

Statistical Survey of Non-Formal Education

Ondřej Nývlt¹ | *University of Economics, Prague, Czech Republic*

Abstract

At present it is not possible to consider the education system only from a point of view of formal education focused on a programme within a regular education system. Labour market flexibility and new requirements on employees create a new domain of education called non-formal education. Is there a reliable statistical source with a good methodological definition for the Czech Republic? Labour Force Survey (LFS) has been the basic statistical source for time comparison of non-formal education for the last ten years. Furthermore, a special Adult Education Survey (AES) in 2011 was focused on individual components of non-formal education in a detailed way. In general, the goal of the EU is to use data from both internationally comparable surveys for analyses of the particular fields of lifelong learning in the way, that annual LFS data could be enlarged by detailed information from AES in five years periods. This article describes reliability of statistical data about non-formal education. This analysis is usually connected with sampling and non-sampling errors.

Keywords

Labour force survey, non-formal education, sampling errors, non-sampling errors

JEL code

I2, J2

INTRODUCTION

Higher or lower education is important for explanation of fundamental social, economic or demographic differences. For example, it is the most important factor for analysis of the labour market. If divided by educational attainment unemployment rate in the Czech Republic according to the ILO definition (in the age group 15–64 years) in 1st quarter of 2011, university graduates permanently show a low unemployment rate (2.5%) and similarly persons who have secondary education with an A-level examination (5.2%). A high unemployment rate grows in the group of persons with basic education (28.4%), and an above average unemployment rate also pertains to the large group of those who have got secondary education without an A-level examination, including persons with apprenticeship certificates (8.2%) (News Releases Employment and Unemployment, 2011).

Significant economic changes bring new requirements on the labour market and force employees to gain new skills. The labour market creates demand for education but does not directly focus on formal educational system. The demand is oriented rather on a particular type of gaining skills. The non-formal education plays an important role in this case. For this reason, the basic indicator of education (life-long learning indicator) used for the international comparison works on the assumption of participation in formal or non-formal education.

¹ University of Economics, Prague, nám. W. Churchilla 4, 130 67 Prague 3, Czech Republic; Czech Statistical Office, Na padesátém 81, 100 82 Prague, Czech Republic, e-mail: ondrej.nyvlt@czso.cz.

Non-formal education covers all taught organised learning activities outside the regular education system. A non-formal learning activity is defined as being organised like a course, conference or seminar for which the interviewee has applied and has participated in. It could be for a short or a longer period with possible minor breaks (EU Labour Force Survey, 2011).

1 METHODOLOGY

Sample surveys are usually connected with sampling and non-sampling errors. The latter are a result, for instance, of administrative drop-outs of dwellings out of the sample, intentional non-response or errors produced by filling in the questionnaire (Employment and Unemployment in the Czech Republic, 2011). With these errors, one cannot determine the deviation of estimate without a rather wide knowledge of the population. On the other hand, the sampling errors which arise by applying characteristics of the sample to the population can be interpreted by means of confidence intervals. The confidence intervals are intervals built around the point estimate in such a way that there is a certain probability that the value of an estimated characteristic is just within this interval. The most widely used is a 95% confidence interval, i.e. an interval within which the actual value of the estimated characteristic is found with 95% probability.

From another point of view it is possible to divide errors into observational and non-observational ones. Observational errors are deviations of the answers of respondents from their true values from the measure; for our purposes, these are measurement errors. Observational errors are conveniently categorized according to different sources – the interviewer, the respondent, the questionnaire, and the mode of data collection. Interviewer's errors are associated with effects on respondents' answers stemming from different ways in which interviewers administer the same survey (Measurement Errors in Surveys, 1991). But once we begin to consider the possibility of alternative wordings, we seem to presume that there is something of which they are alternative versions: an abstract question that we can ask in different ways (The Psychology of Survey Response, 2000). The respondent error shows that different respondents have been found to provide data with different amount of error, because of different cognitive abilities or differential motivation to answer the questions well (Measurement Errors in Surveys, 1991). The meaning that we get from a word or a sentence must be relatively stable across people (The Psychology of Survey Response, 2000). The interpretation of a sentence has to be at least somewhat immune to differences in the amount of knowledge about the concepts.

