en nå i større grad enn tidligere kunne få fra andre statistikkområder ("arkiv-statistiske metoder"). En annen mulighet for å skaffe informasjon om frafallsenheter ved den kommende forbruksundersøkelse i 1973 ville kanskje være å ta et subsampel på de frafallsenheter som "ikke er å treffe".

Ved forbruksundersøkelsen i 1958 hentet SSB inn opplysninger fra likningsmyndighetene om inntekt og familiens størrelse for de husholdninger som falt fra ved undersøkelsen. Korreksjonsmetoden en nyttet, bygget på forutsetningen om at forbrukssammensetningen gjennomgående var lik for husholdninger innen samme inntekts- og familiegruppe, enten disse husholdninger var oppgavegivere til undersøkelsen eller ikke.

Korrigeringen for frafallet ved 1967-undersøkelsen bygde på samme forutsetning som den som ble gjort ved arbeidet i 1958. En så altså bort fra andre forbruksmotiverende faktorer som f.eks. bosted, sesongfaktorer og alder, og dubliserte forbruksoppgavene for et tilfeldig utvalg av de husholdningene som hadde samme inntekt og husholdningsstørrelse som de en ikke hadde regnskap fra.

For 1973-undersøkelsen er det fremkommet et forslag som går i retning av at en ved dubliseringen også skal ta hensyn til faktorene bosted- og sesongvariabel, ved siden av variablene inntekt og husholdningsstørrelse.

VII. OM FRAFALLETS STØRRELSE VED EN DEL NORSKE UNDERSØKELSER

Av de 10 forbruksundersøkelser som er foretatt i Norge etter 1945 (avsnitt VIII, tabell A), har:

a)	3	undersøkelser	(30 %)	et	frafall	på	10 - 20 %
	5	"	(50 %)	"	"	"	21 - 30 %
	og 2	"	(20 %)	"	"	"	31 og mer

Av 21 forskjellige undersøkelser foretatt av SSB's intervjuorganisasjon siden denne ble opprettet i 1967 (avsnitt VIII, tabell B), har:

b)	2	undersøkelser	(10 %)	et	frafall	på	under 10 %
	11	"	(52 %)	"	"	"	10 - 15 %
	3	"	(14 %)	"	"	"	16 - 20 %
	5	"	(24 %)	"	"	"	21 - 30 %

Som en kan se av oppstillingene ovenfor har forbruksundersøkelsene gjennomgående en meget høyere frafallsprosent enn andre intervjuundersøkelser foretatt ved SSB.

Forbruksundersøkelsen 1967

Ved forbruksundersøkelsen i 1967 ble data samlet inn i 3 typer hefter - innledningsintervju, regnskapsbok og årsintervju. Av de 5 008 uttrukne husholdninger ble det oppnådd innledningsintervju av 4 550. Av disse var det 3 918 husholdninger med både regnskap og innledningsintervju og 632 husholdninger med bare innledningsintervju. Det var således 458 husholdninger eller ca. 9 % av de uttrukne en ikke fikk innledningsintervju fra.

Regnskapsbok ble levert av 3 958 husholdninger eller 79 % av utvalget. En del regnskap, i alt 40, var mangelfulle og måtte kasseres. Den effektive oppslutning om undersøkelsen var således 78 %.

Årsintervju ble levert av 4 369 husholdninger eller 87 % av de uttrukne husholdninger. Svarprosenten var her således noe høyere enn ved regnskapsføringen.

Nedenfor er det gitt en skjematisk oppstilling over størrelsen av frafallet for de uttrukne husholdninger ved forbruksundersøkelsen 1967 etter forskjellige kjennetegn. Med en frafalt husholdning menes en uttrukket husholdning uten innledningsintervju eller bare med innledningsintervju. Det prosentvise frafall, 0-15 %, 16-26 % etc. angir den relative andel av husholdninger med frafall av alle husholdninger i samme kjennetegnsklasse - "inntekt under kr. 10 000,-", "inntekt på kr. 10 000 eller mer" etc.

