

Johan Heldal og Audun Rusti

**Om samordning av utvalg ved
bruk av PRN-tall**

Notater

Forord

Dette notatet beskriver hvordan permanente tilfeldige tall kan brukes til å samordne og rullere utvalg, hvordan slike er implementert i tilknytning til Bedrifts- og Foretaksregisteret og hvordan de kan aksesseres av brukere som ønsker å benytte dem. Johan Heldal (seksjon 120) har tatt initiativet og skrevet avsnittene 1-9. Audun Rusti har stått for den tekniske implementeringen i tilknytning til BoF og skrevet avsnitt 10. PRN-tallene vil bli tatt i systematisk bruk i prosjektet ”Oppgavebyrden og samordning av utvalg for næringslivet” (prosj. nr 3015).

Innhold

Forord.....	1
Innhold	2
1. Samordning av utvalg	3
2. Hva er et PRN?	3
3. Hvordan trekke utvalg med PRN?	3
3.1. Bernoulli utvalg.....	4
3.2. Enkelt tilfeldig utvalg.....	4
3.3. Poisson utvalg	5
3.4. Sekvensielt Poissonutvalg.....	5
4. Samordning av utvalg	5
5. Rotasjon	6
6. Fremtiden	6
7. Eksempler på anvendelser.....	7
7.1. Lastebilundersøkelsen. Eksempel på rotasjon.....	7
7.2. Sannsynlighetsdelen til UT-årsutvalg, et eksempel på samordning.....	8
7.3. Prisindeks for transport på vei. Et eksempel på SP trekking og rotasjon.....	9
7.3.1. kort beskrivelse av utvalgsplan	9
7.3.2. Rotasjon av utvalget.....	10
8. Referanser	11
9. Vedlegg 1: Rotasjonsbeskrivelse	12
10. Vedlegg 2: Teknisk beskrivelse av rutine for tildeling av PRN til bedrifter og foretak i BOF....	19
10.1. Generelt.....	19
10.1.1. Databaseobjekter	19
10.1.2. UNIX-miljø	20
10.1.3. prnsrc ser slik ut :	20
10.1.4. SAS-programmet prn_prod.....	22
10.2. Uttak av PRN_BEDRIFTER til SAS-ds	26
10.3. Uttak av VW_PRN_FORETAK til SAS-ds	27
De sist utgitte publikasjonene i serien Notater.....	28

1. Samordning av utvalg

er et viktig tema i en rekke statistiske sentralbyråer rundt i verden, særlig for bedrifts- og foretaksundersøkelser. Spesielt Sverige og Australia har bygget opp omfattende system som skal ivareta slik samordning både på enkelttidspunkt og over tid, men også en rekke andre land har i større og mindre utstrekning benyttet de teknikkene som vil bli nevnt her.

Kjernen i disse systemene er et såkalt "Permanent Random Number" (PRN), faste tilfeldige tall, som genereres tilfeldig og knyttes fast til de statistiske enhetene i registrene. Disse kan så brukes ved trekking av utvalg. Nye enheter får tilknyttet et PRN så snart den blir registrert med en ident, og PRN-et blir hengende ved enheten så lenge den lever. Disse PRN-tallene fungerer deretter som en ressurs som gjør det mulig å koordinere utvalg fra ulike undersøkelser enten negativt for å minimere overlapp, positivt for å fremme overlapp eller som en elegant metode for rotere utvalg og håndtere utvalgene over tid i forhold til fødsler og død i populasjonene. Man kan vedlikeholde utvalg slik at de beholder sine statistiske egenskaper uten å måtte trekke en mengde små oppdateringsutvalg med nye enheter. Det svenske systemet kalles SAMU (SAMordning av Urval) og omfatter alle bedrifts- og foretaksundersøkelser (Ohlsson 1992) og har nå i 2005 rundt 30 års fartstid (Thulin 1976). Den australske varianten kalles Synchronised Sampling og stammer fra begynnelsen av 1980 tallet (McKenzie & Gross 2000). Vi i Norge ligger med andre ord langt etter.

2. Hva er et PRN?

Et PRN er et tall x som er generert av en tilfeldig tall generator uniformt med verdi i intervallet $(0,1)$. Hver statistisk enhet i et register får sitt PRN. Forskjellige enheter må ha forskjellige PRN og de må være statistisk uavhengige. Man kan trekke enkelt tilfeldige utvalg på n enheter med disse tallene ved å sortere en populasjonsfil stigende eller synkende med hensyn på PRN-tallene, velge et startpunkt a i $(0,1]$, og ta de n neste på listen som utvalget. I denne forbindelse må intervallet $(0,1]$ tenkes som en sirkel. Kommer man forbi 1 fortsetter man på begynnelsen. Når man så skal oppdatere utvalget neste gang undersøkelsen skal gå, kan man gjøre likedan. Man vil få et oppdatert utvalg som også er enkelt tilfeldig, og siden PRN tallene er de samme for alle vedblivende enheter i populasjonen, vil man ved samme valg av a få et utvalg som maksimalt overlapper det gamle. Ved å skyve litt på a kan man rotere utvalget så mye som en måtte ønske. For ytterligere beskrivelser av muligheter og detaljer henviser viden interesserte til referansene.

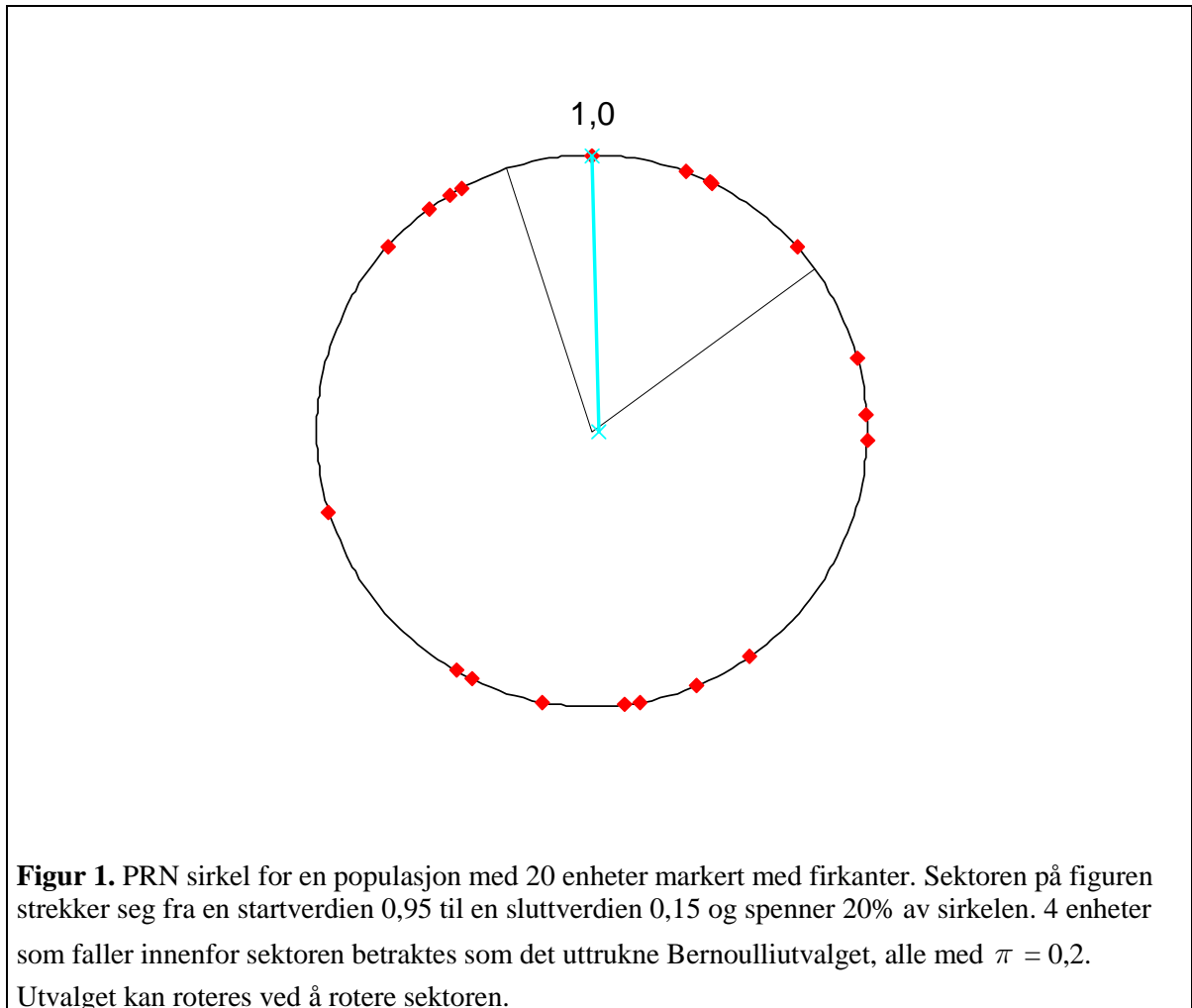
Med dette utgangspunkt har vi etablert PRN for bedrifter og foretak i Norge. De er organisert som en egen database ved siden av BoF hvor alle med tilgang til BoF kan koble seg opp og hente PRN tall for de enheter de måtte ønske. PRN-tallene forvaltes og vedlikeholdes av seksjon 410 og oppdateres hver uke med en kjøring som starter automatisk på samme ukedag og klokkeslett. Basen krever små ressurser å vedlikeholde.

3. Hvordan trekke utvalg med PRN?

Man kan trekke utvalg stratifisert med like eller forskjellige sannsynligheter innen hvert stratum ved hjelp av PRN-tall. Trekker man med like sannsynligheter blir det enten et såkalt Bernoulli utvalg eller et enkelt tilfeldig utvalg. Trekker man med ulike sannsynligheter blir det enten et såkalt Poisson utvalg eller noe som noe misvisende er blitt døpt "seksensiell Poisson" (SP). Bernoulli og Poisson utvalg har den ulempen at utvalgsstørrelsen blir stokastisk og når "forventet" utvalgsstørrelse er liten kan utvalgsstørrelsen bli svært variabel. Det vil kunne skje at ingen enheter blir trukket. Slike utvalg brukes derfor sjelden og bare når de ønskede utvalgsstørrelsene er så store at variasjonen spiller mindre rolle. Enkelt tilfeldig og SP utvalg garanterer at utvalgsstørrelsene blir "faste". Det er derfor disse som primært brukes. SP er benyttet til et utvalg for innsamling av data til den nye indeksen for transport på vei. Denne metoden er siden 1989 også benyttet til utvalget for den svenske konsumprisindeksen.

3.1. Bernoulli utvalg

Vi har en populasjon på N enheter. Vi ønsker å trekke et utvalg der hver enhet har en lik og på forhånd gitt sannsynlighet π for å bli trukket. Prinsippet for Bernoullitrekking er at vi løper nedover en liste med enheter, og for hver enhet gjøres et eksperiment som med sannsynlighet π resulterer i at enheten blir trukket og med sannsynlighet $1 - \pi$ resulterer i at den ikke blir trukket. Trekkene skjer uavhengig av hverandre for de ulike enhetene. Vi kan trekke Bernoulliutvalg ved hjelp av PRN-tall og går frem som følger: Sorter populasjonsfilen økende etter PRN-tallene. Velg en startverdi a , $0 < a < 1$. Ta som utvalg alle enheter med $a \leq x < a + \pi$. Hvis $a + \pi > 1$, betrakt intervallet $[0, 1)$ som en sirkel og gå til $a + \pi - 1$.



Hver enhet i populasjonen vil i utgangspunktet ha en garantert sannsynlighet π for å bli trukket, men ettersom det vil være tilfeldig hvor mange enheter det er med PRN i det aktuelle intervallet i vår populasjon, blir utvalgsstørrelsen n en binomisk stokastisk variabel. $En = N\pi$, $Var(n) = N\pi(1 - \pi)$.

3.2. Enkelt tilfeldig utvalg

Vi ønsker å ha en fast utvalgsstørrelse n . Heller enn å ta alle enhetene i et intervall langs sirkelen, ta de n første etter startverdien a ($x > a$). Hvis $a = 0,95$ som i eksempelet med Bernoullitrekking ovenfor og vi ønsker å trekke 5 enheter, ta de fem første etter 0,95 rundt sirkelen. Dette blir et enkelt tilfeldig utvalg (se Ohlsson 1992. Eksempler på trekking av stratifiserte enkelt tilfeldige utvalg med PRN-tall er gitt i eksemplene i avsnitt 8.1 og 8.2.

