

Quality description of in-service training statistics

Labour force survey data and contents

The statistics on in-service training derive from the Finnish Labour Force Survey carried out in accordance with the EU standards during October, November and December 2001. The first Finnish EU-LFS was carried out in the spring of 1995. The purpose of a Labour Force Survey is to chart out matters relating to the employment and working life of the working age population. Similar surveys are carried out in all EU Member States, in the countries applying for EU membership and in Iceland, Norway and Switzerland in accordance with the directions of Eurostat, the Statistical Office of the European Communities. The aim is to produce optimally comparable data. The Labour Force Surveys of Statistics Finland were adjusted to comply with the EU regulations after Finland joined the EU.

The publication is based on data collected for national use concerning participation in in-service training during the 12-month period prior to the survey. The Labour Force Survey also contains such questions made in accordance with the EU standards that inquire about studies during the four-week period before the survey. The purpose of the questions is to collect internationally comparable data on in-service training, for example. Their formulation will change from the beginning of 2003. These questions are not dealt with in this publication.

The time series tables in In-service Training 2001 contain data from 1982 to 2001. Before the EU standardised Labour Force Survey (from 1982 to 1993), data were collected each autumn by means of the Supplementary Surveys to the Finnish Labour Force Surveys, which included data on issues such as working conditions, pay, employment history and unemployment background and duration. The surveys were implemented as supplementary to the monthly Labour Force Surveys. They were carried out annually between 1982 and 1986, and every second year between 1987 and 1993. The in-service training figures produced from the Labour Force Survey concerning the year 2001 are comparable with the previously produced data.

Sampling design, survey implementation and weighting

The target population for the Labour Force Survey consists of persons aged 15 to 74 permanently resident in Finland. The sample was drawn with systematic sampling from the Central Population Register arranged on the basis of domicile codes, by which the sample is divided regionally in the same way as the target population.

In the survey period, that is, during the last quarter of 2001, the size of the LFS sample was 38,069, of whom 32,239 took part in the interview. Of them 16,578 were employees aged 15 to 64 (the target population of the in-service training statistics). Estimation of in-service training data was made similarly as in the Labour Force Survey.

The sampling design also involves a rotating panel, in which the same person is included at three-month intervals. Thus, there is no overlapping in the sample during the three months and each person is included in the in-service training survey data only once.

The standard errors of the survey parameters were calculated using the CLAN software. CLAN is an SAS macro written by Statistics Sweden that can be used for calculating standard errors according to the sampling and estimation design.

Nonresponse and weighting

In the last quarter of 2001, nonresponse for the LFS was 15 per cent. Error possibly caused by nonresponse was corrected by weighting the response data to the level of the target population using the calibration method.

For calculation of weighting coefficients, the sample was first post-stratified by region (21), gender (2) and 10-year age group (6) into 252 strata. After that the weights were calibrated according to the number of unemployed in the jobseeker register, region, gender and 5-year age groups to correspond with the register data. Calibration produces the correct population distributions for all the auxiliary variables used in the weighting, provided that the calibrated weights differ as little as possible from the original ones (Deville et al. 1992; Deville et al. 1993). Use of additional information increases the accuracy of the results and corrects the bias caused by nonresponse.

If a person to be interviewed was not reached, the information could be inquired of a proxy, who was usually some other person belonging to the household of the target person. As of the 1998 data collection, questions relating to in-service training could also be addressed to proxies. Previously, questions were not asked of proxies, but the data were produced by means of imputing.

Reliability of the figures

In a sample survey, each study variable has a random error, called sampling error. The sampling error of a study parameter can be estimated using the standard error. It shows how closely the findings are centred round the parameter to be measured. Sample size and fluctuations in the values of the variables influence the size of the standard error. Using the standard error, a parameter estimate can be given a confidence interval, defining the area within which the value of the variable will, with given probability, fall.

$(1-\alpha)*100\%$ confidence interval means a closed interval within which the true value of the parameter falls with $(1-\alpha)*100\%$ probability, i.e.

$$Pr\{p \in (p - t_{\alpha} d(p), p + t_{\alpha} d(p))\} = 1 - \alpha$$

where $d(p)$ is the standard error

The t_{α} value for a 95 per cent confidence interval would be 1.96 (the t_{α} value for a 99% confidence interval would be 2.58 and, for a 90% confidence interval, 1.65).

Table 1 shows the standard percentage errors by gender and age.

1. Standard percentage errors by gender and age, Labour Force Survey 2001

Gender Age	Participation percentage in in-service training %	Standard error %
Both Genders	44.8	0.4
15-24	23.7	1.0
25-34	46.6	0.8
35-44	49.9	0.7
45-54	47.3	0.7
55-64	43.4	1.1
Males	41.8	0.5
15-24	23.6	1.4
25-34	45.5	1.1
35-44	45.4	1.1
45-54	43.7	1.1
55-64	37.9	1.7
Females	47.8	0.5
15-24	23.8	1.4
25-34	48.0	1.2
35-44	54.4	1.0
45-54	50.6	1.0
55-64	48.3	1.6

Next is an example of how a confidence interval can be calculated using the standard error table:

According to the 2001 LFS data, 44.8 per cent of 15 to 64-year-old employees participated in in-service training. The standard percentage error of this figure is 0.4 per cent (Table 1). A 95 per cent confidence interval can be calculated as follows: $1.96 \times 0.4 = 0.8$.

With 95 per cent probability, the proportion of those participating in in-service training was between 44.0 and 45.6 per cent ($44.8 \pm 0.8\%$).

References

Deville, J.-C. and Särndal, C.-E. 1992. Calibration Estimators in Survey Sampling. JASA, June 1992. Vol. 87, Theory and Methods, American Statistical Association. pp. 376-382.

Deville, J.-C., Särndal, C.-E. and Sautory O. 1993. Generalized Raking Procedures in Survey Sampling. JASA, September 1993. Vol. 88, Theory and Methods, American Statistical Association. pp. 1013-1020.