

Liv Tove Hytjan

**Lønnsstatistikk for ansatte i privat undervisning
2009**

Klargjøring og analyse

© Statistisk sentralbyrå, november 2010 Ved bruk av materiale fra denne publikasjonen skal Statistisk sentralbyrå oppgis som kilde.	Standardtegn i tabeller	Symbol
ISBN 978-82-537-7971-3 Trykt versjon	Tall kan ikke forekomme	.
ISBN 978-82-537-7972-0 Elektronisk versjon	Oppgave mangler	...
ISSN 1891-5906	Oppgave mangler foreløpig	...
Emne: 06.90	Tall kan ikke offentliggjøres	:
Trykk: Statistisk sentralbyrå	Null	-
	Mindre enn 0,5 av den brukte enheten	0
	Mindre enn 0,05 av den brukte enheten	0,0
	Foreløpig tall	*
	Brudd i den loddrette serien	—
	Brudd i den vannrette serien	
	Desimaltegn	,

Forord

Notatet dokumenterer og beskriver prosessen med å klargjøre, kvalitetssikre, analysere og utarbeide lønnsstatistikken for ansatte i privat undervisning for 2009. Statistikken gir en oversikt over lønnsnivået og endringer i lønn per år for heltidsansatte. Statistikken, i sin nåværende form ble produsert første gang i 1997, og er sammenlignbar tilbake til dette året

Sammendrag

I lønnsstatistikk for privat undervisning benyttes et opplegg kontroll- og revisjonsrutiner som i utgangspunktet er standardiserte for alle lønnsstatistikkene.

Effekten av å kontrollere/revidere rådatamaterialet er i de fleste tilfeller betydelig, spesielt innen enkelte yrkesgrupper. En sammenligning mellom det ferdig reviderte datamaterialet og rådata fra oppgavegiverne viste klare forskjeller både med tanke på gjennomsnittsverdier og standardavvik i materialet. Maskinelt gjennomgås både mikro- og makrobaserte kontroller, og det er makrorutinene som det brukes mest tid og ressurser på i kvalitetssikringen.

Av de innkomne observasjonene i 2009 var det om lag 96 prosent som ble godkjent til frigivning.

Populasjonen utgjør 728 foretak med 15 479 sysselsatte per 3. kvartal 2009. Blant foretakene med 50 eller flere ansatte trekkes alle foretakene (fulltelling), blant foretakene med mellom 13 og 49 ansatte trekkes halvparten og for de foretakene med mellom 5 og 12 ansatte trekkes 20 prosent som kommer med i utvalget.

Innhold

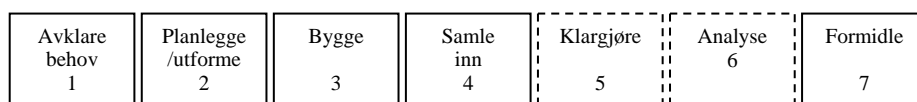
Forord	3
Sammendrag	4
1. Innledning	6
2. Populasjon og utvalg	6
3. Datafangsten	6
3.1. Datainnfang	6
3.2. Brutto inkommet	7
4. Behandling av data fra Oracle-basen til endelig produksjonsfil	8
4.1. Tilrettelegging av endelig produksjonsfil	8
5. Effekten av kontroll- og revisjonsrutinene	13
5.1. Arbeidstid.....	13
5.2. Lønnsvariablene	14
Litteraturliste	17
Vedlegg A: Prosesdiagram lønnsstatistikk	18
Tabellregister	19

1. Innledning

Notatet er i utgangspunktet todelt, der første del tar for seg datafangsten til lønnsstatistikken, mens del 2 tar for seg kontroll- og revisjonsopplegget og den effekt dette har. Årsaken til denne todelingen er at datafangst og kontroll- og revisjonsrutiner henger tett sammen, og at de til en viss grad er betinget av hverandre. Dvs. kvaliteten på datafangsten virker inn på hvordan kontroll- og revisjonen skal legges opp, og motsatt.

I lønnsstatistikken er det anledning til å svare enten elektronisk (Altinn eller på diskett/CD) eller på papirskjema. Bearbeidingen av data frem til ferdig reviderte tall kan deles inn i fire trinn: Mottak, registrering, innlesning til Oracle-tabell, og etablering av produksjonsfil. Dataene blir kontrollert og revidert på hvert av disse trinnene, men dette notatet konsentrerer seg, i tillegg til datafangsten, om den siste delen kontroll- og revisjonsopplegget. Den skjer etter nedlesing av data fra Oracle-tabellen ved etablering av endelig produksjonsfil (se figur vedlegg 1).

Statistisk sentralbyrås virksomhetsmodell for standardiserte arbeidsprosesser i statistikkproduksjonen er en modell med 7 statistikkprosesser som bygger på hverandre:



Dette notatet beskriver prosess 5 og 6, klargjøring og analysering av datagrunnlaget.

2. Populasjon og utvalg

Det trekkes et utvalg av foretak fra populasjonen i Bedrifts- og foretaksregisteret (BoF), (Beyrer, Bråten, Kristoffersen og Tyholt, Interne notater 2009/20). Foretakene blir stratifisert etter antall sysselsatte. Det er totaltelling blant foretak med sysselsetting over en viss grense, mens trekkeandelene er mindre i strata med færre sysselsatte. Foretak med mindre enn fem sysselsatte trekkes ikke. Person-utvalget består av alle sysselsatte i bedriftene i de uttrukne foretakene.