Man kan etablere et panel ved å ha startverdien a konstant over tid. La x_1, \dots, x_n være PRN-tallene til de n uttrukne enhetene. Da vil $a < x_1 < x_2 < \dots < x_n$ regnet rundt sirkelen. Mellom to tidspunkt t_1 og t_2 vil det i en virkelig populasjon være både fødsler og død av enheter med PRN-tall mellom de opprinnelige x_1 og x_n . Hvis det er flere tilganger enn avganger og utvalgsstørrelsen n holdes konstant, vil PRN-verdien til den siste som trekkes på tidspunkt t_2 være mindre enn den tilsvarende verdien på t_1 , altså $x_n^{(2)} < x_n^{(1)}$ og de siste enhetene i det opprinnelige utvalget vil falle ut. Neste gang kan det bli omvendt og de(n) enhet(ene) som falt ut andre gang kommer inn igjen. I det lange løp kan et slikt utvalg derfor bli nokså flossete i enden. Like fullt blir utvalget hele tiden et ekte enkelt tilfeldig utvalg og utvalget beholder under denne forutsetning et maksimalt antall felles enheter over tid. Denne flossetheten kan imidlertid være uheldig av hensyn til respondenter som vil gå inn i og ut av utvalget flere ganger. En måte å avhjelpe problemet er å rotere utvalget noe over tid. Den samme problemstillingen er aktuell også for de andre typene PRN-baserte utvalg. Se avsnitt 6.

3.3. Poisson utvalg

Et Poissonutvalg er som et Bernoulliutvalg, men ulike enheter kan ha kan ha forskjellige sannsynligheter π_i for å bli trukket. Vi modifiserer derfor metoden ved å definere en størrelse

$$\xi_i = \frac{x_i}{\pi_i} \quad (3.1)$$

Nå vil sannsynligheten $\text{Prob}(PRN_i \leq \pi_i) = \pi_i$. Hvis vi velger alle enheter som har $0 < \xi_i \leq 1$ til utvalget vil alle enheter få den ønskede sannsynligheten π_i for å bli trukket. Også her kan man velge et vilkårlig startpunkt a , $0 < a < 1$ og sette

$$\xi_i^* = \begin{cases} \frac{x_i - a}{\pi_i} & \text{hvis } x_i - a > 0 \\ \frac{x_i - a + 1}{\pi_i} & \text{hvis } x_i - a \leq 0 \end{cases} \quad (3.2)$$

som betyr at hvis $x_i - a \leq 0$ erstattes $x_i - a$ med $x_i - a + 1$. Men igjen vil utvalgsstørrelsen bli stokastisk. Metoden brukes derfor lite selv om den er enkel.

3.4. Sekvensielt Poissonutvalg

I stedet for å velge alle enheter med $0 < \xi_i^* \leq 1$, velg enhetene med de n minste ξ_i^* . Vi får da en fast (forutsigelig) utvalgsstørrelse. Ulempen er nå at de nominelle inklusjonssannsynlighetene vi har valgt, π_i , ikke blir de virkelige sannsynlighetene for å bli trukket og som vi kan kalle λ_i . SP er, i motsetning til Poisson trekking, ikke en "streng PPS" prosedyre. Sammenhengen mellom π_i og λ_i er så kompleks at den ikke lar seg uttrykke som en lukket formel. Det kan imidlertid vises (Rosén 1997a,b) at $\lambda_i \rightarrow \pi_i$ asymptotisk, dvs. når n og $N \rightarrow \infty$ under "generelle" betingelser. Det er også vist (samme sted) at Horwitz-Thompson (H-T) estimatoren basert på de nominelle inklusjonssannsynlighetene π_i i stedet for λ_i er "konsistent" hvis den virkelige H-T estimatoren er det. Simuleringer viser at den "nominelle" Horwitz-Thompson estimatoren fungerer bra selv i små utvalg. Dersom en tar sikte på å bruke "modellbaserte" estimeringsmetoder og bare ønsker å trekke med ulike sannsynligheter for å sikre en god overrepresentasjon av "store" enheter i utvalget er SP en egnet trekkemetode og det spiller liten rolle at $\pi_i \neq \lambda_i$. SP er også basis for en generalisering kalt "order sampling", men det vil jeg ikke komme inn på her. Se avsnitt 8.3 for et eksempel på SP utvalgstrekkning.

4. Samordning av utvalg

Det hender ofte at vi ønsker at utvalg til to eller flere undersøkelser på samme populasjon skal samordnes, enten positivt ved at utvalgene i størst mulig grad skal benytte de samme utvalgsenheterne eller negativt ved at utvalgene skal overlappe minst mulig. Det er en viktig utfordring for SSB å redusere svarbyrden til oppgavegiverne i utvalgsundersøkelser blant annet ved å spre utvalgene over flest mulig enheter. Det er et problem når noen små virksomheter opplever å bli trukket ut til 5-6 undersøkelser mens andre slipper å delta i noen.

Positiv samordning kan oppnås ved at utvalgene benytter de samme startverdiene a for sine utvalg. Negativ samordning oppnås ved at startverdiene velges så langt fra hverandre at utvalgene ikke overlapper eller overlapper minst mulig. Dette er lettest å få til ved Bernoulli og enkel tilfeldig trekking. Ved Poisson eller SP trekking vil svært forskjellig inklusjonssannsynligheter i to undersøkelser vanskeliggjøre positiv samordning. For enheter som har svært stor, evt. sikker, sannsynlighet for å bli trukket i den ene eller begge undersøkelsene vil en negativ samordning være vanskelig eller umulig. Det vil også være umulig å helt unngå overlapp mellom to undersøkelser dersom utvalgsandelene for de to undersøkelsene i sum overstiger 1. Eksempel på samordning av to utvalg finnes i avsnitt 8.2.

5. Rotasjon

Dersom vi har en undersøkelse som gjentas over tid og vi ønsker et roterende utvalg der deler av utvalget skiftes ut hver gang, kan dette gjøres ved å forskyve startverdien a litt hver gang en ny undersøkelse skal gjøres. Hvis to samordnede utvalg skal roteres, må de roteres like mye, ellers vil samordningen bli ødelagt etter en viss tid. Dette er vanskelig å få til med mindre stratifiseringen og utvalgsandelene i hvert stratum er identiske i de to undersøkelsene. Rotasjonen bør også være like stor i alle strata. I konstante populasjoner vil man med en slik metode kunne forutsi eksakt når en enhet vil tre inn i eller ut av utvalget. I de aller fleste populasjoner vil det imidlertid være både fødsler og død, og dette vil komplisere forholdet. Ved Poisson og SP utvalg vil enheter ved en slik type rotasjon kunne kastes midlertidig ut av utvalget for så å komme inn igjen. Dette skyldes at enheter med liten π_i vil endre ξ_i^* hurtigere enn enheter med stor π_i når startverdien a endres. I praksis vil også π_i endre seg over tid når populasjonen endrer seg. Dette vil også forsterkes når nye enheter kommer til populasjonen. De prinsipielle statistiske egenskapene ved utvalget vil imidlertid bli bevart uansett. Det finnes en alternativ rotasjonsteknikk, kalt "Random Rotation Cohort method" (Ohlsson 1995) som reduserer forekomsten av uheldige inn- og utkastinger av utvalget. Jeg vil imidlertid ikke komme nærmere inn på denne teknikken her.

For eksempel med rotasjon i strata med enkelt tilfeldige utvalg, se avsnitt 8.1. Se avsnitt 8.3 for et eksempel på rotasjon i SP-utvalg.

6. Fremtiden

Fra 1995 har det vært arbeidet med å organisere ulike tiltak for samordning av utvalgsundersøkelser rettet mot næringslivet. Tiltak som er nødvendige for å få til en slik samordning er foreslått og noen er gjennomført. Blant disse er etableringen av en felles basispopulasjon for økonomiske enheter i et sentralt Bedrifts- og Foretaksregister. Enhetlig populasjonsforvaltning er en nødvendig forutsetning for samordning av utvalg. Et annet tiltak er etableringen av delregistersystemet som holder rede på hvilke enheter som har deltatt i hvilke undersøkelser og når. Dette systemet må tas i bruk for absolutt alle undersøkelser som retter seg mot deler av en basispopulasjon og samordnes med hverandre og med BoF slik at det faktisk blir mulig til enhver tid å holde rede på hvilke enheter som har deltatt i hvilke utvalg og når, på tvers av seksjoner og avdelinger. Den samordnede utvalgsplanen for strukturstatistikkene (Solheim 2001) må også sees på som et resultat av dette arbeidet.

En mest mulig optimal samordning av datainnhenting for å spre oppgavebyrden vil kreve at all trekking av utvalg for økonomisk statistikk flyttes fra fagseksjonene til en sentral enhet som også har ansvar for å forvalte PRN-tallene og føre delregister. Dette må være målet for prosjektet "Oppgavebyrden og samordning av utvalg for næringslivet" (prosj. nr 3015) nevnt i forordet.

7. Eksempler på anvendelser

7.1. Lastebilundersøkelsen. Eksempel på rotasjon.

I denne undersøkelsen trekkes 2400 lastebiler hvert kvartal fordelt på 96 strata. Utvalget skal være enkelt tilfeldig i hvert stratum og roteres slik at en bil som hovedregel blir med i utvalget to kvartal på rad for deretter å ha rett til å "hvile" i minst fire kvartal. En ordning for vedlikehold av PRN-tall i trekkeregisteret er etablert. Hver gang trekkeregisteret oppdateres, følger PRN-tallene for alle biler som var med i registeret forrige kvartal med. Deretter leses en startverdi for tilfeldig tall generatoren og PRN genereres for alle biler som er "nye" i det aktuelle kvartal. Siste nye PRN som ble laget "forrige" kvartal og ble tatt vare på og brukes som startverdi "dette" kvartalet. .

For å administrere at utvalget skal roteres slik at hver bil blir med (høyst) to ganger, trekkes utvalget i to "puljer" som begge hver for seg og til sammen er et enkelt tilfeldig utvalg av hele populasjonen. "Gammel" pulje representerer en de bilene som var med i utvalget første gang forrige kvartal med en oppdatering av nye biler. "Ny" pulje er biler som ikke var med forrige gang. Trekkingen av puljene reguleres ved startverdiene som brukes ved trekkingen.

Trekkingen og rotasjonen foregår uavhengig i hvert av de 96 strataene, så beskrivelsen nedenfor kan referere til et vilkårlig stratum. Utvalgsstørrelsen beregnes for de 96 strataene og fordeles så likt som mulig på "gammel" og "ny" pulje. Hvis den totale utvalgsstørrelsen er et odde tall må den ene puljen få en bil mer enn den andre. Det blir annen hver gang gammel pulje ($n_{gml} = n_{ny} + 1$) og annen hver gang ny pulje ($n_{ny} = n_{gml} + 1$). Ellers $n_{gml} = n_{ny}$. Siste PRN som ble trukket i gammel pulje (**start1**) og siste PRN som ble trukket i ny pulje *forrige* kvartal (**start2**) er tatt vare på og brukes som startverdier for trekkingen nå. **prn2** defineres som

```
prn2 = PRN - start1;  
IF prn2 <= 0 THEN prn2 = prn2 + 1;
```

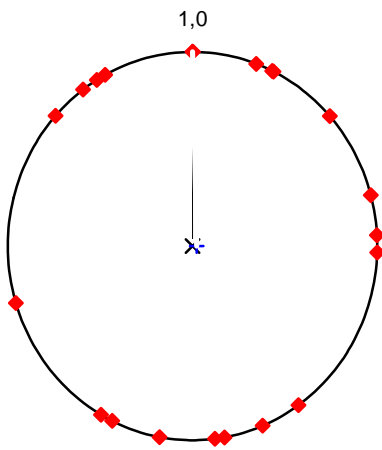
Populasjonen sorteres stigende etter **prn2** og de n_{gml} første med **prn2** > 0 trekkes til gammel pulje. Hvis det er større avgang enn tilgang i populasjonen av enheter med PRN mellom **start1** og **start2**, vil trekkingen til gammel pulje kunne gå forbi start2. Da oppdateres **start2** til verdien til siste PRN som trekkes i gammel pulje. Ellers holdes **start2** uforandret. Vi setter så **start1** = **start2** og gjentar prosedyren for å trekke ny pulje. Årsaken til at trekkingen av gammel pulje kan gå forbi **start2** er delvis hvis $n_{gml} = n_{ny} + 1$ og delvis hvis det er flere avganger enn tilganger i populasjonen blant biler med PRN mellom **start1** og **start2**.

En kunne trekke utvalget i ett uten å bruke **start2**. Da ville en imidlertid risikere at noen enheter som opprinnelig ble trukket som de siste i sin pulje ble med i utvalget tre ganger. Det er dessuten hensiktsmessig å kunne markere på denne måten hvilke biler som skal kunne trekkes en gang til og hvilke som ikke skal det.