Non-observational errors are those arising because of measurement not taken on part of the population and there are viewed as arising from three sources – coverage, non-response and samplings (Measurement Errors in Surveys, 1991). The sample in the Labour Force Survey is based on the evidence of dwellings in the area of the Czech Republic where only registered flats are involved. The problem is that collective households (e.g. lodging house, old people's home) have a totally different demographic and social structure than registered flats (higher number of pensioners and foreigners), so the coverage has a great impact on the basic results of LFS. According to the statistical law, the citizens of the Czech Republic can refuse to participate in household surveys and the non-response significantly influences the quality of LFS data. This description shows that sampling error is only a part of the measurement errors in survey sample.

It is important to determine the time period. However, many surveys items depend on both the interview time and event time. These items include the ubiquitous questions about whether an event happened in the last few weeks, months, or some other time interval. It is necessary to search in memory for the event information and to locate this information with respect to the bounds of the reference period (The Psychology of Survey Response, 2000).

2 DATA

The Adult Education Survey (AES) should take place every 5 years starting from 2011 and is designed to give detailed information on the participation of individuals in education and training activities.

A pilot survey was held in 2007. The reference period for the participation in education and training activities is twelve months prior to the interview. In the Czech Republic the collection of data was implemented in the second half of 2011 on the sample of 9 500 households. Interviewers asked all persons in households at the age of 18–69 years.

The basic source for providing data about non-formal education is now the Labour Force Survey. The LFS is a continuous survey whose results are assessed on quarterly basis. Quarterly, the sample comprised almost 25 thousand dwellings in the area of the Czech Republic (0.6% of all dwellings permanently lived in), including more than 50 thousand respondents aged 15 and over. This sample size allows making estimates of labour market indicators as well as the educational indicator (formal or non-formal education). Weight adjustments have been made to all the sample data according to the age, gender and region structure of the population based on mid-quarter population projection for the Labour Force Sample Survey (Employment and Unemployment in the Czech Republic, 2011).

3 DESCRIPTION OF THE PILOT SURVEY OF NON-FORMAL EDUCATION

Labour Force Survey is a relatively time demanding survey and is primarily focused on data on economic activity or inactivity. The question about non-formal education is at the end of the questionnaire. In addition to that, the determination of non-formal education is a very complicated methodological process. If we express such a complicated process only by one question, it can lead to a relatively high underestimation. In this case it is necessary to describe the non-formal education in a more detailed way.

The pilot survey was conducted on a sample of the standard LFS module in all dwellings in the 4th quarter of 2011. For the 4th quarter of 2011, three basic questions about participation in non-formal education were implemented. Before the 4th quarter of 2011, the information on variable participation in non-formal education was collected only from one question.

Table 1 Question on reference period in non-formal education (standard LFS)

| Filter | Name of question | Question and answers | Code |
|--------|------------------|---|--------|
| 15–69 | VzdNform | Did s/he attend education within non-formal education (in last 4 weeks)? Yes No | 1 2 |

Source: Czech Republic, LFS

In the 4th quarter of 2011 the first question was aimed at any courses, the second on any seminars or workshops and the third on private lessons outside the formal education system. When the person at least once answers the question Yes, this person is afterwards considered as person who attended non-formal education. The questions are associated with a reference period of fourth weeks, both in the pilot survey and standard questionnaire of LFS.

