

*Li-Chun Zhang*



**Dokumentasjonsrapport: Den nye estimeringsmetoden for Arbeidskraftundersøkelsen (AKU)**  
Fylkesvis kalibrering med landsetterstratifiserte vekter som startverdier

## **Forord**

Prosjektet er et samarbeid med Seksjon for arbeidsmarked. Kontaktpersoner har vært Inger Håland, Ole Sandvik og Per Svein Aurdal. Samtidig har Helge Næsheim fra Seksjon for arbeidsmarked, Leiv Solheim og Jan Bjørnstad fra Seksjon for statistiske metoder og standarder vært medvirkende i arbeidet. I tillegg har Ib Thomsen fra Avdeling for personstatistikk bidratt med nyttige kommentarer og diskusjoner.

## 1 Innledning

Etter en kort oppsummering av dagens opplegg, gjennomgås en oversikt over generelle teorier for kalibrering før forslaget og konklusjonen til den nye AKU-estimeringsmetoden presenteres. Tekniske detaljer er blitt lagt til et appendiks. Forøvrig henvises det også til Zhang (1997b).

## 2 Dagens opplegg: Enkel oppblåsning med etterstratifisering på landsnivå

Etterstrata er i utgangspunktet laget ved at man kryss-klassifiserer utvalget ifølge de kjennetegn som var fastsatt på forhånd. Disse består av kjønn, alder, registersyssetning inkl. næring, bostedkommunens sentralitet og uføretrygd (Boije, 1996). Etterpå blir en del av dem slått sammen slik at i alt står man igjen med ca. 90 etterstrata.

Med enkel oppblåsning mener man at alle IO-ene som hører til det samme etterstratum får lik oppblåsningsfaktor (også kalt *vekt*), uavhengig av deres trekk sannsynligheter. Nærmere sagt er vekten bestemt ved

$$\frac{\text{Antall enheter som hører til etterstratum i populasjonen}}{\text{Antall IO-er som hører til etterstratum i utvalget}}.$$

**Eksempel 1**<sup>1</sup> *Etterstratifiser utvalget etter kjønn — (mann, kvinne) og alder — (16-24, 25-64, 65-74), får man etterstratum (a) menn mellom 16 og 24 år, (b) menn mellom 25 og 64 år, (c) menn mellom 65 og 74 år, (d) kvinner mellom 16 og 24 år, (e) kvinner mellom 25 og 64 år, og (f) kvinner mellom 65 og 74 år. Vekt for IO-er fra stratum (a) er gitt ved*

$$w_a = \frac{\text{Antall menn mellom 16 og 24 år i populasjonen}}{\text{Antall menn mellom 16 og 24 år i utvalget}}.$$

*Lignende for stratum (b) - (f).*

Legg da merke til at hvis man bruker de etterstratifiserte vektene til å estimere størrelsene på etterstrata, ville tallene stemme eksakt med populasjonen. Det samme gjelder også for alle marginale størrelser i populasjonen.

**Eksempel 2** Eksempel 1, forts.: *Størrelsen på etterstratum (a) estimeres ved å summere opp vektene til alle IO-ene som hører til etterstratum (a) i utvalget, noe som gir oss*

$$\begin{aligned} w_a \cdot (\text{Antall menn mellom 16 og 24 år i utvalget}) \\ = (\text{Antall menn mellom 16 og 24 år i populasjonen}) \end{aligned}$$

*Samtidig estimeres f.eks. den marginale størrelsen for menn i populasjonen ved å summere opp vektene til alle IO-ene som er menn, dvs. de som hører til etterstratum (a), (b) og (c), noe som ville gi oss antallet menn mellom 16 og 74, altså antallet menn i populasjonen.*

**Kommentar 1** *Denne egenskapen er nødvendig for at estimatoren skal være betinget forventningsrett når trekk sannsynlighetene er konstante innen hvert etterstratum gitt etterstrata-størrelsene i utvalget.*

Ofte blir denne måten å etterstratifisere på referert til som *komplett* etterstratifisering. Den er enkel i bruk, men sårbar når antall strata øker eller størrelsen på utvalget blir mindre, da det

---

<sup>1</sup>Eksempelene er laget for å gjøre begreper/teorier bedre tilgjengelig, spesielt for lesere som fra før ikke er godt kjent med det relevante stoffet.

lett kan resultere i tomme strata i utvalget, som f.eks. når man i AKU skal gjennomføre dagens etterstratifisering innen hvert enkelt fylke.

**Kommentar 2** Uttrykket for den etterstratifiserte estimatoren er gitt som formel (1) i appendiks A.1.

## 3 Kalibrering

### 3.1 Naturlig kalibrering

Vektene til utvalget sies å være (naturlig) kalibrert hvis de gir korrekte tall for de marginale størrelsene i populasjonen mht. kjennetegnene for etterstratifisering. Mao., anta at etterstratifisering er basert på  $k$  kjennetegn som hver har  $c_j$  ( $1 \leq j \leq k$ ) forskjellige kategorier, da har man i utgangspunktet  $c_1 \times \dots \times c_k$  etterstrata og  $c_1 + \dots + c_k$  marginale totaler for kalibrering.

**Eksempel 3** Eksempel 1, forts.: Mht. de 2 kjennetegn (kjønn og alder) som ble brukt får man  $(c_1, c_2) = (2, 3)$ , og i alt 5 marginale størrelser i populasjonen å kalibrere mot, dvs. (i) antall menn, (ii) antall kvinner, (iii) antall personer mellom 16 og 24 år, (iv) antall personer mellom 25 og 64 år, og (v) antall personer mellom 65 og 74 år. Dette er altså et tall færre enn ved etterstratifisering (6 etterstrata). I tillegg minner man om at de etterstratifiserte vektene er kalibrert pga. konstruksjon (Eks. 2).

I de fleste tilfellene finnes det uendelig mange mulige valg for vektene som tilfredstiller kalibreringskravene på utvalgte totaler i populasjonen. For å kunne få et bestemt svar bruker man i tillegg (a) et sett initiale vektorer  $\{a_i; i \in s\}$  og (b) en funksjon til å måle 'avstanden' mellom  $a_i$  og den kalibrerte  $w_i$ . De kalibrerte vektene skal ha minimum 'avstand' til  $\{a_i\}$  samtidig som de tilfredstiller kalibreringskravene.

**Kommentar 3** Det formelle uttrykket for den kalibrerte estimatoren er gitt som formel (7) i appendiks A.2.

Det er både interessant og viktig å notere at, så lenge de initiale vektene er identiske innen hvert etterstratum, vil kalibrering mot stratumtotalene gi oss de etterstratifiserte vektene som kalibrerte vektorer, altså uavhengig av de initiale vektene ellers.

**Eksempel 4** Man kan f.eks. ganske enkelt la  $\{a_i\}$  være etterstratumindeks til den  $i$ -te IO-en, dvs.  $a_i = h$  hvis den  $i$ -te IO-en hører til det  $h$ -te etterstratum. Etter kalibrering får man  $\{w_i\}$  som stemmer med størrelsene på etterstrata i populasjonen, da er  $w_i$  nødt til å være den etterstratifiserte vekten.