Hensynet til at enhetene skal beskyttes mot å trekkes for mange ganger på rad blir ikke automatisk ivarettatt av mekanismen ovenfor dersom det er mulig at enhetene kan skifte stratum over tid. I lastebilundersøkelsen, hvor aldergruppering er en stratifiseringsvariabel, vil dette skje når en bil går fra å være fem til å bli seks år gammel. En bil som har vært med to ganger som femåring og så blir seks vil kunne bli trukket igjen for tredje og fjerde gang. I lastebilundersøkelsen vil dette problemet bli løst ved at slike biler eksplisitt hoppes over i trekkingen. En mer korrekt og elegant måte å løse problemet ville være å bruke samme startverdier i alle strata og rotere alle strata like mye, det vil si like mye som det stratomet som til enhver tid roteres hurtigst. I utvalgplaner der noen strata der noen strata har en stor utvalgsandel i forhold til den tid enhetene skal kunne "hvile" vil dette imidlertid kunne by på andre problem og anbefales helst hvis utvalgsandelen er liten i alle strata. Dersom en ønsker å bruke PRN-tall til å samordne utvalg fra forskjellige undersøkelser, vil dette imidlertid være påkrevd.

Den trekke- og rotasjonsmetoden som er benyttet i lastebilundersøkelsen kan generaliseres til undersøkelser med flere enn to puljer. En slik undersøkelse er AKU hvor hver deltakende familie/person skal

være med åtte ganger. Metoden vil da automatisk hver gang trekke et riktig utvalg for hver pulje uten at det er nødvendig å trekke tilleggsutvalg av nye enheter i populasjonen. Hver pulje blir et enkelt tilfeldig utvalg for hvert stratum og det blir også hele utvalget i hvert stratum. Fusjon og fisjon av familieenheter vil til enhver tid bli behandlet korrekt ved utvalgstrekkningen, helt automatisk. En forutsetning for at dette er imidlertid at PRN implementeres også i personregisteret som er AKUs trekkegrunnlag. Dette er derfor en anbefaling for fremtidig trekking av AKU.

Bil nr.	PRN	Gml	Ny	
1	0,0003	1,8	7,15	
2	0,0560	1,9	8,15	
3	0,0711	1,9	8	
4	0,1335	2,9	1,8	
5	0,2078	2,10	1,9	
6	0,2411	3,10	2,9	
7	0,2551	3,11	2,10	
8	0,4028	3,11	2,10	
9	0,4380	4,11	3,10	
10	0,4723	4,12	3,11	
11	0,4807	5,12	4,11	
12	0,5284	5,13	4,12	
13	0,5716	5,13	4,12	<p>Figur 4.1. Et tenkt stratum med 19 biler. PRN-tallene for de 19 bilene (kolonne 2) er representert ved punkter på en sirkel som symboliserer intervallet fra 0 til 1. Vi tenker oss at fem av de 19 bilene skal være med hver gang. Kolonnen Gml forteller hvilket av de 15 første kvartalene den aktuelle bilen er med i gammel pulje. Kolonnen Ny forteller hvilket kvartal bilen er med i ny pulje. For eksempel er bil nr. 12 med PRN = 0,5284 med i ny pulje gammel pulje 4. og 12. kvartal og i gammel pulje 5. og 13. kvartal. På denne måten roterer utvalget rundt sirkelen.</p> <p>Eksempelet forutsetter en fast populasjon av biler i stratumet. I virkeligheten har populasjonen både avgang og tilgang, gjerne hvert kvartal. Avgangene må representeres ved at punkter på sirkelen forsvinner, mens tilganger blir representert ved at nye punkt blander seg inn. Begge deler vil påvirke rotasjonen.</p>
14	0,5823	6,13	5,12	
15	0,7027	6,14	5,13	
16	0,8670	7,14	6,13	
17	0,8992	7,15	6,14	
18	0,9134	7,15	6,14	
19	0,9216	8,15	7,14	

7.2. Sannsynlighetsdelen til UT-årsutvalg, et eksempel på samordning.

Årsutvalget til UT – Utenrikstransaksjoner, består av to utvalg, ett for kapitalbalanser og ett for tjenestehandel med utlandet. Trekkeenhetene i begge utvalgene er foretak. Populasjonene for de to utvalgene er delvis overlappende, og begge populasjonene ble delt i ”store”, ”middels” og ”små”. Store enheter ble trukket med sikker sannsynlighet. Blant de middels ble det trukket et sannsynlighetsutvalg og blant de små ble det ikke trukket utvalg. Det var så spørsmål om kapitalutvalget og tjenesteutvalget kunne samordnes på noe vis. Det var tre alternativer: Positiv samordning, ingen samordning og negativ samordning. Positiv samordning vil si at vi ønsker *størst* mulig overlapp mellom utvalgene. Negativ

samordning betyr *minst* mulig overlapp, dvs. mest mulig spredning av svarbyrden. Ingen samordning betyr at utvalgene trekkes uavhengig av hverandre.

Det er ikke mulig å samordne deltakelse i utvalgene for enheter som må være med i et eller begge utvalg. Samordningen måtte begrense seg til enheter som var middels store i begge populasjonene. Det var 2522 middels store kapitalenheter og 835 middels store tjenesteenheter. Av disse var bare 178 enheter middels store i begge populasjonene. Det skulle trekkes 1250 middels store kapitalenheter og 421 middels store tjenesteenheter, i begge tilfelle etter stratifiserte planer. Utvalgsandelene kunne variere noe mellom strata.

PRN-tall ble generert for alle enheter i en kopi av bedrifts- og foretaksregisteret. Populasjonene for hvert av de to utvalgene ble ekstrahert og sortert etter stratum x PRN. For å få til samordning mellom to forskjellige utvalgsundersøkelser må *samme* startverdi benyttes i alle strata. Negativ samordning av utvalgene kunne oppnås ved å bruke to startverdier a og b som ligger langt fra hverandre på sirkelen, men like for alle strata. Valg av $a = 0$ for kapitalutvalget og $b = 0,60$ for tjenesteutvalget resulterte i at ingen av de 178 felleseenhetene ble trukket til begge utvalg.

Ingen samordning resulterte i at 60 av de 178 enhetene ble trukket til begge utvalg.

Positiv samordning ble oppnådd ved å bruke samme startverdier for de to undersøkelsene, $a = b (= 0)$. Da ble 91 av de 178 enhetene trukket til begge.

Positiv samordning ble foretrukket da mye av det enhetene må sette seg inn i for å kunne svare i undersøkelsen er likt.

I fremtiden vil PRN-tallene bli brukt til å samordne UT-utvalgene over tid. Dette er en oppgave som er viktigere enn samordningen de to utvalgene imellom.

7.3. Prisindeks for transport på vei. Et eksempel på SP trekking og rotasjon.

7.3.1. kort beskrivelse av utvalgsplan

I denne undersøkelsen ble det trukket 200 foretak med bedrifter som skulle svare på spørsmål om prising av transporttjenester. Populasjonen besto av 9339 foretak med 9380 bedrifter. 13 foretak hadde mer enn en bedrift, og i de 6 av disse som ble trukket ble det trukket ut 8 bedrifter slik at utvalget til slutt ble 202 bedrifter. Vi skal imidlertid her bare se på trekkingen av foretak. 22 av de 200 foretakene ble på grunn av størrelsen (omsetning ifølge moms registeret) satt til side som trukket med "sikker" sannsynlighet. 178 foretak ble så trukket med SP trekking. Populasjonen ble delt i syv strata på grunnlag av den informasjonen som var tilgjengelig om den type transport de fortrinnsvis drev. Denne inndelingen, samt populasjon, utvalgsstørrelse, antall "sikre" og antall som skulle trekkes er gitt i tabell 1.

De 178 foretakene som var igjen fikk tilegnet nominelle inklusjonssannsynligheter

$$\pi_i = n \frac{y_i}{T}$$

der y_i var omsetningstall, T var total omsetning i stratumet utenom de "sikre" og n var utvalgsstørrelsen i stratumet etter at de sikre var fjernet. Alle hadde $\pi_i < 0,75$. Lar $x_i = \text{PRN_foretak}$ og $\xi_i = x_i / \pi_i$. Trekker de n enhetene som har de minste ξ_i – ene i hvert stratum.

Tabell 1. Strataene i utvalgsplanen

Stratum	Definisjon For detaljer henvises til program i vedlegg.	Antall foretak (N)	Utvalgs- størrelse (n)	Derav sikre	Sanns. utvalg
Termo	Alle foretak med bokstavkombinasjonen 'fryse', 'termo' eller 'thermo' i bedrifts eller foretaksnavnet.	49	15	9	6
Tank	Alle foretak med navn som kunne indikere at de spesialiserte seg på tankbiltransport. Dessuten ble alle bedrifter hvor kjøretøygruppe 330 + 335 dominerte den kjente bilparken inkludert i Tank.	275	30	8	22
Kjg_320	Foretak hvor kjøretøy i gruppe 320 dominerer.	1487	30	0	30
Kjg_321	Foretak hvor kjøretøy i gruppe 321 dominerer.	902	30	0	30
Kjg_325	Foretak hvor kjøretøy i gruppe 323 og 325 dominerer.	653	30	3	27
Kjg_340	Foretak hvor kjøretøy i gruppe 340 dominerer.	494	30	1	29
Resten	Foretak som ikke har latt seg klassifisere.	5479	35	1	34

7.3.2. Rotasjon av utvalget

Det er ønskelig at utvalget skulle holdes nokså fast men roteres med jevne mellomrom i tillegg til at det bør oppdateres i forhold til populasjonen. Rotasjon av foretakene finner sted ved at det blir etablert en forskyvning av x_i

$$x_i^* = \begin{cases} x_i - \varepsilon & \text{hvis } x_i > \varepsilon \\ 1 - \varepsilon + x_i & \text{hvis } x_i < \varepsilon \end{cases}$$

der $0 < \varepsilon < 1$. Vanligvis vil vi velge ε som et lite tall. Siden x_i er uniformt fordelt på $[0, 1]$ må x_i^* også være det. La π_i^* være inklusjonssannsynligheten på det nye tidspunktet. Det etableres så en variabel

$$\xi_i^* = \frac{x_i^*}{\pi_i^*}$$

og det roterte utvalget velges som beskrevet i avsnitt 4.4 på basis av ξ_i^* . Enheter som får $x_i^* - \varepsilon < 0$ vil nå bli rotert ut av utvalget med mindre $(1 - \varepsilon + x_i^*)/\pi_i^*$ blir av størrelsesorden 1 eller mindre. Men så lenge π_i^* ellers ikke tillates å være større enn 0,75 vil det ikke skje med de verdier av ε som er aktuelle. Hvis $\pi_i^* > \varepsilon$ vil enheten bli rotert ut senest etter $\lceil \pi_i/\varepsilon \rceil$ rotasjoner der $\lceil \cdot \rceil$ betyr avrunding oppover. Hvis for eksempel $\pi_i = 0,09$ (fast) og $\varepsilon = 0,02$ vil enheten senest være ute etter fem rotasjoner. I forventning vil den være ute etter 2,25 rotasjoner.

En ulempe med metoden er at enheter kan komme inn i utvalget, bli rotert ut ved neste rotasjon og så komme inn igjen i den deretter. Dette kan skje hvis π_i^* varierer mye over tid som følge av at basisen, omsetningstallene (markedsandelene) gjør det eller det er mye inngang og avgang i trekkepopulasjonen. Det kan imidlertid også hende selv om ingenting forandrer seg i trekkegrunnlaget, dvs. $\pi_i^* = \pi_i$ for alle enheter. Dette vil primært skje med "store" enheter som er blant de siste som er trukket med i utvalget, dvs. at enhetene er blant dem som har størst verdi på ξ_i^* blant de n minste i populasjonen. I denne situasjonen er imidlertid rotasjonsmønsteret fastlagt for all fremtid straks det første utvalget er trukket.

For det utvalget som er trukket viser tabell 2 hvor mange som vil skiftes ut i hvert stratum ved hver rotasjon dersom populasjonen er fast og vi velger $\varepsilon = 0,01$. En fullstendig beskrivelse av rotasjonen av de enheter som er inntatt i utvalget de første seks gangene og som tabell 2 bygger på er gitt i Vedlegg.

En ser at utskiftningen blir mye større i noen strata enn i andre. Dette er i noen grad naturlig når strataene i utgangspunktet er store og hver enhet har mange "konkurrenter" om bli trukket. Det kan

imidlertid reguleres ved å variere ε etter stratum. Med $\varepsilon = 0,01$ vil ingen i **Termo** bli skiftet ut i de første fem rotasjonene, men en vil bli skiftet ut i den sjette.

