



*Anne Sundvoll*

**Samordnet levekårsundersøkelse 1999 - panelundersøkelsen**  
Dokumentasjonsrapport

Notater

## **Forord**

Statistisk sentralbyrå har siden 1996 gjennomført leveårsrelaterte undersøkelser etter en modell med to årlige undersøkelser: En temaroterende tverrsnittsundersøkelse i tredje kvartal, og en panelundersøkelse i første kvartal. Panelundersøkelsen ble etablert i 1997. Dette er tredje runde med datainnsamling til denne undersøkelsen, og intervjuarbeidet ble gjennomført i perioden januar - april 1999. Seksjon for leveårsstatistikk var oppdragsgiver for undersøkelsen. Spørreskjemaet er utarbeidet i samarbeid med denne seksjonen, samt eksterne brukere.

Ved seksjon for intervjuundersøkelser var Anne Sundvoll prosjektleder, Thore Nafstad-Bakke IT-ansvarlig og Kari Grøholt intervjuerkontakt.

## **Innhold:**

1.	Hovedresultat.....	4
2.	Utvalg.....	5
3.	Feltarbeidet.....	5
4.	Oversikt over inngangen.....	6
5.	Svar, frafall og avgang i utvalget.....	8
6.	Nekting vanligste frafallsårsak.....	8
7.	Koding av yrke og næring.....	10
8.	Utvalgsskjevhet.....	10
9.	Tiltak for øke oppslutningen og vedlikeholde panelet.....	11
10.	Utvalgsvarians.....	12
11.	Innsamlings- og bearbeidingsfeil.....	13

## **Figurer:**

1.	Oversikt over svarinngangen for reintervjuet.....	6
2.	Oversikt over svarinngangen for førstegangsintervjuet.....	7

## **Tabeller:**

1.	Nøkkeltall 1999.....	4
2.	Svar, frafall og avgang i utvalget.....	8
3.	Frafall etter årsak.....	9
4.	Nye avganger etter årsak 1999.....	9
5.	Akkumulert avgang etter årsak.....	9
6.	Bruttoutvalg, frafall og nettoutvalg etter kjønn, aldersgruppe og landsdel.....	10
7.	Bruttoutvalg, frafall og nettoutvalg for panelet.....	10
8.	Utvalgsvarians.....	10

## **Vedlegg**

Informasjonsbrev nye (N)  
Informasjonsbrev tidligere intervjuede (P)  
Informasjonsbrev tidligere frafall (F)  
Foresattebrev (Fo)  
Spørreskjema  
Kort  
Instruks  
Tilleggsinstruks  
Argumentasjonsliste  
Brosjyre  
Bekreft avtale-brev  
Telefonbrev  
Julekort/resultatrapport

De sist utgitte publikasjonene i serien Notater

I 1996 innførte Statistisk sentralbyrå et nytt system for levekårsrelaterte undersøkelser. Fra og med 1997 har vi gjennomført to levekårsundersøkelser i året; en tverrsnittsundersøkelse på høsten, og en panelundersøkelse i første halvår. Høsten 1998 var helse hovedtemaet i tverrsnittsundersøkelsen. Det var derfor naturlig at spørsmål om helse satte et preg på årets utgave av panelskjemaet.

1999 var tredje runde med datainnsamling til panelundersøkelsen. Undersøkelsen ble gjennomført i samarbeid med seksjon for levekårsstatistikk, og er internt finansiert.

## 1. Hovedresultat

Tabell 1 viser hovedresultatene for undersøkelsen. I 1997 ble 5000 personer i alderen 16-79 år trukket ut til å delta i undersøkelsen. Vi fikk dette året gjennomført intervju med 3890 personer, noe som ga en svarprosent på 79.

I 1998 ble utvalget supplert med til sammen 100 personer, som enten hadde fylt 16 år det siste året, eller som hadde innvandret til Norge i løpet av 1996. Vi fikk i 1998 intervju med 3834 personer, noe som tilsvarer en svarprosent på 78. Etter to runder utgjorde "panelet" (personer med 2-årig tidsserie) 3530 personer.

Før utsending i 1999 ble 18 nye personer kodet som avgang. Samtidig ble noen personer fra gruppen med midlertidige avganger tatt inn i bruttoutvalget igjen. Deretter supplerte vi utvalget med 100 personer etter samme metode som tidligere. I løpet av feltperioden ble ytterligere 67 personer kodet til avgang.

Bruttoutvalget utgjorde da 4985 personer. Vi fikk intervju med 3867 personer - en svarprosent på 78. Det akkumulerte frafallet utgjorde til sammen 1118 personer eller 22 prosent, og bestod i hovedsak av personer som ikke ønsket å delta. "Panelet" (personer med full 3-årig tidsserie) utgjør etter tre runder 3277 personer.

**Tabell 1 Nøkkeltall for samordnet levekårsundersøkelse 1999 - panelundersøkelsen**

Nøkkeltall	Personer	Prosent
Personer trukket ut for intervju totalt	5200	100,0
Nye personer trukket ut for intervju 1999	100	1,9
Ny avgang før utsending 1999	18	0,3
Avgang1999	67	1,3
Bruttoutvalg 1999	4985	100,0
Nettoutvalg (personer oppnådd intervju med)	3867	77,6
Frafall	1118	22,4
Viktigste årsak til frafall:		
Ønsket ikke å delta	696	62,3
Grupper som i forhold til kjennemerkene kjønn, alder og landsdel er over- eller underrepresentert i nettoutvalget med mer enn $\pm 1$ prosentpoeng.		
Alder:		
25-44 år		+ 1,8
67-79 år		-1,3

## 2. Utvalg

I 1997 trakk vi i alt ut 5000 personer i alderen 16-79 år (alder pr. 31.12.96). For at utvalget fortsatt skal være representativt for befolkningen i Norge i alderen 16-79 år, supplerer vi hvert år med 16-åringene og personer innvandret til Norge. Suppleringsutvalget er trukket på samme måte som resten av utvalget, og utvalget vil bli supplert på samme måte hvert år. Personer 80 år og eldre tas ikke ut av utvalget men forsøkes intervjuet hvert år.