Tabell 4

	Stort fracfall (27 % eller mer)	Middels stort fra- fall (16- 26 %)	Lite fracfall (0-15 %)
Inntekt			
Under kr. 10 000	x		
Kr. 10 000 eller mer		x	
Husstandsstørrelse			
1 person	x		
2 personer		x	
3 personer eller flere			x
Hovedpersonens alder			
Under 20 år	x		
20 - 49 år			x
50 - 69 år		x	
70 år og over	x		
Hovedpersonens yrkesstatus			
Selvst. næringsdrivende i jord- bruk, skogbruk, fiske			x
Lønnstakere, selvstendig næringsdrivende ellers		x	
Ikke yrkesaktive	x		
Bosted			
Spredtbygd			x
Tettbygd	x		
Regnskapsmåned			
Januar, april			x
Juli, august	x		
Øvrige måneder i året		x	

VIII. TABELLVEDLEGG

Tabell A. Norske forbruksundersøkelser etter 1945 - og frafallet ved disse undersøkelsene

Nr.	Oppgave- givere (sosial- gruppe)	Oppgavene hentet inn fra:	Utvalgs- type	Under- søkelses- periode	Data- innsam- lings- metode	Observa- sjons- periodens lengde	Tallet på bearbeidde oppgaver	Frafalls- ¹⁾ prosent
1	Industri- arb. + Skogs- og jordbr. arb.	18 byer og 13 land- kommuner	ikke- tilfeldig	mai 1947- april 1948	Regnskap	1 år	452 52	36 39
2	Industri- arb.	17 byer og 10 land- kommuner	ikke- tilfeldig	okt.1951- sept.1952	Regnskap	1 år	100	11
3	Arbeidere	Oslo og Bergen	ikke- tilfeldig	1952	Intervju	1 år	203 ²⁾	28
4	Høyere funk- sjonærer	Oslo, Bergen, Tr.heim og Stavanger	ikke- tilfeldig	april 1952- mars 1953	Regnskap	1 år	74	18
5	Fiskere	Kyst- fylkene	ikke- tilfeldig	1954	Regnskap	1 år	293	26
6	Jord- brukere	Hele landet	ikke- tilfeldig	1954	Regnskap	1 år	585	17
7	Alders- trygdede	9 store byer	ikke- tilfeldig	mai 1955- april 1956	Regnskap	1 år	140	46
8	Alle sosi- algr.unnt. jordbr., fiskere og alders- trygdede	Hele landet	tilfeldig	1958	Regnskap og intervju	Regnskap i 1 mnd. i 3	502	26
9	Alle sosi- algrupper	Hele landet	tilfeldig	1967	Regnskap og intervju	Regnskap i 1 mnd. i 3	918	21
10 ³⁾	Ektepar med 2 barn og med hovedinn- tektstaker i alderen 35-40 år (alle sosi- algrupper)	Oslo, Bergen, og Trondheim	ikke- tilfeldig	28/1-1972- 2/3-1972	Regnskap og intervju	Regnskaps- perioder på 2, 3 og 4 uker x ⁴⁾		ca. 30 ⁵⁾

1) I frafallsmassen er ikke regnet med innkomne regnskap/intervju som er kassert på grunn av ufullstendighet.

2) Ca. 28 % av disse oppgaver stammer fra husholdninger som er plukket ut for å erstatte de frafalte.

3) Prøveundersøkelsen 1972 (før 1973-undersøkelsen).

4) Materialet er ennå ikke bearbeidd (pr. 20/4-72), men utvalget består av 350 husholdn.

5) Foreløpig tall (pr. 20/4-72).

Tabell B. Oversikt over undersøkelser (utvalgsstørrelse, frafall m.v.) foretatt ved Kontoret for intervjuundersøkelser siden intervjuorganisasjonen ble opprettet i 1967.

Undersøkelsene nr. 1-16 (nedenfor), gjelder undersøkelser som er kommet ut i serien "Rapport fra Kontoret for intervjuundersøkelser" og undersøkelser nr. 17-21, er trykt i Norges offisielle statistikk (NOS).

	Type undersøkelse	Type utvalg	Under- søkelses- periode	Data- innsam- lings- metode	Utvalgs- størrelse (enheter)	Frafall Antall enheter	%	Frafallsårsaker
1.	Radio- og fjernsynsundersøkelsen-april 1967	person-tilfeldig und.	16/4-22/4-67	inter- vju	3 023	477	16	13 % av frafallet skyldtes nekting
2.	Forbruksundersøkelsen for studenter og skoleungdom 1967	person-tilfeldig und.	april og oktober 1967	regn- skap og inter- vju	vel 500	150	ca.27	vel 50 % av frafallet skyldtes nekting
3.	Boligundersøkelsen oktober 1967	hush.- og bolig-undersøk.	tilfeldig 2/10-31/10-67	inter- vju	3 126 ¹⁾	311 ¹⁾	10	47 % av frafallet skyldtes nekting og ca. 10 % sykdom
4.	Radio- og fjernsynsundersøkelsen. Sept.-okt.1968	person-tilfeldig und.	21. og 24. sept. 4.okt. 1968	inter- og vju	3 032 ²⁾	652	22	ca. 11 % av frafallet skyldtes nekting
5.	Fritidshusundersøkelse 1967/68	hush.-nings-und.	tilfeldig januar-mars 1968	inter- vju	5 019	ca.650	13	
6.	Radio- og fjernsynsundersøkelsen. Febr. 1969	person-tilfeldig und.	5.,7.,15. og 18. februar 1969	inter- vju	3 048 ³⁾	537	18	20 % av frafallet skyldtes nekting
7.	Ønsker om og behov for sysselsetting blant gifte kvinner	person-tilfeldig und.	4.-30. nov. 1968	inter- vju	2 959 ⁴⁾	319	11	50 % av frafallet skyldtes nekting, og vel 10 % sykdom
8.	Undersøkelse om Forbruker-rapporten, vareundersøk. og reklamasjoner 1969	hush.- og person-und.	tilfeldig 5/5-7/6-1969	inter- vju	2 061 ⁵⁾	244	12	20 % av frafallet skyldtes nekting og 14 % at hush. ikke var å treffe