3. Datafangsten

Datainngangen per 1. oktober 2009 var god, og endelig status før frigivning av statistikk viste at om lag 96 prosent av de trukne foretakene var kommet inn.

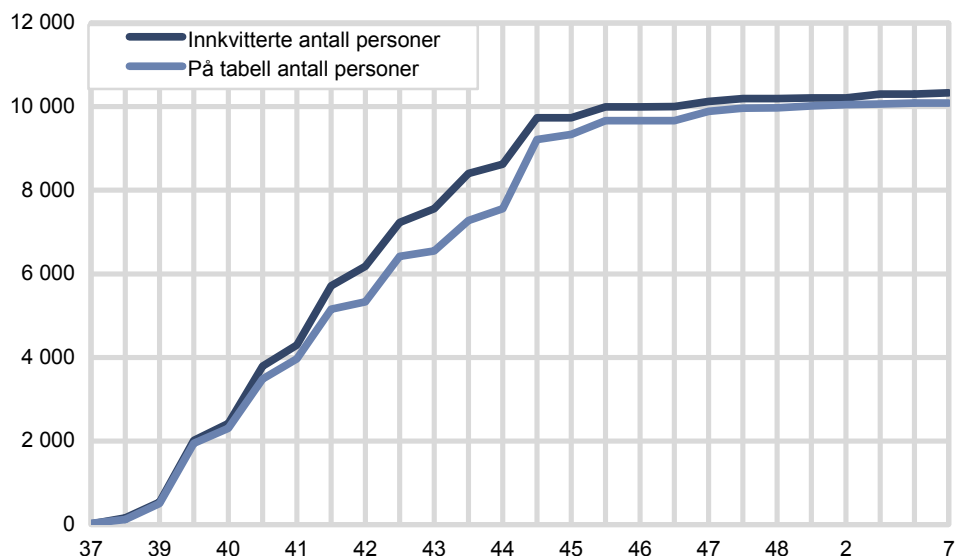
3.1. Datainngang

Datafangsten per 1. oktober 2009

Dato	Hendelse	Antall foretak (trekkeenheter)
18.09.09 (uke 38)	Utsending	312
10.10.09 (uke 41)	Svarfrist	150 (48 prosent)
20.10.09 (uke 47)	Vedtak om tvangsmulkt	100
30.10.09 (uke 49)	Siste frist for å unngå tvangsmulkt	
03.03.10 (uke 9)	Frigivning av statistikk	293 (96 prosent)
Antall fritak gitt	Ingen	

Datainngangen for privat undervisning ved frist uke 41 var om lag 48 prosent av foretakene, 47 prosent av bedriftene og 41 prosent av alle ansatte i næringen innkvittert. Figur 1 viser datainngangen fra uke 37 til uke 7.

Figur 1. Datainngang i privat undervisning. Antall innkvitterte og lest inn i Oracle-tabellen, 2009



Som vist på figuren er det for 2009 et lite avvik mellom innkvitterte og innleste observasjoner.

3.2. Brutto inntekst

I tabell 1 finner vi endelig antall foretak, bedrifter og sysselsatte som er innkvittert i privat undervisning i 2008 og 2009. Andelen innkvitterte foretak og i 2009 var om lag på samme nivå som i 2008.

Tabell 1. Endelig antall foretak, bedrifter og sysselsatte innkvittert i privat undervisning i 2009 og prosentvis andel innkvittert i 2008 og 2009

Enhet	Datainngang		2008 Prosent Innkvittert
	2009		
	Antall Innkvittert	Prosent Innkvittert	
Foretak	299	96	97
Bedrifter	410	93	92
Sysselsatte	10 284	95	97

Tabell 2 er en oversikt over andelen som benytter seg av de ulike innrapporteringsformatene i perioden 2007 til 2009. Den klare hovedvekten ligger på innrapportering via Altinn, mens innrapportering via kravspesifikasjon (CD eller diskett) er avtagende.

Tabell 2. Andel observasjoner etter innrapporteringsformat i perioden 2007 til 2009

KODE	Prosent		
	2009	2008	2007
I alt	100,0%	100,0%	100,0%
Excelfiler	0,5%	0,9%	7,9%
Altinn	66,5%	49,7%	46,8%
Kravspesifikasjon	31,5%	44,2%	34,2%
Manuelt registrert	0,0%	2,0%	2,5%
Optisk lest	1,5%	3,2%	8,6%