Avgang, utvandring fra populasjonen vil redusere ”konkurransen” og bidra til større stabilitet i utvalget for de øvrige enhetene forutsatt at utvalgsstørrelsen holdes konstant. Innvandring vil ha motsatt virkning. Tilsvarende vil reduksjon i π_i ($\pi_i^* < \pi_i$) bidra til at en enhet fases ut forttere enn den ellers ville og en økning vil bidra til det motsatte. Totalt må vi regne med at populasjonsendringer og endringer i størrelsesmål vil bidra til større tendens til at relativt store enheter som er kommet ”sist” inn i utvalget fortrenses for så å komme inn igjen. Dette er lite ønskelig. Enheter som kommer i denne situasjonen bør derfor kunne beholdes i utvalget inntil de fases ut på naturlig måte i andre enden ($x_i^* - \varepsilon < 0$) selv om det skulle bety visse variasjoner i utvalgsstørrelsen. For å unngå for store variasjoner i π_i over tid for samme enhet, bør π_i avhenge av omsetningstall lengre tilbake i tid, f.eks. tre år. Nye enheter i populasjonen vil derved ikke opparbeide seg ”full” sannsynlighet for å bli trukket før det er gått en tid.

Tabell 2. Antall enheter som blir skiftet ut ved hver rotasjon ved statistisk populasjon hvis $\varepsilon = 0,01$.

($\pi_i^* = \pi_i$ for alle enh.)

Stratum	n^*	Rotasjon				
		1	2	3	4	5
Termo	6	0	0	0	0	0
Tank	9	2	1	5	3	1
Kjg_320	30	12	19	8	9	9
Kjg_321	30	7	6	11	6	6
Kjg_325	27	3	8	5	9	3
Kjg_340	29	3	5	3	4	5
Resten	34	16	17	23	17	17
I alt	178	43	56	55	48	41

* n er utvalgsstørrelsen når de sikre er trukket fra.

Såkalt relativt store enheter som er kommet ”sist” inn i utvalget og som ikke fortrenses en kort tid for å komme inn igjen, må forvente å bli lenge i utvalget. En enhet som kommer inn med f.eks. PRN = 0,40, vil med $\varepsilon = 0,01$ måtte forvente å bli i utvalget i 40 rotasjonsperioder. Hvis dette er uheldig, så kan dette aksellereres ved å foreta en mer omfattende rotasjon for eksempel hvert 5. år. Det kan imidlertid ikke loves faste hviletider for relativt store enheter som går ut. I det lange løp må hver enhet være med en andel av tiden som svarer til deres π_i .

8. Referanser

- Heldal, J. (2004). *Prisindekser for transport på vei. Dokumentasjon av utvalgsplan*. Foreløpig notat.
- McKenzie, R. & Gross, B. (2000): *Synchronised Sampling*, Australian Bureau of Statistics.
- Ohlsson, E. (1992): *SAMU, The System for Co-ordination of samples from the Business Register at Statistics Sweden*. R&D Report Statistics Sweden 1992:18
- (1995): *Coordination of Samples Using Permanent Random Numbers*. I ”Business Survey Methods”, Wiley 1995
- (1998): *Sequential Poisson Sampling*. Journal of Official Statistics (JOS), Vol. 14, No. 2.
- Rosén, B. (a) *Asymptotic theory for order sampling*. Journal of Statistical Planning and Inference, 62 (1997) s. 135-158.
- (b) *On sampling with probability proportional to size*. Journal of Statistical Planning and Inference, 62 (1997) s. 159-191.
- Solheim, L. (2001): *En samordnet utvalgsplan. Prinsipper og metoder*. Upublisert notat.
- Thulin, G. (1976): *Co-ordination of samples in enterprise statistics (SAMU). Implementation of the JALES technique*. Statistisk Tidsskrift Vol. 14 pp 85-101 (På svensk med engelsk sammendrag).

9. Vedlegg 1: Rotasjonsbeskrivelse

Rotasjonsbeskrivelse: Tabellen beskriver hvordan enheter går inn i og ut av utvalget over seks ganger. '1' i kolonnen for utv_nr indikerer at enheten er med i utvalget. '.' at den er ute. Enheter som går inn i utvalget flere ganger er markert med rødt.

Obs	stratum	org_nr	Selection	PRN_	ξ	utv_nr1	utv_nr2	utv_nr3	utv_nr4	utv_nr5	utv_nr6
			Prob, π_i	foretak, x_i							
1	tank	939697608	0,23559	0,00206	0,009	1
2	tank	970227849	0,01528	0,00927	0,607	1
3	tank	970511687	0,02286	0,01077	0,471	1	1
4	tank	977546001	0,0274	0,02036	0,743	1	1	1	.	.	.
5	tank	931986732	0,17429	0,02159	0,124	1	1	1	.	.	.
6	tank	871128472	0,0425	0,02426	0,571	1	1	1	.	.	.
7	tank	979933797	0,00751	0,02614	3,483	.	.	1	.	.	.
8	tank	870984812	0,00963	0,02657	2,76	.	1	1	.	.	.
9	tank	983297781	0,31333	0,03065	0,098	1	1	1	1	.	.
10	tank	970165991	0,0154	0,03324	2,158	.	1	1	1	.	.
11	tank	981477898	0,01041	0,04922	4,727	1	.
12	tank	951872288	0,26426	0,05156	0,195	1	1	1	1	1	1
13	tank	957139310	0,04441	0,07267	1,636	1	1	1	1	1	1
14	tank	913521013	0,02662	0,07355	2,763	.	.	.	1	1	1
15	tank	971000767	0,023	0,08312	3,614	1
16	tank	953844931	0,05586	0,09085	1,626	1	1	1	1	1	1
17	tank	926571680	0,03477	0,09321	2,681	1	1
18	tank	967434442	0,08807	0,09444	1,072	1	1	1	1	1	1
19	tank	964989311	0,10669	0,15278	1,432	1	1	1	1	1	1
20	tank	980012832	0,12976	0,16188	1,248	1	1	1	1	1	1
21	tank	984070713	0,08942	0,17969	2,009	.	.	.	1	1	1
22	tank	979971532	0,0975	0,18261	1,873	.	.	.	1	1	1
23	tank	947558773	0,12459	0,1838	1,475	1	1	1	1	1	1
24	tank	976228413	0,1024	0,20422	1,994	1	1
25	tank	933578534	0,39961	0,22146	0,554	1	1	1	1	1	1
26	tank	840758532	0,14942	0,22781	1,525	1	1	1	1	1	1
27	tank	831179252	0,29282	0,25211	0,861	1	1	1	1	1	1
28	tank	935709806	0,4754	0,25247	0,531	1	1	1	1	1	1
29	tank	966025301	0,15206	0,27746	1,825	.	.	.	1	1	1
30	tank	983498922	0,33328	0,29166	0,875	1	1	1	1	1	1
31	tank	960317491	0,17288	0,30381	1,757	1	1	1	1	1	1
32	tank	841369262	0,21043	0,33342	1,584	1	1	1	1	1	1
33	tank	975992756	0,24571	0,44733	1,821	.	.	.	1	.	.
34	tank	982312515	0,41234	0,68213	1,654	1	1	1	1	1	1
35	termo	957992889	0,00755	0,05503	7,286	1	1	1	1	1	1
36	termo	986165142	0,0236	0,11318	4,797	1	1	1	1	1	1
37	termo	959783683	0,08106	0,14579	1,799	1	1	1	1	1	1
38	termo	968542788	0,03099	0,17611	5,682	1	1	1	1	1	1
39	termo	963134649	0,0352	0,20478	5,818	1	1	1	1	1	1
40	termo	970920374	0,06523	0,27676	4,243	1	1	1	1	1	1
41	resten	967371319	0,00626	0,00043	0,068	1
42	resten	976853881	0,00854	0,00046	0,054	1
43	resten	980268179	0,00098	0,00055	0,563	1
44	resten	856741842	0,01036	0,00075	0,073	1
45	resten	979380577	0,00268	0,0008	0,299	1
46	resten	980853381	0,00398	0,00112	0,282	1
47	resten	980762955	0,00832	0,00323	0,388	1
48	resten	974480557	0,13309	0,00334	0,025	1
49	resten	984417071	0,00807	0,00339	0,42	1
50	resten	975921549	0,00618	0,00412	0,668	1
51	resten	931898248	0,06194	0,00433	0,07	1
52	resten	978614590	0,01911	0,00471	0,246	1
53	resten	984511426	0,00505	0,00476	0,942	1
54	resten	981439546	0,00626	0,00661	1,056	1
55	resten	970159177	0,00841	0,008	0,952	1
56	resten	959906270	0,00842	0,00873	1,037	1
57	resten	870899572	0,00125	0,01026	8,202	.	1

58	resten	986109005	0,001	0,0109	10,868	.	1
59	resten	980404587	0,19774	0,01092	0,055	1	1
60	resten	977293499	0,00328	0,01119	3,414	.	1
61	resten	977293500	1,00328	1,01119	4,414	.	1
62	resten	977293501	2,00328	2,01119	5,414	.	1
63	resten	977293502	3,00328	3,01119	6,414	.	1
64	resten	977293503	4,00328	4,01119	7,414	.	1
65	resten	977293504	5,00328	5,01119	8,414	.	1
66	resten	971131160	0,02274	0,01329	0,585	1	1
67	resten	983853129	0,00445	0,01338	3,008	.	1
68	resten	980145549	0,00699	0,01611	2,305	.	1
69	resten	965553576	0,01057	0,01801	1,703	.	1
70	resten	977457149	0,00895	0,01832	2,047	.	1
71	resten	976702972	0,02679	0,01932	0,721	1	1
72	resten	978693288	0,00889	0,01978	2,225	.	1
73	resten	975955877	0,00384	0,02029	5,282	.	.	1	.	.	.
74	resten	953990687	0,05151	0,02052	0,398	1	1	1	.	.	.
75	resten	883366522	0,00308	0,02064	6,7	.	.	1	.	.	.
76	resten	870424612	0,00285	0,02085	7,32	.	.	1	.	.	.
77	resten	984968051	0,00271	0,0212	7,83	.	.	1	.	.	.
78	resten	980435369	0,00222	0,02154	9,718	.	.	1	.	.	.
79	resten	984717865	0,00625	0,02164	3,463	.	.	1	.	.	.
80	resten	979559240	0,00431	0,02267	5,265	.	.	1	.	.	.
81	resten	969427125	0,00362	0,02269	6,262	.	.	1	.	.	.
82	resten	985828350	0,00311	0,02294	7,384	.	.	1	.	.	.
83	resten	986618597	0,00436	0,02321	5,328	.	.	1	.	.	.
84	resten	976048466	0,01932	0,02323	1,202	.	1	1	.	.	.
85	resten	979113269	0,00724	0,02337	3,229	.	.	1	.	.	.
86	resten	983351123	0,00531	0,02442	4,603	.	.	1	.	.	.
87	resten	970130705	0,00445	0,02456	5,522	.	.	1	.	.	.
88	resten	983773818	0,00794	0,02487	3,132	.	.	1	.	.	.
89	resten	984314094	0,01252	0,02499	1,996	.	.	1	.	.	.
90	resten	975850056	0,00626	0,02571	4,104	.	.	1	.	.	.
91	resten	983650724	0,03584	0,0263	0,734	1	1	1	.	.	.
92	resten	977047552	0,01534	0,02653	1,729	.	1	1	.	.	.
93	resten	880356992	0,17077	0,02781	0,163	1	1	1	.	.	.
94	resten	980396479	0,00998	0,0281	2,816	.	.	1	.	.	.
95	resten	959914737	0,04472	0,02827	0,632	1	1	1	.	.	.
96	resten	983559018	0,00049	0,03014	62,042	.	.	.	1	.	.
97	resten	970119876	0,00053	0,0302	56,633	.	.	.	1	.	.
98	resten	969998092	0,00205	0,03029	14,771	.	.	.	1	.	.
99	resten	984772998	0,00116	0,03076	26,48	.	.	.	1	.	.
100	resten	970087508	0,00186	0,0312	16,806	.	.	.	1	.	.
101	resten	970082077	0,002	0,03129	15,638	.	.	.	1	.	.
102	resten	986531416	0,00399	0,03139	7,868	.	.	.	1	.	.
103	resten	958876661	0,00214	0,03162	14,781	.	.	.	1	.	.
104	resten	970079270	0,00249	0,03202	12,874	.	.	.	1	.	.
105	resten	960584600	0,00408	0,03233	7,919	.	.	.	1	.	.
106	resten	978653987	0,00253	0,0324	12,783	.	.	.	1	.	.
107	resten	984156111	0,00416	0,03302	7,935	.	.	.	1	.	.
108	resten	933408337	0,00704	0,03571	5,068	.	.	.	1	.	.
109	resten	876177722	0,05531	0,0358	0,647	1	1	1	1	.	.
110	resten	863179262	0,00537	0,03646	6,792	.	.	.	1	.	.
111	resten	967346136	0,00636	0,03778	5,935	.	.	.	1	.	.
112	resten	984512333	0,01634	0,03874	2,37	.	.	.	1	.	.
113	resten	971015861	0,00204	0,04014	19,672	1	.
114	resten	983889093	0,00309	0,04019	13,004	1	.
115	resten	969242702	0,00131	0,04133	31,632	1	.
116	resten	983216986	0,0026	0,04147	15,949	1	.
117	resten	969961393	0,00226	0,04194	18,561	1	.
118	resten	939330755	0,06765	0,04244	0,627	1	1	1	1	1	.
119	resten	969957728	0,00216	0,0425	19,713	1	.
120	resten	982440483	0,00234	0,04281	18,277	1	.
121	resten	984146051	0,00435	0,04369	10,054	1	.
122	resten	985346216	0,08788	0,04421	0,503	1	1	1	1	1	.
123	resten	983445381	0,50407	0,04642	0,092	1	1	1	1	1	.
124	resten	979803672	0,05698	0,04718	0,828	1	1	1	1	1	.