I forbindelse med at SSB skulle opprette et panel for levekårsundersøkelsene, var det ønskelig å kunne knytte en forbindelse mellom levekårspanelet og Folke- og bolig tellingen 1990 (FoB90) og tidligere folke- og bolig tellinger. Derfor ble FoB90-utvalget rekonstruert, ved at det ble trukket et 8,3 prosent selvveiende «superutvalg» av FoB-populasjonen i 1990 (hver tolvte person født i 1974 eller tidligere). «Superutvalget» ble så brukt som trekkegrunnlag for levekårspanelet 1997, sammen med to tilleggsutvalg trukket etter de samme prinsippene.

«Superutvalget» er et representativt utvalg av personer født i 1974 eller tidligere i 1990. For å kunne bruke utvalget i 1997, ble personene plassert inn i dagens bostedskommune og døde og utvandret i perioden 1991-96 ble fjernet. Fordi personer som har innvandret i perioden 1991-96 og personer født i 1975-80 manglet, var ikke utvalget selvveiende i forhold til dagens populasjonstall i kommunene. Derfor ble det trukket ett 8,3 % ekstrautvalg i hver av disse to gruppene til levekårspanelet etter samme opplegg som ble brukt i FoB90, men der personenes adresse i 1996 ble brukt. Før trekkingen ble personene sortert med hensyn til kommune- og delområdenummer, familiestørrelse, familienummer og de to siste sifrene i fødselsnummeret.

De tre utvalgene ble satt sammen og kontrollert for eventuelle dubletter. Utvalget ble også kontrollert for faktorer som kjønn og alder for å sikre at det var representativt. Deretter trakk vi et selvveiende, stratifisert utvalg med utgangspunkt Statistisk sentralbyrås utvalgsplan.

I utvalgsplanen er hele landet inndelt i et sett av utvalgsområder, som igjen er gruppert i 109 strata. Utvalgsområdene er kommuner eller grupper av kommuner. Kommuner med lavt innbyggertall er slått sammen med andre kommuner, slik at alle utvalgsområder har minst 7 prosent av samlet innbyggertall i det stratomet området tilhører. I en del tilfeller er mindre omegnskommuner til folkerike kommuner slått sammen med den store kommunen til ett område. Alle kommuner med mer enn 30000 innbyggere og en del kommuner med mellom 25000 og 30000 innbyggere er tatt ut som egne strata. De andre utvalgsområdene er stratifisert innen hvert fylke etter næringsstruktur, bosettingstetthet, sentralitet, pendlings- og handelsmønstre, mediadekning og kommunikasjoner. Utvalg som trekkes etter utvalgsplanen trekkes i to trinn: I første trinn trekkes et utvalgsområde fra hvert stratum. Utvalgsområder som utgjør egne strata er trukket ut med 100 prosent sannsynlighet. De resterende er trukket ut med en sannsynlighet proporsjonal med innbyggertallet i utvalgsområdet. I andre trinn er utvalget av personer trukket tilfeldig fra de 109 utvalgsområdene. Trekkingen på andre trinn foregår slik at utvalget er selvveiende når begge trinn ses under ett.

## 3. Feltarbeidet

Den ordinære intervjuperioden var fra 25. januar til 26. februar 1999. I tillegg hadde vi satt av to uker til seniorintervjuernes oppfølging av frafall, og endelig slutt dato var bestemt til 12. mars.

På grunn av forsinkelser i programmeringen, startet ikke feltarbeidet før 1. februar. Vi hadde denne gang valgt å programmere to intervjukjemaer, ett for personer som ikke hadde deltatt før - personer trukket i 1999 og personer som tidligere hadde vært frafall ("**førstegangsintervjuet**") - og ett for intervju av personer allerede etablert i panelet ("**reintervjuet**").

Som følge av denne forsinkelsen, ble den ordinære feltperioden forlenget med én uke. Oppfølgingsarbeidet strakk seg helt til 9. april før vi satte den endelige sluttstreken. Vår erfaring var at intervjuerne gjorde en kjempejobb, men at det også ligger en viss drahjelp i at personer som har svart tidligere stort sett er veldig positive til å delta på nytt.

Opplysningene til reintervjuet ble hovedsakelig samlet inn ved hjelp av telefon, og intervjutiden var beregnet til 30 minutter i gjennomsnitt. Personer som hadde vært med før og som ønsket besøk framfor å svare over telefon, kunne imidlertid besøkes dersom reiseavstanden ikke var for lang.

De som var med for første gang (nye, og personer i gruppen for tidligere frafall), ble i hovedsak intervjuet ved besøk. For disse var intervjutiden beregnet til ca. 45 minutter i snitt.

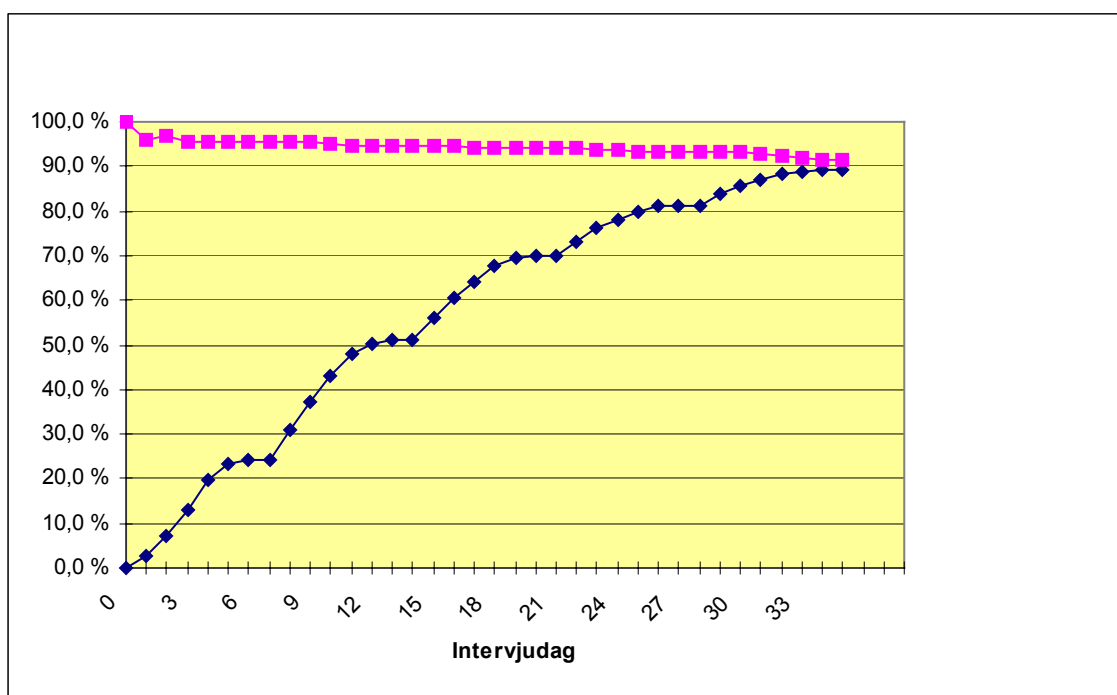
Det ble i alt gjennomført 164 intervju ved besøk og 3703 intervju over telefon.