4. Behandling av data fra Oracle-basen til endelig produksjonsfil

Å innhente lønnsinformasjon på individnivå fra næringslivet har vist seg å medføre at alle opplysninger som kommer inn ikke nødvendigvis er korrekte. Det kan være ulike typer feil i materialet, fra relativt ubetydelige unøyaktigheter, til mer graverende tilfeller som medfører at observasjonen må forkastes. Det er derfor utarbeidet et system med kontroll- og feilrettingsrutiner for lønnsstatistikken, der hovedmålsettingen er å avdekke eventuell systematikk i typen feil og mangler i det som blir innrapportert. I tillegg er det en målsetting å minimere muligheten for at enkeltobservasjoner med åpenbare feil og mangler skal få uønsket innvirkning i statistikken. Systemet er bygget opp med enkle logiske mikrokontroller først (tabell 3-9), før det kontrolleres mer på makronivå etterpå (bl.a. tabulering). Det er arbeidet på makronivå som det klart brukes mest tid på, og det er der vi oppdager systematiske feil/skjevheter i datamaterialet. Allerede ved innlesing av rådata, enten det måtte være fra Altinn-portalen, fra filer basert på kravspesifikasjonen for elektronisk innrapportering, eller fra optisk leste skjema blir det kontrollert for logiske feil og mangler. Dette skjer litt forenklet sagt ved at observasjoner merkes i Oracle-basen, gradert etter betydningen av feilen/mangelen. Ved betydelige feil og mangler blir oppgavegiveren kontaktet med det samme, og oppfordret om å levere nye data.

Denne dokumentasjonen tar for seg kvalitetssikringen av datamaterialet *etter* at den innledende mikro-kvalitetssikringen er gjennomført i databasen. Det er med andre ord allerede ”ryddet opp” i en del av de mest åpenbare feil og mangler.

Tilrettelegging av data fra Oracle-basen fram til endelig produksjonsfil gjøres ved hjelp av fastlagte SAS-rutiner. Rutinene er lagt opp slik at de foregår i 5 trinn eller operasjoner, som i stor grad er standardisert og tilnærmet like fra næringsområde til næringsområde. Opplegget består av:

- Rutiner for *merking* av observasjoner med mulige feil
- Rutiner for *justering/korrigerings* av observasjoner med opplagte feil
- Rutiner for *å holde utenfor* observasjoner med opplagte feil

I korte trekk er rutinene bygget opp som følger:

Trinn 1: Nedlesing fra Oracle og tilrettelegging av sasfil. Observasjoner uten verdier på relevante variable tas ut.

Trinn 2: Grunnleggende kontroll og justering av bakgrunnsvariablene, arbeidstid og avlønningsperiode.

Trinn 3: Etablering av lønnsbegrepene (lønn per måned)

Trinn 4: Godkjenning av yrkeskoder

Trinn 5: Endelig godkjenning av observasjoner. Tar ut observasjoner med ekstremverdier i en eller flere lønnsvariable

Deretter følger trinn 6, der materialet vektet og sjekkes opp mot populasjonen (BoF). Se vedlegg 1 for skjematisk oversikt over disse rutinene.

4.1. Tilrettelegging av endelig produksjonsfil

Program 1 nedlesing fra Oracle-basen

Tabell 3 viser gyldige og ugyldige fødselsnummer for ansatte i privat undervisning i 2008 og 2009. D-nummer er fødselsnummer til utenlandske arbeidere. Alle observasjoner tas med videre i statistikkgrunnet, uavhengig av om de har gyldig fødselsnummer eller ikke.

Tabell 3 Gyldige og ugyldige fødselsnummer. 2008 og 2009

Fødselsnummer	D- nummer	Antall	
		2008	2009
I alt		10 223	10 271
Ugyldige fødselsnummer og Gyldige fødselsnummer og	ikke d-nummer	15	11
Mangelfullt eller ugyldig fødselsnummer og	ikke d-nummer	10 091	10 131
Mangelfullt eller ugyldig fødselsnummer og	ugyldige d-nummer	-	-
Mangelfullt eller ugyldig fødselsnummer og	gyldig d-nummer	109	128
Mangelfullt eller ugyldig fødselsnummer og	ikke d-nummer	8	:

Program 2 arbeidstidsforhold

Program 2 kontrollerer og justerer lønnsperioder og arbeidstid. Kontrollene viser at enkelte observasjoner plasseres i feil lønnsperiode og må omplusses slik at lønnsopplysningene får riktig håndtering videre i kontrollene.

Alle ansatte har en avlønningsperiode:

Lønnsperiode 1: Månedslønn 12 ganger i året
 Lønnsperiode 2: 14-dagerslønn
 Lønnsperiode 3: Lønn hver uke
 Lønnsperiode 4: Månedslønn 12,5 ganger i året
 Lønnsperiode 5: Månedslønn 13 ganger i året
 Lønnsperiode 6: Timelønn
 Lønnsperiode 7: Provisjons- eller prosentlønn

Program 3 beregning av månedslønn

Dette programmet foretar først og fremst en beregning av fastlønn, bonus og uregelmessige tillegg. Det foretas også korreksjon og justering av overtidstimer og overtidsgodtgjørelse.