125	resten	981510283	0,01078	0,04747	4,405	1	.
126	resten	985282595	0,21945	0,04794	0,218	1	1	1	1	1	.
127	resten	970390731	0,00029	0,05006	171,757	1
128	resten	870122772	0,00199	0,05018	25,232	1
129	resten	979626886	0,00094	0,05054	53,857	1
130	resten	970038027	0,00112	0,05078	45,236	1
131	resten	970276084	0,00166	0,05112	30,756	1
132	resten	971015330	0,00223	0,05127	22,944	1
133	resten	981467663	0,00296	0,05186	17,495	1
134	resten	984955103	0,01682	0,05188	3,085	1	1
135	resten	862052382	0,00584	0,05242	8,97	1
136	resten	985786232	0,00312	0,05308	17,028	1
137	resten	982690811	0,01745	0,05331	3,054	1	1
138	resten	980233790	0,00777	0,0536	6,894	1
139	resten	965594299	0,00592	0,0537	9,07	1
140	resten	984065124	0,01251	0,0543	4,34	1	1
141	resten	976390768	0,00574	0,05474	9,536	1
142	resten	985463638	0,01704	0,05534	3,248	1	1
143	resten	978691048	0,00707	0,0562	7,947	1
144	resten	981616332	0,008	0,05635	7,04	1
145	resten	876735652	0,01026	0,05637	5,496	1
146	resten	928531147	0,34675	0,05773	0,166	1	1	1	1	1	1
147	resten	985983224	0,00854	0,05785	6,778	1
148	resten	986146091	0,01588	0,05877	3,702	1	1
149	resten	943295328	0,05877	0,05993	1,02	1	1	1	1	1	1
150	resten	970407065	0,02578	0,06385	2,476	1	1
151	resten	974534851	0,02637	0,07085	2,686	1	1
152	resten	979816588	0,02352	0,07246	3,081	1
153	resten	980837874	0,04705	0,08503	1,807	.	.	.	1	1	1
154	resten	970938664	0,14966	0,1424	0,952	1	1	1	1	1	1
155	resten	936391060	0,09217	0,14684	1,593	.	.	.	1	1	1
156	resten	884048842	0,13583	0,16275	1,198	.	.	.	1	1	1
157	resten	916341474	0,11334	0,18988	1,675	1	.
158	resten	944326731	0,16257	0,20205	1,243	.	.	.	1	1	1
159	resten	911943441	0,34497	0,30201	0,875	1	1	1	1	1	1
160	resten	932725967	0,38593	0,48285	1,251	.	.	.	1	1	.
161	resten	928894657	0,3879	0,49421	1,274	.	.	.	1	1	.
162	resten	984598300	0,57511	0,55348	0,962	1	1	1	1	1	1
163	resten	980405443	0,61911	0,65859	1,064	.	1	.	1	1	1
164	kjg_320	981955633	0,04227	0,00204	0,048	1
165	kjg_320	970219889	0,00676	0,0032	0,473	1
166	kjg_320	936093825	0,0584	0,00334	0,057	1
167	kjg_320	864013732	0,01221	0,00431	0,353	1
168	kjg_320	974332558	0,04346	0,00491	0,113	1
169	kjg_320	976793803	0,00974	0,00528	0,542	1
170	kjg_320	979775040	0,01234	0,00528	0,428	1
171	kjg_320	980246981	0,12939	0,00676	0,052	1
172	kjg_320	959707189	0,01103	0,00722	0,654	1
173	kjg_320	984659504	0,01292	0,00755	0,585	1
174	kjg_320	969924498	0,02249	0,00853	0,379	1
175	kjg_320	970100407	0,008	0,01099	1,374	.	1
176	kjg_320	870334052	0,01627	0,01183	0,727	1	1
177	kjg_320	970960775	0,00427	0,01219	2,856	.	1
178	kjg_320	980047520	0,00398	0,01222	3,073	.	1
179	kjg_320	970088407	0,00674	0,01268	1,882	.	1
180	kjg_320	976940288	0,00663	0,01289	1,943	.	1
181	kjg_320	970391797	0,03078	0,01323	0,43	1	1
182	kjg_320	983298494	0,01065	0,0133	1,249	.	1
183	kjg_320	969943387	0,00545	0,01332	2,445	.	1
184	kjg_320	979171846	0,03201	0,01441	0,45	1	1
185	kjg_320	970064184	0,02645	0,01517	0,574	1	1
186	kjg_320	979645139	0,00826	0,01529	1,851	.	1
187	kjg_320	980453030	0,15136	0,01605	0,106	1	1
188	kjg_320	970295682	0,05366	0,01631	0,304	1	1
189	kjg_320	970913475	0,01372	0,01684	1,228	.	1
190	kjg_320	970267093	0,01427	0,01747	1,224	.	1
191	kjg_320	914351111	0,14942	0,01776	0,119	1	1

192	kjg_320	958706731	0,02077	0,01869	0,9	.	1
193	kjg_320	971163453	0,08764	0,01916	0,219	1	1
194	kjg_320	956755425	0,00339	0,02011	5,932	.	.	1	.	.	.
195	kjg_320	963964412	0,02258	0,02341	1,037	.	1	1	.	.	.
196	kjg_320	958172168	0,00866	0,02344	2,705	.	.	1	.	.	.
197	kjg_320	970259783	0,00483	0,02436	5,048	.	.	1	.	.	.
198	kjg_320	981449576	0,01493	0,02442	1,636	.	.	1	.	.	.
199	kjg_320	970596968	0,00662	0,02482	3,749	.	.	1	.	.	.
200	kjg_320	965900071	0,00901	0,02589	2,871	.	.	1	.	.	.
201	kjg_320	970091483	0,01167	0,02913	2,496	.	.	1	.	.	.
202	kjg_320	951349984	0,02454	0,03118	1,271	.	.	1	1	.	.
203	kjg_320	955306198	0,00531	0,0318	5,99	.	.	.	1	.	.
204	kjg_320	966028882	0,02002	0,03292	1,644	.	.	1	1	.	.
205	kjg_320	970240357	0,00564	0,03376	5,988	.	.	.	1	.	.
206	kjg_320	983065724	0,00828	0,03381	4,086	.	.	.	1	.	.
207	kjg_320	970268332	0,00625	0,03504	5,603	.	.	.	1	.	.
208	kjg_320	970399623	0,01709	0,03711	2,171	.	.	.	1	.	.
209	kjg_320	979451210	0,1219	0,03768	0,309	1	1	1	1	.	.
210	kjg_320	971587016	0,02431	0,03934	1,618	.	.	1	1	.	.
211	kjg_320	932056674	0,07843	0,04111	0,524	1	1	1	1	1	.
212	kjg_320	959148511	0,01216	0,04112	3,381	1	.
213	kjg_320	970171487	0,00552	0,04133	7,485	1	.
214	kjg_320	970183175	0,00738	0,04175	5,656	1	.
215	kjg_320	979664508	0,03848	0,04355	1,132	.	.	1	1	1	.
216	kjg_320	983977979	0,02854	0,04526	1,586	.	.	1	1	1	.
217	kjg_320	975256057	0,07782	0,0466	0,599	1	1	1	1	1	.
218	kjg_320	970028897	0,0084	0,04718	5,616	1	.
219	kjg_320	980872858	0,02565	0,04809	1,875	.	.	.	1	1	.
220	kjg_320	970193316	0,01142	0,05174	4,529	1
221	kjg_320	970260420	0,00431	0,05176	11,999	1
222	kjg_320	970272208	0,00281	0,05201	18,496	1
223	kjg_320	960894022	0,03904	0,05254	1,346	.	.	1	1	1	1
224	kjg_320	955907728	0,03261	0,05357	1,643	.	.	.	1	1	1
225	kjg_320	970137939	0,00826	0,0538	6,513	1
226	kjg_320	883842812	0,02239	0,0546	2,438	1	1
227	kjg_320	970326170	0,03076	0,05815	1,89	1	1
228	kjg_320	975963586	0,03266	0,05837	1,787	.	.	.	1	1	1
229	kjg_320	924197196	0,01114	0,05856	5,258	1
230	kjg_320	974493721	0,01175	0,0598	5,091	1
231	kjg_320	954562182	0,04053	0,06012	1,483	.	.	1	1	1	1
232	kjg_320	954501272	0,02828	0,06546	2,315	1	1
233	kjg_320	946885614	0,07082	0,06561	0,926	.	.	1	1	1	1
234	kjg_320	971240121	0,04272	0,06746	1,579	1	1
235	kjg_320	960270002	0,02364	0,06757	2,858	1
236	kjg_320	952534955	0,1279	0,06842	0,535	1	1	1	1	1	1
237	kjg_320	971032316	0,09155	0,07054	0,771	1	1	1	1	1	1
238	kjg_320	985826447	0,05777	0,07226	1,251	.	.	1	1	1	1
239	kjg_320	976874854	0,13141	0,10231	0,779	1	1	1	1	1	1
240	kjg_320	970047204	0,07994	0,10524	1,317	1	1
241	kjg_320	946662704	0,06948	0,11164	1,607	1
242	kjg_320	981921062	0,10932	0,11696	1,07	.	.	1	1	1	1
243	kjg_320	952016547	0,17632	0,12174	0,69	1	1	1	1	1	1
244	kjg_320	974904535	0,34455	0,13364	0,388	1	1	1	1	1	1
245	kjg_320	916962835	0,39563	0,13497	0,341	1	1	1	1	1	1
246	kjg_320	980719545	0,1501	0,16042	1,069	.	.	1	1	1	1
247	kjg_320	943112517	0,18645	0,1615	0,866	1	.	1	1	1	1
248	kjg_320	931890174	0,25579	0,27954	1,093	1
249	kjg_320	929524039	0,58335	0,4614	0,791	1	1	1	1	1	1
250	kjg_321	970467890	0,00505	0,00223	0,442	1
251	kjg_321	970246576	0,00629	0,00348	0,553	1
252	kjg_321	931019929	0,06282	0,0068	0,108	1
253	kjg_321	970286683	0,00924	0,009	0,975	1
254	kjg_321	970334394	0,00903	0,00946	1,048	1
255	kjg_321	970489304	0,0054	0,01193	2,211	.	1
256	kjg_321	969960729	0,00527	0,01252	2,378	.	1
257	kjg_321	980298450	0,11086	0,01367	0,123	1	1
258	kjg_321	976726685	0,0051	0,01417	2,777	.	1