Table 2 Questions about participation in non-formal education (pilot survey, 4th quarter 2011)

| Filter | Name of question | Question and answers | Code |
|--------|------------------------|---|--------|
| 15+ | H0001 (NFECOURSE) | Within the last 4 weeks did s/he attend any courses? Yes No | 1 2 |
| 15+ | H0003 (NFEWORKSHOP) | Within the last 4 weeks did s/he attend any seminars or workshops? Yes No | 1 2 |
| 15+ | H0005 (NFELESSON) | Within the last 4 weeks did s/he receive private lessons outside the formal education system? Yes No | 1 2 |

Source: Czech Republic, LFS

The important factor of a survey is to define the reference period. In case of non-formal education it is important because the respondent sometimes forgot during the time, that he has attended the non-formal education. Within the framework of the pilot survey one question about attending education within non-formal education in last 12 months was added. This reference period allows the comparison with reference period in AES.

Table 3 Questions about reference period in non-formal education (pilot survey, 4th quarter 2011)

| Filter | Name of question | Question and answers | Code |
|---------------------------------------|------------------|--|--------|
| H0001=2 and H0003=2 and H0005=2 | H0009 | Did s/he attend education within non-formal education (in last 12 months)? Yes No | 1 2 |

Source: Czech Republic, LFS

4 MAIN RESULTS

Generally, it is necessary to define the framework of a sampling and non-sampling error in a survey. The sampling error can be interpreted by means of 95% confidence intervals for the estimate.

The confidence intervals are calculated for the sample size in a given quarter. In order to calculate confidence intervals of aggregates the following formula should be used for the basic aggregate (Employment and Unemployment in the Czech Republic, 2011).

$$95\% \text{ C.I. of estimate } Y = y \mp 1.96 \times s_y, \text{ where } s_y \cong N \times \sqrt{(1-f) \cdot \frac{\frac{y}{N} \times (1 - \frac{y}{N})}{f \times N}}, \quad (1)$$

where: N is the size of the population,

y is the estimate of aggregate Y in the population,

f is the respective proportion of sample, $f = n / N$.

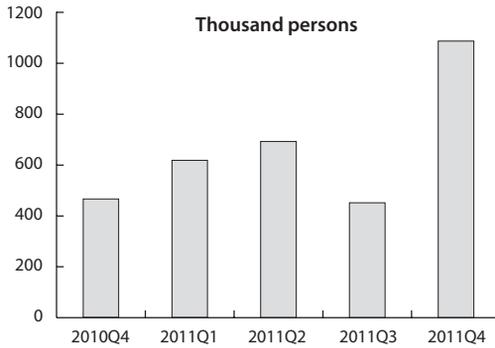
When substituting the variables into the above-mentioned formula we get the confidence intervals of number of participants in non-formal education. The resulting 95% confidence interval for the estimate of the number of persons attending non-formal education in 2nd quarter of 2011 is approx. 684.4 +/- 20.9 thousand, i.e. there is a 95% probability that the actual number of participants in non-formal education in the Czech Republic was not below 663.5 thousand and not above 705.3 thousand. The resulting 95% confidence interval for the estimate of the number of participants in non-formal education in 4th quarter of 2011 is approx. 1 127.6 +/- 26.0 thousand, i.e. there is a 95% probability that the actual number of participants in non-formal education in the Czech Republic was not below 1 101.6 thousand and not above 1 153.6 thousand.

It is a fact that sampling error itself cannot explain differences between the 2nd quarter of 2011 and the 4th quarter of 2011. It is clear that asking three separate questions about non-formal education instead of one has a significant impact on data and time-series. Such change in methodology influences the respondent's ability to better recall previous education activities. Therefore, the number of non-formal education participants in the 4th quarter of 2011 more than doubled in comparison with the 2nd quarter of 2011.

² A course is defined as "a planned series of single learning activities in a particular range of subject-matters offered by a provider". Courses are typically subject oriented and they are taught by one or more persons specialised in the field(s) of education and training. They may take place in one or more settings/environments.

³ Sessions combining theoretical instruction with "hands-on" training provided during a conference or congress.