Månedslønnsbegrepet består av en summering av avtalt lønn, uregelmessige tillegg og bonus. For å beregne månedslønn må det derfor først gjøres en beregning av de øvrige lønnslementer som månedslønna bygger på:

Avtalt lønn

Avtalt lønn er beregnet lønn siste måned for alle, både fastlønte, timelønte og provisjonslønte. Dette for å få sammenlignbare størrelser uavhengig av lønnsperiode. Innrapportert lønn siste avlønningsperiode knyttet til variabel for månedslønn benyttes i utregningen av avtalt lønn:

Avlønnsperiode 1	=	Opgitt månedslønn siste periode
Avlønnsperiode 2	=	Opgitt månedslønn per 14-dager multiplisert med 2.165 (antall 14-dagers-perioder i en måned).
Avlønnsperiode 3	=	Opgitt månedslønn per uke multiplisert med 4.33 (antall uker i en måned).
Avlønnsperiode 4	=	Opgitt månedslønn per uke multiplisert med (12,5/12) .
Avlønnsperiode 5	=	Opgitt månedslønn per uke multiplisert med (13/12) .
Avlønnsperiode 6	=	Hovedregel er timesats (enten beregnet eller direkte oppgitt) multipliser med antall timer i en måned (avtalt arbeidstid per uke multiplisert med antall uker i en måned).
Avlønnsperiode 7	=	Hovedregel: (bonus hittil i år + bonus 4. kvartal i fjor)/12 + evt. oppgitt fastlønn siste periode. Bonusfeltene tømmes etter beregning av avtalt lønn.

Uregelmessige tillegg

Innhentes som aggregert størrelse fra 1. januar til tellingstidspunktet. Gjennomsnittsberegning per måned for denne perioden.

<i>Bonus</i>	Innhentes både for 4. kvartal i fjor, samt som aggregert størrelse fra 1. januar til tellingstidspunktet. Gjennomsnittsberegning per måned for hele denne 12-måned-perioden.
<i>Månedslønn</i>	Månedslønna framkommer ved en summering av de beregnede månedsverdiene for avtalt lønn, uregelmessige tillegg og bonus.
<i>Overtidsgodtgjørelse</i>	Inngår ikke i månedslønnsbegrepet. Innhentes som aggregert størrelse fra 1. januar til tellingstidspunktet. Gjennomsnittsberegning per måned for denne perioden.

For overtid gjøres en spesiell kontroll av forholdet mellom innrapporterte overtidstimer i forhold til overtidsgodtgjørelsen.

Tabell 4. Oversikt over overtidstimer og overtidbetaling

Status observasjoner	Frekvens	Prosent
I alt	10 271	100,0
Observasjoner uten overtidstimer eller betaling	9 397	91,5
Observasjoner med flere overtidstimer enn betaling	:	0,0
Observasjoner med verdier på betaling	842	8,2
og av disse mangler det timer på	31	0,3

På basis av denne kontrollen blir overtidsgodtgjørelsen justert på observasjoner som åpenbart bare har fått innrapportert *overtidstillegg*, og ikke hele godtgjørelsen, for utført overtidarbeid. Dette gjøres ved at det beregnes en lønn per time innen normal arbeidstid for den aktuelle observasjonen, som så *legges til* overtidsgodtgjørelsen per time. Alt dette gjøres maskinelt uten tidsbruk for saksbehandleren.

Program 4 yrkeskoding

I program 4 gjennomføres en logisk kontroll av gyldighet av innrapporterte yrkeskoder. Oppgavegiverne skal rapportere yrkeskoder i henhold til standard for yrkesklassifisering (styrk). I kvalitetssikringsarbeidet før frigivning brukes relativt mye tid på å kvalitetssikre yrkeskodingen. Blant annet sjekkes forholdet mellom yrkeskode og yrkesbetegnelse, og ved feil og mangler kontaktes oppgavegiverne i så stor grad som mulig. 1.siffer i yrkeskoden betegnes som yrkesfelt:

1. Lederyrker
2. Akademiske yrker
3. Høyskoleyrker
4. Kontor- og kundeserviceyrker
5. Salgs- og serviceyrker
6. Yrker innen jordbruk, skogbruk og fiske (ikke i bruk)
7. Håndverkere o.l.
8. Operatører og sjåfører
9. Yrker uten krav til utdanning

Innrapporteringa av yrkeskoder så slik ut i 2008 og 2009:

Tabell 5. Fordeling av observasjoner på 1-siffer styrk, 2008 og 2009

St1	2008		2009	
	Frekvens	Prosent	Frekvens	Prosent
I alt	10 223	100	10 271	100
1	622	6,1	602	5,9
2	4 146	40,6	4 212	41
3	2 906	28,4	3 028	29,5
4	674	6,6	671	6,5
5	1 109	10,9	1 082	10,5
6	42	0,4	27	0,3
7	36	0,3	32	0,3
8	34	0,3	13	0,1
9	581	5,7	552	5,4
X	73	0,7	52	0,5

Som en kontroll av innrapporterte yrkeskoder til lønnsstatistikken kobles det på 7-sifret yrkeskode fra AA-registeret (Arbeidstaker-/arbeidsgiverregisteret) med oppdatering nærmest mulig tellingstidspunktet. Det viser seg å være avvik mellom hva som ligger i AA-registeret og hva som rapporteres til lønnsstatistikken (se tabell 6).

Tabell 6. Fordeling av antall med 1-siffer styrk i lønnsstatistikken sammenlignet med AA-registeret

Lønns- statistikken	I alt	AA- registeret									
		1	2	3	4	5	6	7	8	9	X
I alt	10 271	693	3 429	2 744	568	980	33	60	14	1 497	253
1	602	467	58	38	5	7	:	:	:	16	9
2	4 212	131	2 964	437	26	73	:	3	4	430	142
3	3 028	48	317	2 114	57	135	3	13	6	281	54
4	671	31	39	53	454	22	:	:	:	63	7
5	1 082	14	45	77	16	702	:	21	:	181	24
6	27	:	:	:	:	:	21	:	:	4	:
7	32	:	:	4	:	4	:	17	:	:	5
8	13	:	:	5	:	:	:	:	:	:	5
9	552	:	3	9	8	33	5	:	:	487	6
X	52	:	:	6	:	4	:	:	:	34	:

Det er foreløpig et problem at mange ansatte ligger uten yrkeskode i AA-registeret. Mens hele 253 observasjoner ligger uten kode i AA-registeret, er det i lønnsstatistikken 52 observasjoner.