#### 4. Oversikt over inngangen

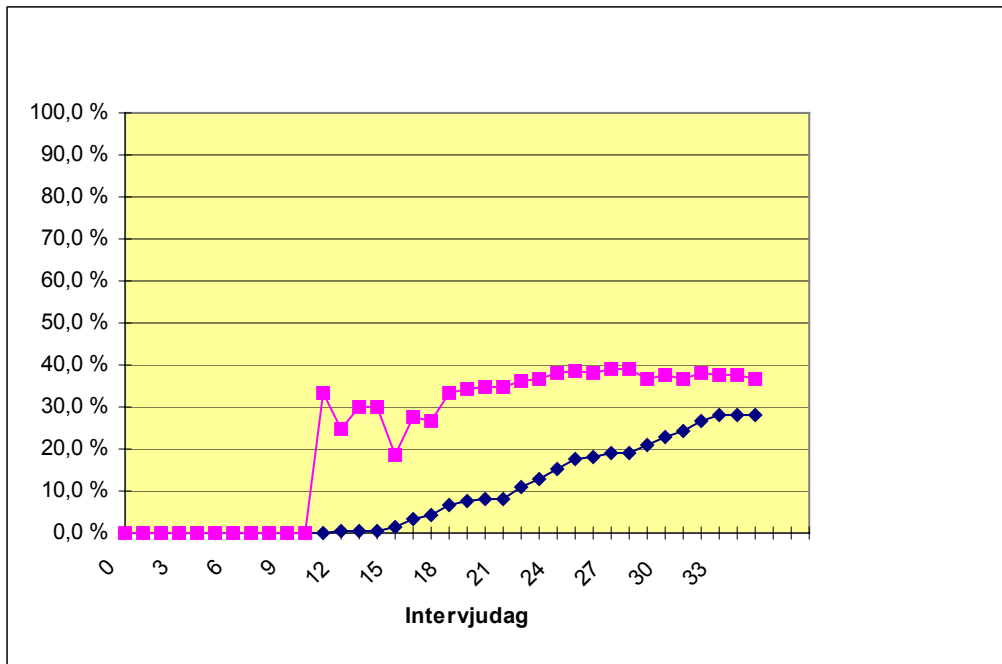
Figur 1 og 2 viser hvordan svarprosenten utviklet seg i løpet av den ordinære intervjuperioden for hhv. reintervjuet og førstegangsintervjuet. Hver av figurene viser en linje over akkumulert svarprosent av bruttoutvalget (minus avganger) for hver dag i feltperioden.

Figur 1 går i bølger oppover, og kurven flater ut i helgene. Etter hvert flater kurven mer og mer ut, noe som er normalt utover i feltperioden.

**Figur 1 Oversikt over svarinngangen for reintervjuet**



**Figur 2 Oversikt over svarinngangen for førstegangsintervjuet**



Figur 2 har et litt mer merkelig utseende. På grunn av forsinkelser i programmeringen, kom ikke intervjuarbeidet knyttet til førstegangsintervjuet i gang før 11 dager ut i den nye feltperioden. Etter hvert flater kurven ut på i underkant av 30 % svar.

Figurene viser også grafer over svarprosenten av det som til enhver tid er sendt inn fra intervjuerne. Svarprosenten er svært høy i begynnelsen, fordi intervjuerne hovedsakelig sender inn intervju i starten. Etter hvert som frafallet begynner å komme inn, faller kurven noe. Til slutt møtes de to kurvene og stabiliserer seg rundt det som blir den endelige svarprosenten.

Vi hadde som mål å nå en samlet svarprosent på 78 prosent, som sist. For å nå dette målet måtte vi ha en svarprosent på godt over 90 % blant personene som skulle reintervjues. Blant personer intervjuet for første gang, var det på forhånd vanskelig å anslå utfallet. En kunne forvente 80 % svar blant de helt nye (personer supplert i 1999), mens blant personer i gruppen av tidligere frafall hadde vi et berettiget håp om å rekruttere hver fjerde.

Bruttoutvalget utgjorde 4985 personer, og av disse ble 478 personer merket med frafallskode sentralt forut for intervjuperioden og ble derfor i realiteten aldri sendt ut til intervjuerne. Disse personene utgjør en kjerne av "harde nektere", som foreløpig ikke skal kontaktes på noen år. Vi vil komme tilbake med vurderinger av hvordan vi eventuelt kan kontakte personer i denne gruppen.

Det viste seg at noen av personene med frafallskode som *ble sendt ut*, også burde vært merket med kode for "harde nektere". Noen av intervjuerne reagerte på at personer de kontaktet ved flere anledninger uttrykkelig hadde gitt beskjed om at de ikke var interesserte i å delta. Her balanserer vi på en fin linje mellom å gjøre anstrengelser for å rekruttere, og det å godta et "nei" fra personer vi iherdig forsøker å overtale.

