R&D Report **2002:4**
Research – Methods – Development

Statistiska centralbyrån Statistics Sweden

The Sampling- and the Estimation Procedure in the Swedish Labour Force Survey

by

Hassan Mirza and Jan Hörngren

INLEDNING

TILL

R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.
Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.

Föregångare:

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

Efterföljare:

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

R & D Report 2002:4. The sampling- and the estimation procedure in the Swedish Labour Force Survey / Hassan Mirza; Jan Hörngren.
Digitalt skapad fil, anpassad efter de digitaliserade delarna i serien. Statistiska centralbyrån (SCB) 2016.

urn:nbn:se:scb-2002-X101OP0204

R&D Report **2002:4**
Research – Methods – Development

The Sampling- and the Estimation Procedure in the Swedish Labour Force Survey

by

Hassan Mirza and Jan Hörngren

R&D Report 2002:4

Research - Methods - Development

The Sampling- and the Estimation Procedure in the Swedish Labour Force Survey

Från trycket December 2002
Producent Statistiska centralbyrån, *Statistics Sweden,* metodenheten
Box 24300, *SE*-104 51 STOCKHOLM

Förfrågningar Hassan Mirza
hassan.mirza@scb.se
Telefon 08- 506 946 72

J an Hörngren
j an.horngren@scb.se
Telefon 019- 17 60 82

ISSN 0283-8680

Printed in Sweden
SCB-Tryck, Örebro 2002

The Sampling- and the Estimation Procedure in the Swedish Labour Force Survey

By
Hassan Mirza
Jan Hörngren
Statistics Sweden

Abstract

The Swedish Labour Force Surveys (LFS) have been performed by Statistics Sweden since 1961. During the years the sampling- and estimation procedures have changed several times. In this paper we will cover the procedures since 1993, when a new estimation procedure was introduced, which meant that register data on employment and unemployment were used as auxiliary information in the estimation stage. The estimates were based on two different poststratification systems; one for estimating the number of unemployed and its constituent subgroups and the other one for estimating the number of employed.

With respect to the estimates of unemployment and employment, the use of auxiliary information significantly improved the national estimates by reducing sampling errors and nonresponse bias.

However, it was not possible to produce the regional estimates from the system developed for national estimates. A third poststratification system was used for these estimates.

Thus, national and regional estimates were based on different auxiliary information and different poststratification systems. This led to inconsistencies between these estimates and the request for further regional classification could therefore not be met. In order to eliminate the inconsistencies between national and regional estimates, the estimation method was modified beginning with the October 1999 survey. General regression estimation (GREG), was used in this procedure.

A comparison of the new and old estimators on the first quarter 1999 indicated that the new regional estimator gave estimates with lower sampling error for all counties except one, where the new estimator produced a slightly larger sampling error.

Acknowledgements:
We would like to thank Claes Andersson and Karl-Erik Kristiansson who contributed with valuable comments.

Contents

1. Introduction and summary

In connection with the October 1999 survey, the estimation procedure was changed for the Swedish Labour Force Surveys (LFS).

This paper describes the sampling plan and the new estimation method for the standard LFS survey. The report also provides a brief review of earlier estimation procedures.

The main purpose of the new estimation method was to integrate national and regional estimates into the same system without causing a break in the time series in the national estimates.

In brief, the new estimation method involves:
1) A transition from two different poststratification estimators to two different general regression estimators.
2) The assumption that LFS sample is drawn from poststrata is abandoned. Since 1993, the estimation method in LFS has been based on the assumption that the sample at the time of the survey can be considered a stratified sample drawn from poststrata with simple random sampling without replacement (SI) in the poststrata. The assumption was based on the judgment that the differences in the existing inclusion probabilities had a marginal effect. A further consequence of the assumption was that the apparatus of formulas for estimation was simplified.
3) The assumption that inclusion probabilities have a marginal effect on estimates was abandoned and theoretically correct estimators were introduced.

Compared with the methodology prior to October 1999, the new methodology has the following advantages:

1) Consistency between regional estimates and national estimates is obtained, which in turn increases the accessibility of LFS data.
2) A reduction in sampling errors and nonresponse bias in the estimates on regional level. Appendix 6 describes these effects.

2. Sampling plan

During a quarter, the sample in LFS consists of three separate samples, one for each month of the quarter. LFS is a panel survey with a rotating sample, which refers to a procedure where individuals in the sample are included in the survey on several, but limited, number of occasions. The rotation schedule is so constructed that 7/8 of each of the three samples recur after three months and 1/8 of the sample is replaced with new individuals in the sample. This means that each sample person is included in the survey a total of eight times, one time per quarter, during two years.
Appendix 1 provides a schematic description of the sample and rotation systems.

The sample rotation scheme implies that:
1- Positive correlations are obtained between surveys three months (13 weeks) apart. It is possible at a given cost to obtain estimates of differences (estimates of three 3-months, quarterly and annual changes) with better precision than would otherwise

have been the case without such a rotating sample system. LFS's sample design prioritises estimates of quarterly averages and estimates of change between adjacent quarters.

2 A subset of the sample on a specific survey occasion (7/8, 6/8, 5/8, 4/8, 3/8, 2/8 and 1/8, respectively) recurs in later surveys at a distance of 3, 6, 9, 12, 15, 18, 21 and 24 months. This enables comparisons of values of variables from different survey occasions, for each individual in the common sample, and thus the possibility of estimating gross changes between two points in time.

2.1 Annual sample selection

The sample for LFS is drawn once a year (at the end of the year). The size[1] of the annual sample is sufficient to meet the need for individuals in the sample for the next 12 months (April-March). This need is related to the condition that 1/8 of the sample on each survey occasion is replaced by new individuals in the sample. It is also related to the fact that there is a gradual replacement of individuals who have reached age 65. These individuals leave the survey, regardless of the number of times they participated in it.

2.1.1 Sampling frame

The sampling frame for LFS is a version of Statistics Sweden's register of the total population (RTB), sorted by personal identification number. It is supplemented with information on whether each person is gainfully employed or not according to Statistics Sweden's employment register (SREG) for year *t-2,* where *t* is the year the sample is drawn. The sampling frame is generated by processing the RTB version for week 04, year *t,* together with SREG for year *t-2*.

2.1.2 Stratification

The population is stratified by the variables sex, region[2], citizenship (Swedish/non-Swedish) and gainfully employed (gainfully employed/not gainfully employed) according to SREG, which results in 2x24x2x2=192 sample strata.

2.1.3 Sample selection scheme

The individuals in each stratum are sorted by personal identification number. Then a systematic sample is drawn in each stratum. Since the sampling frame is sorted by age, a more equal distribution of ages is achieved in the systematic sample than compared with stratification by age classes and SI in strata. The sample fraction of the subpopulation foreign citizens is somewhat higher than for the subpopulation Swedish citizens. The inflated sample size for foreign citizens is due to the requirement of separate reporting for this group.

2.1.4 Sample coordination

There is a rule in LFS that individuals who have been included in LFS's sample at any time in the last five years shall not be included in the sample for the coming year. When the year's sample is drawn, it is matched with the earlier years' samples from the most recent five years. Individuals who were in the earlier samples are thus excluded from the sample.

---

[1] The 1999 annual sample contained approximately 25 600 individuals.

[2] Regional divisions in sample stratification are based on the division of counties in 1996.

2.1.5 Partition of sample into panels and random cells

The rotation schedule entails that the sample for each survey occasion (month) consists of eight panels that originate from the two or three year's samples. Each panel in turn is randomly divided into two sub-panels. In addition, each panel is divided into 40 random groups. Thus, the combination of the sub-panel and random groups provides 80 random cells per panel. Each random cell has an identity that consists of four digits. The first two digits indicate which sub-panel it belongs to, while the last two digits indicate the random group membership.

The year's sample shall cover the need for new individuals in the sample for 12 months (April-March) and replace the exiting 1/8 part each month. This implies that the year's sample is distributed into 12 panels, one for each month, which consists of a total of 12*80=960 random cells. The distribution of the year's sample into 960 random cells is done in the following way:

1- The year's sample is sorted by stratum membership (stratum number) and personal identification number.
2- The first 960 individuals are distributed randomly among the 960 random cells. These individuals are assigned '01' as an operational number.
3- The next 960 individuals on the list are distributed randomly among the 960 random cells. These individuals are assigned '02' as an operational number.

This procedure continues until all individuals in the sample on the list have been distributed among the 960 random cells. The combination random cell number (four digits) and operational number (two digits) give each sample person a unique sample unit number (Uenr). Consequently, a Uenr consists of six positions/digits.

The division of the year's sample into panels, random groups and random cells is partly related to the fact that LFS has a rotating sample system. The purpose of the division is also to enable the drawing of a sub sample from the LFS sample in a simple manner. This occurs frequently with supplementary surveys (supplementary questions) in association with the regular LFS. These surveys are often based on a subset of the total sample for the month or quarter.