1) Antall husholdninger. 144 av de opprinnelig uttrukne leiligheter (husholdninger) ble erstattet i undersøkelsesperioden. De erstattede husholdninger er ikke regnet som frafall.

2) 233 av de opprinnelig uttrykte personer ble erstattet på grunn av at de ikke lenger befant seg i utvalgsområdene. De erstattede personer er ikke regnet som frafall.

3) 288 av de opprinnelig uttrukne personer ble erstattet på grunn av at de ikke lenger befant seg i utvalgsområdene. De erstattede personer er ikke regnet som frafall.

4) 341 av de opprinnelig uttrukne personer ble erstattet, hovedsakelig på grunn av flytting. De erstattede personer er ikke regnet som frafall.

5) Antall husholdninger. 180 av de opprinnelig uttrukne husholdninger ble erstattet. De erstattede husholdninger er ikke regnet som frafall.

	Type- under- søk- else	Type utvalg	Under- søkelses- periode	Data- innsam- lings- metode	Utvalgs- størrelse (enheter)	Frafall		Frafallsårsaker
	und.	til- feldig			Antall enheter	%		
9. Valgprogram- mer i radio og fjernsyn i samband med Stortings- valget 1969	person- und.	til- feldig	9.-23. sept.1969	inter- vju	2 999 ¹⁾	297	10	Storparten av fra- fallet skyldtes at intervjupersonene nektet å svare.
10. Reklamesen- dinger i post- kassene, 1969	hush. und.	til- feldig	18.-31. okt. og 2.-15. nov.1969	intervju og dag- bokfør- ing	2 080 ²⁾	198	10	52 % av frafallet skyldtes nekting, 12 % sykdom og 30 % at hushold- ningen ikke var å treffe.
11. Pensjons- alderunder- søkelsen 1969	person- und.	til- feldig	11.-30. aug.1969	inter- vju	6 019 ³⁾	577	10	De viktigste år- saker til frafall var at intervju- personene ikke var å treffe eller at de nektet å la seg intervjue
12. Lytterunder- søkelse om ungdommens radioavis, april 1970	person- und.	til- feldig	20.og 27. april 1970	inter- vju	798 ⁴⁾	74	9	Den vanligste årsak til frafall var at intervjupersonen ikke var å treffe.
13. Radio- og fjernsyns- undersøkelsen, juni/august 1970	person- und.	til- feldig	19.-21. juni og 17.-20. august 1970	inter- vju	2 973 ⁵⁾	860	30	Storparten av fra- fallet skyldtes at intervjupersonene var fraværende, ca. 12 % av fra- fallet skyldtes nekting.
14. Radio- og fjernsyns- undersøkelsen, febr. 1971	person- und.	til- feldig	6.-12. febr. 1971	inter- vju	2 822 ⁶⁾	703	25	Storparten av fra- fallet skyldtes at intervju- personene var fra- værende. Ca. 12 % av frafallet skyldtes rekting.

- 1) 170 av opprinnelig uttrukne personer ble erstattet, hovedsakelig på grunn av flytting. De erstattede personer er ikke regnet som frafall.
- 2) 80 av de opprinnelig uttrukne husholdninger ble erstattet. De erstattede husholdninger er ikke regnet som frafall.
- 3) 299 av de opprinnelig uttrukne personer ble erstattet. De erstattede personer er ikke regnet som frafall. Erstatningene kom vesentlig istand p.g.a. flyttinger.
- 4) 293 av de opprinnelig uttrukne personer ble erstattet. Erstatningsårsakene var vesentlig flytting og fravær fra hjemmet i forbindelse med arbeid, skolegang etc. De erstattede personer er ikke regnet som frafall.
- 5) 217 personer blant de opprinnelig uttrukne ble erstattet, da de ikke lenger befant seg i utvalgsområdene. De erstattede personer er ikke regnet som frafall.
- 6) 181 av de opprinnelig uttrukne personer ble erstattet. De erstattede personer er ikke regnet som frafall.