259	kjg_321	969019205	0,00628	0,01457	2,319	.	1
260	kjg_321	958588003	0,01006	0,01506	1,498	.	1
261	kjg_321	965259783	0,09794	0,02363	0,241	1	1	1	.	.	.
262	kjg_321	970170553	0,00554	0,02371	4,28	.	.	1	.	.	.
263	kjg_321	932896036	0,05297	0,02474	0,467	1	1	1	.	.	.
264	kjg_321	957028179	0,10403	0,0252	0,242	1	1	1	.	.	.
265	kjg_321	974341565	0,20921	0,02567	0,123	1	1	1	.	.	.
266	kjg_321	970279555	0,04649	0,02617	0,563	1	1	1	.	.	.
267	kjg_321	858480272	0,0185	0,02645	1,43	.	.	1	.	.	.
268	kjg_321	977542545	0,02076	0,02681	1,291	.	1	1	.	.	.
269	kjg_321	910593633	0,09479	0,02729	0,288	1	1	1	.	.	.
270	kjg_321	970139400	0,01973	0,02758	1,398	.	.	1	.	.	.
271	kjg_321	979787758	0,09705	0,028	0,288	1	1	1	.	.	.
272	kjg_321	968115146	0,08425	0,03162	0,375	1	1	1	1	.	.
273	kjg_321	967988677	0,04075	0,03231	0,793	1	1	1	1	.	.
274	kjg_321	976691350	0,15112	0,03342	0,221	1	1	1	1	.	.
275	kjg_321	936917801	0,03249	0,03416	1,051	.	1	1	1	.	.
276	kjg_321	916406541	0,10028	0,03712	0,37	1	1	1	1	.	.
277	kjg_321	983650856	0,02576	0,03764	1,461	.	.	1	1	.	.
278	kjg_321	971127600	0,01784	0,04091	2,293	.	.	.	1	1	.
279	kjg_321	980063437	0,03627	0,04229	1,166	.	.	1	1	1	.
280	kjg_321	984689942	0,0091	0,043	4,725	1	.
281	kjg_321	970934464	0,00458	0,04314	9,428	1	.
282	kjg_321	953900696	0,01311	0,04551	3,471	.	.	.	1	1	.
283	kjg_321	881513412	0,01749	0,04971	2,842	.	.	.	1	1	.
284	kjg_321	979678843	0,00753	0,05245	6,966	1
285	kjg_321	960811615	0,01293	0,0549	4,247	1	1
286	kjg_321	980052753	0,1285	0,05635	0,439	1	1	1	1	1	1
287	kjg_321	936869513	0,31167	0,0587	0,188	1	1	1	1	1	1
288	kjg_321	984276591	0,11033	0,06366	0,577	1	1	1	1	1	1
289	kjg_321	983434746	0,02285	0,07074	3,096	1
290	kjg_321	976482441	0,08666	0,07079	0,817	1	1	1	1	1	1
291	kjg_321	913519132	0,03959	0,0723	1,826	.	.	.	1	1	1
292	kjg_321	978611613	0,06574	0,07258	1,104	.	.	.	1	1	1
293	kjg_321	975259455	0,0455	0,07275	1,599	.	.	.	1	1	1
294	kjg_321	979537298	0,02323	0,07424	3,196	1
295	kjg_321	948321386	0,09325	0,07754	0,832	1	1	1	1	1	1
296	kjg_321	923144099	0,22449	0,07807	0,348	1	1	1	1	1	1
297	kjg_321	877073602	0,0342	0,07809	2,283	1	1
298	kjg_321	971223863	0,03127	0,07904	2,528	1
299	kjg_321	961386233	0,05648	0,08054	1,426	.	.	.	1	1	1
300	kjg_321	933654397	0,08589	0,08171	0,951	1	.	1	1	1	1
301	kjg_321	979615973	0,09404	0,08175	0,869	1	1	1	1	1	1
302	kjg_321	912991946	0,12994	0,08543	0,657	1	1	1	1	1	1
303	kjg_321	933641317	0,35657	0,09421	0,264	1	1	1	1	1	1
304	kjg_321	960397339	0,19509	0,12409	0,636	1	1	1	1	1	1
305	kjg_321	920634060	0,10132	0,12422	1,226	.	.	.	1	1	1
306	kjg_321	830170642	0,10697	0,13811	1,291	.	.	.	1	1	1
307	kjg_321	979844174	0,40738	0,15319	0,376	1	1	1	1	1	1
308	kjg_321	956971799	0,13281	0,20807	1,567	1
309	kjg_321	819992762	0,1903	0,26556	1,396	1	1
310	kjg_321	925503444	0,20163	0,30341	1,505	1
311	kjg_321	979672969	0,42386	0,43954	1,037	1	.	.	1	1	1
312	kjg_321	966742089	0,38746	0,44157	1,14	.	.	.	1	1	1
313	kjg_321	941084109	0,34495	0,46597	1,351	1	1
314	kjg_325	963022336	0,01219	0,00509	0,418	1
315	kjg_325	936444229	0,02199	0,00533	0,242	1
316	kjg_325	979574487	0,01558	0,01015	0,652	1	1
317	kjg_325	970175636	0,0014	0,01102	7,88	.	1
318	kjg_325	957304095	0,01165	0,01286	1,104	1	1
319	kjg_325	957914985	0,02269	0,01329	0,586	1	1
320	kjg_325	983856993	0,02191	0,01407	0,643	1	1
321	kjg_325	983449794	0,04118	0,01547	0,376	1	1
322	kjg_325	969981335	0,00778	0,01607	2,066	.	1
323	kjg_325	910197959	0,04067	0,01965	0,483	1	1
324	kjg_325	953154013	0,01367	0,02183	1,596	.	1	1	.	.	.
325	kjg_325	980909336	0,02619	0,02475	0,945	1	1	1	.	.	.

326	kjg_325	970036601	0,04148	0,02584	0,623	1	1	1	.	.	.
327	kjg_325	981438132	0,01832	0,02948	1,61	.	.	1	.	.	.
328	kjg_325	979513348	0,01388	0,03038	2,189	.	.	1	1	.	.
329	kjg_325 980020223	0,00597	0,03132	5,247	.	.	.	1	.	.	.
330	kjg_325	970015051	0,01409	0,03281	2,328	.	.	1	1	.	.
331	kjg_325 977675006	0,01037	0,03309	3,19	.	.	.	1	.	.	.
332	kjg_325	976212460	0,0175	0,03382	1,932	.	.	1	1	.	.
333	kjg_325	964376417	0,11428	0,03434	0,301	1	1	1	1	.	.
334	kjg_325	971595221	0,02205	0,03691	1,674	.	.	1	1	.	.
335	kjg_325 953987635	0,01315	0,03934	2,991	.	.	.	1	.	.	.
336	kjg_325 970262016	0,0112	0,03953	3,53	.	.	.	1	.	.	.
337	kjg_325	981237110	0,01527	0,04505	2,95	1	.
338	kjg_325	980524914	0,0105	0,04584	4,368	1	.
339	kjg_325	919450444	0,05964	0,04678	0,784	1	1	1	1	1	.
340	kjg_325	943474133	0,05193	0,05146	0,991	1	1	1	1	1	1
341	kjg_325	951381640	0,05628	0,05166	0,918	1	1	1	1	1	1
342	kjg_325	959091188	0,07933	0,05252	0,662	1	1	1	1	1	1
343	kjg_325	922007861	0,10218	0,05253	0,514	1	1	1	1	1	1
344	kjg_325	971063580	0,01794	0,05537	3,087	1	1
345	kjg_325	960289579	0,0658	0,05543	0,842	1	1	1	1	1	1
346	kjg_325 951228788	0,02889	0,05574	1,93	.	.	.	1	1	1	1
347	kjg_325	969073935	0,05901	0,05745	0,974	1	1	1	1	1	1
348	kjg_325	870552742	0,02361	0,05954	2,522	1	1
349	kjg_325	961939577	0,03138	0,06272	1,999	1	1
350	kjg_325	916231105	0,027	0,06285	2,328	1	1
351	kjg_325	976861604	0,02577	0,0668	2,592	1
352	kjg_325	978661718	0,07013	0,0674	0,961	1	1	1	1	1	1
353	kjg_325	977293367	0,09102	0,0701	0,77	1	1	1	1	1	1
354	kjg_325	945741473	0,04975	0,07147	1,437	.	.	1	1	1	1
355	kjg_325	932581000	0,02708	0,07579	2,798	1
356	kjg_325	974690896	0,03562	0,08221	2,308	1
357	kjg_325	974340496	0,10352	0,09179	0,887	1	1	1	1	1	1
358	kjg_325	944446400	0,28003	0,09241	0,33	1	1	1	1	1	1
359	kjg_325	937564007	0,42971	0,09455	0,22	1	1	1	1	1	1
360	kjg_325	951120766	0,06764	0,10508	1,554	1	1
361	kjg_325	929055276	0,14379	0,12275	0,854	1	1	1	1	1	1
362	kjg_325	975783073	0,13015	0,1498	1,151	.	.	1	1	1	1
363	kjg_325	977049369	0,15815	0,1607	1,016	1	1	1	1	1	1
364	kjg_325	951141836	0,14008	0,17339	1,238	1	1
365	kjg_325 970989765	0,1838	0,20076	1,092	1	.	1	1	1	1	1
366	kjg_325	979913818	0,38608	0,22595	0,585	1	1	1	1	1	1
367	kjg_325 958944314	0,35818	0,37535	1,048	1	1	1	.	1	1	1
368	kjg_325	916830890	0,49052	0,424	0,864	1	1	1	1	1	1
369	kjg_340	974357887	0,00808	0,00046	0,056	1
370	kjg_340	981897005	0,03865	0,0014	0,036	1
371	kjg_340	980555631	0,01298	0,00341	0,263	1
372	kjg_340	870931832	0,12677	0,01063	0,084	1	1
373	kjg_340	965064338	0,00931	0,01749	1,878	.	1
374	kjg_340	977182409	0,13808	0,01868	0,135	1	1
375	kjg_340	979156464	0,01144	0,01924	1,681	.	1
376	kjg_340	982692520	0,03962	0,01968	0,497	1	1
377	kjg_340	976466519	0,01893	0,02228	1,177	1	1	1	.	.	.
378	kjg_340	980281051	0,01918	0,0276	1,439	.	1	1	.	.	.
379	kjg_340	961174171	0,04551	0,02786	0,612	1	1	1	.	.	.
380	kjg_340	965320989	0,03584	0,03199	0,893	1	1	1	1	.	.
381	kjg_340	964434670	0,14201	0,03204	0,226	1	1	1	1	.	.
382	kjg_340	920240712	0,03555	0,03354	0,943	1	1	1	1	.	.
383	kjg_340	974432420	0,01764	0,03993	2,264	.	.	1	1	.	.
384	kjg_340	971062851	0,01871	0,04335	2,317	.	.	.	1	1	.
385	kjg_340	980386864	0,14793	0,04406	0,298	1	1	1	1	1	.
386	kjg_340	952599852	0,16602	0,04595	0,277	1	1	1	1	1	.
387	kjg_340	945937513	0,20604	0,04862	0,236	1	1	1	1	1	.
388	kjg_340	970343385	0,09883	0,04985	0,504	1	1	1	1	1	.
389	kjg_340	933734781	0,00694	0,05067	7,304	1
390	kjg_340	951324361	0,11381	0,05697	0,501	1	1	1	1	1	1
391	kjg_340	956893070	0,01696	0,06075	3,583	1	1
392	kjg_340	970995560	0,01012	0,06513	6,434	1

393	kjg_340	976594177	0,13253	0,06633	0,5	1	1	1	1	1	1
394	kjg_340	982556104	0,01769	0,06889	3,894	1
395	kjg_340	974692244	0,06101	0,07251	1,189	1	1	1	1	1	1
396	kjg_340	980505618	0,05534	0,07339	1,326	.	.	1	1	1	1
397	kjg_340	980496597	0,03286	0,08055	2,451	1	1
398	kjg_340	933741443	0,24482	0,08082	0,33	1	1	1	1	1	1
399	kjg_340	984774346	0,05724	0,08924	1,559	.	.	.	1	1	1
400	kjg_340	976466748	0,06235	0,09199	1,475	.	.	1	1	1	1
401	kjg_340	967776793	0,03688	0,09529	2,584	1
402	kjg_340	951273880	0,18312	0,10125	0,553	1	1	1	1	1	1
403	kjg_340	952115146	0,25956	0,11043	0,425	1	1	1	1	1	1
404	kjg_340	976012089	0,04852	0,11644	2,4	1
405	kjg_340	979657781	0,09647	0,14952	1,55	.	.	.	1	1	1
406	kjg_340	950001593	0,22782	0,16524	0,725	1	1	1	1	1	1
407	kjg_340	932256738	0,13574	0,17726	1,306	.	.	1	1	1	1
408	kjg_340	980899756	0,16229	0,18201	1,121	1	1	1	1	1	1
409	kjg_340	980383105	0,2008	0,20438	1,018	1	1	1	1	1	1
410	kjg_340	929017668	0,58527	0,26699	0,456	1	1	1	1	1	1
411	kjg_340	920118895	0,25457	0,3087	1,213	.	.	1	1	1	1
412	kjg_340	966233753	0,22285	0,35041	1,572	1	1
413	kjg_340	948482789	0,22439	0,356	1,587	1	1
414	kjg_340	958324014	0,55732	0,40645	0,729	1	1	1	1	1	1
415	kjg_340	981660714	0,56028	0,47362	0,845	1	1	1	1	1	1

10. Vedlegg 2: Teknisk beskrivelse av rutine for tildeling av PRN til bedrifter og foretak i BOF.

10.1. Generelt

For en faglig beskrivelse av PRN vises til notatet jhe, 9. september 2005 **Om samordning av utvalg ved bruk av PRN-tall** (Fra side 9)

Rutine for tildeling av PRN (Permanent Random Numbers), faste tilfeldige tall, til bedrifter og foretak i BoF kjøres ukentlig, og PRN lagres permanent i oracle-basen BOF. PRN-systemet bruker egne tabeller ved siden av bedrifter og foretak. Koblingsnøkkel mot enheter i BoF er bedrifts_nr og foretaks_nr.