Etter første uke hadde vi intervju med 999 personer. De tre neste ukene fikk vi hhv. 1093, 799 og 505 intervju. De siste ukene ble brukt til oppfølging av frafall og personer vi ikke hadde fått kontakt med, og antall intervju var naturligvis langt lavere (henholdsvis 356 og 115). Fordi dette for det meste var oppfølging av frafall, bidro det likevel til å øke svarprosenten, som altså ble 78 til slutt (3867 personer).

Blant dem som ble forsøkt intervjuet, ble imidlertid svarprosenten høyere - 86 prosent. Fordi dette er et panel, har vi dratt med oss en del frafall fra tidligere. Disse tilhører populasjonen og utvalget, selv om vi *ikke* har prøvd å kontakte dem. Derfor bidrar de til å svekke svarprosenten (se neste avsnitt).

## 5. Svar, frafall og avgang i utvalget

Tabell 2 viser bevegelser i utvalget fra 1997 til 1999. Første linje viser akkumulert trukket utvalg før avganger på hhv. 5000, 5100 og 5200 personer. Det årlige bruttoutvalget påvirkes av suppleringer og nye avganger. Årsaken til at tallet for akkumulerte avganger er såvidt likt i 1998 og 1999, skyldes at man i 1998 glemte å inkludere mulige kandidater fra gruppen for midlertidige avganger i bruttoutvalget.

I 1997 deltok 3890 personer i undersøkelsen. Av disse deltok 3530 personer også i 1998, mens 292 personer gikk til frafall. Imidlertid fikk vi i 1998 intervju med 231 personer blant frafallet i 1997. En slik utvikling fører til at det blir noen "huller" i paneldataene, og vi har foreløpig ikke tatt stilling til hvor lenge en person kan være frafall for han/hun eventuelt kan innlemmes i panelet igjen.

Med datainnsamlingen i 1999 har vi fått et noe mer komplisert bilde. Vi vil nå illustrere utviklingen ved å skille personer med full (3-årig, 2-åring og 1-årig) tidsserie fra dem som har delvis tidsserie (dvs. der et har oppstått ett eller flere "hull"). 3277 personer har etter 1999-innsamlingen full 3-årig tidsserie. Fordi vi årlig supplerer med ca. 100 nye personer, vil den "fulle" tidsserien for en del personer (blant annet personer som på intervjutidspunktet er 16 år) strekke seg over et noe kortere tidsrom.

Tabell 2 Svar, frafall og avgang i utvalget

	Fordeling 1997	Fordeling 1998	Fordeling 1999
Uttrukne personer i alt	5000	5100	5200
Permanente avganger akkumulert	62	136	180
Midlertidige avganger saldo	27	49	35
Bruttoutvalg	4911	4915	4985
Svar i alt	3890	3834	3867
Derav			
Full tidsserie	3890	3530	3277
Tidsserie -1 år	-	73	67
Tidsserie -2 år	-	-	65
Delvis tidsserie	-	231	458
Frafall i alt	1021	1081	1118
Derav			
Fullt frafall	1021	789	695
Delvis frafall	-	292	423

Fordi vi anser at frafall i utvalgsundersøkelser ofte er situasjonsbetinget, valgte vi også i 1999 å kontakte noen av de som gikk til frafall i 1998 og 1997. Dette medvirker til at det akkumulerte frafallet holder seg så vidt stabilt, selv om vi årlig får et nytt bidrag til frafallet.

Personer som vi har vurdert ikke bør kontaktes, er også med i frafallsberegningene. Til sammen utgjorde denne gruppen 478 personer, eller 43 prosent av frafallet i 1999. I samråd med seksjon 350, har vi bestemt oss for å vente noen år før vi forsøker å intervju disse personene på nytt.

Resultatene viser i all hovedsak at det er stor svarvillighet blant dem som har vært med tidligere, og det lover godt for levekårspanelet i framtiden med hensyn på oppslutningen. Etterhvert må man imidlertid regne med en viss tretthetseffekt, noe som kan gi utslag på svarprosenten.

## 6. Nekting vanligste frafallsårsak

Tabell 3 viser at den vanligste årsaken til frafall er at intervjupersonen ikke ønsker å delta, med 62,3 prosent. Denne frafallsgruppen utgjør 22,4 prosent av bruttoutvalget. Deretter følger "ikke oppnådd kontakt" (19,7 prosent), "forhindret fra å delta" (11,0 prosent) og til sist "andre grunner til å ikke delta" (7,0 prosent).



**Tabell 3 Frafall etter årsak. Bruttoutvalg 1999.**

	<b>Antall personer</b>	<b>Prosent av frafall</b>	<b>Prosent av bruttoutvalg</b>
<b>Frafall i alt</b>	<b>1118</b>	<b>100,0</b>	<b>22,4</b>
Ønsker ikke å delta	696	62,3	14,0
Forhindret fra å delta	123	11,0	2,5
Ikke oppnådd kontakt	220	19,7	4,4
Andre grunner til å ikke delta	79	7,0	1,6

Tabell 4 og 5 viser fordelingen på avgangsårsaker for 1999-innsamlingen separat (tabell 4) og akkumulerte avganger for undersøkelsen i sin helhet (tabell 5). De fleste avgangene skyldes at respondenten er død, har flyttet til utlandet, bor midlertidig i utlandet eller at respondenten er fast bosatt i institusjon. Årsaken til at vi her har skilt mellom personer som bor midlertidig i utlandet og de som har flyttet permanent til utlandet, er at vi *ikke* anser de som bor der midlertidig som permanente avganger. Denne gruppen omfatter studenter og andre som oppholder seg i utlandet i tidsbegrensede perioder, og som vil flytte tilbake til Norge i løpet av et par år. Vi ønsker derfor å innlemme dem i panelet når de flytter tilbake.