Figure 1 Non-formal education, population 15–69, Czech Republic



Source: Czech Republic, LFS

5 REFERENCE PERIOD

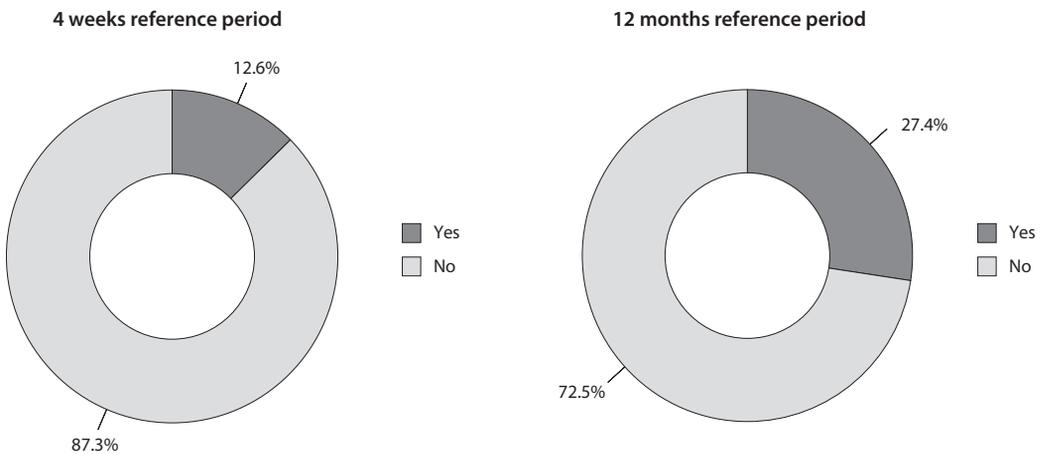
The reference period in standard LFS (the last 4 weeks) is defined differently than in AES (the last 12 months). From this perspective, the impact on the total number of participants in non-formal education was, therefore, determined.

On average, more than 1 out of 10 respondents participated in non-formal education in the last 4 weeks. Considering the last 12 months, it was more than 1 out of 4. The effect of a different reference period is clear and obvious.

6 DETAILED RESULTS

The main results show underestimation of non-formal education from a standard questionnaire

Figure 2 Participation in non-formal education, reference period comparison



Source: Czech Republic, LFS

in the Labour Force Survey. From this result, we cannot obtain a detailed explanation of this underestimation. But it is clear that the underestimation is created mainly by non-sampling errors. A detailed analysis according to the regions of the Czech Republic shows the basic differences. During the year, there is not a preference of a specific mode of questioning and therefore the model does not count with the mode effect. The results express the bias from the part of the interviewer, respondent or questionnaire errors. For analytic purposes the comparison of the 2nd quarter and the 4th quarter of 2011 was used. The 3rd quarter generally is seasonally influenced by the summer holiday. In this quarter the number of participants within non-formal education is lower.

The interregional comparison shows the divergence in results. In some regions the differences were high, in others insignificant. Generally, in the Czech Republic in the comparison between the 2nd and 4th quarter of 2011, there was a double difference. In the 2nd quarter of 2011, of 9 013.0 thousand persons aged 15 or more 684.4 thousand attended the non-formal education, in the 4th quarter of 2011

of 9 013.5 thousand persons aged 15 or more 1 127.6 thousand attended the non-formal education. Relatively, in the 2nd quarter of 2011 the proportion of participants in non-formal education was 7.6%, in the 4th quarter of 2011 already 12.5%. We can compare relative values between the 2nd quarter of 2011 and the 4th quarter of 2011 according to regions (NUTS3). We express the comparison by index in percent. The high value of this index was in Jihomoravský region, (where the index was 298.8%), in Praha (212.1%) and in Středočeský region (198.0 %). On the contrary, the lowest value of index was in Karlovarský region (100.9%), then in Zlínský region (109.4%) and region Vysočina (117.8%). Differences between regions were huge and they indicate the problem of quantification of non-formal education in the Czech Labour Force Survey. The comparison between the regions of the Czech Republic was used because the organizational structure of regions is relatively independent of the central level. It is obvious from the results that each region approaches this problem differently. Different approaches influence the final results for the Czech Republic.