**Kommentar 4** Den formelle begrunnelsen finner man i appendiks A.2.

Samtidig er det slik at, så lenge de initiale vektene er identiske innen hvert etterstratum, vil de kalibrerte vektene også være identiske innen hvert etterstratum, altså uavhengig av hvilke totaler i populasjonen man kalibrerer mot.

**Eksempel 5** Eksempel 1, forts.: I stedet for å kalibrere mot de 5 naturlige marginale totalene (Eks. 3), er det fullt lovlig at man f.eks. velger følgende 3 totaler: (i) antall menn, (ii) antall kvinner mellom 16 og 64 år, og (iii) antall kvinner mellom 65 og 74 år. De kalibrerte vektene vil allikevel være like innen hvert etterstratum (a) - (f) så lenge man starter med identiske initiale vektorer innen hvert av disse etterstrata.

**Kommentar 5** Den formelle begrunnelsen finner man i appendiks A.2.

### 3.2 Kalibrering med tomme etterstrata

Mens etterstratifisering bryter sammen i tilfellet tomme etterstrata, er kalibrering gjennomførbar så lenge det ikke finnes tomme marginaler i utvalget.

**Eksempel 6** Eksempel 1, forts.: Etterstratifisering er udefinert hvis noen av de 6 etterstrata er tomme i utvalget. Mens naturlig kalibrering (Eks. 3) slutter å fungere bare hvis ingen i utvalget er (i) menn, eller (ii) kvinner, eller (iii) mellom 16 og 24 år, eller (iv) mellom 25 og 64 år, eller (v) mellom 65 og 74 år — altså tomme marginaler.

**Kommentar 6** Forskjellen i risiko vokser fort når antall kjennetegn eller/og antall kategorier for hvert kjennetegn øker. Mens etterstratifisering må dekke  $c_1 \times \dots \times c_k$  etterstrata, trenger kalibrering bare å dekke  $c_1 + \dots + c_k$  marginaler.

Ofte brukt i tilfellet tomme etterstrata er å slå disse sammen med andre slik at utvalget ikke lenger inneholder tomme etterstrata. I sammenligningen med kalibrering virker denne metoden mer tilfeldig rett og slett fordi de to dekker tomme etterstrata på forskjellige måter.

**Eksempel 7** Eksempel 1, forts.: Vi er nødt til å innføre litt notasjon her. Identifiser etterstratum  $(a) - (f)$  med  $(i, j)$  der  $i = 1$  betegner menn og  $i = 2$  kvinner, samtidig som  $j = 1$  betegner 16 - 24 år og  $j = 2$  for 25 - 64 år og  $j = 3$  for 65 - 74 år. La  $N_{ij}$  være antall enheter som hører til etterstratum  $(i, j)$  i populasjonen, og  $n_{ij}$  antall i utvalget. På denne måten er f.eks.  $N_{12}$  antall menn mellom 25 og 64 år i populasjonen, og  $n_{23}$  antall kvinner mellom 65 og 74 år i utvalget. De etterstratifiserte vektene for IO-ene fra etterstratum  $(i, j)$  er gitt som  $q_{ij} = N_{ij}/n_{ij}$ . Anta f.eks. at  $n_{11} = 0$ , dvs. det finnes ingen menn mellom 16 og 24 år i utvalget, slik at etterstratifisering bryter sammen.

Vurder først å slå etterstratum  $(1, 1)$  med f.eks. etterstratum  $(2, 1)$ , slik at nye etterstratifiserte vektene  $v_{ij}$  er gitt som

$$v_{21} = \frac{N_{11} + N_{21}}{n_{21}} \quad v_{22} = \frac{N_{22}}{n_{22}} \quad v_{23} = \frac{N_{23}}{n_{23}} \quad v_{12} = \frac{N_{12}}{n_{12}} \quad v_{13} = \frac{N_{13}}{n_{13}}$$

I tilfellet treksannsynligheten er konstant innen hvert etterstratum gitt etterstrata-størrelsene i utvalget, vil den nye etterstratifiserte estimatoren ikke være skjev kun hvis  $\bar{Y}_{11} = \bar{Y}_{21}$ , dvs. gjennomsnittet til den variabelen man ønsker å måle er lik i etterstratum  $(1, 1)$  og  $(2, 1)$  i populasjonen.

Betrakt nå naturlig kalibrering mot de 5 marginale totalene (Eks. 3) langs retning  $i = 1, 2$  og  $j = 1, 2, 3$ : (i) for at den marginale totalen for  $j = 1$ , dvs. personer mellom 16 og 24 år, skal være riktig er det nødvendig at den kalibrerte vekten  $w_{21} = v_{21}$  ettersom  $n_{11} = 0$ ;

(ii) la  $w_{12} = q_{12} + b$ , for at totalen for  $j = 2$  skal være riktig må man da ha

$$w_{22} = [(N_{12} + N_{22}) - w_{12}n_{12}]/n_{22} = (N_{22} - n_{12}b)/n_{22};$$

(iii) samtidig mht. totalen for  $i = 1$  må man i tillegg ha

$$w_{13} = [(N_{11} + N_{12} + N_{13}) - w_{12}n_{12}]/n_{13} = (N_{11} + N_{13} - n_{12}b)/n_{13};$$

og (iv) til slutt  $w_{23} = (N_{23} - N_{11} + n_{12}b)/n_{23}$ . Den forventede skjevheten er gitt som

$$\begin{aligned} \delta(b) &= \sum_{i,j:(i,j) \neq (1,1)} w_{ij}n_{ij}\bar{Y}_{ij} - \sum_{i,j} q_{ij}n_{ij}\bar{Y}_{ij} \\ &= b \cdot n_{12}(\bar{Y}_{12} - \bar{Y}_{13} + \bar{Y}_{23} - \bar{Y}_{22}) + N_{11}(\bar{Y}_{13} + \bar{Y}_{21} - \bar{Y}_{23}) - N_{11}\bar{Y}_{11}, \end{aligned}$$

som forsvinner hvis

$$b = N_{11}(\bar{Y}_{13} + \bar{Y}_{21} - \bar{Y}_{23} - \bar{Y}_{11})/[n_{12}(\bar{Y}_{12} - \bar{Y}_{13} + \bar{Y}_{23} - \bar{Y}_{22})].$$

Det viktigste her er å legge merke til at kalibrering dekker tomme etterstrata i utvalget ved å utjevne dem over alle ikke-tomme etterstrata, istedenfor å legge alt opp i noen få ikke-tomme etterstrata på tilfeldig vis — slik som sammenslåing av etterstrata gjør.

**Kommentar 7** En grundigere diskusjon finner man i appendiks A.5.

## 4 Kalibrering for AKU

### 4.1 Omlegging i AKU

I forbindelsen med omleggingen i AKU (1996-1997) vurderer man å lage statistikk på fylkesnivå. I tillegg kommer krav (Draft council regulation (EEC) no. (1996): On the organization of labour force sample survey in the Community.) om å dele alder i femårsgrupper — 12 til sammen. Det blir dermed umulig å fortsette med dagens opplegg innen hvert fylke ettersom dette vil resultere i et stort antall tomme etterstrata.

**Kommentar 8** Boije (1996) fremmet noen forslag på endringer, eller mer korrekt redusering i antall etterstrata, slik at man kunne fortsette med etterstratifisering selv når fylke tas inn som et av stratifiseringskriteriene. Men da har man ikke tatt hensyn til inndeling i femårige aldersgrupper.