**Tabell 4 Nye avganger etter årsak 1999**

	<b>Antall personer</b>	<b>Prosent</b>
<b>Avgang i alt</b>	<b>85</b>	<b>100,0</b>
Respondenten er død	21	24,7
Respondenten er flyttet perm. til utlandet	20	23,5
Respondenten bor midlertidig i utlandet	30	35,3
Respondenten er bosatt i institusjon	7	8,2
Annen avgangsgrunn	7	8,2

**Tabell 5 Akkumulerte avganger etter årsak**

	<b>Antall personer</b>	<b>Prosent</b>
<b>Avgang i alt</b>	<b>215</b>	<b>100,0</b>
Respondenten er død	64	29,8
Respondenten er flyttet til utlandet	65	30,2
Respondenten bor midlertidig i utlandet	35	16,3
Respondenten er forpleid i institusjon	44	20,4
Annen avgangsgrunn	7	3,3

## 6. Koding av yrke og næring

I denne undersøkelsen har vi to ulike yrkeskodinger. Intevjuerne koder yrke under intervjuet med 3-siffer-kode (yrkesgruppe) etter Standard for yrkesgruppering i offentlig norsk statistikk (1965). I tillegg skal de gi så nøyaktige opplysninger at yrke også kan kodes manuelt i ettertid. Etter at datainnsamlingen er ferdig blir yrke kodet sentralt etter ny standarden med fire siffers kode.

Det er bare personer som svarer for første gang og personer som har skiftet arbeidssted eller yrke i bedriften som har gitt opplysninger om yrke og kan kodes. For resten av utvalget har vi disse opplysningene fra tidligere.

Næring kodes med tre sifre etter NACE. Bare de som er med for første gang og de som har skiftet arbeidssted har gitt opplysninger om bedriften og kan kodes. For de øvrige har vi opplysninger fra 1997 og 1998. De manuelle kodingen gjøres av seksjon 450 i Kongsvinger.

## 7. Utvalgsskjevhet

Frafall fører til utvalgsskjevhet når fordelingen av et bestemt kjennemerke er annerledes blant dem som svarte (nettutvalget) enn blant dem det var mulig å intervjuet (bruttoutvalget). Utvalgsskjevhet i forhold til ett kjennemerke medfører ikke nødvendigvis at nettoutvalget er skjevt i forhold til andre kjennemerker. Omvendt innebærer godt samsvar mellom fordelingene i netto- og bruttoutvalget for ett eller flere kjennemerker ikke noen garanti for at utvalget ikke er skjevt på andre kjennemerker.

Tabell 6 og 7 viser hvordan kjennemerkene kjønn, alder og landsdel er fordelt i bruttoutvalget, frafallet og nettoutvalget for 1999-innsamlingen og for den 3-årige tidsserien. Det er små forskjeller mellom brutto- og nettoutvalget når det gjelder kjønn og landsdel både totalt og for den fulle tidsserien.

**Tabell 6 Bruttoutvalg, frafall og nettoutvalg totalt 1999 etter kjønn, aldersgruppe og landsdel. Prosent**

	<b>Bruttoutvalg</b>	<b>Frafall</b>	<b>Nettoutvalg</b>
<b>I alt</b>	100,0	100,0	100,0
<b>Kjønn:</b>			
Menn	48,7	49,6	48,4
Kvinner	51,3	50,4	51,6
<b>Aldersgrupper:</b>			
16 - 24 år	14,1	13,7	14,2
25 - 44 år	39,4	33,5	41,2
45 - 66 år	32,7	33,2	32,6
67 - 79 år	12,4	17,2	11,1
80 - år	1,3	2,5	1,0
<b>Landsdel:</b>			
Oslo og Akershus	22,5	25,6	21,6
Hedmark og Oppland	8,3	8,3	8,3
Østlandet ellers	19,7	21,6	19,2
Agder og Rogaland	13,6	10,8	14,4
Vestlandet	17,1	16,5	17,3
Trøndelag	8,7	7,1	9,1
Nord-Norge	10,1	10,2	10,1
<b>Tallet på personer</b>	<b>4985</b>	<b>1118</b>	<b>3867</b>

Aldersmessig viser tabell 6 en overrepresentasjon av personer 25-44 år på 1,8 prosentpoeng i nettutvalget, og en underrepresentasjon av personer 67-79 år på 1,3 prosentpoeng. Skjevheten i utvalget er så liten at vi ikke har funnet grunn til å vekte resultatene.

**Tabell 7 Bruttoutvalg, fullt frafall, delvis tidsserie og full tidsserie (1, 2 og 3 år) etter kjønn, alder og landsdel. Prosent**

	<b>Bruttoutvalg</b>	<b>Fullt frafall</b>	<b>Delvis tidsserie</b>	<b>Full tidsserie</b>
<b>I alt</b>	100,0	100,0	100,0	100,0
<b>Kjønn:</b>				
Menn	48,7	51,9	46,4	48,6
Kvinner	51,3	48,1	53,6	51,4
<b>Aldersgrupper: (pr. 31.12.98)</b>				
16 - 24 år	14,1	6,9	19,8	14,1
25 - 44 år	39,4	31,2	40,3	40,9
45 - 66 år	32,7	39,6	26,7	32,9
67 - 79 år	12,4	19,3	12,0	11,1
80- år	1,3	3,0	1,2	1,0
<b>Landsdel:</b>				
Oslo og Akershus	22,5	23,5	24,5	21,8
Hedmark og Oppland	8,3	9,2	7,9	8,2
Østlandet ellers	19,7	22,4	18,8	19,4
Agder og Rogaland	13,6	10,9	13,6	14,1
Vestlandet	17,1	16,1	16,2	17,5
Trøndelag	8,7	6,3	8,7	9,1
Nord-Norge	10,1	11,5	10,1	9,9
<b>Tallet på personer</b>	<b>4985</b>	<b>695</b>	<b>881</b>	<b>3409</b>

Aldersmessig viser den fulle tidsserien en overrepresentasjon av personer i alderen 25-44 på 1,5 prosentpoeng og en underrepresentasjon av personer i alderen 67-79 år på 1,3 prosentpoeng.

#### **4. Tiltak for øke oppslutningen og vedlikeholde panelet**

Etter tre innsamlingsrunder ser vi at frafallet har stabilisert seg, og at vi har en kjerne med personer som det fortsatt vil være mulig å følge i noen år framover. Vi kan ikke snakke om slitasjeeffekter etter tre år, men det vil være viktig allerede på et så tidlig tidspunkt å forsøke å finne tiltak som kan motivere for videre deltakelse.