Table 4 Participation in non-formal education by region in the Czech Republic (2nd quarter 2011, 4th quarter 2011)

| Region | 2011q2 (1 question) | | | 2011q4 (3 questions) | | |
|-----------------|---------------------|---------------------------------------|-------------------|----------------------|---------------------------------------|-------------------|
| | Total | Participation in non-formal education | | Total | Participation in non-formal education | |
| | | Absolutely (thousand) | Relatively (in %) | | Absolutely (thousand) | Relatively (in %) |
| Czech Republic | 9 013.0 | 684.4 | 7.6 | 9 013.5 | 1 127.6 | 12.5 |
| Praha | 1 097.4 | 100.9 | 9.2 | 1 099.9 | 2 14.5 | 19.5 |
| Středočeský | 1 073.0 | 79.4 | 7.4 | 1 079.5 | 158,1 | 14.7 |
| Jihočeský | 545.6 | 36.6 | 6.7 | 545.2 | 60.2 | 11.0 |
| Plzeňský | 491.2 | 51.8 | 10.5 | 491.4 | 65.4 | 13.3 |
| Karlovarský | 262.0 | 20.7 | 7.9 | 261.5 | 20.9 | 8.0 |
| Ústecký | 706.8 | 62.7 | 8.9 | 705.8 | 87.0 | 12.3 |
| Liberecký | 373.4 | 27.8 | 7.4 | 373.3 | 43.3 | 11.6 |
| Královéhradecký | 473.6 | 40.7 | 8.6 | 473.0 | 66,8 | 14.1 |
| Pardubický | 440.9 | 39.3 | 8.9 | 440.7 | 55.6 | 12.6 |
| Vysočina | 439.4 | 36.3 | 8.3 | 438.8 | 42.7 | 9.7 |
| Jihomoravský | 991.4 | 31.1 | 3.1 | 990.9 | 92.9 | 9.4 |
| Olomoucký | 548.9 | 30.4 | 5.5 | 547.8 | 44.8 | 8.2 |
| Zlínský | 506.8 | 40.9 | 8.1 | 505.7 | 44.7 | 8.8 |
| Moravskoslezský | 1 062.6 | 85.9 | 8.1 | 1 059.9 | 130.8 | 12.3 |

Source: Czech Republic, LFS

These results indicate that the original model of the questionnaire (before 4th quarter of 2012) was inconvenient due to its vagueness. The original model did not specify the individual forms of non-formal education. The total length of the questionnaire of the Czech Labour Force Survey causes generally simplifying of questions which are methodologically very difficult. Furthermore, the situation with a relatively long questionnaire can often lead to satisfaction with a simple answer No in the question about non-formal education. When the question about non-formal education is asked in an exhaustive form it is possible that the original answer No will be changed to Yes. In this case we can talk about a combination of an interviewer and respondent error. This result is documented in Table 5. In this table 95% confidence intervals to estimates of non-formal education for individual regions are calculated. It is necessary to say that in case of regions (NUTS3) the estimates of non-formal education are relatively low. For this reason the sampling errors are relatively higher than in case of the whole Czech Republic. In five regions are relative 95% confidence intervals to estimates of non-formal education higher than 10%, for the Czech Republic it is only 2.3%. In spite of the exception of two regions (Karlovarský and Zlínský region), there is an absolute and relative difference between the 2nd quarter of 2011 (one question) and the 4th quarter of 2011 (three questions) higher than 95% confidence intervals to es-

timates of non-formal education. For example in Olomoucký region the difference between the 2nd and 4th quarter of 2011 was 14.4 thousands persons (47.6% growth), 95% confidence intervals to estimates of non-formal education was only 5.6 thousands (relatively 12.6%). From these results it is obvious that non-sampling error plays an important role in the interpretation of the results of non-formal education from the Labour Force Survey data.