**Kommentar 9** I AKU kan tomme etterstrata innen fylke variere fra måned til måned og samtidig fra fylke til fylke, slik at sammenslåing av tomme etterstrata vil over tid bli helt tilfeldig.

Studier (Boije, 1996; Zhang, 1997a) peker ut følgende 3 kjennetegn som etterstratifiseringskriterier: (i) kjønn — (mann, kvinne) — i alt 2 kategorier, (ii) alder — (16-19, 20-24, 25-30, ..., 65-69, 70-74) — i alt 12 kategorier, og (iii) registersyssetning inkl. næring — (ikke-sysset, syssel.-primærnæringer, syssel.-sekundærnæringer, syssel.-tertiærnæringer) — i alt 4 kategorier. Til sammen gir dette  $96 (= 2 \times 12 \times 4)$  etterstrata, og  $18 (= 2 + 12 + 4)$  marginale totaler å kalibrere mot innen hvert fylke.

### 4.2 Opplegg

Kalibrering for AKU består nå av følgende to trinn:

1. Etterstratifiser først utvalget etter kjønn, alder og registersyssetning inkl. næring på landsnivå. (I alt 96 etterstrata.)
2. Bruk deretter de etterstratifiserte vektene på landsnivå som startverdier, og kalibrer dem fylkesvis mot de marginale totalene for de samme kjennetegn innen hvert fylke. (I alt 18 marginale totaler for hvert av de 19 fylkene.)

### 4.3 AKU-estimering

AKU-estimering baseres på de fylkesvis kalibrerte vektene. Resultatene kan summeres slik:

1. Tallene på landsnivå har små endringer i forhold til dagens opplegg, samtidig som usikkerheten i disse tallene ligger på det samme nivået som før.
2. Tallene på fylkesnivå er betydelig forbedret i forhold til dagens opplegg: de er tilnærmet forventningsrette samtidig som de omtrent halverer usikkerheten i sysselsettingsestimatene.

**Kommentar 10** Detaljene finner man i Zhang (1997a), appendiks B og tabell 1-3.

## A Teori

### A.1 Den etterstratifiserte estimatoren

Formelt kan man sammenfatte den etterstratifiserte estimatoren slik: La  $U = \{1, \dots, N\}$  betegne populasjonen. La  $y_i$  betegne objekt variabelen for enhet  $i$ , og  $x_i$  dens tilleggsvariabel. Videre er  $U$  delt i  $H$  etterstrata  $\{U_1, \dots, U_H\}$  ifølge  $x$ , med størrelsene  $(N_1, \dots, N_H)$ . La  $s$  betegne utvalget, som igjen er delt i  $H$  grupper  $\{s_1, \dots, s_H\}$ , med størrelsene  $\mathbf{n} = (n_1, \dots, n_H)$ . Den etterstratifiserte estimatoren for totalen  $Y = \sum_{i=1}^N y_i$  er definert som

$$\hat{Y}_{est} = \sum_{h=1}^H N_h \left( \sum_{i \in s_h} y_i / n_h \right) = \sum_h \left( \sum_{i \in s_h} q_h y_i \right) = \sum_h \left( \sum_{i \in s_h} y_i / f_h \right), \quad (1)$$

der  $q_h = N_h / n_h = 1 / f_h$  er den etterstratifiserte vekten til  $i \in s_h$ .

Anta at de betingede treksannsynlighetene gitt  $\mathbf{n}$ ,  $\pi_{i|\mathbf{n}}$ , er konstante innen hvert etterstratum, dvs.  $\pi_{i|\mathbf{n}} = n_h / N_h$  for  $i \in U_h$ . (Dette holder f.eks. når utvalget er enkelt tilfeldig.) Da er estimatoren (1) betinget forventningsrett og dermed også ubetinget. I tilfellet stratifisert enkelt tilfeldig utvalg (etu.) gitt  $\mathbf{n}$ , har  $\hat{Y}_{est}$  betinget varians

$$Var(\hat{Y}_{est} | \mathbf{n}) = \sum_{h=1}^H N_h (q_h - 1) \sigma_h^2, \quad \sigma_h^2 = \sum_{i \in U_h} \frac{(y_i - \bar{Y}_h)^2}{N_h - 1} \text{ og } \bar{Y}_h = \sum_{i=1}^{N_h} \frac{y_i}{N_h}. \quad (2)$$

### A.2 Kalibrering: Definisjonen

Med dummy-indeksing for tilleggsvariablene  $x$ , kan naturlig kalibrering uttrykkes ved

$$\hat{x} = \sum_{i \in s} w_i x_i = X, \quad (3)$$

der  $X = (X_1, \dots, X_k)$  er de marginale totalene for kalibrering. Generelt kan man fritt velge totalene for kalibrering ut ifra alle sammensetninger av  $(U_1, \dots, U_H)$ . La  $R = \{r_1, \dots, r_m\}$  betegne en indeksmengde som består av  $m$  ( $\leq H$ ) elementer der  $1 \leq r_j \leq H$  for  $1 \leq j \leq m$ . La  $N_R = \sum_{r \in R} N_r$  betegne den tilsvarende marginale totalen. Vektene kan sies å være kalibrert mht. bl.a.  $N_R$  hvis

$$\sum_{r \in R} \sum_{i \in s_r} w_i = N_R. \quad (4)$$

Den vanligste metriske avstandsfunksjonen mellom initiale vektorer  $\{a_i; i \in s\}$  og kalibrerte vektorer  $\{w_i; i \in s\}$  er, ifølge Deville og Särndal (1992),

$$D_s = \sum_{i \in s} G(r_i) = \sum_{i \in s} \frac{1}{2} a_i (r_i - 1)^2 \quad r_i = w_i / a_i. \quad (5)$$

Minimering av  $D_s$  med betingelse (3) med Lagrange metoden gir  $r_i - 1 - x_i \lambda = 0$ , slik at

$$w_i = a_i (1 + x_i \lambda), \quad (6)$$

der  $\lambda$  er en kolonnevektor. Lagrange multiplikatoren som tilfredstiller (3) er gitt ved

$$\lambda = \left( \sum_{i \in s} a_i x_i^T x_i \right)^{-1} (X - \tilde{x}),$$

der  $\tilde{x} = \sum_{i \in s} a_i x_i$ . La  $\tilde{y} = \sum_i a_i y_i$ . Det kalibrerte estimatet for  $Y$  er da identisk med det såkalte generaliserte regresjonsestimatet ( $\hat{y}_{greg}$ ), dvs.

$$\hat{y}_{kal} = \tilde{y} + (X - \tilde{x}) \left( \sum_{i \in s} a_i x_i^T x_i \right)^{-1} \left( \sum_{i \in s} a_i x_i^T y_i \right). \quad (7)$$

I det tilfellet at de initiale vektene er konstante innen hvert etterstratum, kan de representeres ved  $\{a_h\}_{h=1}^H$ . Kalibrering mot totalene i etterstrata gir (5) som

$$D_s = \sum_h D_{s,h} \propto \sum_h a_h \sum_{i \in s_h} \left( \frac{w_i}{a_h} - 1 \right)^2 = \sum_h \left( a_h^{-1} \sum_{i \in s_h} w_i^2 - 2 \sum_{i \in s_h} w_i + a_h n_h \right).$$

Ettersom  $\sum_{i \in s_h} w_i = N_h$ , er  $D_s$  minimert ved å sette  $w_i = q_h$  for  $i \in s_h$ . Samtidig er det klart fra (6) at de naturlig kalibrerte vektene også vil være identiske innen hvert etterstratum. Mer generelt ser man fra (4) at et sett av kalibreringsligninger aldri kan defineres på et dypere nivå enn etterstratifisering, eller skjære gjennom et etterstratum. Dette betyr at man fritt kan velge  $w_i$  for  $i \in s_h$  uten å forstyrre hvilken som helst kalibreringsligning så lenge disse summerer opp til det samme  $W_h = \sum_{i \in s_h} w_i$  som tilfredstiller de kalibreringsligningene. Dette igjen dekomponerer  $D_s$  i  $\sum_h D_{s,h}$ , slik at for enhver  $W_h$  er  $D_{s,h}$  minimert ved å sette  $w_i = W_h/n_h$  for  $i \in s_h$ .