10.1.1. Databaseobjekter

DSBBASE.SEED_BEDRIFT	
SEED_BEDRIFT	VARCHAR2 (9)
PRN_DATO	DATE

DSBBASE.SEED_FORETAK	
SEED_FORETAK	VARCHAR2 (9)
PRN_DATO	DATE

DSBBASE.PRN_BEDRIFTER	
BEDRIFTS_NR	VARCHAR2 (11)
GYLDIG_FRA_DATO	DATE
FORETAKS_NR	VARCHAR2 (11)
STATUSKODE	VARCHAR2 (1)
ORGNR_BEDRIFT	VARCHAR2 (11)
ORGNR	VARCHAR2 (11)
PRN_BEDRIFT	NUMBER
PRN_FORETAK	NUMBER
PRN_TILDELT	DATE

DSBBASE.PRN_FORETAK	
FORETAKS_NR	VARCHAR2 (11)
ORGNR	VARCHAR2 (11)
PRN_FORETAK	NUMBER
PRN_TILDELT	DATE

Tabellen SEED_BEDRIFT inneholder kolonnene SEED_BEDRIFT og PRN_DATO. SEED_BEDRIFT er startverdi for generering av PRN. En ny rad lagres med sist genererte startverdi for neste kjøring. Raden med nyeste PRN_DATO hentes ut som startverdi ved tildeling av PRN. Det samme gjelder for tabellen SEED_FORETAK.

Tabellen PRN_FORETAK inneholder PRN for foretak, mens PRN_BEDRIFTER inneholder PRN for bedrifter.

Kolonna STATUSKODE i PRN_BEDRIFTER inneholder bedriftens status i BoF.

GYLDIG_FRA_DATO er bedriftens gyldighetsdato fra BoF.

STATUSKODE kan ha følgende verdier :

B = Ikke slettet

D = Slettet som dublett

F = Slettet for sammenslåing

S = Slettet

Det er også opprettet et databaseview (utsnitt) av tabellen PRN_FORETAK som inneholder statuskode og gyldighetsdato for statuskode for alle foretaka. Dette utsnittet inneholder ellers de samme kjennemerkene som PRN_FORETAK.

Kjennemerkene statuskode og statuskode_gdato oppdateres i sanntid i forhold til foretaka i BoF. Det vil si at verdien på kjennemerkene til enhver tid er den samme som for foretaket i BoF. Statuskode for bedriftene oppdateres en gang i uka sammen med tildeling av PRN.

Utsnittet (databaseviewet) heter VW_PRN_FORETAK og kan behandles på samme måte som PRN_FORETAK vha. SAS. Kodelista for statuskoder for foretak er den samme som for bedrifter.

10.1.2. UNIX-miljø

Tildeling av PRN kjører på UNIX-server Sarepta, og startes ved hjelp av crontab kl. 15.04 hver tirsdag. Crontab tilhører brukerident aru på unix. Cron-entry ser slik ut :

```
04 15 * * 2 /ssb/sarepta/a1/dsb/sbdsbbase/prog/prn/prnsrc > /dev/null 2>&1
```

Program - katalog er : /ssb/sarepta/a1/dsb/sbdsbbase/prog/prn/

Katalogen inneholder følgende filer :

-rw-r--r--	1	aru	dsbw	229	Oct	18	15:02	body.txt
-rwx-----	1	aru	dsbw	7203	Oct	17	14:57	prn_prod.sas
-rwx-----	1	aru	dsbw	2620	Oct	18	15:02	prnsrc

body.txt inneholder teksten som sendes i rapport - mailen fra kjøringa.

prn_prod.sas inneholder SAS-programmet som tildeler PRN.

prnsrc inneholder UNIX-scriptet som starter SAS-jobben.

10.1.3. prnsrc ser slik ut :

```
#!/bin/sh
#####
# Kjører programmet prn_prod.sas #
# #
# scriptet startes på sarepta #
# #
# Opprettet: 17.10.2005 (aru) #
# Endret: #
#####

HOME=/ssb/sarepta/h1/aru ; export HOME
DSB=/ssb/sarepta/a1/dsb ; export DSB
PROG_DIR=$DSB/sbdsbbase/prog/prn ; export PROG_DIR
PROG_FILE=$PROG_DIR/prn_prod.sas ; export PROG_FILE
LOGG_FILE=$HOME/prn_prod.log ; export LOGG_FILE
LST_FILE=$HOME/prn_prod.lst ; export LST_FILE
KJORE_RAPP=$HOME/prn_rapp.err ; export KJORE_RAPP
BODY=$PROG_DIR/body.txt ; export BODY
MAIL_MOTT='aru@ssb.no jhe@ssb.no' ; export MAIL_MOTT
ANTERR=

#####
# Setter Oracle miljø: #
#####

ORACLE_BASE=/u01/app/oracle ; export ORACLE_BASE
ORACLE_HOME=$ORACLE_BASE/product/8.1.7 ; export ORACLE_HOME
PATH=$ORACLE_HOME/bin:/local/bin:/usr/bin:/bin ; export PATH
ORA_NLS33=$ORACLE_HOME/ocommon/nls/admin/data ; export ORA_NLS33
NLS_LANG=norwegian norway.we8iso8859p1 ; export NLS_LANG
LD_LIBRARY_PATH=$ORACLE_HOME/lib ; export LD_LIBRARY_PATH
TNS_ADMIN=/local/oracle/net ; export TNS_ADMIN

#####
# Kjører SAS programmet #
#####
/local/bin/sas $PROG_FILE > /dev/null 2> $KJORE_RAPP

#####
# Sender SAS-logg i vedlegg på mail til MAIL_MOTT #
#####
if /usr/bin/test -s $LOGG_FILE
then
    ANTERR=`grep -c 'ERROR:' $LOGG_FILE`
    export ANTERR

    if [ $ANTERR -gt 0 ]
    then
        /local/bin/mutt -a $LOGG_FILE \
        -s "Logg (med feil!) for SAS program prn_prod.sas `date "+%d.%m.%Y"``" \
```

```
    $MAIL_MOTT < /dev/null
else
    /local/bin/mutt -a $LOGG_FIL -a $LST_FIL \
    -s "Logg (OK) for SAS program prn_prod.sas `date +%d.%m.%Y`" \
    -i $BODY \
    $MAIL_MOTT < /dev/null
fi

fi

exit 0
```

10.1.4. SAS-programmet prn_prod.

```
/* ***** */
/*
/*  Program som lager PRN-tall
/*  Johan Heldal 23 februar 2005
/*  Audun Rusti 17 oktober 2005
/*  Modifisert for lagring av data i BOF-basen
/*  Automatisk ubetjent kjøring ukentlig
/*
/* ***** */
/* Setter ERRORABEND for å tvinge SAS i kne hvis noe går galt */
OPTIONS ERRORABEND;
OPTIONS LS=128 PS=max NODATE NOCENTER NONUMBER;
LIBNAME bof ORACLE USER=xxx PASSWORD=yyyyyy SCHEMA=dsbbase PATH='BOF';

PROC SQL;
  CREATE TABLE bof_bedr AS
  SELECT bedrifts_nr
        ,bedrifts_nr_gdato AS gyldig_fra_dato
        ,foretaks_nr
        ,statuskode
        ,orgnr AS orgnr_bedrift
  FROM bof.ssb_bedrift;
QUIT;
PROC SQL;
  CREATE TABLE bof_frtk AS
  SELECT foretaks_nr
        ,orgnr
  FROM bof.ssb_foretak;
QUIT;
PROC SORT DATA=bof_bedr;
  BY foretaks_nr;
PROC SORT DATA=bof_frtk;
  BY foretaks_nr;
RUN;

DATA bof_bedrifter;
  MERGE bof_bedr (IN=bed) bof_frtk (IN=FOR);
  BY foretaks_nr;
  IF bed THEN OUTPUT;
RUN;
/*
Ser først på foretak.
Selekterer ut en fil med en linje per foretak.
*/
PROC SORT DATA=bof_frtk OUT=foretaks_nr;
  BY foretaks_nr;
RUN;
/*
Henter inn foretak som har fått PRN fra før.
Sorterer dem stigende på foretaks_nr (eller orgnr).
Kobler dem til "foretaks_nr" og markerer dem der.
Sorterer slik at foretak som alt har fått PRN kommer først.
*/
DATA foretak_med_PRN;
  SET BOF.PRN_foretak;      * Datasett med tidligere genererte PRN_foretak. ;
  KEEP foretaks_nr orgnr PRN_foretak;
RUN;
PROC SORT DATA=foretak_med_PRN;
  BY foretaks_nr;
RUN;

DATA foretaks_nr_2;
```

```

MERGE foretaks_nr (IN=in1) foretak_med_PRN (IN=in2);
  BY foretaks_nr;
  IF in2 THEN harprn = 1; ELSE harprn=0; * Markerer enheter som har PRN fra
før ;
PROC SORT DATA=foretaks_nr_2; * og sorterer slik at de kommer
først. ;
  BY DESCENDING harprn foretaks_nr;
RUN;
/*
Setter PRN_foretak på nye enheter.
*/
TITLE1 'Setter PRN_foretak på nye enheter.';
%GLOBAL SEED_FORETAK;
* Henter inn startverdi for generering av nye PRN. ;
* Disse startverdiene er lagret på slutten av forrige oppdatering. ;
PROC SQL;
  CREATE TABLE seed_f AS
  SELECT seed_foretak FROM bof.seed_foretak
  WHERE prn_dato = (SELECT MAX(prn_dato) FROM bof.seed_foretak);
QUIT;
DATA seed_foretak;
  SET seed_f;
  CALL SYMPUT('SEED_FORETAK',seed_foretak); * Kopierer startverdi til
makrovariabel. ;
RUN;
%PUT SEED_FORETAK = &SEED_FORETAK;
DATA PRN_foretak;
  SET foretaks_nr_2 END=siste;
  BY DESCENDING harprn foretaks_nr;
  IF _N_=1 THEN seed_foretak = &SEED_FORETAK;
  IF harprn=0 THEN
    DO;
      CALL RANUNI(seed_foretak,PRN_foretak);
      tildelt + 1;
      RETAIN tildelt;
    END;
  IF siste THEN
    DO;
      CALL SYMPUT('SEED_FORETAK',seed_foretak);
      PUT 'Antall nye PRN = ' tildelt;
      END;
    RETAIN seed_foretak;
    KEEP foretaks_nr orgnr PRN_foretak;
RUN;
%PUT SEED_FORETAK = &SEED_FORETAK;
PROC SORT DATA=PRN_foretak;
  BY foretaks_nr;
RUN;

PROC SQL;
  CREATE TABLE nye AS
  SELECT * FROM prn_foretak
  WHERE foretaks_nr NOT IN
  (SELECT foretaks_nr FROM foretak_med_PRN);
QUIT;
DATA nye;
  SET nye;
  FORMAT prn_tildelt DATETIME20.;
  prn_tildelt = DHMS(DATE(),0,0,0);
RUN;

PROC APPEND BASE=bof.PRN_foretak DATA=nye;
RUN;

PROC PRINT DATA=nye (OBS=50);

```



```

RUN;
PROC CONTENTS DATA=bof.PRN_foretak;
RUN;

/*
Lagrer sist genererte startverdi til bruk neste gang.
*/
DATA seed_f;
    SET seed_foretak;
    seed_foretak = &SEED_FORETAK;
    FORMAT prn_dato DATETIME20.;
    prn_dato = DHMS( DATE(), 0, 0, 0 );
RUN;
PROC PRINT DATA=seed_f;
RUN;
%PUT SEED_FORETAK = &SEED_FORETAK;

PROC SQL;
    INSERT INTO bof.seed_foretak
    SELECT * FROM seed_f;
QUIT;

PROC PRINT DATA=bof.seed_foretak;
RUN;

/*
Bedrifter.
Sorterer bof_bedrifter etter bedriftsnummer.
*/
PROC SORT DATA=bof_bedrifter;
    BY bedrifts_nr;
RUN;

/*
Henter inn bedrifter som allerede har fått PRN.
Sorterer dem stigende på bedriftsnummer (eller orgnr_bedrift).
Kobler dem til "bof_bedrifter" og markerer dem der.
Sorterer slik at foretak som alt har fått PRN kommer først.
*/
DATA bedrifter_med_PRN;
    SET BOF.prn_bedrifter;
    KEEP foretaks_nr orgnr orgnr_bedrift bedrifts_nr PRN_bedrift;
RUN;
PROC SORT DATA=bedrifter_med_prn;
    BY bedrifts_nr;
RUN;
DATA bedrifter;
    MERGE bof_bedrifter (IN=innr) bedrifter_med_prn (IN=inbedr);
    BY bedrifts_nr;
    IF inbedr THEN harprn=1; ELSE harprn=0;
RUN;
PROC SORT DATA=bedrifter;
    BY DESCENDING harprn bedrifts_nr;
RUN;

/*
Setter PRN_bedrift på nye enheter;
*/
TITLE1 'Setter PRN_bedrift på nye enheter';
%GLOBAL SEED_BEDR;
* Henter inn startverdi for generering av nye PRN. ;
* Disse startverdiene er lagret på slutten av forrige oppdatering. ;

PROC SQL;
    CREATE TABLE seed_b AS
    SELECT seed_bedrift FROM bof.seed_bedrift