Table 5 Differences and 95% confidence intervals to estimate the number of participations in non-formal education in the Czech Republic (in thousands, in %)

| Region (NUTS3) | Growth (2011q4, 2011q2) – absolutely | Growth (2011q4, 2011q2) – relatively | 95% confidence interval (abs. +/-) 2011q4 | 95% confidence interval (rel. +/-) 2011q4 |
|-----------------|--------------------------------------|--------------------------------------|---|---|
| Czech Republic | 443.2 | 64.7% | 26.0 | 2.3% |
| Praha | 113.6 | 112.1% | 14.3 | 6.8% |
| Středočeský | 78.8 | 98.0% | 9.4 | 6.0% |
| Jihočeský | 23.5 | 64.4% | 5.3 | 8.8% |
| Plzeňský | 13.6 | 26.2% | 5.5 | 8.5% |
| Karlovarský | 0.1 | 0.9% | 2.9 | 14.1% |
| Ústecký | 24.3 | 39.0% | 8.2 | 9.5% |
| Liberecký | 15.5 | 55.9% | 4.8 | 11.1% |
| Královéhradecký | 26.1 | 64.2% | 6.4 | 9.6% |
| Pardubický | 16.3 | 41.6% | 5.3 | 9.6% |
| Vysočina | 6.4 | 17.8% | 4.6 | 11.0% |
| Jihomoravský | 61.8 | 198.8% | 7.9 | 8.4% |
| Olomoucký | 14.4 | 47.6% | 5.6 | 12.6% |
| Zlínský | 3.7 | 9.4% | 5.1 | 11.3% |
| Moravskoslezský | 44.9 | 52.7% | 9.0 | 7.0% |

Source: Czech Republic, LFS

CONCLUSION

The pilot project provided much more detailed information on participation in non-formal education than standard core LFS. It proved the fact that a change in methodology leads to different, better quality results. The pilot survey indicated that the indicator of non-formal education is significantly underestimated. This error can be due to the interviewer, respondent or methodological side. The interviewers often do not gain relevant information on participation in non-formal education. Furthermore, the respondents react negatively to the survey and reduce their answers. Finally we can say that sampling error does not play as important role as a non-sampling error, and this conclusion is valid even in the regions (NUTS3) of the Czech Republic. The pilot survey indicated a possible solution to the problem. Based on the output of the pilot survey, the wording was changed for the variable expressing non-formal education as follows:

Before the pilot (2011Q3):

Did s/he attend education within non-formal education (in last 4 weeks)?

After the pilot (2012Q1):

Did s/he attend course, seminar, workshop or private lesson within the last 4 weeks?

References

- Analysis of Survey Data*. Southampton: University of Southampton, 2004.
- CZSO. *Employment and Unemployment in the Czech Republic as Measured by the Labour Force Sample Survey* [online]. Prague: Czech Statistical Office, 2011. <<http://www.czso.cz/eng/edicniplan.nsf/aktual/ep-3?opendocument#31>>.
- CZSO. *News Releases Employment and Unemployment* [online]. Prague: Czech Statistical Office, 2011. <<http://www.czso.cz/csu/csu.nsf/engkalendar/2011-zam>>.
- EUROSTAT. *EU Adult Education Survey. Explanatory Notes*. Brussels: Eurostat, European Commission, 2012.
- EUROSTAT. *EU Labour Force Survey. Explanatory Notes*. Brussels: Eurostat, European Commission, 2011.
- Measurement Errors in Surveys*. New Jersey: John Wiley & Sons. 2004.
- The Psychology of Survey Response*. Cambridge: Cambridge University Press, 2005.
- THOMPSON, S. K. *Sampling*. New York: Pennsylvania State University, 2002.