Et viktig poeng i denne diskusjonen er at vi foreløpig har forespeilet respondenten deltagelse i 5 år. På ett tidspunkt må vi fortelle at vi er interessert i å forlenge tidshorizonten, med fare for dette kan påvirke motivasjonen til å delta.

For å holde på interessen, må man forsøke å gjøre belastningen det er å delta så liten som mulig. Vi vil forsøke.....

- at samme intervjuer kommer tilbake hvert år, dersom relasjonen mellom IO og intervjuer er god.
- å ta nye tema inn i undersøkelsen, slik at spørreskjemaet ikke oppfattes som identisk år etter år.
- å skreddersy den årlige resultatrapporten til ulike grupper, slike at alle skal ha relevant og interessant lesestoff.
- å bedre begrunne behovet for tidsserier overfor både intervjuere og respondenter.

Forut for denne meget viktige tredjerunden la vi vekt på følgende:

- Intervjuerne fikk god informasjon om undersøkelsen, blant annet ved innlegg i "Intervjuer'n".
- Informasjonsbrevet (IO-brevet) var skreddersydd til de ulike gruppene, et introduksjonsbrev til de helt nye, et brev med spesialinformasjon til de som hadde deltatt tidligere, og et brev for å rekruttere tidligere frafall. Foreldre til personer under 18 år får alltid eget brev om undersøkelsen.
- Alle som har spørsmål om undersøkelsen kunne ringe oss gratis på grønt nummer.
- Vi sendte ut en egen informasjonsbrosjyre om undersøkelsen, med mer utførlig informasjon enn det som kom fra i IO-brevet.
- Vi laget en egen liste med forslag til argumenter intervjueren kunne bruke for å overtale personer som i utgangspunktet ikke ønsket å være med.
- Alle som var trukket ut fikk tilsendt resultater fra undersøkelsen i desember 1998 sammen med en julehilsen. Sendingen inneholdt også et adresseforandringskort som kunne returneres ved eventuell skifte av adresse.
- Alle som svarte ble med i en trekning av i alt 21 gavekort; ett på kr 10 000 og 20 på kr 1 000.
- Intervjuerne fikk et brev de kunne sende til respondentene tre til fire dager før avtalt tidspunkt for å bekrefte avtalen når de laget avtale om intervju fram i tid.
- Intervjuerne fikk et brev de kunne sende når de ikke klarte å spore opp et telefonnummer, hvor respondenten ble oppfordret til å ta kontakt.
- Eget oppfølgingsbrev til seniorintervjuerne.

## 5. Utvalgsvarians

Fordi resultatene bygger på opplysninger om et utvalg av den befolkningen som undersøkelsen dekker, er det knyttet en viss usikkerhet til dem. Denne usikkerheten kalles utvalgsvarians. Siden utvalget er trukket etter reglene for tilfeldig utvalg, er det mulig å beregne størrelsen på utvalgsvariansen.

Et hyppig brukt mål på usikkerheten i resultatet for et kjennemerke, er standardavviket til den observerte verdien av dette kjennemerket. Størrelsen på dette standardavviket avhenger av tallet på observasjoner i utvalget, måten utvalget er trukket på og av fordelingen til det aktuelle kjennemerket i befolkningen. Fordelingen i befolkningen kjenner vi ikke, men det er mulig å anslå standardavviket ved hjelp av observasjonene i utvalget.

Det er ikke foretatt egne beregninger av slike anslag for denne undersøkelsen, men tabell 8 viser størrelsen på standardavviket for observerte prosentandeler ved ulike utvalgsstørrelser. Av tabellen går det fram at usikkerheten øker når antall observasjoner minker og når prosenttallet nærmer seg 50.

**Tabell 8 Forventet standardavvik for observerte prosentandeler ved ulike utvalgsstørrelser**

Antall observasjoner	5/95	10/90	20/80	30/70	40/60	50/50
25	5,4	7,5	10,0	11,5	12,2	12,5
50	3,8	5,2	7,0	8,0	8,6	8,7
100	2,7	3,7	4,9	5,6	6,0	6,2
200	1,9	2,6	3,5	4,0	4,3	4,3
300	1,5	2,1	2,8	3,2	3,5	3,5
500	1,2	1,6	2,2	2,5	2,7	2,7
1000	0,8	1,2	1,5	1,8	1,9	1,9
1200	0,8	1,1	1,4	1,6	1,7	1,8
1500	0,7	0,9	1,3	1,4	1,5	1,6
2000	0,6	0,8	1,1	1,3	1,3	1,4
2500	0,5	0,7	1,0	1,1	1,2	1,2
3000	0,5	0,7	0,9	1,0	1,1	1,1
4000	0,4	0,6	0,8	0,9	0,9	1,0
5000	0,4	0,5	0,7	0,8	0,8	0,9

Ved hjelp av standardavviket er det mulig å beregne et intervall som med en bestemt sannsynlighet inneholder den sanne verdi av en beregnet størrelse (den verdien vi ville ha fått dersom vi hadde foretatt en totaltelling i stedet for en utvalgsundersøkelse). Slike intervaller kalles konfidensintervaller hvis de er konstruert på en bestemt måte: La  $M$  være den beregnede størrelsen og  $S$  være et anslag for standardavviket til  $M$ . Konfidensintervallet med grensene  $(M-2S)$  og  $(M+2S)$ , vil med omtrent 95 prosent sannsynlighet inneholde den sanne verdien.

Følgende eksempel illustrerer hvordan vi kan bruke tabell 8 for å finne konfidensintervaller: Anslaget på standardavviket til et observert prosenttall på 70 er 2,5 når antall observasjoner er 500.

Konfidensintervallet for den sanne verdi får grensene  $70 \pm 2 \times 2,5$ . Det vil si at intervallet som strekker seg fra 65 til 75 prosent med 95 prosent sannsynlighet inneholder den tallstørrelsen vi ville fått om hele befolkningen hadde vært med i undersøkelsen.

