

**QUALITY REPORT**  
**on the Structure of Earnings Survey 2002**  
**in Hungary**

Compiled by: György Lázár

Head of Labour Market Information Department

in the National Employment Office

with the contribution of:

Ágnes Engler Schödel senior statistician and

Sándor Grad senior IT specialist

Phone: 36-1210-4032

Fax: 36-1-210-1129

e-mail: [LazarG@lab.hu](mailto:LazarG@lab.hu)

[GradS@lab.hu](mailto:GradS@lab.hu)

[SchodelA@lab.hu](mailto:SchodelA@lab.hu)

Attachments:

Attachment A: Tabular analyses

Attachment B: Coefficients of variation

Attachment C: Coherence of SES-2002 with LFS-2002

This report is based on Commission Regulation (EC) No 72/2002 of 16 January 2002, implementing Council Regulation (EC) No 530/1999 as regards quality evaluation of structural statistics on earning

## **Preface**

Hungary has a long tradition (almost 30 years) in conducting similar surveys. Since 1992 it has been yearly survey with the reference month of May.

Part of this tradition is that this survey belongs to the Ministry of Labour in Hungary and it is conducted by a background institution of the Ministry (and not the Central Statistical Office). This institution is now called the National Employment Office which is the headquarters of the Public Employment Service of Hungary and – together with the PES – belongs to the Ministry of Employment Policy and Labour.

The main uses of the survey:

- It serves as a common database for the social partners participating in the wage bargaining process;
- The government and all the interested parties can analyze the wage rates of different groups of employees;
- It helps to prepare important decisions concerning the development of different wage scale systems (e.g. for public employees and civil servants);
- It provides a sound basis for model computations to prepare minimum wage decisions and other wage agreements on the macro and medium levels;
  
- Some international statistical obligations can be fulfilled only from these databases (like the October Inquiry of ILO and some Eurostat and OECD data requests);
- It is an extremely good basis for many different research projects.

Its use for the wage negotiations explains why it is so important to keep May as the reference month: it makes possible the use of the latest figures as early as in October and November of the same year to the wage agreements for the next year.

It was very plausible that the requirements of SES 2002 can be fulfilled by some modifications and the development of our traditional survey.

The most important modifications were:

- The scope of the survey was extended to non-profit organizations;
- It was also extended to part-time workers (earlier they were excluded);
- Some new variables were introduced including working time, overtime, length of service, paid vacation etc.

The Hungarian authorities decided to fulfill all the requirements of Council Regulation (EC) No 530/1999 and Commission Regulation (EC) No 1916/2000. Two problems remained:

- With the reference month of May it is not possible to ask the total yearly earnings in the current year. The solution: to use the non-regular payments of the previous year;
- The main activity of the local units can not be perfect (if it is different from the NACE code of the whole company), because the business register of the CSO does not contain the local units by their own main activity.

During the preparations of SES 2002 we discussed these questions with the international experts provided by Eurostat and made the following decisions:

- As we have a yearly survey, with the next year's survey (May 2003) we shall collect the non-regular payments of 2002, so we can try to exchange the non-regular parts of earnings in 2001, with the new figures for SES 2002 by matching individual persons or very small groups in the May 2002 and May 2003 surveys;
- With the help of CSO and by asking the local units themselves we shall try to identify the true main activities of the local units in the 2002 survey, and the CSO promised to solve this problem before the next survey (SES 2006).

The results of these will be discussed in Annex # 1 of this quality report.

## **Part A of the quality report**

All the required grossed up results for tabular analyses are attached in Attachment A.

The distribution of full-time (FT) employees and part-time (PT) employees are provided separately according to all the required breakdowns, except the breakdown by NUTS level 1, which is not applicable for Hungary.

Distribution of employees by	Full-time	Part-time
	employees	
- band of hourly gross earnings and by sex:	Table 1 FT	and Table 1 PT
- band of monthly gross earnings and by sex:	Table 2 FT	and Table 2 PT
- band of annual gross earnings and by sex:	Table 3 FT	and Table 3 PT
- band of annual holidays and by sex:	Table 4 FT	and Table 4 PT
- band of monthly hours paid and by sex:	Table 5 FT	and Table 5 PT
- NACE Rev. 1 section and by sex:	Table 6 FT	and Table 6 PT
- ISCO-88 at the 1 digit level and by sex:	Table 7 FT	and Table 7 PT
- ISCED 0 to 6 and by sex:	Table 8 FT	and Table 8 PT
- age band and by sex:	Table 9 FT	and Table 9 PT
- length of service and by sex:	Table 10 FT	and Table 10 PT
- size of enterprise in terms of no of employees:	Table 11 FT	and Table 11 PT
.		
.		
.		
.		
.		
.		
.		