En direkte sammenligning av den 4-sifrede yrkeskoden fra den ferdige reviderte fila fra lønnsstatistikken og fra AA-registeret, viser at vel 7 200 observasjoner hadde identisk yrkeskode i disse to kildene.

Program 5 godkjenning av observasjoner

I program 5 lages en oversikt over godkjente observasjoner og observasjoner som forkastes i statistikken. Lønnsopplysninger har maksimums- og minimumsgrenser. Disse grensene er ikke blitt endret i 2009 i forhold til 2008. Hovedtankegangen bak dette kontrollopplegget er som følger:

- Det skal settes romslige grenser for å sikre at observasjoner med antatt korrekte verdier beholdes.
- Grensene skal likevel være satt slik at observasjoner med åpenbart ekstreme verdier, som med stor grad av sikkerhet er feilaktig, holdes utenfor datagrunnlaget.
- Grenseverdiene settes på grunnlag av en rekke forhold, som blant annet erfaring fra tidligere år, kommunikasjon med forhandlingsverden (for eksempel minstelønn innen de ulike tariffområdene) og oppgavegivere, samt statistiske mål på spredning. Grenseverdiene gjennomgås og justeres år for år.
- Underveis i kvalitetssikringsarbeidet sjekkes observasjoner som faller ut med feilkoder, i så stor grad som mulig.

Tabell 7. Antall observasjoner fordelt etter koder for grenseverdier (ok-koder) i 2009

Ok-kode	Frekvens	Prosent
I alt	10 271	100
Ok=0 - Godkjente observasjoner	9 680	94,2
Ok=1 - Lave verdier på lønnsvariabler	5	0,1
Ok=3 - Timelønnte med utilstrekkelig info om timer	1	0,0
Ok=4 - Alder missing	1	0,0
Ok=10 - Heltidsansatte med i alt<10 500	78	0,8
Ok=12 - Heltidsansatte, timelønnte med fmlonn over 65 000	1	0,0
Ok=14 - Heltidsansatte, st1=1 og i alt>200 000	2	0,0
Ok=15 - Heltidsansatte, st1=1 og i alt<20 000	2	0,0
Ok=16 - Heltidsansatte, st1 = 2,3 og i alt over 110 000	15	0,2
Ok=17 - Heltidsansatte, st1 = 4,5 og i alt over 75 000	5	0,1
Ok=18 - Heltidsansatte, st1>5 og i alt over 65 000	3	0,0
Ok=19 - Heltidsansatte, overtid (ot2) over 20 000	2	0,0
Ok=20 - Heltidsansatte, tillegg (tl) over 65 000	2	0,0
Ok=30 - Deltidsansatte med arbtid<=1	70	0,7
Ok=31 - Deltidsansatte, st1=1 og timefortjeneste >600	1	0,0
Ok=34 - Deltidsansatte, st1=1 og i alt<20 000	17	0,2
Ok=36 - Deltidsansatte, st1=2,3 og timefortjeneste<80	228	2,2
Ok=37 - Deltidsansatte, st1=4,5 og timefortjeneste>250	52	0,5
Ok=38 - Deltidsansatte, st1=4,5 og timefortjeneste<60	24	0,2
Ok=41 - Deltidsansatte og timefortjeneste>250	21	0,2
Ok=42 - Deltidsansatte og timefortjeneste<50	14	0,1
Ok=90 - Heltidsdubletter	28	0,3
Ok=91 - Deltidsdubletter	4	0,0
Ok=92 - Heltid/deltidsdubletter	15	0,2

Den vanligste ”feiltypen” er for lav lønn i forhold til arbeidstiden. Dette skyldes i mange tilfeller at foretak, som leverer elektronisk, ikke har ”renset” lønnsystemet sitt for personer, som ikke er ansatt på tellingstidspunktet. Disse har fått en eller annen form for godtgjørelse i løpet av året. Samlet sett holdes 94 prosent av observasjonene innen rammene av øvre og nedre grenseverdier.

Program 5b, etablering av heltidsekvivalenter

Inkludering av heltidsekvivalenter i de næringsspesifikke frigivingene er nytt av året, og program 5b regner ut ekvivalentvekter og ekvivalentlønsarter for disse.

Oppgavegivere som leverer data elektronisk ved å sende inn fil ifølge kravspesifikasjonen for innrapportering av lønnsstatistikk eller via Altinn rapporterer en kode for heltid eller deltid. Vi ser denne koden i sammenheng med rapportert avtalt arbeidstid og lønnsperiode, se pkt. 4.1.2. Vi kontrollerer også alle lønsarter for både heltids- og deltidsansatte i forhold til maksimums- og minimumsgrenser, se pkt 4.1.5.