	Type	Type	Under-	Data-	Utvalgs-	Frafall		Frafallsårsaker
	under-	utvalg	søkelses-	innsam-	størrelse	Antall	%	
	søk-		periode	lings-	(enheter)	enheter		
	else			metode				
15. Finansiering av nye boliger 1968	hush.- og bolig-und.	til-feldig	febr.-mars 1970	inter-vju	2 824	363	13	41 % av frafallet skyldtes nekting, og omlag 23 % at intervjuobjektene ikke var å treffe.
16. Offentlig informasjon og publikumsinformasjonsbehov 1970	hush.-und.	til-feldig	9.-30. nov.1970	inter-vju	2 895	457	16	34 % av frafallet skyldtes nekting, og 24 % at intervjupersonene ikke var å treffe
17. Forbruksundersøkelsen 1967	hush.-und.	til-feldig	1967	regn-skap og inter-vju	5 008	1 090	21	De viktigste årsaker til frafallet var nekting, sykdom i familien og at husholdningene ikke var å treffe.
18. Helseundersøkelse 1968	person- og hus-holdn.-und.	til-feldig	7/10-2/11-1968	inter-vju	11 130 ¹⁾	637	6	75 % av frafallet skyldtes at intervjupersonene var borte på skole, arbeid etc.
19. Ferieundersøkelsen 1968	hush.-und.	til-feldig	mai 1968	inter-vju	2 999	306	10	49 % av frafallet skyldtes nekting, og 18 % at husholdningene ikke var å treffe.
20. Stortingsvalget 1969	person-und.	til-feldig	9.-23. sept.1969	inter-vju	2 999 ²⁾	297	10	Den viktigste årsak til frafallet var nekting.
21. Ferieundersøkelsen 1970	person-und.	til-feldig	1/9-15/10-1970	inter-vju	3 053 ³⁾	372	12	29 % av frafallet skyldtes nekting og 27 % at intervjupersonene var borte på arbeid, skole e.l.

1) Antall personer.

2) 170 av de opprinnelig uttrukne personer ble erstattet. Erstatningsgrunn var hovedsakelig flytting. De erstattede personer er ikke regnet som frafall.

3) 239 av de opprinnelig uttrukne personene ble erstattet på grunn av at de ikke lenger befant seg i utvalgsområdene. De erstattede personer er ikke regnet som frafall.

KORREKSJON FOR FRAFALL VED FORBRUKSUNDERSØKELSEN 1973.
SPESIELT OM ESTIMERING AV SVARSSANNSYNLIGHETER

Av Odd Skarstad

Innhold

	Side
I Innledning	60
II Korreksjonsmetoder	60
II.a Bør en "korrigerer" for frafall?	60
II.b Dublisering	61
II.c Middelveidiimputering	61
II.d Veining med estimerte svarsannsynligheter	61
III Estimering av svarsannsynligheter. Kort skisse av opplegget	62
IV Nærmere om estimering av svarsannsynligheter	64
IV.a Variable i relasjonen	64
IV.b Spesifikasjon og klassifisering av variablene i relasjonen	65
V Henvisninger	69

I. Innledning

I dette notatet behandles først tre måter ved korreksjon for frafall: dublisering, middelverdiimputering og veining. Deretter gjøres det rede for en måte for estimering av svarssannsynligheter ved veining.

Notatet er hastverkspreget og fremstillingen er mangelfull på mange punkter. Jeg håper imidlertid at idéene er kommet noenlunde klart frem og muligens kan være til noen nytte ved eventuelle senere arbeider på dette område.

II. Korreksjonsmetoder

II.a Bør en "korrigerer" for frafall?

En har åpenbart de to mulighetene, enten:

1. "ta hensyn" til frafallet ved en eller annen form for "korreksjonsmetoder", eller
2. presentere materialet "ukorrigert", men samtidig med en fyldig redegjørelse for hvordan frafallet fordeler seg i forhold til variable som husholdningstype, geografisk område, bokføringsperiode m.v.

Resonnementet bak alternativ 2 er at enhver form for korrigerer må bygge på helt bestemte forutsetninger, og at en mener at brukerne av statistikken selv bør få avgjøre hvilke forutsetninger de anser forsvarlig og hvilke korrigeringer som eventuelt kan gjøres. En viktig innvending til dette resonnement er imidlertid at de fleste brukere hverken har noen mening om hva som er en god korrigeringsmetode eller har anledning til å utføre korreksjonene. Derfor må en korrigerer for frafall på en eller annen måte.