2.1.6 "Sample bank"

The year's sample is put into a "sample bank". The sample bank contains individuals in the sample who come from the most recent two or three annual samples. The sample bank contains the Personal identification number, Uenr and register variables from the sampling frame(RTB and SREG). The sample bank is updated continuously with matches with the current RTB versions. These updates aim partly to remove those individuals in the sample who are no longer found in RTB, and partly to update information on the sample person's city of residence and citizenship.

2.2 The monthly sample

A "monthly sample " is selected from the sample bank with the aid of the sub-panel number and date of birth. The sample for each survey occasion (= month) consists of 8 panels consisting of 16 sub-panels each. Each panel consists of 40 random groups, which means that the monthly sample extends to 8*2*40=640 random cells. The monthly sample on each survey occasion originates from the last two or three year's sample. Appendix 2 provides a schematic description of the sample for a survey month during the second quarter in year *t* with respect to the sample's origin and its distribution of panels and random groups.

2.2.1 Partition of monthly sample by reference week

Individuals in the samples were asked about their labour market situation during a specific week, which means that the reference period in LFS is a week, called the reference week. Since January 1993, all weeks are surveyed during the year. Each survey month contains four or five reference weeks. During the first and second months of each quarter, the monthly sample is distributed among the four reference weeks and during the third month it is distributed among the five reference weeks. Every seventh year, when the year contains 53 calendar weeks, October is assigned five reference weeks instead of four. This means that the monthly sample is divided randomly into four or five reference weeks. Each panel is randomly divided into 40 random groups (see section 2.1.5), which implies that the monthly sample is also divided into 40 random groups. These random groups are numbered 10-49.

Distribution of the monthly sample among four reference weeks during the two first months of each quarter proceeds in the following manner. The first reference week is assigned to the 10 first random groups, i.e. random groups 10-19. The second reference week is assigned to the next 10 random groups (random groups 20-29), etc. In a similar way, each of the five reference weeks is assigned to 8 random groups in the last month of the quarter. The first reference week is assigned to the 8 first random groups (random groups 10-17). The second reference week is assigned to the 8 next random groups (random groups 18-25), etc. The division of the monthly sample by reference week is illustrated in Appendix 2.

2.3 Inclusion probabilities

Inclusion probabilities for individuals in a monthly sample vary primarily with the share of the current sample that consists of the total sample in a month. The sample from a given year's sample is represented in general as 1/8, 2/8, 3/8 or 4/8 of the monthly sample. Furthermore, the calculations of inclusion probabilities are affected somewhat by the matches that are made with earlier samples to avoid a quick reinclusion of recently participating individuals. They are also affected by continuous updates of the sample bank, which means that individuals not found in the current RTB version are excluded from the sample bank.

The inclusion probability for individual *k* on a certain survey occasion is calculated approximately as:

$$\pi_k = \frac{n_{h_j}}{N_{h_j}} \text{, where}$$

n_{h_j} =The number of individuals in the sample that belong to sample stratum *h* and to the year's sample *j*.

N_{h_j} =Number of individuals in the age group 16-64 in sample stratum *h* at sample selection occasion *j*.

3. The previous estimation procedure

During the period January 1993-September 1999, *national estimates* in LFS were based on two different poststratification procedures with two different types of auxiliary information. In the first procedure, **poststratification (I)**, the sample and population were distributed among 16 groups in the dimension with the current auxiliary variable (*aux1*)[3]. Poststratification I was used for estimating the number of unemployed and its constituent subgroups. In the second procedure, **poststratification (II),** the sample and population were distributed among 143 groups in the dimension with the auxiliary variable (*aux2*). Poststratification II was used for estimating the number of employed and LF-status 1. The estimation of LF-status 5[4] was obtained as a function of both the systems.

The use of auxiliary information in these systems improved the national estimates in LFS with[5]:

- Precision gains up to 18% in the estimate of the number of unemployed.
- Precision gains up to 30% in the estimate of the number of employed for by industry.
- Reduction of nonresponse bias for estimates of the number of unemployed and the number of employed.

It is well known that in poststratification, the number of observations in each poststratum should not be too few. This is why two estimators were constructed. It is also why the regional estimates could not be integrated into national estimates. A third poststratification system, **poststratification (III),** was used for these estimates, in which the sample and population were distributed among 300 groups in combinations of sex (2), age group (10) and region (15). The national and regional estimates were thus based on different auxiliary information and different poststratification procedures. This led to insufficient consistency between the estimates thereby reducing the possibility of meeting requests for further regional reports.

For an orientation in the estimation method prior to 1993, see Japec (1992).

4. The estimation method from October 1999

The main reasons for the new estimation method were to avoid earlier theoretical assumptions (see section 1) and to integrate national and regional estimates in the same system without causing breaks in the time series in the national estimates. A transition from a poststratified estimator to regression estimator made this integration possible.
The assumption, which was introduced in 1993, that inclusion probabilities have marginal effects on estimates was abandoned in the new methodology. This enables changes in the sample size that can be handled simply and that obtains truly theoretical estimators.

[3] For a description of aux1, aux2 and aux3, see section 4.3.
[4] LF-status 1 = students, working from home, pensioners and military conscripts.
LF-status 5 = incapacitated and individuals living outside Sweden.
[5] See Hörngren (1992).

The estimates since October 1999 build on two regression estimators (GREG) with two different types of auxiliary information. The first GREG estimator is used in the estimate of the number of unemployed (LF-status = 4) and the other GREG estimator is used in the estimation of the number of individuals in labour force status (1,2, 3) and the number of individuals in the population (labour force status = 1-5). The estimate of the number of individuals with labour force status =5 is obtained as a function of both GREG estimators. It should be stressed that it is theoretically and practically possible to construct a single regression estimator with the simultaneous use of auxiliary variables: aux1, aux2 and aux3. However, this estimator causes breaks in the time series in the national estimates. The reason for using an estimation method with two regression estimators (instead of one) is to avoid breaks in the time series in the national estimates[6].

This chapter provides first a short theoretical insight into the regression estimator. This is followed by a brief description of the registers that are used for creating auxiliary information in LFS. The construction of the different auxiliary variables is also described, and lastly the two different estimators GREG (I) and GREG (II) and their constituent functions are described.

4.1 General REGression estimation (GREG)

The regression estimator is an estimation method that utilises auxiliary information in the estimation stage. The idea of using auxiliary information is based on the covariation of the auxiliary variables with the survey variable. The use of auxiliary information aims to reduce sampling and nonresponse errors. Regression estimation means that for element *k* in the sample, $(y_k, \boldsymbol{x}_k)$ are observed, where y_k is the observed value of y (variable of interest), while $\boldsymbol{x}_k$ is a vector of auxiliary information. The methodology also requires that the total population for the $\boldsymbol{x}$ -vector is known.

For a more detailed description of regression estimators, see *Särndal C, Swensson B and Wretman J,* (1992). The following provides a brief formalization of a GREG estimator.

A random sample *s* of size n_s is drawn from a population *U* consisting of *N* individuals, according to the sampling design *p*(.) which is such that all individuals have a probability > 0 of being included in the sample. Due to nonresponse, data on the y -variable can be collected only for a subset *r* of size m_r. The sampling design *p*(.) in LFS implies, for example, that the population is divided into *H* strata, where strata *h* contains N_h individuals. In each stratum *h*, a random sample of size n_h is drawn so that all individuals have the same probability of being included in the sample.

A regression estimator for the total $t_y = \sum_U y_k$ can be written in the formula

$$\hat{t}_y = \sum_r w_k y_k \qquad (4.1.1)$$

where

$\hat{t}_y$ = Estimator of a total, for example, the number of employed.

---

[6] However, the new estimating procedure has caused breaks in the time series in the regional estimates. LFS has published reprocessed regional estimates, based on the new methodology, beginning with first quarter 1997.

y_k = The value of variable *y* for element *k*.

w_k = The weight depends on the design of the sample, the auxiliary vector $\boldsymbol{x}_k$ and the model used for nonresponse adjustment.

$w_k = g_k \times d_k$,

$d_k = 1/(\pi_k \hat{\theta}_k)$,

π_k = The inclusion probability for individual *k*, in LFS $\pi_k = \frac{n_h}{N_h}$, for all individuals *k* that belong to stratum *h*.

$\hat{\theta}_k$ = The estimated response probability for individual *k*, $\hat{\theta}_k = \frac{m_h}{n_h}$ when *k* belongs to stratum *h*. This estimate is based on the assumption that individuals respond independently of each other and with the same probability in stratum *h*.

$$g_k = 1 + (\mathbf{t}_x - \hat{\mathbf{t}}_x)' \left(\sum\nolimits_r \frac{\mathbf{x}_k \mathbf{x}'_k q_k}{\pi_k \hat{\theta}_k} \right)^{-1} \mathbf{x}_k q_k \qquad (4.1.2)$$

g_k can be seen as a correction factor reflecting the contribution of auxiliary information to reducing nonresponse and random error.