Grenseverdiene sees også i sammenheng med hvilket yrke den ansatte har. Observasjoner som faller ut blir kontrollert for om mulig å kunne rettes opp eller bli bekreftet riktig slik at de kan godkjennes og benyttes i statistikken.

Følgende oversikt viser resultatet etter kontrollene vi gjør, altså hvor mange heltids- og deltidsansatte som ikke blir godkjent.

	Heltidsansatte	Deltidsansatte
Totale obs i næringen	5 890	4 381
Godkjente obs	5 768	3 739
Ikke godkjente obs	122	642
% ikke godkjente obs	0,0	14,7

Vi får en unik forståelse av heltidsekvivalenten for næringen i og med at vi beregner gjennomsnittlig antall avtalte timer per uke for heltidsansatte. Dette gjennomsnittet blir utgangspunktet for videre behandling og omregning av deltidsansattes lønn.

Det er dermed kun de deltidsansattes månedslønn vi regner om. Utgangspunktet for omregning er å multiplisere lønna til den enkelte deltidsansatte til hva den ville vært dersom han eller hun jobbet heltid.

Program 6 vekting

Programmet stratifiserer populasjonen etter antall ansatte i foretakene og beregner vekter. Vektemetoden er basert på justert etterstratifisert invers trekkesannsynlighet.

Tabell 8. Antall foretak i populasjon og utvalg og invers trekkesannsynlighet (Vekt)

Stratum	Antall			Vekt 2009
	Foretak i utvalget	Foretak i pop.	Ansatte i pop.	
8501	51	57	5 165	1,11765
8502	127	225	5 596	1,77165
8503	65	402	3 182	6,18462
8551	65	65	1 522	1,00000

Program 7 Vektjustering

Prinsippet for justering av vektene er at sum oppblåst antall i lønnsstatistikken skal reflektere sum antall ansatte fra BoF for hver næring (Grini, Notater 2003/74). Dette betyr at gjennomsnittlig lønn for et næringshovedområde skal reflektere et veid gjennomsnitt av hver av næringene som området består av. For å oppnå dette må hver av næringenes relative andel av den estimerte sysselsettinga i undersøkelsen, gjenspeile populasjonsandelene fra BoF. I tillegg må sammenhengen mellom de to siste årganger av lønnsstatistikken kontrolleres. Denne justeringen er spesielt viktig i områder med store lønnsforskjeller mellom næringer, der avvik vil få større betydning for det endelige lønnsnivået.

I privat undervisning har det for 2009 ikke vært nødvendig å justere vektene, da hver delnærings andel av oppblåst antall stemmer overens med BOF.

Frafall

Vi har to hovedtyper frafall, enhetsfracfall og partielt frafall. Enhetsfracfall defineres her som frafall av bedrifter, mens partielt frafall er frafall av personer. Enhetsfracfall skyldes at bedrifter ikke har innrapportert til oss, eller at det er feil i datamaterialet vi har fått innrapportert. Under enhetsfracfall av bedrifter kommer også de bedriftene vi har valgt å holde ute manuelt. Totalt sett var enhetsfracfallet på 7 foretak med 9 bedrifter i 2009.

Partielt frafall kan blant annet skyldes mangelfull innrapportering, feil i data eller at personer ikke er blitt innrapportert. Denne type frafall er omtalt i kapittel 4.1.5 hvor det er blitt satt minimums- og maksimumsgrenser for de ulike lønnsvariablene. I privat undervisning er de fleste observasjonene som ikke ble godkjent, deltidsansatte med arbeidstid mindre eller lik 1. Det partielle frafallet var totalt sett på mindre enn 3 prosent av observasjonene i 2009.

5. Effekten av kontroll- og revisjonsrutinene

Måling av effekten av kontrollopplegget gjøres her ved å sammenligne datagrunnlaget før og etter at alle sas-rutinene er kjørt. Dette kapitlet tar for seg hvilken virkning rutinene har på de mest sentrale variablene i lønnsstatistikken.

5.1. Arbeidstid

Arbeidstiden er grunnleggende bestemmende for to forhold: a) om en person er hel- eller deltidsansatt, og b) arbeidstiden benyttes for å beregne månedslønn for en representativ måned. Det legges derfor relativt mye ressurser inn i arbeidet med denne variabelen, og det er et absolutt krav i innrapporteringen at feltet må være utfylt. Mye av kontrollene går av den grunn ut på sjekke sammenheng mellom arbeidstid og lønn.

Tabell 9. Effekt av kontroll og revisjon av avtalt arbeidstid per uke

	Antall observasjoner		Gjennomsnittlig avtalt arbeidstid	
	Før revisjon	Etter revisjon	Før revisjon	Etter revisjon
I alt	10 271	9 507	29,6	30,1
Heltidsansatte	5 890	5 768	37,6	37,6
Deltidsansatte	4 380	3 739	17,8	18,5

Som det går fram av tabellen er det flere observasjoner på den ureviderte fila (rådata) enn på den reviderte fila (publiseringsfila). Grunnen er at alle ikke-godkjente observasjoner her er holdt utenfor etter kriteriene beskrevet i kapittel 4.1.5. I tillegg til dette er også observasjoner helt uten innrapportert arbeidstid holdt utenfor den ferdig reviderte fila. Disse observasjonene registreres som deltidsansatte inntil de sjaltes ut, derav nedgangen i antall deltidsansatte fra urevidert til revidert fil.