Konfidensintervall kan konstrueres for ulike sikkerhetsnivå. Et konfidensintervall med grensene  $(M+S)$  og  $(M-S)$  vil med 68 prosent sannsynlighet inneholde populasjonens sanne verdi. Et konfidensintervall med grensene  $(M+2,6S)$  og  $(M-2,6S)$  vil med 99 prosent sannsynlighet inneholde populasjonens sanne verdi.

## 6. Innsamlings- og bearbeidingsfeil

I alle undersøkelser, både i totaltellinger og utvalgsundersøkelser, vil det forekomme svar som er feil. Feilene kan oppstå både i forbindelse med innsamlingen og under bearbeidingen.

All datainnsamling i denne undersøkelsen har foregått ved hjelp av bærbare pc-er. Intervjuerne leser opp spørsmålene fra dataskjermen og registrerer svarene til respondentene direkte på pc.

En viktig fordel med denne innsamlingsmetoden er at alle hopp i spørreskjema og tildeling av spørsmål til undergrupper programmeres på forhånd. Dermed reduseres mulighetene for at intervjuerne stiller feil spørsmål til feil personer. Tildelingen av spørsmål kan defineres på grunnlag av opplysninger vi har om respondentene på forhånd, men også ved å bruke de svarene som blir gitt underveis i spørreskjemaet. Det kan imidlertid også være ulemper ved at hoppene er programmert på forhånd. Dersom programmet inneholder feil i hopprutinene som ikke blir oppdaget, vil enkelte spørsmål eller sekvenser kunne bli verdiløse fordi de ikke blir stilt til riktige personer.

Vi tok lærdom at det kompliserte skjemaet i 1998-undersøkelsen og programmerte derfor to varianter for å lette hoppstrukturen. Programmeringen ble mer oversiktlige og skjemaene ble dessuten lettere å teste.

En annen fordel med edb-assistert intervjuing er mulighetene for direkte kontroll av svarene som blir registrert. For hvert spørsmål er det lagt inn grenser for gyldige verdier. I tillegg kan det bygges inn et omfattende feilmeldingssystem dersom intervjueren taster inn et svar som er inkonsistent i forhold til tidligere svar. Slike svarkontroller er lagt inn på to nivå. En del feilmeldinger er absolutte, og kan ikke overstyres. Disse brukes f.eks. ved aldersgrenser. Andre ganger legges det inn advarsler som intervjuerne kan velge å se bort fra, f.eks. ved svært høye inntekter, lange arbeidsdager (over 60 timer) og når to av samme kjønn oppgis å være samboere. Dette skjemaet inneholdt færre kontroller av denne typen enn det som er vanlig, hovedsakelig fordi programmet var så stort at et omfattende feilmeldingssystem kunne ført til at programmet ikke kunne kjøres på intervjuernes pc-er.

Videre gir edb-assistert intervjuing muligheter for å tilpasse ordlyden i spørsmålene avhengig av intervjuerpersonenes situasjon eller svar på tidligere spørsmål. Dette gir intervjusituasjonen et mer personlig preg, og det er lettere for intervjuerne å få til en muntlig samtale.

Alle disse sidene ved edb-assistert intervjuing fører til økt datakvalitet. Det blir færre innsamlingsfeil som skyldes intervjuerens feilregistreringer enn ved bruk av papirskjema. Vi unngår registrering av ugyldige verdier og vi oppnår et redusert frafall på enkeltspørsmål ved at mulighetene for hoppfeil reduseres.

I en panelundersøkelse, som denne, gir også edb-assistert intervjuing mulighet til å lagre svarene og bruke dem i den neste intervjusituasjonen. En godt egnet metode ved retrospektive spørsmål er å stille spørsmålet på følgende måte: "Sist fortalte du at du jobbet ved .....Gjelder dette fortsatt? Denne metoden er tatt i bruk her for å kartlegge endringer med hensyn til blant annet bosted, bolig og arbeidsplass.

Det kan imidlertid også være problemer knyttet til edb-assistert intervjuing. Hvis det først gjøres feil som ikke oppdages i tide, vil disse få langt større konsekvenser enn i papirundersøkelser. I denne undersøkelsen viste det seg dessverre at skjemaet inneholdt en hoppfeil:

1) Det var en feil i datafilen som genererte opplysninger om ektefelles/samboers arbeidsforhold. Intervjuobjektet (IO) ble under intervjuet konfrontert med opplysninger om at ektefelle/samboer hadde arbeid ved forrige intervju, selv om dette ikke stemte. Feilen ble oppdaget tidlig, og vi botet på problemet ved å sende intervjuerne en oversikt over IO som hadde fått feil ektefelle-/samboeropplysninger samt et forslag til hvordan man da kunne manøvrere i skjemaet.

Innsamlingsfeil kan også komme av at respondenten avgir feil svar. Det kan skyldes vansker med å huske forhold tilbake i tiden, eller misforståelser av spørsmål. Når vi spør om forhold som folk erfaringsmessig finner kompliserte, må vi regne med noen svar som ikke stemmer med terrenget:

1) På spørsmålet om boligens strømforbruk svarte noen at de hadde felles måler for våningshus, driftsbygning og korntørke. Personer som leide ut deler av huset, f.eks. en sokkelleilighet, hadde også problemer med å gi anslag for boligen (dvs. de rom som husholdningen disponerer til eget bruk). Vi vil ut fra intervjuemailer sendt oss i løpet av feltperioden vurdere hvordan vi skal behandle/kode hvert enkelt tilfelle.

Vi fikk dessuten flere meldinger om at mengde-/volummålene for ved var vanskelige å forholde seg til.

2) IO som regnet seg som ansatt i eget firma, oppga i noen tilfeller timeverk både som ansatt og som selvstendig næringsdrivende. Dermed ble timeverkstallet for høyt. Misforståelsen ble oppdaget tidlig og korrigert ved en tilleggsinstruks.