$\mathbf{x}_k = (x_{1k}, \ldots, x_{jk}, \ldots x_{Jk})'$ is a vector of length *J*. Where *J* is the number of auxiliary variables.
q_k is a known constant (see section 5).
$\mathbf{t}_x = (t_{x1}, \ldots, t_{xj} .., t_{xJ})$ is a vector of length *J* that contains known totals from the register.

$\hat{\mathbf{t}}_x = (\hat{t}_{x1}, \ldots, \hat{t}_{xj} .., \hat{t}_{xJ})$ is a vector that contains estimates of elements in vector $\mathbf{t}_x$. Where the estimation of each element t_{xj} is given by:

$$\hat{t}_x = \sum\nolimits_r d_k x_k$$

The variance for $\hat{t}_y$ is estimated with

$$\hat{V}(\hat{t}_{yGREG}) = \sum \sum \frac{\pi_{kl} \hat{\theta}_{kl} - \pi_k \hat{\theta}_k \pi_l \hat{\theta}_l}{\pi_{kl} \hat{\theta}_{kl}} w_k e_k w_l e_l \qquad (4.1.3)$$

where
π_{kl} is the second order inclusion probability,
$\hat{\theta}_{kl}$ is the estimated probability that *k* and *l* belong to *r* (i.e. the respondents),

$$\hat{\theta}_{kl} = \frac{m_h}{n_h} \frac{m_h - 1}{n_h - 1}, \; k \neq l$$

and

$$e_k = y_k - \mathbf{B}' \mathbf{x}_k, \; \mathbf{B} = \left(\sum\nolimits_r \frac{\mathbf{x}_k \mathbf{x}'_k q_k}{\pi_k \hat{\theta}_k} \right)^{-1} \sum\nolimits_r \frac{\mathbf{x}_k y_k q_k}{\pi_k \hat{\theta}_k} \qquad (4.1.4)$$

4.2 Auxiliary information

Auxiliary information in LFS consists of register variables, or derived variables from register variables, which are correlated with the survey variable. LFS is a survey that describes the labour market conditions of the population. Thus, the auxiliary information of interest to LFS consists of register variables that have a relatively high correlation with the survey variable *labour force status*[7] in LFS.

4.2.1 The register of total population (RTB)

The RTB is the primary source for information on Sweden's population and it provides the basic information for LFS in both the sampling and estimation stage. RTB is used primarily as a base source for current population statistics. LFS uses data from RTB on sex, age, region and citizenship. RTB is continuously (daily) updated on changes in the composition of the population with the notifications from the tax authorities. The notifications contain information on births, deaths, change of address and immigration and emigration. At the estimations stage LFS uses data of RTB population from two months earlier. One can say that LFS is calibrating the current population totals at the time of estimation.

4.2.2 The employment register (SREG)

The employment register (SREG) was established in 1985 and is produced annually by Statistics Sweden. The register contains data on the population's "employment conditions". These data are obtained from the processing/derivation of data from six different registers.

The primary source is income statements from employers. The reference period for SREG is the month of November and the register is completed a little over one year later. The information used in LFS is data on "employment conditions" on an industry basis according to the Swedish industrial classification system (SNI 92).

The "employment variables" in SREG are coded according to the following divisions:

1 = gainfully employed
5 = has a wage statement, but is not gainfully employed
6 = does not have a wage statement and is not gainfully employed
4 = child, age 0-15

According to SREG, gainfully employed refers to an individual age 16 or older who has income from work. The income shall represent at least one hour's work per week during November. Gainful employment is delimited by a model-based method described in *Registerbaserad arbetsmarknadsstatistik* (Statistics Sweden 1998).

In LFS, *employment* is defined according to ILO's definitions and recommendations. The classification, whether an individual is employed or not at a certain reference week, is based on the individual's answers in LFS's questionnaire.

---

[7] Labour force status has the following variable values: 1 and 5= not in the labour force , 2 and 3=employed, 4= unemployed.

4.2.3 AMV job seeker register (SOK)

The aim of job-seekers register (SOK) at the National Labour Market Board's (AMV) is to describe job seekers and to provide a basis for planning and monitoring of public employment office activities. SOK is updated continuously with data from Sweden's public employment offices. LFS is primarily interested in data on job seeker categories. In addition, data on the date for leaving the SOK register is also necessary in this context.

.

Individuals registered at the employment offices are classified in different applicant categories. Appendix 3 provides a list of these categories[8].
The category that is closest to LFS's unemployment definition is applicant category (11). For a more detailed description of differences between categorisations of the unemployed by LFS and AMV, see *Olofsson* (1992).

4.3 Construction of auxiliary variables

4.3.1 Use of job seekers in SOK as an auxiliary variable

At the time of the survey, RTB is matched with the current SOK register, which refers to the end of the survey month, and the register RTB/SOK[9] is obtained. In this register, the variable *aux1* is formed as a combination of
(Job seekers in SOK/not in SOK), sex, and age groups according to the following:

SOK:	applicant category 11/ not 11	2
Sex:		2
Age groups:	16-24, 25-34,35-44 and 45-64	4

This gives 2*2*4=16 groups.

Thus, the variable has the set value $\{1, 2, 3,....,16\}$. This variable is entered into the monthly sample by a matching with the RTB/SOK register. The population number in each of the 16 groups that is defined by *aux1* is obtained from a total computation from the RTB/SOK/SREG register.

4.3.2 Use of employment status in SREG as a auxiliary variable

The RTB/SOK/SREG register is obtained by matching the RTB/SOK register with SREG. The second auxiliary variable (*aux2*) is formed here in combinations of sex, SNI-group according to SREG, and age group. This gives 2*8*10= 160 groups. For some combinations of SNI-group and sex, the age groups are included in larger age classes, whereby a total of 143 groups are obtained. Thus, this auxiliary variable has the value set $\{1, 2, 3,....,143\}$.

Appendix 4 describes the division of the register into the 143 groups. The auxiliary variable (*aux2*) and the variable SNI-group are entered into the monthly sample by a matching with the RTB/SOK/SREG register. The population number in each of the 143 groups defined by *aux2* is obtained from a total computation from the RTB/SOK/SREG register.

---

[8] Mona Henriksson (2000) provides a detailed description of different definitions and concepts in AMV's statistics.
[9] In practice, only the register RTB/SOK/SREG is created.

At best, the auxiliary information in SREG is 15 months old, and in the worst case, it is 26 months old. However, the auxiliary information has proved to be very effective, especially with respect to estimates of the number of employed by industries.

4.3.3 Use of county/municipality membership in RTB as a auxiliary variable

The third auxiliary variable (*aux3*) is formed in the register RTB/AMS/SREG. *Aux3* is divided into 27 groups as described in Appendix 5 and has the value set $\{1, 2, 3, ..., 27\}$. This auxiliary variable is also entered into the sample. The population number in each of the 27 groups defined by *aux3* is obtained from a total computation from the RTB/SOK/SREG register.

4.4 Estimation system

The new situation with two different GREG estimators involves the following estimation system:

I) GREG (I) with auxiliary information from SOK and RTB (*aux1* and *aux3*)
- *Number of unemployed*

II) GREG (II) with auxiliary information from SREG and RTB (*aux2* and *aux3*)
- *Number of employed*
- *Number of persons at work*
- *Number of persons absent* (sickness, vacation or on leave)
- *Degree of labour market attachment*
(E.g. *the number of permanent employees* and *temporary employees*)
- *Number of hours worked*
- *Number of persons with LF-status 1*[10] (i.e. certain categories of *individuals not in the labour force*)
- *Number of persons in the population*

III) Estimation of totals based on (I) and (II)
- *Number of persons in the labour force = number of employed (II) +number of unemployed (I)*
- *Number of persons not in the labour force = the population (II) – number of persons in the labour force (III)*
- *Number of persons with LF-status 5 = Number of persons not in labour force (III) - number of persons with LF-status 1 (II)*

IV) Ratio estimators
- *Unemployment rate = number of unemployed (I) / the number of persons in the labour force (III)*
- *Absence rate = number absent (II) / number employed (II)*

[10] AK-status 1 = students, working from home, pensioners and military conscripts.
AK-status 5 = incapacitated and individuals living outside Sweden.
Not in the labour force = AK-status 1 + AK-status 5

- Labour force rate = number of persons in the labour force (III) / population (II)
- Employment intensity = number employed (II) / population (II)
- Average working time = actual number hours worked (II) / number of individuals at work (II)

Where the number in the *population* is usually known totals.