## Part B of the quality report

### 1. Relevance

The main objective of the survey is to create a common database for the social partners for the reconciliation of interests. It means that the main users are the three parties participating in wage negotiations on the macro level, i.e. the representatives of the government, the employees and the employers.

The biggest user is the **Ministry of Employment Policy and Labour**, but most of the computations made for the Ministry serve the other two parties at the same time.

The main uses are:

- Analyses of wage rates
- Minimum wage computations
- Model computations for the development of wage scale systems to determine the expected impacts of the different modifications

Another important user is the **Central Statistical Office**. It takes over the whole database and uses it to fulfill different international statistical reporting obligations. The **County Labour Centres** use the average earnings by occupations to determine the unemployment benefit in special cases and to pass judgment on work permit applications by foreign citizens.

The **National Directorate of the Pension Fund** also uses the average earning by occupations to determine pensions in special cases according to the international agreements.

The **juries of justice** also use the data in cases of claims for damages.

**Organizations of employees and employers:** they can also influence the contents of the survey, their opinion is important and taken into consideration in evaluating the results of the survey.

There is a huge set of standard tables, agreed upon by the three parties, which is given for all participating confederations every year, free of charge.

In some cases the special requests of trade unions and employers' organizations can also be satisfied using this database, provided that data protection is respected.

**Other requests:** universities, research institutions, students working on their thesis, embassies, individual companies, etc. Most of these are fulfilled free of charge, in some cases with a solid reimbursement of costs.

In most cases we can satisfy the specific requests, the only occasional limits are lack of capacity and the rules of data protection.

## 2. Accuracy

### 2.1 Sampling errors

#### 2.1.1 Probability sampling

Two different methods are used for sampling

- All employers over 50 employees are obliged to report about a sample of their employees: those who were born on specific dates in any

month, in any year.\* (One date means a 3,29% perfect probability sample, two dates 6,57% and three dates 9,86%)

- From those employers who have less than 50 employees we choose a 20% random sample using the business register of the CSO, and those who were chosen have to report about each of their employees.

For budgetary institutions which report individually, the method is the same, but most of these institutions are incorporated in different centrally handled payroll systems and for these institutions we get all the necessary data of each individual employee. (Now about 70% of public employees and 50% of government employees are reported this way, not only for the samples.)

#### 2.1.1.1 Bias

Due to the probability sampling methods, there is no systematic bias in the estimations.

#### 2.1.1.2 Variance

The coefficients of variation for monthly earnings of full-time and part-time employees, according to the breakdowns required were determined and they are given in Attachment B (Table 12).

---

\* We use three birthdates: the 5th, 15th and 25th day of any month in any year. Three times twelve days in a year results in a  $36/365=0,0986$  sampling ratio.

The coefficients of variation for monthly hours of full-time employees, according to the required breakdowns, are also given in Attachment B (Table 13).

### **2.1.2 Non-probability sampling**

This is not applicable in our case.

## **2.2 Non-sampling errors**

### **2.2.1 Coverage errors**

Possible under-coverage may happen because of new units: birth, mergers and demergers of old units. To avoid this to the possible rate, we use the latest, most up to date version of the sampling frame, the business register of the CSO.

Under- or over-coverage may happen because of misclassification by bands of the number of employees. In most cases it does not cause any problem, because for grossing up we use the actual total number of employees, given on the cover-sheet of the report by each employer. (This grossing up method is used for firms over 50 employees.)

Misclassification can also happen if the main activity (NACE code) of the local unit is wrong. We tried to correct these, but probably some misclassifications remained. For the SES 2006 hopefully they can be corrected with the development of the frame, the CSO business register.

### **2.2.2 Measurement errors**

In our case the measurement errors are equal to the reporting errors. The most important sources of these are:

- The misinterpretation of the questionnaire (the reason in most cases is that the respondents do not read the instructions);
- Erroneous coding of the identifier of the firm or the occupations
- Data entry errors on the respondents' side (if they use the data entry program which anybody can use taken from the internet);
- Data entry errors during the central data entry of reports sent on paper;
- Possible errors during data transmission or transformation;
- Possible errors during data processing at NEO;
- Possible errors during the transformation of national codes of education into ISCED codes.

Measurement errors are checked and amended through a lengthy and thorough editing process.