```

```

WHERE prn_dato = (SELECT MAX(prn_dato) FROM bof.seed_bedrift);
QUIT;
DATA seed_bedrift;
SET seed_b;
CALL SYMPUT('SEED_bedr',seed_bedrift);
RUN;
%PUT SEED_BEDR = &SEED_BEDR;
DATA bedrifter_2;
SET bedrifter END=siste;
BY DESCENDING harprn bedrifts_nr;
IF _N_=1 THEN seed_bedrift = &SEED_BEDR;
IF harprn=0 THEN
DO;
CALL RANUNI(seed_bedrift,PRN_bedrift);
tildelt + 1;
RETAIN tildelt;
END;
IF siste THEN
DO;
CALL SYMPUT('SEED_BEDR',seed_bedrift);
PUT 'Antall nye PRN = ' tildelt;
END;
RETAIN seed_bedrift;
KEEP foretaks_nr orgnr bedrifts_nr orgnr_bedrift PRN_bedrift;
RUN;
%PUT SEED_BEDR = &SEED_BEDR;
/*
Lagrer sist genererte startverdi til bruk neste gang.
*/
DATA seed_b;
SET seed_bedrift;
seed_bedrift = &SEED_BEDR;
FORMAT prn_dato DATETIME20.;
prn_dato = DHMS(DATE(),0,0,0);
RUN;
PROC PRINT DATA=seed_b;
RUN;
%PUT SEED_BEDRIFT = &SEED_BEDR;

PROC SQL;
INSERT INTO bof.seed_bedrift
SELECT * FROM seed_b;
QUIT;

PROC PRINT DATA=bof.seed_bedrift;
RUN;
/*
Kobler PRN_foretak og bedrifter_2 sammen til en fil.
*/
PROC SORT DATA=PRN_foretak;
BY foretaks_nr;
PROC SORT DATA=bedrifter_2;
BY foretaks_nr bedrifts_nr;
RUN;
DATA PRN_bedrift;
MERGE bedrifter_2 PRN_foretak;
BY foretaks_nr;
RUN;
PROC SORT DATA=PRN_bedrift;
BY foretaks_nr bedrifts_nr;
RUN;
PROC SORT DATA=bof_bedrifter;
BY foretaks_nr bedrifts_nr;
RUN;
DATA PRN_bedrifter;

```

```

MERGE bof_bedrifter PRN_bedrift;
  BY foretaks_nr bedrifts_nr;
RUN;

PROC SQL;
  CREATE TABLE nye AS
  SELECT * FROM prn_bedrifter
  WHERE bedrifts_nr NOT IN
    (SELECT bedrifts_nr FROM bedrifter_med_PRN);
QUIT;

DATA nye;
  SET nye;
  FORMAT prn_tildelt DATETIME20.;
  prn_tildelt = DHMS(DATE(), 0, 0, 0);
RUN;

PROC APPEND BASE=bof.PRN_bedrifter DATA=nye;
RUN;

PROC PRINT DATA=nye (OBS=50);
RUN;

PROC CONTENTS DATA=bof.PRN_bedrifter;
RUN;

```

10.2. Uttak av PRN_BEDRIFTER til SAS-ds

Oracle tabellene kan hentes inn som SAS-datasett dersom bruker har tilgang til BoF-basen.

Tilordning av LIBREF i SAS :

```
LIBNAME bof ORACLE USER=xxx PASSWORD=yyyyyy SCHEMA=dsbbase PATH='BOF';
```

DATA-steg som leser oracle-tabellen

```

DATA bedrifter_med_PRN;
  SET BOF.prn_bedrifter;
  KEEP foretaks_nr orgnr orgnr_bedrift bedrifts_nr PRN_bedrift;
RUN;

```

Eller uttak vha. PROC SQL i SAS :

```

PROC SQL;
  CREATE table bedrifter_med_prn as
  SELECT foretaks_nr
         ,orgnr
         ,orgnr_bedrift
         ,bedrifts_nr
         ,PRN_bedrift
  FROM bof.prn_bedrifter;
QUIT;

```

Tilsvarende uttak kan gjøres for de andre databaseobjektene beskrevet i kapittel 2 Databaseobjekter. PRN kan så kobles til populasjonene før kjøring av trekkeprogram.

10.3. Uttak av VW_PRN_FORETAK til SAS-ds

```
DATA vw_prn_foretak;
  SET BOF.vw_prn_foretak;
RUN;
```

Resultatet av dette datasteget vil fra SAS-explorer se slik ut.

SAS windows NT, Statistisk sentralbyrå - [VIEWTABLE: Workunix.Vw_prn_foretak]

	FORETAKS_N	ORGNR	PRN_FORETAK	PRN_TILDELT	STATUSKODE	STATUSKODE_GDATO
1	00000175		0.9967940077	28FEB2005:00:00:00	S	01JAN1979:00:00:00
2	00000183		0.3713485787	28FEB2005:00:00:00	S	26JAN1983:00:00:00
3	00000264		0.8262678123	28FEB2005:00:00:00	S	25MAY1981:00:00:00
4	00000639	862715152	0.986828194	28FEB2005:00:00:00	S	22MAR1999:00:00:00
5	00001309	854181262	0.2896921906	28FEB2005:00:00:00	S	03APR2002:00:00:00
6	00001341		0.1103826426	28FEB2005:00:00:00	S	01MAR1990:00:00:00
7	00001988		0.751067466	28FEB2005:00:00:00	S	01JAN1982:00:00:00
8	00002631	970085939	0.3435434375	28FEB2005:00:00:00	S	01JUN1994:00:00:00
9	00003875		0.0748210117	28FEB2005:00:00:00	S	01JAN1987:00:00:00
10	00004413		0.2159462558	28FEB2005:00:00:00	S	31DEC1992:00:00:00
11	00006033		0.8912547705	28FEB2005:00:00:00	S	01JAN1989:00:00:00
12	00006408	929026829	0.7048712707	28FEB2005:00:00:00	F	21MAR2000:00:00:00
13	00007536		0.3811440041	28FEB2005:00:00:00	S	01JAN1992:00:00:00
14	00008702	913883284	0.044030725	28FEB2005:00:00:00	B	31DEC1965:00:00:00
15	00008915		0.0386023456	28FEB2005:00:00:00	S	12NOV1981:00:00:00
16	00009253		0.7427101004	28FEB2005:00:00:00	S	01JAN1984:00:00:00
17	00009423		0.1275724602	28FEB2005:00:00:00	S	16DEC1980:00:00:00
18	00010200	947025430	0.187017101	28FEB2005:00:00:00	S	22AUG1997:00:00:00
19	00010324		0.7318835988	28FEB2005:00:00:00	S	01JAN1987:00:00:00
20	00011665		0.7238501705	28FEB2005:00:00:00	S	31DEC1979:00:00:00
21	00011754		0.4920089899	28FEB2005:00:00:00	S	01JAN1979:00:00:00
22	00011770		0.4944356165	28FEB2005:00:00:00	S	31DEC1981:00:00:00
23	00012092		0.6585467144	28FEB2005:00:00:00	S	15APR1985:00:00:00
24	00012211		0.3447646123	28FEB2005:00:00:00	S	01JAN1985:00:00:00
25	00012467	937674287	0.2973016548	28FEB2005:00:00:00	B	31DEC1965:00:00:00
26	00012637	970461175	0.2479286321	28FEB2005:00:00:00	B	30JAN1997:00:00:00
27	00012785		0.181345441	28FEB2005:00:00:00	S	01JUL1990:00:00:00
28	00013595	974908298	0.0304395901	28FEB2005:00:00:00	S	31DEC1996:00:00:00
29	00014400	954012212	0.7946584503	28FEB2005:00:00:00	B	30JAN1997:00:00:00
30	00014753		0.0295132483	28FEB2005:00:00:00	S	01JAN1983:00:00:00
31	00015016		0.2789733029	28FEB2005:00:00:00	S	01JUL1981:00:00:00
32	00015067		0.6541901144	28FEB2005:00:00:00	S	01JAN1980:00:00:00

De sist utgitte publikasjonene i serien Notater

- | | | | |
|---------|--|---------|---|
| 2006/1 | S. Abonyo og T. Hagen:
Tidsbruksundersøkelse - hvor lang tid bruker oppgavegiver på rapportering til kvartalsvis lønnsstatistikk. 24s. | 2006/14 | K.A. Kjesbu: Dokumentasjon av tidsseriebase for FoU-data. 47s. |
| 2006/2 | H. Hungnes: Hvitevarer 2006. Modell og prognose. 12s. | 2006/15 | B.O. Lagerstrøm og M. Høstmark: Kultur- og mediebruksundersøkelsen 2004. Dokumentasjonsrapport. 55s. |
| 2006/3 | O. Villund: Evaluering av omkodingen fra stillingskode til yrkeskode i Statens sentrale tjenestemannsregister (SST). 26s. | 2006/16 | H. Skullerud: Metanutslipp fra norske avfallsfyllinger. Reviderte beregninger av deponert avfall 1945 - 2004*. 15s. |
| 2006/4 | S.W. Bogen: Håndbok for rapportering av regnskapsdata for helseforetak og regionale helseforetak 2005. 59s. | 2006/17 | S.K. Boateng og S. Ferstad: Dokumentasjonsnotat for FylkesKOSTRA videregående opplæring. Publisering av 2004-tallene. 312.s |
| 2006/5 | T.A. Steinset, H. Brenna, L. Solheim og J.E. Wålberg: Dokumentasjon av Landbruksundersøkelsen 2004. 75s. | 2006/18 | K.I. Bøe, S. Lien og Ø. Sivertstøl: Fd-Trygd. Dokumentasjonsrapport. Demografi revidert. 1992-2003. 130s. |
| 2006/6 | P.E. Tønjum: Teknisk dokumentasjon av FAME-rutiner for indikatorberegningene i utenriksregnskapet (UR). 46s. | 2006/19 | A. Holmøy og B.O. Lagerstrøm: Interkommunalt legevaktsamarbeid - en forundersøkelse. 19s. |
| 2006/7 | L. Andreassen og G.H Bjertnæs: Tallfesting av faktoreterspørsel i MSG6. 31s. | 2006/20 | H. Tønseth: Årsrapport for kontaktutvalget for helse- og sosialstatistikk. 19s. |
| 2006/8 | T. Nygård Evensen: Oljenæringen - dokumentasjon av kilder, beregninger mv. (SKA-prosjekt). Rapport fra prosjektgruppen. 68s. | 2006/21 | D. Gronna og S. Todsen: Nasjonalregnskap: Beregning av olje- og gassnæringene. 31s. |
| 2006/9 | M. Aasgaard Walle, S. Mæland, A.B. Dahle, T. Meggison og L. Høgseth: Oljenæringen - dokumentasjon av kilder, beregninger mv. (SKA-prosjekt). Vedlegg til prosjektrapport. 53s. | 2006/22 | D. Gronna, S. Todsen og K. Erlandsen Kolshus: Beregning av olje- og gassnæringene i KNR. 24s. |
| 2006/10 | O. Villund: Klassifisering ved hjelp av tekst - noen resultater fra yrkeskodingen i Arbeidskraftundersøkelsen. 31s. | 2006/23 | T. Tveiekrem Sæter og I. T. Holmen: Prisindeks for bilutleie. 31s. |
| 2006/11 | E.C. Rauan og R. Nyggård Johnsen: Forventningsindikator - Konsumprisene. November 2005 - mai 2006. 18s. | 2006/24 | Ø. Linnestad og G.M. Molseth: Forprosjekt "Godstransport på kysten". 66s. |
| 2006/12 | S. Lien og Ø. Sivertstøl: Veier ut av langtidsmottak av sosialhjelp. 47s. | 2006/25 | K. Loe Hansen: Indikatorer på kjemikalieområdet- Risiko for skade på helse og miljø grunnet bruk av kjemiske stoffer. 46s. |
| 2006/13 | M. Hansen-Møllerud, A. Kalvøy, G. M. Pilskog og A-H. Sølverud: Informasjonssamfunnet 2005. 49s. | 2006/26 | A. Akselsen, S. Lien og Ø. Sivertstøl: FD-Trygd. Variabelliste. 58s. |
| | | 2006/28 | C. Nordseth og Ø. Sivertstøl: FD - Trygd. Dokumentasjonsrapport. Fødsels- og sykepenges, 1992-2003. 134s. |