5. Estimation of totals - monthly

The first stage in LFS estimations procedure is the estimations of different totals at a certain point in time, a month. These estimates are based on the monthly sample. Estimates of all other parameter types provided by LFS can be described as a function[11] of the estimated monthly totals.

The sample for a given survey month is derived from two or three different, independent annual samples. The monthly estimates are obtained by inflating each such subset of the monthly sample to known population totals with consideration of the sample individuals' inclusion probabilities. Estimates from the different annual samples are subsequently weighted together and a monthly estimate is obtained. This can be formalized in the following way:

The estimator of a total t_y, e.g. the number of unemployed, in the month is given by

$$\hat{t}_y = \sum_j c_j \hat{t}_{y_j} \qquad (5.1)$$

The variance for $\hat{t}_y$ is estimated with

$$\hat{V}(\hat{t}_y) = \sum_j c_j^2 \; * \hat{V}(\hat{t}_{y_j}) \qquad (5.2)$$

where
c_j is a constant that states the share of the entire monthly sample that comes from the annual sample *j,* i.e. (the number of persons from the annual sample *j* in the monthly sample) / (the number of persons in the entire monthly sample).

$\hat{t}_{y_j}$ = The estimate of a total based on the part of the monthly sample that is derived from the annual sample *j*. In the calculation of $\hat{t}_{y_j}$ and $\hat{V}(\hat{t}_{y_j})$, the regression estimator in section 4.1 is used, where the constant q_k in formula 4.1.2 and 4.1.4 is set to 1.

In practice, estimates in LFS are obtained by inflating the entire monthly sample in a single step using a direct application of formulas 4.1.1- 4.1.4 in section 4.1. Where q_k is set to c_j. The calculation of weights and variances applicable to point estimates (month), are carried out with the program package CLAN.

[11] For an orientation in estimation according to the function principle, see Rosén (2000).

5.1 Estimation of the number of unemployed - GREG (I)

Estimation of the number of unemployed in different domains is obtained from this system. Estimates according to GREG (I) can be described with formulas (4.1.1) - (4.1.4) under the following conditions:

1. The constant q_k in formula 4.1.2 and 4.1.4 is set to c_j (see section 5).
2. The auxiliary vector $\mathbf{x}_k$ represents the variable values for *aux1* and *aux3* for individual *k*. However, *aux1* and *aux3* are categorical variables and indicate the individual's group membership. This means that $\mathbf{x}_k = (x_1, x_2 ..., x_{16}; x_{17}, x_{18}, ..., x_{43})'$ is a vector with 16+27= 43 elements, where the first 16 elements, $x_1 - x_{16}$, state the individual's group membership according to the grouping from *aux1*, while the last 27 elements, $x_{17} - x_{43}$, state the individual's group membership according to *aux3*. Consequently, the vector $\mathbf{x}_k$ consists of 41 elements with the value 0 and 2 cells with the value 1. The first '1' states the individual's group membership in *aux1*, while the second '1' states group membership in *aux3*.
3. The vector $\mathbf{t}_x$ contains population totals for those categories defined in *aux1* and *aux3*.
   $\mathbf{t}_x = (N_1, N_2, ..., N_{16}; N_{17}, ..., N_{43})$

5.2 Estimation of the number of employed and number of individuals with labour force status 1 – GREG (II)

The following holds for the employment concept in LFS:

Number of employed = number of persons at work + number of persons absent from work (e.g. on leave)

Number of employed = permanent employees + temporary employees + self-employed + assisting family members. In addition, there are several different types of domains in the above categories.

Total employment and subgroup of employed are estimated accordance with GREG (II) and can be described with formulas (4.1.1)-(4.1.4) under the following conditions:

1 The constant q_k in formulas 4.1.1 and 4.1.4 is set to c_j (see section 5).

2 The auxiliary vector $\mathbf{x}_k$ represents the variable values for *aux2* and *aux3* for individual *k*. $\mathbf{x}_k = (x_1, x_2 ..., x_{143}; x_{144}, x_{145}, ..., x_{170})'$ is a vector with 143+27= 170 elements, where the first 143 elements, $x_1 - x_{143}$, give the individual's group membership according to groupings from *aux2* while the last 27 elements, $x_{144} - x_{170}$ give the individual's group membership according to *aux3*. The vector $\mathbf{x}_k$ consists of 168 elements with the value 0 and two elements with the value 1. The first '1' give the individual's group membership in *aux2* while the second '1' gives group membership in *aux3*.

3 The vector $\mathbf{t}_x$ contains population totals for the categories defined by *aux2* and *aux3*.
$\mathbf{t}_x = (N_1, N_2, \ldots, N_{143}; N_{144}, \ldots, N_{170})$

5.3 Estimation of number of persons in the labour force

Estimates of the number of persons in the labour force (active population) are based on GREG (I) and GREG (II). An estimator of the total number in the labour force can be written as:

$$\hat{t}_A = \hat{t}_{EM} + \hat{t}_{UN} \qquad (5.3.1)$$

Where
$\hat{t}_A$ = The estimator of the number in the labour force
$\hat{t}_{EM}$ = The estimator of the number employed according to (5.1) and auxiliary variables in GREG (II)
$\hat{t}_{UN}$ = The estimator of the number of unemployed according to (5.1) and auxiliary variables in GREG (I)
The variance for $\hat{t}_A$ can be written as:

$$\hat{V}(\hat{t}_A) = \hat{V}(\hat{t}_{EM}) + \hat{V}(\hat{t}_{UN}) + 2\hat{C}(\hat{t}_{EM}, \hat{t}_{UN}) \qquad (5.3.2)$$

$\hat{V}(\hat{t}_{EM})$ and $\hat{V}(\hat{t}_{UN})$ are calculated according to (5.2) with the auxiliary variables in GREG (I) and GREG (II), respectively. The calculation of the covariance, $\hat{C}(\hat{t}_{EM}, \hat{t}_{UN})$ is based on GREG(I) and GREG(II)

All variance calculations for different point estimates (month) are carried out with the program package CLAN. An estimate of the number in the labour force for different domains can be written as:
$\hat{t}_{d_A} = \hat{t}_{d_{EM}} + \hat{t}_{d_{UN}}$, where *d* represents the domain. (5.3.3)
The variance for $\hat{t}_{d_A}$ is estimated in a similar way as for $\hat{t}_A$.

5.4 Estimation of the number of persons not in the labour force

Estimates of the number of persons not in the labour force (not active=$\hat{t}_{NA}$) are also based on estimated quantities from GREG (I) and (II).
$\hat{t}_{NA} = N - \hat{t}_A$, where N, the number of individuals in the population is known. (5.4.1)
and the variance for $\hat{t}_{NA}$:

$$\hat{V}(\hat{t}_{NA}) = \hat{V}(N - \hat{t}_A) = \hat{V}(\hat{t}_A) \qquad (5.4.2)$$

For the domains:

$$\hat{V}(\hat{t}_{d_{NA}}) = \hat{V}(\hat{N}_d) + \hat{V}(\hat{t}_{d_A}) - 2\hat{C}(\hat{N}_d, \hat{t}_{d_A}) \qquad (5.4.3)$$

Where $\hat{N}_d$ is an estimate of the population in domain *d.*

In those cases $\hat{N}_d$ is known, i.e. $\hat{N}_d = N_d$ so is

$$\hat{V}(\hat{t}_{d_{NA}}) = \hat{V}(\hat{t}_{d_A}) \qquad (5.4.4)$$

5.4.1 Estimation of the number of persons with labour force status 5

The number of persons with labour force status 5 ($\hat{t}_{st5}$) is estimated according to the following:

$$\hat{t}_{st5} = \hat{t}_{NA} - \hat{t}_{st1} \qquad (5.4.1.1)$$

$\hat{t}_{st1}$ is the estimator of the number of individuals with labour force status 1

since

$$\hat{t}_{st5} = N - \hat{t}_{A} - \hat{t}_{st1} = N - \hat{t}_{EM} - \hat{t}_{UN} - \hat{t}_{st1} \qquad (5.4.1.2)$$

then

$$\begin{aligned}\hat{V}(\hat{t}_{st5}) &= \hat{V}(\hat{t}_{EM}) + \hat{V}(\hat{t}_{UN}) + \hat{V}(\hat{t}_{st1}) + \\ &2\hat{C}(\hat{t}_{EM}, \hat{t}_{UN}) + 2\hat{C}(\hat{t}_{EM}, \hat{t}_{st1}) + 2\hat{C}(\hat{t}_{UN}, \hat{t}_{st1})\end{aligned} \qquad (5.4.1.3)$$

For domains *d*:

$$\begin{aligned}\hat{V}(\hat{t}_{d_{st5}}) &= \hat{V}\left[\hat{N}_d - (\hat{t}_{d_{EM}} + \hat{t}_{d_{UN}} + \hat{t}_{d_{st1}})\right] \qquad (5.4.1.4)\\ &= \hat{V}(\hat{t}_{d_{EM}}) + \hat{V}(\hat{t}_{d_{UN}}) + \hat{V}(\hat{t}_{d_{st1}}) + \\ &2\hat{C}(\hat{t}_{d_{EM}}, \hat{t}_{d_{UN}}) + 2\hat{C}(\hat{t}_{d_{EM}}, \hat{t}_{d_{st1}}) + 2\hat{C}(\hat{t}_{d_{UN}}, \hat{t}_{d_{st1}}) + \\ &\left[\hat{V}(\hat{N}_d) - 2\hat{C}(\hat{N}_d, \hat{t}_{d_{EM}}) - 2\hat{C}(\hat{N}_d, \hat{t}_{d_{UN}}) - 2\hat{C}(\hat{N}_d, \hat{t}_{d_{st1}})\right]\end{aligned}$$

If the size of the population for the domain is known, i.e. $\hat{N}_d = N_d$, then the last term in (5.4.1.4) disappears.