- missing variables: sometimes they are asked again from the respondents, in other cases they are imputed
- slipped fields: they can be detected through the checking process and in most cases they can be adjusted
- detecting and excluding duplicate reports or records
- checking the possible minimum and maximum values for each field containing elements of earnings
- checking the identifiers and code values, whether they are existing codes (using control tables and control digits): erroneous identifiers are amended using the register
- logical checks and amendments by comparing different fields within the individual records, like education, occupation and wage category.

A special feature of the survey in Hungary is that the majority of data in the budgetary sector come from central payroll computing systems. It means that theoretically measurement errors are not possible, unless some variables are missing from these central systems or there are errors in these systems.

These central payroll systems use the same terms and definitions that are determined and used by the Central Statistical Office for statistical purposes. (The CSO establishment based wage and earnings surveys are also based on these central payroll systems.)

At the time of the SES 2002 survey the variable of paid holidays was missing from these payroll systems. The problem was solved by imputing these figures for each individual by using the length of holiday determined in the law for each category. This deficiency will be amended by the time of the 2006 SES.

### **2.2.3 Non-response errors**

The survey is part of the yearly National Programme of (obligatory) Statistical Reports for all firms over 50 employees, for all budgetary institutions and for those firms in the competitive sector with less than 50 employees which were chosen for the sample.

In spite of obligatory reporting, some of the bigger companies do not respond to the survey. (Each gets the questionnaires and the instructions by mail, escorted with a letter and some of the results of the previous survey.)

By using arrival lists we try again the most important respondents by calling them on the phone. In most cases we succeed with the bigger firms. But some remain missing from the survey.

The lack of capacity limits these efforts.

For smaller firms, especially under 20 employees, the non-response rate is much higher, but it can be partly corrected through using a different grossing up method. (Where the 20% sample is chosen from among the complete list of the smaller companies, the grossing up factors are determined for the individual size bands by using the total number of employees in the appropriate size band according to the CSO'S establishment based labour statistics.)

The response rate in the competitive (enterprise) sector is changing from 90 to 62 per cent (from the biggest to the smallest companies). The smallest are often simply not found at their mail address officially given in the register.

In the budgetary sector the response rate is excellent (99,9%) in the case of local budgetary institutions, thanks to the well functioning central payroll systems. It is much smaller in the case of central budgetary institutions, because their separate payroll computing system is relatively new and in some cases the individual institutions were already not able to report, and the central system was still not able to fulfill the reporting for them. (It will certainly improve substantially by the time of SES 2006.)

The non-response rate certainly causes biases concerning the total number of employees and their distribution by size bands, but these biases are much smaller concerning the estimates of monthly or annual earnings or for hours paid, which are the most important results of the survey.

#### **2.2.4 Model assumption errors**

It was already explained above why the month of May is used for the Hungarian survey.

The month of May is reasonable, because in most cases the wage raises of the year are implemented before May. The length of this month is somewhat bigger in the number of normal working hours than the average month, but this difference can be corrected during the data processing.

There is no need to adjust the fiscal year to the calendar year, because in Hungary the fiscal year is the same as the calendar year.

In Hungary our survey covers all sections of NACE Rev. 1, including sections A, B, L, and it covers all firms over 4 employees (including the 5-9 size band).

For the data base sent to Eurostat we left out those sections which are excluded from the SES 2002 survey (sections A, B and L).

The combination of administrative sources (central payroll systems) and the representative survey do not cause any problem because these systems use the same statistical concepts and definitions (determined by the Central Statistical Office) which are used in the survey.

### **3. Timeliness and punctuality**

The reference month of the Hungarian survey is the month of May.

According to this, key data collection dates are as follows:

- the questionnaires and instructions are sent out to the respondents at the end of April
- the deadlines to send back the questionnaires:
  - = for companies with less than 300 employees: 5<sup>th</sup> July 2002
  - = for companies over 300 employees: 12<sup>th</sup> July 2002
  - = for all budgetary institutions: 17<sup>th</sup> July 2002
- post collection phases:
  - = receiving the reports, manual checks, recall of missing responses using computerized arrival lists: continues until the end of August
  - = data entry and compiling the reports received via e-mail and on CDs: parallel with the previous phase
  - = logical checks, data corrections using the phone where programmed corrections are not possible: September – October 2002
  - = data entry for late arrivals and data corrections: October 2002
  - = quality check before data processing: November 2002
- Preliminary results given to the social partners to use during wage negotiations: December 2002
- Final results: February 2003 (on CD)
- Four volume publication: August 2003  
(For the experiment to change the non-regular payments of 2001 with the non-regular payments of 2002 for Eurostat we had to wait until February 2004, when our 2003 survey results were finalized. The results of this experiment are described in Appendix #1.)
- The experiment was finished in May 2004
- Transforming the database of SES 2002 according to the requirements: July 2004
- Sending the database to Eurostat through the CSO): August 2004 (the second and final version)