5.2. Lønnsvariablene

Responsvariablene i lønnsstatistikken er lønnsvariablene som frigis i den offisielle statistikken, og er derfor også de vi først og fremst ønsker å måle revisjonseffekten av. Dette gjelder først og fremst månedslønn, som består av avtalt lønn, uregelmessige tillegg og bonus. I tillegg kommer tall for overtidsgodtgjøring, men dette inngår ikke i månedslønna.

Revisjonseffekten på disse variablene defineres her som forskjellen mellom revidert og urevidert verdi, dvs. forskjellen på verdi på endelig publiseringsfil og verdi på rådatafil. Problemet i dette tilfellet er at responsvariablene ikke finnes direkte på rådatafila fordi de ikke innrapporteres direkte fra oppgavegiverne. De er avledede variable som beregnes i løpet av kontrollrutinene, på grunnlag av innrapporterte variable.

For å forenkle vurderingen av revisjonsmålingen tar vi bare for oss heltidsansatte i det følgende. Det er relativt sett få deltidsansatte i denne næringa, og i tillegg blir resultatene noe vanskeligere å tolke når lønna for de deltidsansatte skal inkluderes.

Først ser vi næringa som helhet, og tar med alle lønnsvariable, med verdier for gjennomsnitt, median, nedre og øvre kvartil, samt minimum og maksimum.

Tabell 10. Effekt av kontroll og revisjon av lønnsvariable. Heltidsansatte. Kroner

	Revidert	Urevidert	Avvik			
Antall	5 768	5 990	-222			
	Gjennomsnitt			Median		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	35 448	36 103	-655	33 458	33 445	13
Avtalt lønn	34 766	35 416	-650	33 317	33 233	84
Uregelmessige tillegg	437	444	7	0	0	0
Bonus	245	243	0	0	0	0
Overtid	284	285	-13	0	0	0
	1. kvartil			3. kvartil		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	29 167	29 150	17	39 196	39 183	13
Avtalt lønn	28 973	28 860	113	38 713	38 713	0
Uregelmessige tillegg	0	0	0	0	0	0
Bonus	0	0	0	0	0	0
Overtid	0	0	0	0	0	0
	Min			Max		
	Revidert	Urevidert	Avvik	Revidert	Urevidert	Avvik
Månedslønn	10 966	0	10 996	175 001	610 296	-435 295
Avtalt lønn	10 966	0	10 996	175 001	610 296	-435 295
Uregelmessige tillegg	0	0	0	55 653	79 239	-23 586
Bonus	0	0	0	57 608	57 608	0
Overtid	0	0	0	15 275	34 567	-19 292

Denne tabellen viser kun tall på aggregert nivå, dvs. næringshovedområdet som helhet. Kontroll- og revisjonsopplegget er imidlertid bygget opp og tilpasset til at den endelige lønnstatistikken skal utarbeides på et mer disaggregert nivå. Det er derfor tatt høyde for at verdi på lønnsvariablene vil variere etter gitte kjennemerker, f. eks. næring og yrke.

Ved å se på forskjellen mellom revidert og urevidert verdi for lønnsvariablene fordelt på næring og yrke, vil man få et bedre bilde av effekten av kontroll- og revisjonsopplegget.

For å forenkle tar vi i tabell 11 bare for oss månedslønn.

Tabell 11. Effekt av kontroll og revisjon av månedslønn etter næring og yrke. Antall og kroner

	Antall		Avvik	Revidert	Gjennomsnitt	
	Revidert	Urevidert			Urevidert	Avvik
I alt	5 768	5 990	-222	35 448	36 103	-655
Førskole- og grunnskoleundervisning	1 006	1 022	-16	30 846	30 972	-126
Av dette:						
Akademiske yrker	271	272	-1	32 374	32 580	-206
Høgskoleyrker	530	540	-10	30 149	30 433	-284
Undervisning på videregående skolenivå	1 497	1 615	-11	36 802	36 664	138
Av dette:						
Lederyrker	121	128	-7	49 353	49 676	-323
Akademiske yrker	556	621	-65	37 790	37 785	5
Høgskoleyrker	450	472	-22	34 496	34 442	54
Undervisning på universitets- og høgsolenivå	1 600	1 629	-29	39 388	39 665	-277
Av dette:						
Lederyrker	146	150	-4	51 080	51 281	-201
Akademiske yrker	1 152	1 171	-19	40 053	40 369	-316
Høgskoleyrker	180	183	-3	33 377	33 536	-159
Voksenopplæring og annen undervisning	1 665	1 724	-59	35 088	36 489	-1 401
Av dette:						
Lederyrker	194	196	-2	45 271	47 023	-1 752
Akademiske yrker	381	399	-18	38 437	38 947	-510
Høgskoleyrker	565	586	-23	34 686	36 061	-1 375