6. Estimation of ratios

6.1 The unemployment rate

The unemployment rate ($\hat{R}_{d_{UN}}\%$), in a certain domain *d*, describes the share of unemployed as a percentage of the number of persons in the labour force in the group, i.e.:

$$\hat{R}_{d_{UN}}\% = \hat{R}_{d_{UN}} * 100 = \frac{\hat{t}_{d_{UN}}}{\hat{t}_{d_A}} * 100 \qquad (6.1.1)$$

and the variance can be calculated according to:

$$\hat{V}(\hat{R}_{d_{UN}}\%) = \left(\frac{100}{\hat{t}_{d_A}}\right)^2 \left[\hat{V}(\hat{t}_{d_{UN}}) + \hat{R}^2_{d_{UN}}\hat{V}(\hat{t}_{d_A}) - 2\hat{R}_{d_{UN}}\hat{C}(\hat{t}_{d_{UN}}, \hat{t}_{d_A})\right] \qquad (6.1.2)$$

6.2 Absence rate

The absence rate ($\hat{R}_{d_{AB}}\%$), in a certain domain, describes the share absent (whole reference week) as a percentage of the number of employed.

$$\hat{R}_{d_{AB}}\% = \hat{R}_{d_{AB}} * 100 = \frac{\hat{t}_{d_{AB}}}{\hat{t}_{d_{EM}}} * 100 \;; \qquad (6.2.2)$$

Where $\hat{t}_{d_{AB}}$ is the estimator of the number of persons absent in the domain *d*.

The variance for $\hat{R}_{d_{AB}}$ is estimated in a similar way as for $\hat{R}_{d_{UN}}$, i.e.

$$\hat{V}(\hat{R}_{d_{AB}}\%) = \left(\frac{100}{\hat{t}_{d_{EM}}}\right)^2 \left[\hat{V}(\hat{t}_{d_{AB}}) + \hat{R}^2{}_{d_{AB}}\hat{V}(\hat{t}_{d_{EM}}) - 2\hat{R}_{d_{AB}}\hat{C}(\hat{t}_{d_{AB}}, \hat{t}_{d_{AB}})\right] \quad (6.2.3)$$

6.3 The labour force rate

The labour force rate, ($\hat{R}_{d\,A}\%$), in a certain domain *d,* is the share of persons in the labour force as a percentage of the population, i.e.:

$$\hat{R}_{d_A}\% = \hat{R}_{d_A} * 100 = \frac{\hat{t}_{d_A}}{\hat{N}_d} * 100, \qquad (6.3.1)$$

and the variance can be calculated according to:

$$\hat{V}(\hat{R}_{d_A}\%) = \left(\frac{100}{\hat{N}_d}\right)^2 \left[\hat{V}(\hat{t}_{d_A}) + \hat{R}^2{}_{d_A}\hat{V}(\hat{N}_d) - 2\hat{R}_{d_A}\hat{C}(\hat{t}_{d_A}, \hat{N}_d\right] \qquad (6.3.2)$$

In those cases $\hat{N}_d$ is known, i.e. $\hat{N}_d = N_d$ the following are obtained

$$\hat{V}(\hat{R}_{d_A}\%) = \left(\frac{100}{N_d}\right)^2 \hat{V}(\hat{t}_{d_A}) \text{ where} \qquad (6.3.3)$$

$\hat{V}(\hat{t}_{d\,A})$ is calculated as in formula (5.3.3.2).

6.4 The employment rate

Employment rate ($\hat{R}_{d_{EM}}\%$), in a certain domain *d,* is the share of employed as a percentage of the population in the group and is estimated by:

$$\hat{R}_{d_{EM}}\% = \hat{R}_{d_{EM}} * 100 = \frac{\hat{t}_{d_{EM}}}{\hat{N}_d} * 100, \qquad (6.4.1)$$

$\hat{V}(\hat{R}_{d_{EM}}\%)$ is calculated in a similar way as in (6.3.2) and (6.3.3) by replacing $\hat{t}_{d_A}$ with $\hat{t}_{d_{EM}}$.

7 Quarterly and annual estimation and estimation of change

Thus far in the report, we have limited the description to estimates concerning a given month. LFS's reporting also includes the following parameter types:

- *change between adjacent months (e.g. between January and February)*

- *change between months over a year (e.g. between January 2000 and January 1999)*

- *quarterly averages*

- *change between adjacent quarters*

- *change between quarters over a year*

- *annual averages*

- *change between adjacent years*

The LFS system with rotating samples means that the variance structure for estimation of certain averages and changes becomes rather complicated. For example, an estimator of change between adjacent years involves in theory some 84 different covariance terms. However, certain assumptions are made in this example that reduce the number of different covariance terms to 28.

7.1 Change between adjacent months

Estimation of change between adjacent months is not complicated since we do not have any overlapping samples between the two consecutive monthly estimates. This type of a difference estimator can be written as:

$$\hat{t}_j - \hat{t}_{j-1} \qquad (7.1.1)$$

where

$\hat{t}_j$ is the estimator for an arbitrary month *j* (e.g. February 2000)

$\hat{t}_{j-1}$ is the estimator for an arbitrary month *j-1* (e.g. January 2000)

and the estimator of the variance can be written as:

$$\hat{V}(\hat{t}_j - \hat{t}_{j-1}) = \hat{V}(\hat{t}_j) + \hat{V}(\hat{t}_{j-1}) \qquad (7.1.2)$$

However, the following approximation is made in the numerical calculations:

$$\hat{V}(\hat{t}_j) = \hat{V}(\hat{t}_{j-1}) = \hat{V}(\hat{t}_{j-2}) = \hat{V}(\hat{t}_{j-3}) = \hat{V}(\hat{t}_{j-4}) = \hat{V}(\hat{t}_{j-5})$$
$$= \hat{V}(\hat{t}_{j-6}) = \hat{V}(\hat{t}_{j-7}) = \hat{V}(\hat{t}_{j-7}) = \hat{V}(\hat{t}_{j-8}) = \hat{V}(\hat{t}_{j-9}) \qquad (7.1.3)$$
$$= \hat{V}(\hat{t}_{j-10}) = \hat{V}(\hat{t}_{j-11}) = \hat{V}(\hat{t}_{j-12})$$

I.e. approximate variance estimation for point estimates (monthly) are made equal for all months. The variance estimator for (7.1.1) is calculated as:

$$\hat{V}(\hat{t}_j - \hat{t}_{j-1}) = \hat{V}(\hat{t}_j) + \hat{V}(\hat{t}_{j-1}) = 2\hat{V}(\hat{t}_j) \qquad (7.1.4)$$

7.2 Changes between months over a year

Changes between two months over a year are one of the most common comparisons made in LFS. Approximately 50 per cent of the sample in this type of estimates is overlapping. This has a variance-reducing effect compared with the case where we have independent samples between the points in time, as when comparisons are made between adjacent months in (7.1.1). A difference estimator between months over a year can be written as:

$$\hat{t}_j - \hat{t}_{j-12} \qquad (7.2.1)$$

where

$\hat{t}_j$ is the estimator for an arbitrary month *j* (e.g. February 2000)

$\hat{t}_{j-12}$ is the estimator for an arbitrary month *j-12* (e.g. February 1999)

and the estimator for the variance can be written as:

$$\hat{V}(\hat{t}_j - \hat{t}_{j-12}) = \hat{V}(\hat{t}_j) + \hat{V}(\hat{t}_{j-12}) - 2\hat{C}(\hat{t}_j, \hat{t}_{j-12}) \qquad (7.2.2)$$

Where $\hat{C}(\hat{t}_j, \hat{t}_{j-12})$ is approximated by

$$\hat{C}(\hat{t}_j, \hat{t}_{j-12}) = \lambda_{j,j-12} \sqrt{\hat{V}(\hat{t}_j).\hat{V}(\hat{t}_{j-12}).\hat{\rho}_{j,j-12}} \qquad (7.2.3)$$

and where

$\lambda_{j,j-12}$ = The share of responding individuals in the sample at both times j and j-12

$\hat{\rho}_{j,j-12}$ = An estimate of the correlation for t between j and j-12.