### A.3 Kalibrering: Betingede egenskaper

Med transformasjonen

$$y_i = x_i \beta + \epsilon_i, \quad (8)$$

kan det kalibrerte estimatet  $\hat{y}_{kal}$ , (7), ifølge (3) omskrives som

$$\begin{aligned} \hat{y}_{kal} &= \hat{y}_{est} + \sum_{i \in s} v_i (x_i \beta + \epsilon_i) & v_i &= w_i - q_h \text{ for } i \in s_h \\ &= \hat{y}_{est} + \sum_{i \in s} v_i \epsilon_i & \sum_{i \in s} w_i x_i &= \sum_h q_h n_h x_h = X. \end{aligned}$$

I tilfellet (a)  $\pi_{i|\mathbf{n}} = n_h/N_h$  og (b)  $w_i = w_h$  for  $i \in U_h$ , forenkles den betingede skjevheten til  $\hat{Y}_{kal}$

$$\sum_h E \left[ \sum_{i \in s_h} v_i \epsilon_i | \mathbf{n} \right] = \sum_h n_h v_h \left( \sum_{i=1}^{N_h} \epsilon_i / N_h \right) = \sum_h n_h v_h \bar{E}_h, \quad (9)$$

slik at den er forventningsrett uavhengig av initiale vektene  $\{a_i\}$  hvis,  $\forall 1 \leq h \leq H$ ,

$$\sum_{i=1}^{N_h} \epsilon_i = 0. \quad (10)$$

**Kommentar 11** Tilsynelatende definerer (8) en regresjonsmodell. Men i virkeligheten er den ikke noe mer enn en omparametrisering for  $(y_1, \dots, y_N)$ , siden den kan entydig skrives for enhver endelig populasjon. Ved å fastsette  $\beta$ , kan man transformere en populasjon av  $y_i$  til en populasjon av  $\epsilon_i$ . Samme som  $\hat{Y}_{greg}$  svarer  $\hat{Y}_{kal}$  til det vanlige minste-kvadraters kriterium ved bestemmelsen av  $\beta$  slik at  $\sum_{i=1}^N \epsilon_i^2$  minimert, dvs.  $\sum_{i=1}^N x_i \epsilon_i = 0$ , noe som er nødvendig konsekvens av (10) men ikke medfører den siste.

**Kommentar 12** Betingelsen (b) kan generaliseres til (b)'  $\{w_i; i \in s_h\}$  er konstant gitt  $\mathbf{n}$ , uten å ha særlig praktisk betydning.



Dersom (a) stratifisert etu. gitt  $\mathbf{n}$  og (b)  $w_i = w_h$  for  $i \in U_h$ , har man, for  $\sigma_h^2$  i (2) og  $f_h$  i (1),

$$\text{Var}(\hat{Y}_{kal}|\mathbf{n}) = \sum_h n_h(1 - f_h)w_h^2\sigma_h^2. \quad (11)$$

#### A.4 Variansbasert vurdering mht. utvalg

Variansevaluering er ofte sterkt avhengig av de antagelsene som muliggjør en slik evaluering. Betrakt det selvsveide estimatet for  $Y$ , dvs.

$$\hat{y}_{slv} = \sum_{i \in s} (N/n)y_i = \sum_{i \in s} (1/f)y_i. \quad (12)$$

I tilfellet stratifisert etu. gitt  $\mathbf{n}$ , har  $\hat{Y}_{slv}$  betinget varians, med  $\sigma_h^2$  og  $f_h$  som i (2),

$$\text{Var}(\hat{Y}_{slv}|\mathbf{n}) = \left(\frac{N}{n}\right)^2 \sum_{h=1}^H (1 - f_h)n_h\sigma_h^2, \quad (13)$$

som faller sammen med (2) hvis  $n_h/N_h = n/N$ . For et gitt utvalg kan man derfor måle hvor balansert utvalget er mht. etterstratifisering ved

$$\gamma_{slv,est} = \log \text{Var}(\hat{Y}_{slv}|\mathbf{n}) - \log \text{Var}(\hat{Y}_{est}|\mathbf{n}). \quad (14)$$

Desto nærmere  $\gamma_{slv,est}$  er lik 0, desto mer balansert er utvalget. På lignende måte har man

$$\gamma_{kal,est} = \log \text{Var}(\hat{Y}_{kal}|\mathbf{n}) - \log \text{Var}(\hat{Y}_{est}|\mathbf{n}). \quad (15)$$

På den andre siden kan man på en måte bruke

$$\rho_{est,slv} = \log\left(\sum_h q_h^2 n_h\right) - \log(N^2/n) \quad \text{og} \quad \rho_{kal,slv} = \log\left(\sum_{i \in s} w_i^2\right) - \log(N^2/n) \quad (16)$$

til å måle hvor robust etterstratifisering og kalibrering er.

**Kommentar 13** De to siste faktorene er definert på forholdstallene mellom tilsvarende varianser under modellen av homogen populasjon, og er dermed ofte relevante selv med andre utvalgsplaner.

#### A.5 Kalibrering: Betingede egenskaper med tomme etterstrata

La  $R_0 \cup R_0^c = \{1, \dots, H\}$ , slik at  $R_0 \cap R_0^c = \emptyset$  og  $n_h = 0$  for  $h \in R_0$  og  $n_h > 0$  for  $h \in R_0^c$ . La igjen  $w_i = q_h + v_i$  for  $i \in s_h$  og omskriv det kalibrerte estimatet for  $Y$  som  $\hat{y}_{kal} = \sum_{h \in R_0^c} q_h(\sum_{i \in s_h} y_i) + \sum_{i \in s} v_i y_i$ . La  $X_0 = \sum_{h \in R_0} \sum_{i=1}^{N_h} x_i$ ,  $E_0 = \sum_{h \in R_0} \sum_{i=1}^{N_h} \epsilon_i$  og  $E_h^c = \sum_{i=1}^{N_h} \epsilon_i$  for  $h \in R_0^c$ . Legg merke til at  $\sum_{i \in s} w_i x_i = X$  og  $\sum_{h \in R_0^c} q_h n_h x_h = X - X_0$ . I tilfellet (a)  $\pi_{i|\mathbf{n}} = n_h/N_h$  og (b)  $w_i = w_h$  for  $i \in U_h$ ,

$$\begin{aligned} E[\hat{Y}_{kal} - Y|\mathbf{n}] &= \sum_{h \in R_0^c} \left( \sum_{i \in s_h} v_i \right) \left( \sum_{i=1}^{N_h} y_i / N_h \right) - \sum_{h \in R_0} \sum_{i=1}^{N_h} y_i = \sum_{h \in R_0^c} V_h^c \bar{Y}_h - Y_0 \\ &= \left( \sum_{h \in R_0^c} V_h^c x_h \right) \beta + \sum_{h \in R_0^c} V_h^c \bar{E}_h^c - Y_0 \\ &= (X_0 \beta + \sum_{h \in R_0^c} V_h \bar{E}_h) - (X_0 \beta + E_0), \end{aligned}$$