- Data processing for the quality report: October-November 2004 (in cooperation with the CSO, concerning the LFS tabulation for the cohesion study)
- Quality report writing: December 2004
- Translating, checking, correcting the report: January 2005
- Sending the quality report to Eurostat: February 2005

The reasons for being late with the data transmission and with the quality report are as follows:

- a) The experiment mentioned above and described in detail in Annex #1 was quite time consuming
- b) In 2003 and 2004 we had to implement our May 2003 and May 2004 surveys with the same small team which was responsible for preparing the data transmission and the quality report of SES 2002
- c) All correspondence, data transmission, error messages and exchange of information between NEO and Eurostat go through the Central Statistical Office. We understand that Eurostat considers the National Statistical Institutes as their only partners in the Member States (and earlier in the Candidate Countries), but it has to be taken into consideration that this procedure is time consuming, causes delays and sometimes misunderstandings too.

Maybe this is a special case with SES 2002 in Hungary, that it is not the CSO which is responsible for the survey, but the National Employment Office.

We hope that this special situation will be taken into consideration for SES 2006.

#### **4. Accessibility and clarity**

- a) For everybody
  - Some of the most important results are accessible for everybody, free of charge, on the website of the Hungarian Public Employment Service ([www.afsz.hu](http://www.afsz.hu))
  - Some of the aggregated results are sent out to the respondents of the new survey together with the questionnaires of the next year's survey
- b) For the interested ministries and confederations of workers and employers organizations
  - A menu driven collection of all "standard" tables on CDs, agreed with the representatives of social partners (the required tables by sections, wage categories, occupations, education, etc. can be easily found and printed if required)

- the results of model computations to prepare negotiations on minimum wage decisions and macro level wage agreements
- c) For selected important users of the results (social partners, Directorates of CSO, County Labour Centres, Universities, Research Institutes, ministries, libraries) a four-volume publication every year on basic wages and earnings (cca 4\*200 pages)
- About the Hungarian economy as a whole
  - About the competitive (enterprise) sector
  - About the budgetary sector
  - By regions (counties, NUTS 3 level)
- d) Users of the whole database
- Each year the CSO takes over the whole database to fulfill some important international reporting obligations (ILO, Eurostat, OECD)
  - Some Hungarian research institutes take over the whole database to use them for research purposes (without the individual identifiers of the employers)
- e) Special requests  
Trade unions, companies, potential investors etc. may order for special data processing on a cost reimbursement basis taken into consideration the rules of data protection.
- f) Methodological documents  
Methodological comments and explanations are given to the users with the CD containing the standard tables and with the four-volume publication of basic wages and earnings.  
The four-volume publication and the CD will be sent to Eurostat separately (via surface mail).

## **5. Comparability**

### **5.1 Geographical comparability**

The national concepts, terms and definitions are equivalent with the European ones.

The industrial classification is equivalent with NACE generally used in the EU.

The Hungarian classification of occupations (FEOR-93) is similar to ISCO-88.

The structure and principles are the same, but they are equivalent only on the 1<sup>st</sup> digit level. Transcoding is possible on the two and three-digit levels and not possible on the 4-digit level.

There is no generally used classification of education in Hungary, but the level of education may be transcoded to ISCED (0 to 6).

The statistical unit in this survey is the local unit.

The industrial classification comes from the register of employers taken care of by the Central Statistical Office (KSH). The quality of the business-register is the responsibility of CSO. It is updated monthly, but this a very difficult task in the case of small enterprises (with less than 20 employees). The birth of the new companies can be followed easily (because each of them have to ask for an identifier), but it is much more difficult if they cease to exist. Up till now the business register does not contain the local units with their own (in some cases different) NACE codes.

We do hope that for the time of SES-2006 this problem will be solved.

## **5.2 Comparability over time**

There were no previous structure of earning surveys in this form in Hungary. Our yearly survey (since 1992) was similar, but not completely equivalent. Earlier we did not collect wages and earnings of part-time workers and some variables were missing from the surveys before 2002. So the earlier results are not fully comparable with the results of SES-2002. (See in Attachment C.)

## **6. Coherence**

### **6.1 Coherence with the distributions of employees in the labour force survey in the same reference period**

In close cooperation with the CSO we prepared the same tables on the distributions of employees from LFS and SES-2002.