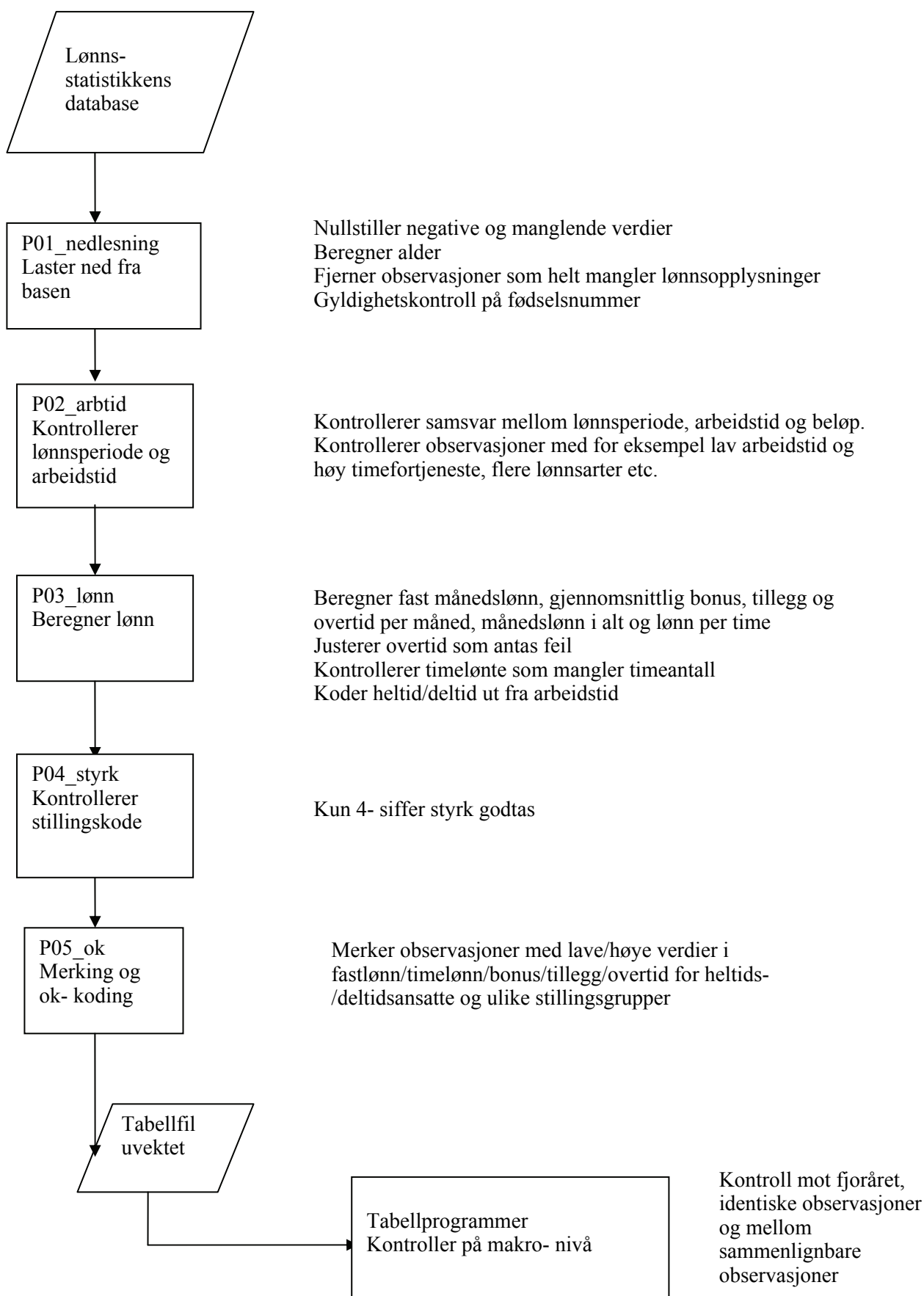
Antall observasjoner som får endret månedslønna i løpet av revisjonsprosessen er vel 200 stykker. Dette utgjør om lag 3 prosent av det totale antallet heltidsansatte.

Litteraturliste

Beyrer, Kristoffersen, Tyholt (2008): Lønnsstatistikk 2008. Etablering av populasjon og utvalg. Statistisk sentralbyrå, interne notater 2008/27.

Grini, Knut Håkon (2003): Lønnsstatistikk privat sektor 1997 – 2001. Dokumentasjon av utval og berekning av vekter. Statistisk sentralbyrå, notater 2003/74.

Vedlegg A: Prosesdiagram lønnsstatistikk



Tabellregister

1.	Endelig antall foretak, bedrifter og sysselsatte innkvittert i privat undervisning i 2009 og prosentvis andel innkvittert i 2008 og 2009	7
2.	Andel observasjoner etter innrapporteringsformat i perioden 2007 til 2009	7
3.	Gyldige og ugyldige fødselsnummer. 2008 og 2009	9
4.	Oversikt over overtidstimer og overtidsbetaling	10
5.	Fordeling av observasjoner på 1-siffer styrk, 2008 og 2009.....	11
6.	Fordeling av antall med 1-siffer styrk i lønnsstatistikken sammenlignet med AA- registeret	11
7.	Antall observasjoner fordelt etter koder for grenseverdier (ok-koder) i 2009.....	12
8.	Antall foretak i populasjon og utvalg og invers trekkesannsynlighet (Vekt).....	13
9.	Effekt av kontroll og revisjon av avtalt arbeidstid per uke	14
10.	Effekt av kontroll og revisjon av lønnsvariable. Heltidsansatte. Kroner.....	15
11.	Effekt av kontroll og revisjon av månedslønn etter næring og yrke. Antall og kroner.....	16