Mao. er  $\hat{Y}_{kal}$  forventningsrett uavhengig av initiale vektene  $\{a_i\}$  hvis

$$\sum_{h \in R_0} \sum_{i=1}^{N_h} \epsilon_i = 0 \quad \text{og} \quad \sum_{i=1}^{N_h} \epsilon_i = 0 \text{ for } h \in R_0^c. \quad (17)$$

**Kommentar 14** Det at (17) følger fra (10) betyr at kalibrering faktisk kan være 'immun' mot tomme etterstrata for slike populasjoner.

Dersom (a) stratifisert etu. gitt  $\mathbf{n}$  og (b)  $w_i = w_h$  for  $i \in U_h$ , er den betingede variansen til  $\hat{Y}_{kal}$  gitt som

$$Var(\hat{Y}_{kal} | \mathbf{n}) = \sum_{h \in R_0^c} n_h (1 - f_h) w_h^2 \sigma_h^2. \quad (18)$$

**Kommentar 15** I tilfellet  $\exists h \in R_0^c$  slik at  $n_h = 1$ , kan man ikke lage  $\hat{\sigma}_h^2$  i disse etterstrata. En måte å gå fram på er f.eks. å slå sammen alle, si,  $n_e$  entallstrata og bruke

$$\hat{\sigma}_e^2 = \sum_{h \in R_0^c: n_h=1} [y_i - (\sum_{h \in R_0^c: n_h=1} y_i / n_e)]^2 / (n_e - 1)$$

som den felles  $\hat{\sigma}_h^2$  for disse entallstrata. Vi skal kalle denne estimatoren for fattigmannsestimatoren.

## B Resultatene basert på utvalget 4. kvartal 1996

Med dagens opplegg estimeres tallene innen hvert fylke ved å bruke de etterstratifiserte vektene på landsnivå, altså  $\hat{Y}_{est}$ . La  $\hat{Y}_{kal}$  være den fylkesvis kalibrerte estimatoren med de nye landsetterstratifiserte startverdier. Her finnes det i alt  $2 \times 12 \times 4 = 96$  etterstrata og  $2 + 12 + 4 = 18$  marginaler. I tillegg la  $\hat{Y}_{ka2}$  være den fylkesvis kalibrerte estimatoren med dagens landsetterstratifiserte vekter som startverdier (ca. 90 etterstrata), og  $\hat{Y}_{ka0}$  den fylkesvis kalibrerte estimatoren med konstante initiale vekter innen hvert fylke.

### B.1 Skjevhet

For å studere skjevheten i kalibrering kan man sammenligne  $\hat{y}_{kal}$  og  $\hat{y}_{ka0}$  — i tilfellet (17) er kalibrering forventningsrett uavhengig av initiale vekter, slik at store avviker mellom  $\hat{y}_{kal}$  og  $\hat{y}_{ka0}$  kan tyde på at (17) ikke holder for AKU-populasjonen.

Innen hvert fylke beregnes estimatene for antall sysselsatte innen følgende næringer (a) jordbruk, skogbruk, fiske og fangst, (b) oljeutvinning, bergverksdrift, industri, kraft- og vannforsyning, (c) bygge- og anleggsvirksomhet, (d) varehandel, hotel- og restaurant virksomhet, (e) transport, lagring, post og telekommunikasjon, (f) bank- og finansiering, forsikring, eiendomsdrift og forretningsmessig tjenesteyting, og (g) offentlig, sosial og privat tjenesteyting.

Histogram relat.bias.11 til relat.bias.31 viser, i rekkefølgen (a) - (g), tilsvarende  $(\hat{y}_{ka0} - \hat{y}_{kal}) / \hat{y}_{kal}$  på tverss over alle 19 fylkene. I tillegg viser relat.bias.32 det tilsvarende for sysselsettingstallene og relat.bias.33 for ledighetstallene over alle fylkene. Disse histogrammene tyder på at kalibrering her gir tilnærmede forventningsrette estimatorene. Samtidig viser relat.bias.41 til relat.bias.43 tilsvarende  $(\hat{y}_{ka2} - \hat{y}_{kal}) / \hat{y}_{kal}$  for antall (i) sysselsatte, (ii) arbeidssøkere og (iii) utenfor arbeidsstyrken, og relat.bias.51 til relat.bias.53 for  $(\hat{y}_{est} - \hat{y}_{kal}) / \hat{y}_{kal}$  på samme måten. Spesielt kan det synes at de relative skjevhetene i  $\hat{Y}_{est}$  er flere ganger større enn  $\hat{Y}_{kal}$ .

## B.2 Varians

Ettersom man ikke har variansform for  $\hat{Y}_{est}$  og  $\hat{Y}_{ka2}$  gitt dagens utvalgsplan, har man brukt den såkalte jack-knife metoden (Wolter, 1985) til å estimere variansen. Først deles månedsutvalget fra hvert fylke i  $k$  tilfeldige grupper av mest mulig lik størrelse. På denne måten får man i alt  $k$  delutvalg for hvert kvartal. For hvert delutvalg beregnes nå estimatet basert på de resten  $k - 1$  delutvalgene på akkurat samme måten som på det fullstendige utvalget. Dette gir oss  $\hat{\theta}_{(1)}, \dots, \hat{\theta}_{(k)}$  samt  $\hat{\theta}$  basert på hele utvalget. Da er jack-knife variansestimaten til  $\hat{\theta}$  gitt som

$$\hat{Var}_{jck}(\hat{\theta}) = \frac{k-1}{k} \sum_{j=1}^k (\hat{\theta}_{(j)} - \hat{\theta})^2.$$

Siden tomme etterstrata i utvalget igjen vil være tomme i alle  $k$  delutvalgene, estimerer  $\hat{Var}_{jck}(\hat{\theta})$  variansen til  $\hat{\theta}$  betinget på tomme etterstrata.

Samtidig beregnes fattigmannsvariansen  $\hat{Var}_{ftm}(\cdot)$  for  $\hat{Y}_{ka2}$  ved å erstatte  $n_h w_h^2$  i (18) med  $\sum_{i \in s_h} w_i^2$ , i tillegg til sammenslåing av entallstrata. Til slutt tar man med variansen (2) til den *ad hoc* etterstratifiserte estimatoren ved å slå sammen alle entall- og tomme etterstrata.

Histogram log.rat.11 viser estimerte (14) for sysselsetting tverrsover de 19 fylkene. Log.rat.12 og log.rat.13 viser estimerte (15) og (16) for  $\hat{Y}_{ka2}$  der  $\hat{Var}_{ftm}(\hat{Y}_{ka2})$  er brukt for  $Var(\hat{Y}_{ka2}|\mathbf{n})$ , samtidig som log.rat.21 og log.rat.22 viser det samme for  $\hat{Y}_{kal}$ . Ut fra dem kan utvalget synes å være rimelig balansert mht. etterstratifisering.