3) IO over 18 år som bodde på folkehøyskole hadde problemer med å svare på husholdnings-spørsmålene. Dette gjaldt først og fremst i forhold til hustype, eieforhold, husleie og antall personer i husholdningen.

Det er noen ganger vanskelig å finne gode operasjonelle mål på de teoretiske begrepene vi ønsker å si noe om. Vi mener at disse undersøkelsene vil gi gode levekårsindikatorer, men det vil alltid være noen personer vi intervjuer som mener at det mangler viktige aspekter. Det kan for eksempel være at vi utelater å spørre om en bestemt utgiftsbyrde IO sliter med. I denne innsamlingsrunden fikk vi tilbakemelding om at personer med stor bidragsgjeld ikke syntes de fikk fram hvor dårlig levekårene deres egentlig var.

Innsamlingsfeil kan også oppstå fordi spørsmålene blir oppfattet som ømtålige av enkelte. I slike tilfeller kan respondentene svare feil med vilje. De vurderinger som ligger til grunn for svaret kan også bli påvirket av hva respondenten oppfatter som sosialt ønskelig. I intervjuarbeidet bruker vi to forskjellige metoder - telefon og besøk. Vi har imidlertid ingen dokumentasjon på at vi her har større intervjuereffekt i besøkssituasjonen sammenliknet med telefonintervjuet.

Bearbeidingsfeil er avvik mellom den verdien som registreres inn og den verdien som til slutt rapporteres ut. Slike feil kan oppstå for eksempel under omkodinger. Gjennom maskinelle kontroller har vi søkt å finne feil og rette opp disse.

Når vi har rettet opp feil så langt det er mulig, er erfaringen at de statistiske resultatene i de fleste tilfeller påvirkes forholdsvis lite av både innsamlingsfeil og bearbeidingsfeil. Virkningen av feil kan likevel være av betydning i noen tilfeller, spesielt hvis feilene ikke blir oppdaget.

## De sist utgitte publikasjonene i serien Notater

- 99/34 E. Birkeland (red.): Forskjeller i levekår: Hefte 3: Bruk av velferdsordninger. 126s.
- 99/35 E. Birkeland (red.): Forskjeller i levekår: Hefte 4: Regionale forskjeller. 118s.
- 99/36 M. Stålnacke, J-A. Sigstad Lie og L. Solheim: En analyse av SSBs generelle utvalgsplan fra 1995 basert på næringsvise sysselsettingstall. 83s.
- 99/37 B.O. Lagerstrøm: Trivsels- og arbeidsmiljøundersøkelse blant intervjuere i Statistisk sentralbyrå. 155s.
- 99/38 K.J. Einarsen: Evalueringsrapport for pilotforsøket for FylkesKOSTRA-utdanning. 55s.
- 99/39 L. Rogstad: FoB2000: Adressesamsvar mellom folkeregister og adresseregister i GAB: – rapport fra Lysebu-seminar 8. og 9. desember 1998, – tiltaksplan for bedre adressesamsvar. 39s.
- 99/40 D. Roll-Hansen: Samordnet levekårsundersøkelse 1998 – tverrsnittsundersøkelsen: Dokumentasjonsrapport. 102s.
- 99/41 R. Johannessen: Kommunale gebyrer knyttet til bolig. Januar 1999. 30s.
- 99/42 M. Stålnacke, A.G. Hustoft og L. Solheim: Vurdering av kvalitet i statistikk: En oversettelse av notater fra Eurostat om kvalitetsrapportering. 77s.
- 99/43 E. Engelién, K. Myklebust, J.A. Paulsen og L. Rogstad: FoB2000: Stedfesting av bedrifter – forprosjekt. 40s.
- 99/44 I. Hauge, C. Hendriks, Ø. Hokstad og A.G. Hustoft: Standard for begreper og kjennermerker knyttet til familie- og husholdningsstatistikken. 37s.
- 99/45 E. Rønning: Omnibusundersøkelsene 1998: Dokumentasjonsrapport. 123s.
- 99/46 C. Torp: Situasjonsuttak fra Bedrifts- og foretaksregisteret. 33s.
- 99/47 T.N. Evensen: Utlendingers konsum i Norge: En vurdering av eksporttallene for reisetrafikk i nasjonalregnskapet på bakgrunn av statistikk som belyser forbruket til utenlandske turister i Norge. 28s.
- 99/48 H. Hartvedt (red.): Definisjonskatalog for grunnskoleopplæring for barn og voksne. 14s.
- 1999/49 K. Bjønnes og J. Johansen: FD - Trygd: Dokumentasjonsrapport. Attføringspenger, 1992-1997. 126s.
- 1999/50 E. Høydahl: FoB2000: Rapport fra seminar 4. juni 1999 om kommuneprodukter fra Folke- og bolig tellingen 2000. 32s.
- 1999/51 P.E. Tønjum: Teknisk dokumentasjon av beregningsopplegget for kvartalsvis nasjonalregnskap (KNR). 91s.
- 1999/52 F. Gundersen: Statistikk over etterforskede lovbrudd: Dokumentasjon. 46s.
- 1999/53 N. Arnesen og Ø. Skullerud: Statistikk over emballasjeavfall: Beregningsresultater for 1997. 36s.
- 1999/54 Ø. Kleven: Bruk av kreftundersøkelsen PSA blant menn i alderen 50 til 65 år. 19s.
- 1999/55 P.M. Holt og L. Wiker: Inntekts- og formuesundersøkelsen for aksjeselskaper 1996: Dokumentasjon. 30s.
- 1999/56 B.O. Lagerstrøm: Små og mellomstore bedrifters vurdering av kostnader ved lover og regelverk: Hovedresultater. 129s.
- 1999/57 L.H. Thingstad: Regnskapsstatistikk for varehandel 1996: Dokumentasjon av produksjonsrutiner. 36s.
- 1999/58 P.E. Tønjum: Teknisk dokumentasjon av det årlige realregnskapets FAME-databaser og rutiner. 53s.
- 1999/59 E.J. Fløttum: Konsumgrupperinger i offisiell statistikk. 103s.
- 1999/60 R. Johannessen: Kvalitetssikring av korttidsstatistikk. 26s.