Noen kjente metoder for korrigerings for frafall: En deler inn husholdningene (enhetene) i grupper og kan korrigere ved:

- a) dublisering,
- b) imputering av middelveier for husholdninger i gruppen,
- c) veining med den estimerte svarsansynlighet for husholdninger i gruppen.

Det er foretatt en jamføring mellom de tre metodene av G. Elofson ved Statistiska Centralbyråen i Sverige når det gjelder forventning og varians på estimatorene [1].

II. b. Dublisering

Ved forbruksundersøkelsen i Norge i 1967 brukte en dublisering ved korrigerings for frafall. Dette ble også brukt ved forbruksundersøkelsen i Sverige i 1969.

Fordelen med dubliseringsmetoden er vel først og fremst at den vel programmeringsmessig er enkel ved tabellutkjøringer.

Ulempen er vesentlig at metoden medfører en viss øking i variansen i forhold til f.eks. imputering av middelveier og at en kan "ta hensyn" til bare et nokså begrenset antall variable.

II. c. Middelveiimputering

Dette består i at frafalte husholdninger gis verdier lik gjennomsnittsverdien for ikke-fracfalte husholdninger innen hver husholdningsgruppe. En vil vel få en del mer programmeringsarbeid ved denne korrigeringsmetoden enn ved dublisering.

Det er også her et nokså begrenset antall variable en kan ta hensyn til ved den størrelse på utvalget en har.

(Det kan kanskje nevnes at det ikke gir god mening med middelveiimputering av kvalitative variable hvis det skulle være aktuelt (yrke, sivilstand, osv.).)

II. d. Veining med estimerte svarsansynligheter

Dette vil si at den enkelte husholdning ved estimering av f.eks. total- og gjennomsnittstall får vekt lik den inverse av den estimerte svarsansynlighet for husholdningen.

Svarsansynligheten for den enkelte husholdning estimeres gjerne ved svarshyppigheten innen den husholdningsgruppe husholdningen tilhører. I likhet med ved de forannevnte metoder vil de med den aktuelle størrelse på utvalget kunne konstrueres bare i et begrenset antall grupper (relativt

få variable og grove inndelinger av disse). Dette gjør det vanskelig å konstruere homogene grupper, noe som er en ulempe.

I neste avsnitt (iii) gjøres det rede for en alternativ metode for estimering av svarsansynligheter.

Veiningsmetoden vil medføre en del programmeringsarbeid vel noe avhengig av hvordan estimering av vektene foregår.

(Når det gjelder forbruksundersøkelser er det planlagt undersøkelser hvert år f.o.m. 1974, og et program vil kunne brukes om igjen hver gang. Derfor bør en vel ikke ta altfor mye hensyn til programmeringsarbeidet ved valg av korreksjonsmetode.)

III. Estimering av svarsansynligheter. Kort skisse av opplegget

Her skisseres kort en måte for estimering av svarsansynligheter ved variansanalyse (utformet som regresjonsanalyse).

En innfører en binær variabel Y hvor

$$Y_i = \begin{pmatrix} 1 & \text{hvis hush.nr. } i & \text{ikke har falt fra} \\ 0 & \text{" " " " " har falt fra.} \end{pmatrix}$$

Denne variabel betraktes som avhengig variabel (venstresidig variabel i analysen).

Videre innføres eksogene binære variable (nevnt som eksempel) Z_1 , der

$$Z_{1i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i & \text{bor i Vestre handelsfelt} \\ 0 & \text{ellers,} \end{pmatrix}$$

$$Z_2 \text{ der } Z_{2i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i & \text{bor i Midtre handelsfelt} \\ 0 & \text{ellers,} \end{pmatrix}$$

$$\text{og } Z_3 \text{ der } Z_{3i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i & \text{bor i Nordre handelsfelt} \\ 0 & \text{ellers.} \end{pmatrix}$$

(Her er antatt Østre handelsfelt som referansegruppe.)

Tilsvarende kan en innføre binære variable for hush.type, inntektsgruppe, alder på hovedperson osv. En danner regresjonen

$$Y = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \dots + \text{restledd.}$$

[Det er selvsagt mange problemer som vil oppstå i forbindelse med oppstillingen av regresjonslikningen. Det gjelder bl.a. hvilke variable skal inngå. Videre er det spørsmål om evt. "samspill" mellom eksogene variable; f.eks. kan det hende at virkningen av ulik inntekt ikke kan antas å være den samme i de forskjellige landsdeler. Videre må det også avgjøres

hvordan kvantitative variable skal inngå (inntekt f.eks.) som en "vanlig" variabel eller som binære variable (dvs. inndelt i inntektsklasser). Se senere.]