Log.rat.23 viser forholdstall mellom  $\hat{Var}_{jck}(\hat{Y}_{ka2})$  (med  $k = 64$  grupper) og  $\hat{Var}_{ftm}(\hat{Y}_{ka2})$  på log-skala for sysselsetting innen fylkene, og log.rat.31 for arbeidssøkere. Mens log.rat.32 og log.rat.33 viser de tilsvarende for  $\hat{Var}_{ftm}(\hat{Y}_{ka2})/\hat{Var}_{ftm}(\hat{Y}_{kal})$ . Mens de to siste ligger på omtrent samme nivået, virker jack-knife variansestimaten noe mer varierende.

Var.rat.41 til var.rat.43 viser direkte forholdstallene mellom  $\hat{Var}_{jck}(\hat{Y}_{ka2})$  og  $\hat{Var}_{jck}(\hat{Y}_{est})$  for antall sysselsatte, arbeidssøkere og utenfor arbeidsstyrken: Mens variansen til ledighetstallet er omtrent det samme før og nå, har kalibrering stort sett halvert variansen til sysselsettingstallet. Forholdstallene mellom  $\hat{Var}_{ftm}(\hat{Y}_{kal})$  og  $\hat{Var}(\hat{Y}_{slv})$ , dvs. var.rat.51 til var.rat.53, viser lignende resultater, som kan kalles estimerings-effekt (EEFF)<sup>2</sup> i motsetning til designeffekt (DEFF).

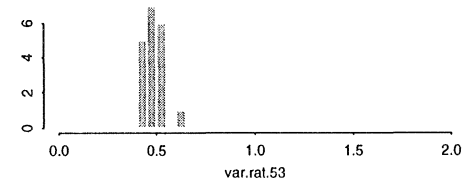
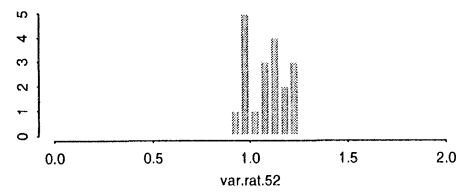
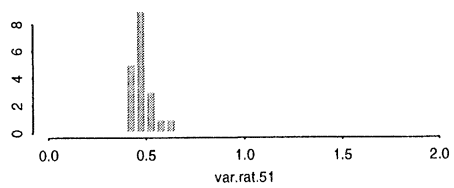
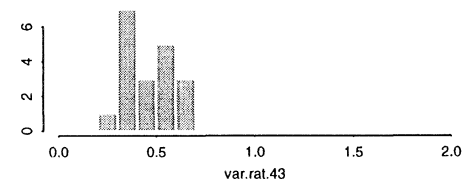
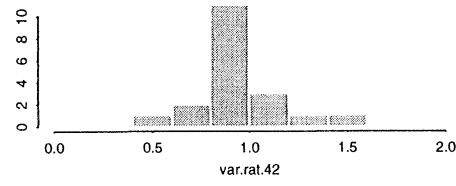
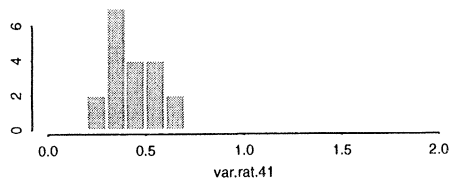
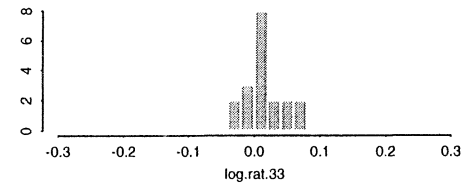
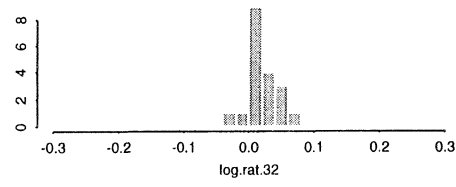
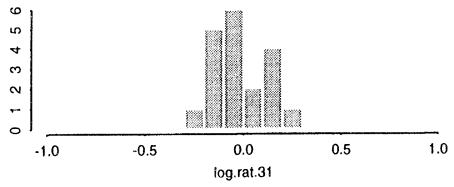
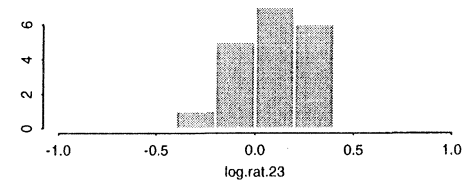
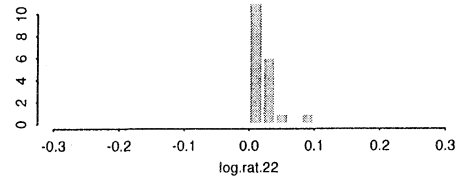
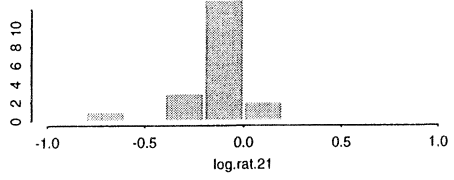
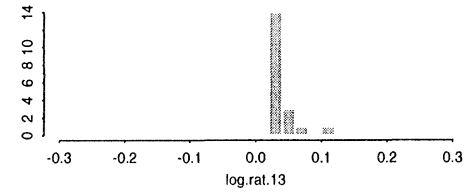
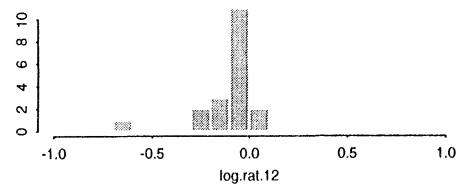
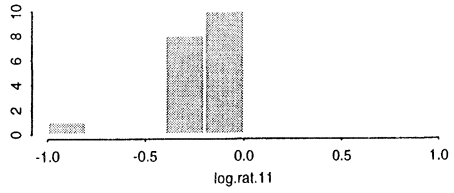
Til slutt viser histogram relat.err.11 til relat.err.13 tilsvarende  $\sqrt{\hat{Var}_{ftm}(\hat{Y}_{kal})}/\hat{y}_{kal}$  for antall sysselsatte, arbeidssøkere og utenfor arbeidsstyrken. I tillegg viser relat.err.21 til relat.err.41 det tilsvarende for næringstallene i samme rekkefølgen som tidligere nevnte. Mens verdiene til disse  $\hat{Y}_{kal}$  både i total og i prosent er gitt i Tabell 1 og 2.