The CSO made the tabulations for the whole year of 2002 and the SES tables contain the distributions in May 2002 (which is the reference period of our survey).

The tables contain the following:

- The number and distribution of employees by NACE sections, for the required age-groups and for both genders

- The number and distribution of employees by major occupational groups (ISCO-88), for the required age-groups and for both genders
- The number and distribution of employees by levels of education (ISCED 0 to 6), for the required age-groups and for both genders<sup>1/</sup>

All the tables were made for all employers and separately for full-time and part-time workers, both from the LFS and from the SES.

We can find quite substantial differences between the distributions of employees in SES and in LFS.

There are of course some obvious reasons: the different statistical units, data collecting methods, sampling ratio, coverage etc., but in spite of all these factors it is surprising that the total number of employees is higher by almost 50% in the LFS than in SES-2002. It can be explained only partly with the fact that the smallest employers (with 1-4 employees) are missing from SES.

We are planning to analyze the reasons of such big differences in greater detail together with the experts of the CSO.

---

<sup>1/</sup> The ISCED codes used in the SES tables were more detailed than those used for LFS tabulation.

## Annex #1

### A) The non-regular payments

One disadvantage of the fact, that we use in our SES the month of May as the reference period, that in the middle of the year we can not ask the total yearly non-regular payments in the same year. So we always have to ask the total yearly non-regular payments paid for the same individual in the previous year.

During the preparations of SES 2002, we decided to make an experiment. As we have a yearly survey, we collected the non-regular payments of 2002 with our next survey (SES 2003). The essence of the experiment was, that we tried to exchange the non-regular payments of 2001 (what we collected with SES-2002), with the non-regular payments of 2002, collected one year later.

We knew that this is not an easy task, because in our survey the individual employees can not be identified because of the strict data protection rules.

So we decided, that for the 2003 survey we shall not change the sample of the 2002 survey and we shall try two different methods.

- a) First we tried to match the individual records in the two surveys by using several important characteristics of the individuals, (like sex, age, education, occupation) working at the same company or local unit.

The rate of success of finding the pairs at the same individuals in the two surveys was very low, for many reasons:

- in some cases the company was not found in one of the two surveys
- same individuals leave the company before the next years survey and others are new in the second one
- occupations, positions and educational levels are not necessarily coded exactly the same way in the two years: in some cases they actually may change and in others they were just coded differently
- in the case of small companies in spite of using the same samples, part of them which reported in the first year, can not be found in the second.

At the end, we had to accept, that this method can not be used for our purposes.

- b) The second experiment was to make many small groups (by age, gender, occupation, education, section etc.) and to determine indexes to multiply the non-regular payments of each individual who belongs to the same group with the same index determined earlier for each group.

We tried out many different methods and levels of grouping the individuals, but none of these resulted in satisfactory figures for the indexes. (The dispersion of the indexes was so wide, that we were not able to accept this method as well.)

At the end we decided to use the original data of the SES-2002 (with the non-regular payments for 2001) because after these experiments and trials we were convinced that this will cause much smaller errors than any other method.

The reasons: the non-regular payments (elements of earnings not paid in every month) give only about 3-10% of the total earnings of the individuals. The use of the amounts paid to the same individuals in the previous year may cause a 5-10% smaller amount (mostly because the inflation), which means 0.15-1.0% reduction in the total earnings for 2002, what is much smaller in most cases than the sampling error for any specific small group, and it is about 0.5% for the whole population.

Taking into consideration that the direction of this error must be the same for each group (because of the inflation rate) the ratios of earnings between sections, occupational groups etc. are shown correctly.

For SES-2006 this type of error will be even smaller because the inflation rate is going down.

## **B) The main activity of the local unit**

The Hungarian Central Statistical Office in its business register does not deal with the local units (up till now).

So, as the sampling frame was this business register, we were not able to determine the main activity of the local units, especially if it is different from the main activity of the whole company.

In our survey we are collecting the information by local units, what is very important from the point of view of regional statistics. So all the data of wages and earnings of individuals will appear at the right region or even at the right settlement.

The only problem is the actual main activity of these local units, which may be different from the main activity of their whole company.

During the SES-2002, we tried to identify the actual main activity of each local units (with the help of CSO), but we can not be sure that this was successful in each and every case.

For the time being this was acceptable, but for the next survey (SES-2006) we have to find a better solution, with the help of the HCSO.

We do hope, that sooner or later the business register of the CSO will contain also the local units and their main activities. This is the only possible perfect solution of this problem.