Minste kvadraters estimatorer på koeffisientene i regresjonsligningen estimeres, altså $\hat{\beta}_0, \hat{\beta}_1, \dots$, osv. Når estimatene er funnet, beregnes svarsansynligheten for husholdning nr. i ved

$$\hat{p}_i = \hat{\beta}_0 + \hat{\beta}_1 Z_{1i} + \hat{\beta}_2 Z_{2i} \text{ osv.}$$

Gjennomsnittsverdien for en variabel V (f.eks. forbruk av en vare) i populasjonen estimeres så ved

$$\hat{\bar{V}} = \frac{\sum \frac{1}{P_i} V_i}{\sum \frac{1}{P_i}}.$$

(Formelen gjelder ved rent tilfeldig utvalg.)

På tilsvarende måte estimeres gjennomsnitt (og totaler) for delpopulasjoner.

Egenskapene ved estimatoren $\hat{\bar{V}}$ (forventningsrettet, varians) avhenger av i hvilken grad

1. frafalte og ikke-frafalte husholdninger med samme bakgrunnkjennetegn har samme forbruk,
2. relasjonen til estimering av svarsansynligheter er "riktig".

Når det gjelder problemet hvorvidt frafalte og ikke-frafalte husholdninger (innenfor en hush.gruppe) er "like", kan dette (hvis praktisk mulig) belyses ved evt. etterundersøkelser på et utvalg av den frafalte masse av husholdninger. For en del av husholdningsmassen vil en få frafall for en del av undersøkelsen, men ikke for hele undersøkelsen (f.eks. frafall ved regnskapsføring, men ikke ved intervju). Det er mulig dette vil kunne brukes til å si noe om frafalte og ikke-frafalte husholdningers forbruk er sammenliknbart.

Relasjonen til estimering av svarsansynligheter behandles i det følgende.

IV. NÆRMERE OM ESTIMERING AV SVARSANNSYNLIGHETER

IV.a. Variable i relasjonen

Det første problemet som reiser seg er hvilke variable skal inngå i relasjonen.

Det er for det første nærliggende å tenke at de variable som skal være med må antas å være av betydning for forbruket. Alternativt kan det være aktuelt av praktiske hensyn i stedet å ha med variable som er sterkt korrelert med variable som er av betydning for forbruket (klime er f.eks. sikkert av betydning for forbruket, men av praktiske hensyn foretrekkes kanskje en bostedsvariabel).

For det andre må selvsagt variablene være av betydning for frafallet for at det skal ha noen hensikt å ta dem med ved estimering av svarsansynligheter. Alternativt kan en også her bruke variable som er korrelert med andre variable som spiller en rolle for frafallet. (Sykelighet har sikkert betydning for frafall. Av praktiske grunner må en kanskje bruke alder i stedet - idet alderen i noen grad har sammenheng med sykelighet.)

I sitt notat (se [2]) har Petra Vestbye stilt opp tabeller som viser marginale fordelinger av frafall etter

husholdningens inntekt
husholdningsstørrelse
hovedpersonens alder
hovedpersonens yrkesstatus
handelsfelt
spredtbygd/tettbygd strøk

Det er også stilt opp en del fordelinger i to variable.

Det er vanskelig ut fra marginale fordelinger å si noe særlig om hvordan hver enkelt av disse variable influerer på frafallet. En ser f.eks. av tabell 5 at frafall varierer med inntekten. Men inntekt er korrelert med andre variable (bosted, husholdningsstørrelse) som kanskje er variable som særlig sterkt influerer på frafallet.

Uten å spekulere mer på hvilke variable som bør være med, settes her opp de variable en kan få både for frafalte og ikke-frafalte husholdninger:

1. Bosted (utv.område, tettbygd/spredtbygd)
2. Alder på personene i husholdningen
3. Husholdningstype (herunder antall personer)
4. Sivilstand på personene i husholdningen
5. Bokføringsperiode
6. Yrkesstatus
7. Inntekt

(For nyinnflyttede, frafalte husholdninger vil de fleste av disse variable mangle (hvis en uttrukket husholdning er flyttet, skal den nyinnflyttede være med i stedet).)

IV.b. Spesifikasjon og klassifisering av variablene i relasjonen

1. Bosted

Ut fra utvalgsområde kan en konstruere en variabel for type bosted og geografisk område. Dette kan f.eks. gjøres på følgende måte:

Type bosted

- i) Oslo, Bergen, Trondheim
- ii) Andre tettbygde strøk
- iii) Spredtbygde strøk

Handelsfelt

- i) Østre
- ii) Vestre
- iii) Midtre
- iv) Nordre

Petra Vestbye viser i sitt notat hvordan frafallet varierer med type bosted og handelsfelt. Det er imidlertid vanskelig å vurdere hvilke årsakssammenhenger som gjelder m.h.t. frafall ut fra dette. Det er videre et problem om en her kan anta at virkningen av forskjellige "type strøk" er den samme i alle handelsfelt m.h.t. frafallssannsynligheter¹⁾. Det er vanskelig å vurdere dette, men jeg vil anta det.