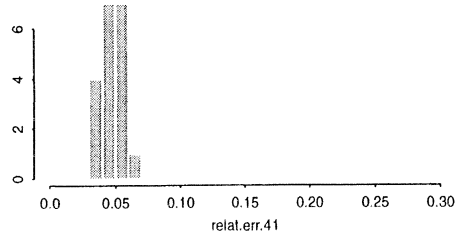
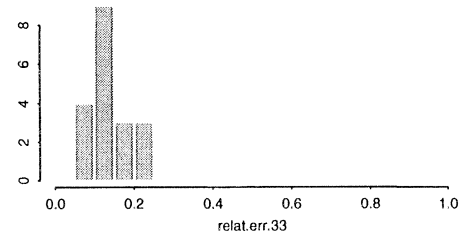
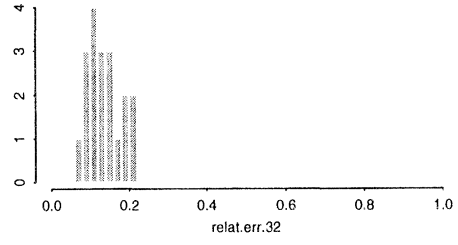
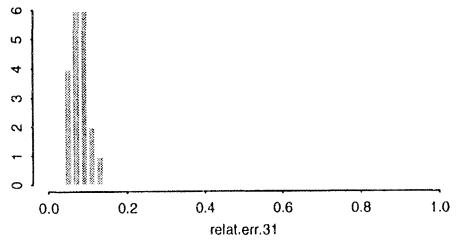
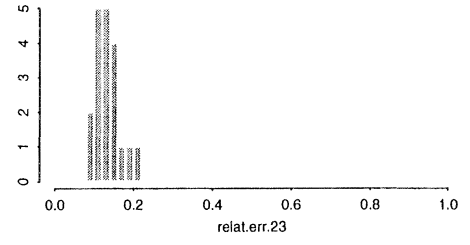
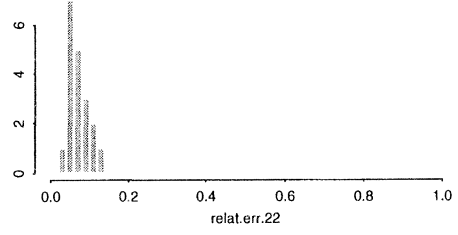
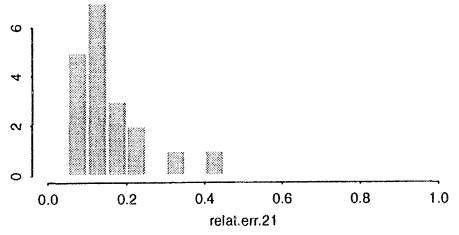
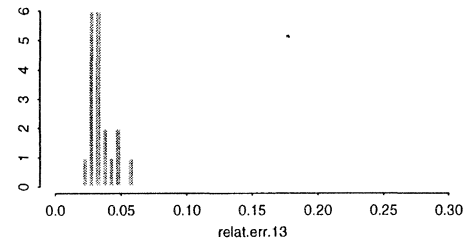
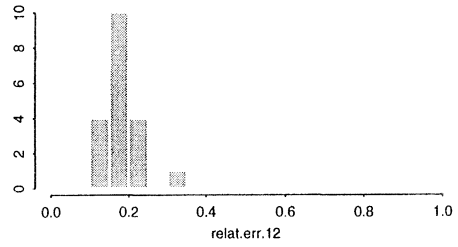
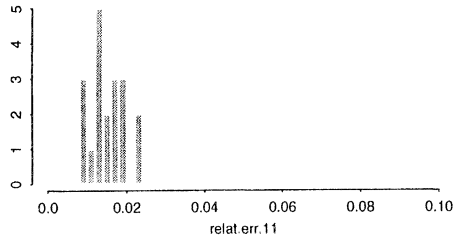
## Referanser

- Boije, L. (1996). Estimering og Etterstratifisering i AKU (boi, 25.03.96). Upublisert notat.
- Deville, J.-C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *J. Am. Statist. Assoc.*
- Wolter, K.M. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.
- Zhang, L.-C. (1997a). Progresjonsrapport for AKU-estimering (lcz, juni 1997). Upublisert notat.
- Zhang, L.-C. (1997b). Notes on post-stratification and calibration (lcz, juni 1997). Upublisert notat.

<sup>2</sup>Personlig kommunikasjon med Ib Thomsen.







Tabell 1. Personer i alderen 16-74 år etter arbeidsstyrkestatus innen de 19 fylkene. Sysselsatte gruppert etter følgende næringer: (a) jordbruk, skogbruk, fiske og fangst, (b) oljeutvinning, bergverksdrift, industri, kraft- og vannforsyning, (c) bygge- og anleggsvirksomhet, (d) varehandel, hotel- og restaurant virksomhet, (e) transport, lagring, post og telekommunikasjon, (f) bank- og finansiering, forsikring, eiendomsdrift og forretningsmessig tjenesteyting, og (g) offentlig, sosial og privat tjenesteyting. 4. kvartal 1996. (I 100)

Fylke	Næring							Total		
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Syssel	Ledig	Resten
Østfold	37	270	68	196	71	80	414	1137	51	566
Akershus	35	266	110	497	209	377	897	2391	77	773
Oslo	9	279	102	499	212	470	1027	2597	131	924
Hedmark	75	138	56	152	51	52	343	867	34	439
Oppland	90	141	72	145	52	53	327	880	31	408
Buskerud	38	245	73	211	95	116	380	1158	42	466
Vestfold	26	199	64	195	80	86	334	984	45	450
Telemark	29	174	56	138	39	59	255	749	43	373
Aust-Agde	13	93	41	70	32	37	178	464	29	218
Vest-Agde	27	149	42	146	57	48	250	719	27	316
Rogaland	98	440	83	305	127	177	542	1772	63	672
Hordaland	85	370	139	367	175	196	738	2070	104	828
Sogn og F	54	102	42	84	42	20	190	533	12	199
Møre og R	88	298	64	185	100	70	351	1155	43	496
Sør-Trønd	86	183	75	239	95	102	480	1260	38	552
Nord-Trønd	88	84	43	113	33	31	220	612	23	258
Nordland	94	132	99	188	111	60	444	1127	60	512
Troms	38	67	35	161	51	61	303	716	42	332
Finnmark	37	44	19	59	27	18	164	370	34	147
I alt i sammenligning med AKU-4/96. (1000)										
Ny	100	367	128	395	166	211	783	2154	93	893
AKU-4/96	100	367	130	401	163	208	781	2154	94	894

Tabell 2. Personer i alderen 16-74 år etter arbeidsstyrkestatus innen de 19 fylkene. Sysselsatte gruppert etter følgende næringer: (a) jordbruk, skogbruk, fiske og fangst, (b) oljeutvinning, bergverksdrift, industri, kraft- og vannforsyning, (c) bygge- og anleggsvirksomhet, (d) varehandel, hotel- og restaurant virksomhet, (e) transport, lagring, post og telekommunikasjon, (f) bank- og finansiering, forsikring, eiendomsdrift og forretningsmessig tjenesteyting, og (g) offentlig, sosial og privat tjenesteyting. 4. kvartal 1996. (I prosent)