By using the approximation in (7.1.3) and the reformulation in (7.2.3), we gain the following expressions:

$$\hat{V}(\hat{t}_j - \hat{t}_{j-12}) = \hat{V}(\hat{t}_j) + \hat{V}(\hat{t}_{j-12}) - 2\hat{C}(\hat{t}_j, \hat{t}_{j-12})$$
$$2\hat{V}(\hat{t}_j) - 2.\lambda_{j,j-12}.\hat{V}(\hat{t}_j).\hat{\rho}_{j,j-12} = \qquad (7.2.4)$$
$$2\hat{V}(\hat{t}_j).(1 - \lambda_{j,j-12}.\hat{\rho}_{j,j-12})$$

7.3 Point estimation - Quarterly

Quarterly estimates are based on the three including months (estimates) of the quarter. Since January 1993, all weeks of the year are included in LFS. Each survey month contains four or five reference weeks. This implies that the monthly sample is normally[12] distributed over four reference weeks in the first two months of each quarter and over five weeks in the third month.

Given the reference week system in LFS the quarterly estimates needs to be calculated in such a way that the including months are proportionally weighed in relation to the number of reference weeks in each month. The estimates for a quarter ($\hat{t}_q$) in a normal year are calculated[13]according to:

$$\hat{t}_q = \frac{4}{13}\left(\hat{t}_j + \hat{t}_{j+1}\right) + \frac{5}{13}\hat{t}_{j+2}, \qquad (7.3.1)$$

$$q = 1, 2, 3, 4 \quad and \quad j = 1, 4, 7, 10$$

Three independent monthly samples are included in quarterly estimates and an estimator of the variance can be written as:

$$\hat{V}(\hat{t}_q) = \hat{V}\left[\frac{4}{13}\left(\hat{t}_j + \hat{t}_{j+1}\right) + \frac{5}{13}\hat{t}_{j+2}\right] \qquad (7.3.2)$$

By using the assumption (7.1.3), the variance estimate is calculated as:

$$\hat{V}(\hat{t}_q) = \frac{16}{169}\left[2\hat{V}(\hat{t}_j)\right] + \frac{25}{169}\hat{V}(\hat{t}_j) \approx 0.337\hat{V}(\hat{t}_j) \approx \frac{\hat{V}(\hat{t}_j)}{3} \qquad (7.3.3)$$

$\hat{V}(\hat{t}_j)$ is the variance estimator for the middle month in the numerical calculations for the relevant quarter.

---

[12] For years with 53 weeks, which occur every seventh year, the monthly sample in the fourth quarter is distributed over five weeks for October and December and four weeks for November.

[13] For a detailed description, see Mirza (1995).

7.4 Changes between adjacent quarters

When we estimate changes between adjacent quarters, 7/8 of the sample is overlapping. The estimator can be written as:

$$\hat{t}_q - \hat{t}_{q-1} = \left(\frac{4}{13}\left(\hat{t}_j + \hat{t}_{j+1}\right) + \frac{5}{13}\hat{t}_{j+2} \right) - \left(\frac{4}{13}\left(\hat{t}_{j-3} + \hat{t}_{j-2}\right) + \frac{5}{13}\hat{t}_{j-1} \right) \qquad (7.4.1)$$

The result of (7.3.3) is used to estimate the variance in (7.4.1). We also make the approximation

$$\hat{V}(\hat{t}_q) = \hat{V}(\hat{t}_{q-1}) \qquad (7.4.2)$$

and

$$\lambda_{q,q-1} = \lambda_{j,j-3} = \lambda_{j+1,j-2} = \lambda_{j+2,j-1} \qquad (7.4.3)$$

and

it is assumed that the correlation of the three-month differences is the same, i.e.:

$$\hat{\rho}_{j,j-3} = \hat{\rho}_{j+1,j-2} = \hat{\rho}_{j+2,j-1} \qquad (7.4.4)$$

The correlation between adjacent quarters is made approximately the same as for the correlation for three-month differences, i.e.:

$$\hat{\rho}_{q,q-1} = \hat{\rho}_{j,j-3} = \hat{\rho}_{j+1,j-2} = \hat{\rho}_{j+2,j-1} \qquad (7.4.5)$$

The estimator of the variance for (7.4.1) has the following form:

$$\begin{aligned} &\hat{V}(\hat{t}_q - \hat{t}_{q-1}) = \hat{V}(\hat{t}_q) + \hat{V}(\hat{t}_{q-1}) - 2\hat{C}(\hat{t}_q, \hat{t}_{q-1}) \approx \\ &2\hat{V}(\hat{t}_q) - 2\lambda_{q,q-1} \cdot \hat{V}(\hat{t}_q) \cdot \hat{\rho}_{q,q-1} \approx \qquad (7.4.6) \\ &\frac{2}{3}\hat{V}(\hat{t}_j) \cdot (1 - \lambda_{j,j-3} \cdot \hat{\rho}_{j,j-3}) \end{aligned}$$

7.5 Changes between quarters over a year

An estimator of the differences between quarters over a year has the same structure as the difference between adjacent quarters and is written as:

$$\hat{t}_q - \hat{t}_{q-4} = \left(\frac{4}{13}\left(\hat{t}_j + \hat{t}_{j+1}\right) + \frac{5}{13}\hat{t}_{j+2} \right) - \left(\frac{4}{13}\left(\hat{t}_{j-12} + \hat{t}_{j-11}\right) + \frac{5}{13}\hat{t}_{j-10} \right) \qquad (7.5.1)$$

We also use here the result from (7.3.3) in estimating the variance and the corresponding assumptions as in (7.4.3), (7.4.4) and (7.4.5). We consequently make the approximation

$$\lambda_{q,q-4} = \lambda_{j,j-12} = \lambda_{j+1,j-11} = \lambda_{j+2,j-10} \qquad (7.5.2)$$

and

$$\hat{\rho}_{j,j-12j} = \hat{\rho}_{j+1,j-11} = \hat{\rho}_{j+2,j-10} = \hat{\rho}_{q,q-4} \qquad (7.5.3)$$

An estimator of the variance for (7.5.1) can be written as:

$$\begin{aligned} &\hat{V}(\hat{t}_q - \hat{t}_{q-4}) = \hat{V}(\hat{t}_q) + \hat{V}(\hat{t}_{q-4}) - 2\hat{C}(\hat{t}_q, \hat{t}_{q-4}) \approx \\ &2\hat{V}(\hat{t}_q) - 2\lambda_{q,q-4} \cdot \hat{V}(\hat{t}_q) \cdot \hat{\rho}_{q,q-4} \approx \qquad (7.5.4) \\ &\frac{2}{3}\hat{V}(\hat{t}_j) \cdot (1 - \lambda_{j,j-12} \cdot \hat{\rho}_{j,j-12}) \end{aligned}$$

7.6 Annual point estimation

The point estimators for a certain year ($\hat{t}_y$) are based on the year's including Monthly estimates in a way similar to that for quarterly estimates. The 12 monthly estimates are weighed together proportionally in relation to the number of reference weeks for each monthly estimate and ($\hat{t}_y$) is obtained. Thus, the estimator for a normal year with 52 weeks is:

$$\hat{t}_y = \frac{4}{52}\left(\hat{t}_j + \hat{t}_{j+1} + \hat{t}_{j+3} + \hat{t}_{j+4} + \hat{t}_{j+6} + \hat{t}_{j+7} + \hat{t}_{j+9} + \hat{t}_{j+10}\right) + \frac{5}{52}\left(\hat{t}_{j+2} + \hat{t}_{j+5} + \hat{t}_{j+8} + \hat{t}_{j+11}\right) \quad (7.6.1)$$

$$\hat{V}(\hat{t}_y) = \hat{V}\left[\frac{4}{52}\left(\hat{t}_j + \hat{t}_{j+1} + \hat{t}_{j+3} + \hat{t}_{j+4} + \hat{t}_{j+6} + \hat{t}_{j+7} + \hat{t}_{j+9} + \hat{t}_{j+10}\right) + \frac{5}{52}\left(\hat{t}_{j+2} + \hat{t}_{j+5} + \hat{t}_{j+8} + \hat{t}_{j+11}\right)\right] \quad (7.6.2)$$

According to the expression in (7.3.3), an approximation of (7.6.2) is

$$\hat{V}(\hat{t}_y) = \hat{V}\left(\frac{\hat{t}_j + \hat{t}_{j+1} + \hat{t}_{j+2} + \dots\dots + \hat{t}_{j+11}}{12}\right) \quad (7.6.3)$$