Innfører følgende binære variable:

$$X_{1i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i \text{ bor i Oslo, Bergen eller Trondheim} \\ 0 & \text{ellers} \end{pmatrix}$$

$$X_{2i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i \text{ bor i andre tettbygde strøk} \\ 0 & \text{ellers} \end{pmatrix}$$

Referansegruppe: spredtbygde strøk.

Videre:

$$X_{3i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i \text{ bor i Østre handelsfelt} \\ 0 & \text{ellers} \end{pmatrix}$$

$$X_{4i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i \text{ bor i Vestre handelsfelt} \\ 0 & \text{ellers} \end{pmatrix}$$

$$X_{5i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i \text{ bor i Midtre handelsfelt} \\ 0 & \text{ellers} \end{pmatrix}$$

Referansegruppe: Nordre handelsfelt.

1) Hvis en tror at så ikke er tilfelle, bør en innføre såkalte samspillvariable ("spredtbygde strøk i Østre handelsfelt" osv.).

2. Alder på personene i husholdningen

Det foreligger her mange muligheter for hvordan alder på personene skal inngå, f.eks. alder på person nr. 01 (dvs. den person som inntektsmessig bidrar mest til familiens underhold), gjennomsnittsalder på alle personene, alder på eldste person osv.

Jeg velger her alder på person 01 som variabel. Petra Vestbye viser i sitt notat hvordan frafallet varierer (marginalt) med alder på hovedpersonen. Videre viser hun hvordan frafallet varierer simultant med hovedpersonens alder og antall personer pr. husholdning. Selv om det er vanskelig å vurdere årsakssammenhenger, tyder materialet på at frafallet er stort med hovedperson i alderen over 70 år, noe mindre i alderen 50-70 år og minst i alderen 20-50 år. Med hovedpersonen under 20 år er frafallet stort, men disse utgjør en meget liten del av populasjonen (0,26 % av husholdningene i utvalget i 1967).

Følgende gruppering foreslås:

- i) Hovedperson over 70 år
- ii) " mellom 50 og 70 år
- iii) " under 50 år

En innfører de binære variable

$$X_{6i} = \begin{cases} 1 & \text{hvis person 01 i hush. nr. i er over 70 år} \\ 0 & \text{ellers} \end{cases}$$

$$X_{7i} = \begin{cases} 1 & \text{hvis person nr. 01 i hush. nr. i er i alderen 50-70 år} \\ 0 & \text{ellers} \end{cases}$$

Referansegruppe: Person 01 under 50 år.

3. Antall personer i husholdningen

I Petra Vestbyes notat er det vist hvordan frafallet varierer med antall personer i husholdningen. Frafallet var i 1967 42,9 % i énpersons-hushold, 25,8 % i topersonshushold, mens det f.eks. bare var 8,7 % i husholdninger med 6 eller flere personer.

Foreslår følgende gruppering:

- i) énpersonshushold
- ii) topersonshushold
- iii) hushold med 3 eller flere personer

Innfører variablene:

$$X_{8i} = \begin{cases} 1 & \text{hvis énpersonshushold} \\ 0 & \text{ellers} \end{cases}$$

$$X_{9i} = \begin{cases} 1 & \text{hvis topersonshushold} \\ 0 & \text{ellers} \end{cases}$$

Referansegruppe: Hushold med 3 eller flere personer.

Alternativt til "alder på personer" eller "antall personer" kunne en trolig bruke "husholdningstype", f.eks. enslig, ektepar med og uten barn osv.

4. Bokføringsperiode

Ved forbruksundersøkelsen 1967 varierte frafallet betydelig med årstiden. Frafallet var lite om vinteren og stort om sommeren. På grunnlag av resultatene fra 1967 foreslås følgende gruppering (året er delt i 26 perioder á 14 dager):

- i) Periodene 1 - 9 (januar, februar, mars, april)
- ii) " 10 - 13 og 19 - 26 (mai, juni, september, oktober, november, desember)
- (iii) " 14 - 18 (juli og august)

Innfører følgende variable:

$$X_{10i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i \text{ er trukket ut til å føre regnskap} \\ & \text{i periodene 1 - 9} \\ 0 & \text{ellers} \end{pmatrix}$$

$$X_{11i} = \begin{pmatrix} 1 & \text{hvis hush.nr. } i \text{ er trukket ut til å føre regnskap} \\ & \text{i periodene 10 - 13 eller 19 - 26} \\ 0 & \text{ellers} \end{pmatrix}$$

Referansegruppe: husholdninger som er trukket ut til å føre regnskap i periodene 14 - 18.