Fylke	Næring							Total		
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Syssel	Ledig	Resten
Østfold	3,3	23,8	6,0	17,2	6,2	7,1	36,4	64,8	4,3	32,3
Akershus	1,5	11,1	4,6	20,8	8,7	15,8	37,5	73,8	3,1	23,9
Oslo	0,4	10,7	3,9	19,2	8,2	18,1	39,5	71,1	4,8	25,3
Hedmark	8,6	15,9	6,5	17,5	5,9	6,1	39,5	64,7	3,8	32,8
Oppland	10,2	16,0	8,1	16,5	5,9	6,1	37,1	66,7	3,4	30,9
Buskerud	3,3	21,1	6,3	18,2	8,2	10,0	32,8	69,5	3,5	28,0
Vestfold	2,7	20,2	6,5	19,8	8,2	8,7	33,9	66,5	4,4	30,4
Telemark	3,8	23,2	7,4	18,4	5,2	7,9	34,0	64,3	5,4	32,0
Aust-Agde	2,8	20,0	8,8	15,2	7,0	7,9	38,4	65,3	5,8	30,7
Vest-Agde	3,8	20,7	5,9	20,3	7,9	6,7	34,8	67,7	3,6	29,7
Rogaland	5,5	24,8	4,7	17,2	7,2	10,0	30,6	70,7	3,4	26,8
Hordaland	4,1	17,9	6,7	17,7	8,5	9,5	35,6	69,0	4,8	27,6
Sogn og F	10,2	19,1	7,8	15,8	7,8	3,7	35,6	71,7	2,1	26,7
Møre og R	7,6	25,8	5,5	16,0	8,6	6,1	30,3	68,2	3,6	29,3
Sør-Trønd	6,8	14,5	6,0	19,0	7,5	8,1	38,1	68,1	2,9	29,8
Nord-Trønd	14,3	13,8	7,0	18,5	5,4	5,1	36,0	68,5	3,7	28,9
Nordland	8,4	11,7	8,7	16,7	9,8	5,3	39,4	66,3	5,1	30,2
Troms	5,4	9,4	4,9	22,4	7,1	8,6	42,3	65,7	5,6	30,4
Finnmark	10,1	12,0	5,2	16,0	7,4	4,9	44,5	67,3	8,3	26,6



Tabell 3. Sysselsatte i alt etter alder og kjønn i sammenligning med AKU-4/96. 4. kvartal 1996. (I 1000)

Alder	I alt		Menn		Kvinner	
	Ny	AKU-4/96	Ny	AKU-4/96	Ny	AKU-4/96
I alt	2154	2154	1164	1164	990	990
16-19	78	78	41	41	37	37
20-24	201	199	106	106	94	93
25-29	275	275	149	147	127	128
30-34	288	291	157	159	131	132
35-39	272	273	146	142	125	131
40-44	273	270	144	146	129	125
45-49	261	262	140	142	121	120
50-54	231	228	125	127	105	101
55-59	145	145	79	81	66	64
60-64	93	94	52	51	41	43
65-66	21	21	11	11	10	10
67-69	10	10	7	7	3	3
70-74	7	8	5	6	2	2

## De sist utgitte publikasjonene i serien Notater

- 97/39 E. Holmøy og Ø. Thøgersen: Virkninger av strukturpolitiske reformer: Forslag til konkrete forskningsprosjekter. 67s.
- 97/40 E. Gulløy: Journalister og personvern - om personvernproblematikk og Datatilsynet. 37s.
- 97/41 E. Holmøy: En presisering av hva som skal menes med tilbudskurven for arbeid i en generell likevektsmodell. 10s.
- 97/42 A. Sundvoll og H.M. Teigum: IT i skolen 1997. Del 1: Tilstandsundersøkelse i skolene. Hovedresultater og dokumentasjon. 65s.
- 97/43 P. Schønning og K. Jonassen: Sammenligning av foreslåtte nasjonale tilpasninger av Corine Land Cover med andre arealklassifikasjons-systemer og systemer for arealstatistikk. 39s.
- 97/44 A. Holmøy: Inntekts- og formuesundersøkelsen for personlig næringsdrivende 1995: Dokumentasjon. 45s.
- 97/45 A. Katz, B.M. Larsen, K.S. Eriksen og T. Jensen: Transport og makroøkonomi - en samkjøring av GODMOD-3 og MSG-6. 62s.
- 97/46 S. Todsens: Nasjonalregnskap: Beregning av oljenæringene. 23s.
- 97/47 O.F. Vaage: Undersøkelse om voksenopplæring i Rogaland: Dokumentasjonsrapport. 33s.
- 97/48 E. Gulløy: Undersøkelse om personvern: Holdninger og erfaringer 1997. 76s.
- 97/49 H.M. Edvardsen, J. Mønnesland og K.Ø. Sørensen: Regional arbeidsdeling: Sogn og Fjordanes plass i norsk verdiskaping. 35s.
- 97/50 O. Rognstad: SSBs forslag til landbrukstelling 1999. 65s.
- 97/51 J.E. Sivertsen: Flyktninger og arbeidsmarkedet 4. kvartal 1996. 38s.
- 97/53 S. Hansen og T. Skoglund: Sammenligning av data for sysselsetting og lønn fra ulike kilder. 30s.
- 97/54 S. Blom: Holdning til innvandrere og innvandringspolitikk: Spørsmål i SSBs omnibus i mai/juni 1997. 39s.
- 97/55 K. Mork: SSB-AVLØP: Fylkeshefte 1996. 203s.
- 97/56 Opplysninger om inntekt, formue og skatt i forløpsdatabasen Trygd-fobhistorie: Tilrådinger fra et utvalg. 52s.
- 97/57 E.J. Fløttum: Ordliste og definisjoner i økonomisk statistikk: Engelsk - bokmål - nynorsk. 166s.
- 97/58 T. Dale: Samordnet levekårsundersøkelse 1997 - panelundersøkelsen: Dokumentasjonsrapport. 87s.
- 97/59 H. Høie og A. Grønlund: Driftstypemodellen: Modell for tilrettelegging av jordbruksstatistikk for beregning av tap av næringsstoffer fra jordbruksarealene: Dokumentasjon. 37s.
- 97/60 A. Sundvoll: Undersøkelse om mødre med nyfødte barn. 36s.
- 97/61 S. Todsens: Nasjonalregnskap: Beregning av realkapitalbeholdninger og kapitalslit. 34s.
- 97/62 K. Mork: Utslepp og rensing av avløpsvatn: Datakvalitet og beregningsmåter. 64s.
- 97/63 S. Stamnes og B.L. Western: Inntekts- og kostnadsundersøkelse for privatpraktiserende psykologer 1996: Dokumentasjon. 26s.
- 97/64 H.M. Teigum: Barns helse og velferd 1996: Dokumentasjon og frafallsanalyse. 39s.
- 97/65 F. Gjertsen: Dødsårsaksregistret i Statistisk sentralbyrå: Rapport om virksomheten i 1996. 56s.
- 97/66 B. Olsen: Prøveundersøkelse om 1-3 dagers sykefravær i sentral sykefraværstatistikk: Dokumentasjon. 15s.
- 97/68 R. Johansen: REGARD - Modell for regional analyse av arbeidsmarked og demografi: Teknisk dokumentasjon. 212s.

## Notater



Tillatelse nr.  
159 000/502

**B** *Returadresse:*  
Statistisk sentralbyrå  
Postboks 8131 Dep.  
N-0033 Oslo

### Statistisk sentralbyrå

*Oslo:*  
Postboks 8131 Dep.  
0033 Oslo

Telefon: 22 86 45 00  
Telefaks: 22 86 49 73

*Kongsvinger:*  
Postboks 1260  
2201 Kongsvinger

Telefon: 62 88 50 00  
Telefaks: 62 88 50 30

ISSN 0806-3745



**Statistisk sentralbyrå**  
Statistics Norway