There are three different types of correlations for annual point estimation:
- *The same individuals for 3 months' term.*
- *The same individuals for 6 months' term.*
- *The same individuals for 9 months' term.*

An estimator of the variance can be written analogous to earlier approximations with respect to λ och $\hat{\rho}$, i.e. λ och $\hat{\rho}$ are "constants" that vary only with the interval between two survey occasions:

$$\hat{V}(\hat{t}_y) = \hat{V}\left(\frac{\hat{t}_j + \hat{t}_{j+1} + \hat{t}_{j+2} + \dots\dots + \hat{t}_{j+11}}{12}\right) = \frac{12 \cdot \hat{V}(\hat{t}_j) + 2 \cdot \left[9 \cdot \hat{C}\left(\hat{t}_j, \hat{t}_{j+3}\right) + 6 \cdot \hat{C}\left(\hat{t}_j, \hat{t}_{j+6}\right) + 3 \cdot \hat{C}\left(\hat{t}_j, \hat{t}_{j+9}\right)\right]}{144} = \frac{\hat{V}(\hat{t}_j) \cdot \left(6 + 9 \cdot \lambda_{j,j+3} \cdot \hat{\rho}_{j,j+3} + 6 \cdot \lambda_{j,j+6} \cdot \hat{\rho}_{j,j+6} + 3 \cdot \lambda_{j,j+9} \cdot \hat{\rho}_{j,j+9}\right)}{72} \quad (7.6.4)$$

7.7 Difference between adjacent years

The estimator can be written as:

$$\hat{t}_y - \hat{t}_{y-1} \quad (7.7.1)$$

where $\hat{t}_y$ and $\hat{t}_{y-1}$ is given in (7.6.1)

This case involves correlations with the following possible intervals:
- *The same individuals for 3 months' term* (t-3)

- *The same individuals for 6 months' term* (t-6)
- *The same individuals for 9 months' term* (t-9)
- *The same individuals for 12 months' term* (t-12)
- *The same individuals for 15 months' term* (t-15)
- *The same individuals for 18 months' term* (t-18)
- *The same individuals for 21 months' term* (t-21)

With the support of assumptions used earlier, we can write a variance estimator for (7.7.1) as:

$$\hat{V}\left[\left(\frac{\hat{t}_j+\hat{t}_{j+1}+\hat{t}_{j+2}+......+\hat{t}_{j+11}}{12}\right)-\left(\frac{\hat{t}_{j-1}+\hat{t}_{j-2}+\hat{t}_{j-3}+......+\hat{t}_{j-12}}{12}\right)\right]=$$

$$\Big[3\hat{V}(\hat{t}_j)\cdot\big(8+2\cdot(6\lambda_{j,j-3}\cdot\hat{\rho}_{j,j-3}-\lambda_{j,j-3}\cdot\hat{\rho}_{j,j-3}+4\lambda_{j,j-6}\cdot\hat{\rho}_{j,j-6}-$$
$$2\lambda_{j,j-6}\cdot\hat{\rho}_{j,j-6}+\lambda_{j,j-9}\cdot\hat{\rho}_{j,j-9}-3\lambda_{j,j-9}\cdot\hat{\rho}_{j,j-9}-4\lambda_{j,j-12}\cdot\hat{\rho}_{j,j-12}- \qquad (7.7.2)$$
$$3\lambda_{j,j-15}\cdot\hat{\rho}_{j,j-15}-2\lambda_{j,j-18}\cdot\hat{\rho}_{j,j-18}-\lambda_{j,j-21}\cdot\hat{\rho}_{j,j-21}\big)/144\Big]$$

References

1- Andersson, Claes and Nordberg, Lennart (1998). *A User's Guide to **CLAN 97** - a SAS-program for computation of point- and standard error estimates in sample surveys.*
2- Hörngren, Jan (1992). *The use of register as auxiliary information in the Swedish Labour Force Survey*. R&D Report 1992:13. Statistics Sweden.

3- Hörngren, J and Mirza, H (1998).*Sikte mot ett modifierat estimationsförfarande i SCB:S arbetskraftsundersökningar (AKU) med hjälp av CLAN*. Pm 980109. (In Swedish).
4- Henriksson Mona (2000). *Definitioner inom AMV: s statistik*. Uin 2000:2. The National Labour Market Board. (In Swedish).
5- Japec, Lilli (1992). *Förändringar i arbetskraftsundersökningarna*. Bakgrundsfakta till arbetsmarknads- och utbildningsstatistiken. 1992.1. Statistics Sweden. (In Swedish).
6- Mirza, Hassan (1995). *Kvartals – och årsskattningar AKU*. Pm 1995-10-10. (In Swedish).
7- Mirza, Hassan (1999). *Instruktioner inför det nya estimatonsförfarandet i AKU från oktoberundersökningen 1999. Pm 1999-08-23*. (In Swedish).
8- Olofsson, Anita (1996) *De svenska arbetskraftsundersökningar*. Bakgrundsfakta till arbetsmarknads- and utbildningsstatistiken 1996.2 Statistics Sweden. (In Swedish).
9- Olofsson, Anita (1998) *Arbetslöshetssiffror från AMS och SCB*. Pm 1998-03-26. (In Swedish)
10- Rosén, Bengt (Våren 2000). *Teori för urvalsundersökningar*. Statistics Sweden. (In Swedish)
11- Sixten, Lundström (2000). *Kalibrering av vikter*. R&D Report 2000:1. Statistics Sweden. (In Swedish)
12- Särndal, C, Swensson, B and Wretman, J.H. (1992). *Model Assisted Survey Sampling*. New York: Springer- Verlag.
13- Statistics Sweden (1998). *Registerbaserad arbetsmarknadsstatistik (SREG).* Description of statistics. (In Swedish)
14- Person, Arne. *Årsurval i AKU*, pm 000809. (In Swedish)
15- Person, Arne. *Uenr i AKU*, pm 990108. (In Swedish)

Appendix 1. Schematic description of the rotation system in LFS

Survey occasion	Number of occasions in panel (A_i, B_i, C_i) included in the survey							
	1	2	3	4	5	6	7	8
.								
.								
January	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
February	B_1	B_2	B_3	B_4	B_5	B_6	B_7	B_8
March	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
April	A_9	A_1	A_2	A_3	A_4	A_5	A_6	A_7
May	B_9	B_1	B_2	B_3	B_4	B_5	B_6	B_7
June	C_9	C_1	C_2	C_3	C_4	C_5	C_6	C_7
July	A_{10}	A_9	A_1	A_2	A_3	A_4	A_5	A_6
August	B_{10}	B_9	B_1	B_2	B_3	B_4	B_5	B_6
September	C_{10}	C_9	C_1	C_2	C_3	C_4	C_5	C_6
October	A_{11}	A_{10}	A_9	A_1	A_2	A_3	A_4	A_5
November	B_{11}	B_{10}	B_9	B_1	B_2	B_3	B_4	B_5
December	C_{11}	C_{10}	C_9	C_1	C_2	C_3	C_4	C_5
January	.	.	.	.	.	.	.	A_4
.	.	.	.	.	.	.	.	B_4
.	.	.	.	.	.	.	.	C_4
	.	.	.	.	.	.	.	A_3
								.

Appendix 2: April surveys year *t* distributed by random cells (panel * random group) and sample drawing occasion (year *t, year t-1* and year *t-2*)

Reference Week	Random group (Positions 3-4 in UB-nr.)	Panels 1		2		3		4		5		6		7		8	
		Sub-panel number (pos. 1-2 in UB-nr.)															
		32	36	31	35	24	28	23	27	22	26	21	25	14	18	13	17
1	10	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	11	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	12	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	13	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	14	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	15	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	16	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	17	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	18	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	19	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
2	20	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	21	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	22	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	23	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	24	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	25	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	26	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	27	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	28	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	29	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
3	30	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	31	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	32	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	33	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	34	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	35	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	36	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	37	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	38	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	39	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
4	40	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	41	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	42	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	43	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	44	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	45	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	46	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	47	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	48	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*
	49	*t*	*t*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-1*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*	*t-2*

Appendix 3

Applicant categories

11	Unemployed, public employment service
12	Unemployed, vocational counselling service
13	Unemployed, awaiting labour market measure
14	Other, registered in public employment service
21	Part-time unemployed
22	Employed by the hour
31	Temporary Employed
34	EES-applicant
35	Job-Changers, Samhall
41	Job-Changers
42	Wage subsidy (disabled persons)
43	Public Sheltered Employment
46	Start-up grant
47	General Recruitment Incentive
48	Extended Recruitment Incentive
49	Special Recruitment Incentive
50	Extended Recruitment Incentive
54	Work Placement Scheme
56	Public Temporary Work
64	Computer /Activity Centre
65	Municipal Youth Programmes (age below 20)
66	Youth Guarantee (ages 20-24)
71	Employability Rehabilitation Programme
72	Supported Employment Programme (disabled persons)
73	Activities within counselling guidance and placement service
74	Job-seeking Activities
75	Projects with Employment Policy orientation
81	Employment Training