5. Yrkesstatus

Petra Vestbye viser i sitt notat at frafallet (marginalt) varierer betydelig med hovedpersonens yrkesstatus. Frafallet er f.eks. lite blant selvstendig næringsdrivende i jordbruk, skogbruk og fiske, mens det er stort blant ikke-yrkesaktive. Hun har foretatt følgende inndeling etter "hovedperson" i husholdningen:

- i) Selvstendig næringsdrivende i jordbruk, skogbruk og fiske
- ii) Selvstendig næringsdrivende ellers
- iii) Lønnstakere
- iv) Ikke-yrkesaktive

Foreslår å bruke denne grupperingen, og innfører variablene:

$$X_{12i} = \begin{pmatrix} 1 & \text{hvis person 01 i hush.nr. } i \text{ er selvstendig nærings-} \\ & \text{drivende i jordbruk, skogbruk eller fiske} \\ 0 & \text{ellers} \end{pmatrix}$$

$$X_{13i} = \begin{pmatrix} 1 & \text{hvis person nr. 01 i hush.nr. } i & \text{er annen selvstendig} \\ & \text{næringsdrivende} \\ 0 & \text{ellers} \end{pmatrix}$$

$$X_{14i} = \begin{pmatrix} 1 & \text{hvis person nr. 01 i hush.nr. } i & \text{ikke er yrkesaktiv} \\ 0 & \text{ellers} \end{pmatrix}$$

Referansegruppe: person nr. 01 i hush. er lønnstaker.

6. Inntekt

Petra Vestbye viser at frafallet varierer (marginalt og simultant med antall personer pr. husholdning) med inntekten. Det synes som om frafallet er størst blant husholdninger med særlig lave eller særlig høye inntekter. Det er imidlertid vanskelig å vurdere årsakssammenhenger ut fra dette. Det er heller ikke intuitivt lett å forstå at inntekten i seg selv er særlig bestemmende for frafallet. Det er imidlertid ikke noe i veien for at vi kan ta med også denne variable ved estimering av frafalls-sannsynlighetene og foreslår følgende inndeling:

- i) Under kr. 15 000 pr. år
- ii) Kr. 15 000 - kr. 70 000 pr. år
- iii) Over kr. 70 000 pr. år

Innfører variablene:

$$X_{15i} = \begin{pmatrix} 1 & \text{hvis inntekt i hush.nr. } i & \text{er under kr. 15 000} \\ 0 & \text{ellers} \end{pmatrix}$$

$$X_{16i} = \begin{pmatrix} 1 & \text{hvis inntekten i hush.nr. } i & \text{er mellom kr. 15 000} \\ & \text{og kr. 70 000} \\ 0 & \text{ellers} \end{pmatrix}$$

Referansegruppe: inntekten i hush.nr. i er over kr. 70 000.

7. Andre variable?

Det er spesifisert en del høyresidige variable, og jeg foreslår at vi nøyer oss med disse (det er 16 variable). Utplukkingen av variablene og klassifiseringen av dem er foretatt ut fra resultatene i Petra Vestbyes notat og etter "beste skjønn". Det kan bli aktuelt å foreta justeringer senere etter at det er gjort forsøk på å estimere svarsannsynlighetene på et materiale. Det kan være at en da f.eks. finner at en del av de her foreslåtte variable ikke behøver å være med eller at en bør foreta andre grupperinger av variablene.

Relasjonene som danner grunnlaget ved estimering av frafalls-sannsynlighetene får følgende utseende ved en lineær utforming:

$$Y_i = \beta_0 + \sum_{i=1}^{16} \beta_i X_i + U_i$$

hvor U_i betegner restleddet.

Noen merknader:

1. Den venstresidige (avhengige) variable er her en binær (dummy) variabel. Dette medfører visse problemer m.h.t. estimeringen idet f.eks. variansen på restleddet ikke vil være konstant (for alle verdier på X_i -ene) hvis en også forutsetter at det har forventning null (se f.eks. [3] kap. 8.3.).
2. Ved forbruksundersøkelsen 1973 vil en få 5 000 observasjoner hvorav ca. 1 000 vil være frafalte. Dette skulle danne et brukbart grunnlag for estimering av de 17 parametrene.

V. Henvisninger

- 1 G. Elofson: Teoretisk jämförelse mellan tre metoder att behandla bortfall.
- 2 Petra Vestbye: Arbeidsnotat IO 70/6: Metodisk vurdering av forbruksundersøkelsen 1967.
- 3 J. Johnston: The Theory of Econometrics