Appendix 4: Division of RTB/SREG to 143 group and construction of *aux2*

Employed according to SREG		Sex/Age group	
SNI-group	Industry code SNI 92	Men	Women
1	**A+B**: 01-02, 05	16-24	16-24
	C: 10-14	25-29	25-29
	E: 40-41	30-34	30-34
	F: 45	35-39	35-39
		40-44	40-44
		45-49	45-49
		50-54	50-54
		55-59	55-59
		60-64	60-64
2	**DA+DB+BC+DD+DE+DF+DG+DH+DI:**	16-24	16-24
	15-26	25-29	25-29
	part of **DJ:** 27	30-34	30-34
	part of **DN:** 36	35-39	35-39
		40-44	40-44
		45-49	45-49
		50-54	50-54
		55-59	55-59
		60-64	60-64
3	part of **DJ:** 28	16-24	16-24
	DK+DL+DM: 29-35	25-29	25-29
		30-34	30-34
		35-39	35-39
		40-44	40-44
		45-49	45-49
		50-54	50-54
		55-59	55-59
		60-64	60-64
4	part of **G:** 51-52	16-24	16-24
	In: 60-64	25-29	25-29
		30-34	30-34
		35-39	35-39
		40-44	40-44
		45-49	45-49
		50-54	50-54
		55-59	55-59
		60-64	60-64
5	**J:** 65-67	16-24	16-24
	part of **K:** 70-72, 74	25-29	25-29
		30-34	30-34
		35-39	35-39
		40-44	40-44
		45-49	45-49
		50-54	50-54
		55-59	55-59
		60-64	60-64

Appendix 4 continued

Employed according to SREG		Sex/Age group	
SNI-group	Industry code SNI 92	Men	Women
6	part of **K:** 73	16-24	16-24
	L: 75	25-29	25-29
	M: 80	30-34	30-34
	N: 85	35-39	35-39
	part of **O:** 91	40-44	40-44
		45-49	45-49
		50-54	50-54
		55-59	55-59
		60-64	60-64
7	part of **DN:** 37	16-24	16-24
	part of **G:** 50	25-29	25-29
	H: 55	30-34	30-34
	part of **O:** 90, 92-93	35-39	35-39
	P: 95	40-44	40-44
	Q: 99		45-49
		45-54	50-54
		55-64	55-64
8	00 = Not employed and/or SNI lacking	16-19	16-19
		20-24	20-24
		25-29	25-29
		30-34	30-34
		35-39	35-39
		40-44	40-44
		45-49	45-49
		50-54	50-54
		55-59	55-59
		60-64	60-64

Appendix 5: County/municipality as an auxiliary variable (*aux 3*)

1) Stockholm county excl. Stockholm municipality
2) Stockholm municipality
3) Uppsala county
4) Södermanland county
5) Östergötland county
6) Jönköping county
7) Kronoberg county
8) Kalmar county
9) Gotland county
10) Blekinge county
11) ex-Kristianstad county
12) ex-Malmöhus county excl. Malmö municipality
13) Halland county
14) ex-Göteborg and Bohus county excl. Göteborg municipality
15) ex-Älvsborg county
16) ex-Skaraborg county (beginning 1998 excl. Mullsjö and Habo)
17) Värmland county
18) Örebro county
19) Västmanland county
20) Kopparberg county
21) Gävleborg county
22) Västernorrland county
23) Jämtland county
24) Västerbotten county
25) Norrbotten county
26) Malmö municipality
27) Göteborg municipality

Appendix 6: The effects of the new estimation procedure

In this appendix the effect of the new (estimation) procedure on the estimates of number of employed persons on a regional level is accounted for. In the table below the estimates using the new procedure (GREG) are being compared with the ones where the old procedure was used. This means that the results using GREG (II) are compared with the use of poststratification (III). The comparison refers to the first quarter 1999.

Notation:
$\hat{t}_{PST}$ = An estimate based on the old procedure, which means poststratification (III). See chapter 3: The previous estimation procedure.
$\hat{\sigma}_{PST}$ = An estimate of the standard deviation for $\hat{t}_{PST}$.

$\hat{t}_{GREG}$ = An estimate based on the new procedure, which means GREG (II). See chapter 5 point 2.
$\hat{\sigma}_{GREG}$ = An estimate of the standard deviation for $\hat{\sigma}_{PST}$.

$$\hat{\vartheta} = \left[1 - \left[\frac{\hat{\sigma}_{GREG} / \hat{t}_{GREG}}{\hat{\sigma}_{PST} / \hat{t}_{PST}}\right]\right] * 100$$; is the estimated gain or loss in precision

Gains in precision have thus been calculated in terms of coefficient of variation (c.v.)

Table: Estimates of Employed Persons by Region/Counties. LFS first quarter 1999. Thousands.

County	$\hat{t}_{PST}$	$\hat{t}_{GREG}$	$\hat{\sigma}_{PST}$	$\hat{\sigma}_{GREG}$	$\hat{t}_{GREG} - \hat{t}_{PST}$	$\hat{\vartheta}$
Stockholm	908.3	898.2	4.7	4.8	-10.1	-3.1
Uppsala	142.8	139.0	3.2	2.0	-3.8	35.8
Södermanland	106.7	110.9	3.0	1.7	4.2	46.2
Östergötland	172.8	174.1	2.2	2.1	1.3	5.3
Jönköping	153.3	152.4	3.6	1.8	-0.9	49.8
Kronoberg	82.4	84.9	2.9	1.3	2.5	56.8
Kalmar	100.9	102.9	3.1	1.7	2.0	46.6
Gotland	24.1	25.4	1.8	0.8	1.3	56.9
Bleking	65.6	64.3	2.5	1.3	-1.3	46.9
Skån	477.0	476.2	4.5	3.8	-0.8	14.5
Hallands	114.8	121.4	3.0	1.7	6.6	46.8
Västra Götaland	672.1	664.7	5.0	4.3	-7.4	13.0
Värmland	117.0	118.4	3.4	1.8	1.4	46.9
Örebro	123.2	122.1	3.0	1.8	-1.1	40.3
Västmanland	110.1	112.3	3.0	1.7	2.2	44.3
Dalarna	116.2	116.2	3.4	1.7	0.0	49.9
Gävleborg	120.9	121.6	3.4	1.7	0.7	50.3
Västernorrland	107.1	109.0	2.6	1.7	1.9	36.0
Jämtland	55.4	54.5	2.4	1.3	-0.9	44.5
Västerbotten	112.4	111.2	3.0	1.7	-1.2	43.1
Norrbotten	103.5	107.8	2.9	1.8	4.3	40.2
Total	3986.4	3987.5	9.5	9.5	1.1	0.1

From this table it can be seen that strong gains in precision valid for the estimates on number of employed persons on a regional level. Only for the county of Stockholm there is no gain in precision. To attain the same gains in precision without using the new procedure, integrating of the estimates for the whole country regional estimates through GREGF, large increases in LFS-sample is needed. For the counties of Jönköping, Kronoberg, Kalmar and Dalarna samples of at least 4 times larger are needed.

It must be noted that also the levels of the estimates of the number of employed persons using the new procedure are lower in the big counties The other counties show unchanged or higher levels compared with the former procedure of regional estimation. This is most probable due to the fact that the new procedure gives better adjustment of the sampling error on a regional level. These results are in accordance with earlier studies of non-response in the Swedish LFS.

The calculations in the table above have been performed for several periods. These calculations give mainly the same results as the ones for the first quarter 1999.

Regarding estimation of the unemployed on a regional level, only marginal positive effects have been noticed.

Förteckning över utkomna R&D Reports

R&D Reports är en för IT-enheten och Metodenheten gemensam publikationsserie, som 1988-01-01 ersatte de tidigare "gula" och "gröna" serierna. I serien ingick fram till årsskiftet 1992-93 även **Abstracts** (sammanfattning av metodrapporter från SCB).

Reports published during 1999 and onwards:

ISSN 0283-8680

Tidigare utgivna **R&D Reports** kan beställas genom Katarina Klingberg, SCB, MET, Box 24 300, 104 51 STOCKHOLM (telefon 08-506 942 82, fax 08-506 945 99, e-post katarina.klingberg@scb.se).
R&D Reports from 1988-1998 can - in case they are still in stock - be ordered from Statistics Sweden, attn. Katarina Klingberg, MET, Box 24 300, SE-104 51 STOCKHOLM (telephone +46 8 506 942 82, fax +46 8 506 945 99, e-mail katarina.klingberg@scb.se).