



# Quality report of the 4th European working conditions survey

*Foreword*

*Introduction*

*Relevance*

*Accuracy*

*Comparability*

*Coherence*

*Annex 1: Sample Evaluation tests*

*Annex 2: Error rates as reported by the national institutes*

*Annex 3: Duration of Interviews*

*Annex 4: Edit rules*

*Annex 5: Questionnaire Concordance*

This report is available in electronic format only and has not been submitted to the standard Foundation editorial procedures.

# Foreword

The Foundation is strongly committed to providing consistent and reliable information about living and working conditions in Europe, in order to assist the European institutions and social partners, and feed the European debate on social and economic issues. For that reason, the Foundation puts a strong emphasis in the continuous improvement of our research methods, especially (because of its prominence) of the various Foundation surveys.

The latest edition of the European Working Conditions Survey involved a strong effort of planning, design and implementation of a new quality framework, and a considerable increase in the available documentation and methodological information about the survey. As part of this commitment to the quality and transparency of research, the Foundation commissioned the present report on the quality of the fourth European Working Conditions Survey, using the Eurostat framework of statistical data quality. This report is a valuable complement to the reports based on the results of the survey, allowing users to understand better the weaknesses and strengths of the data.

**Karppinen Jorma**  
Director

**Willy Buschak**  
Deputy Director

# Introduction

The European Foundation for Living and Working Conditions is carrying out three surveys, the European Working Conditions Survey, the Quality of life survey and the Establishment Survey. Of these three the EWCS has been implemented since 1991 at approximately 5-year intervals. EWC surveys provide an overview on the state of working conditions throughout Europe, and indicate the nature and content of changes affecting the workforce and the quality of work.

In order to be successful in its mission, the information that the Foundation provides to its users should be of high quality so as to form a basis of common understanding that facilitates efficient tripartite negotiations as well as policy making at European level.

During the design and implementation phases of the EWCS there are three layers of quality checks that ensure that data of high quality are collected and processed:

- Self checks by national companies that actually do the data collection. They need to conform to the current ESOMAR standards, additional (e.g. ISO) quality standards they have volunteered for, and also the requirements of the contract with the Foundation. They also need to protect their good name, something quite important in this industry.
- Checks by the coordinating organization that needs to make sure that all members of its consortium perform their tasks by the book.
- Checks by the Foundation or independent auditors hired for this purpose that evaluate the quality of delivered data and reports and may also perform on site audits (with or without prior notice) during data collection and processing.

This can be a quite effective and at times cumbersome quality control system that makes the data collection and processing to conform to the standards.

The ISO 9000 definition states that ‘quality is a composite of all the characteristics, including performance, of an item, product or service that bear on its ability to satisfy stated or implied needs’. In the European Union, Eurostat’s framework on statistical data quality is usually adopted and recognises six dimensions. This framework will be applied to the survey and its results:

- **Relevance:** the correspondence between the information conveyed by statistics and the information required by the users of statistics.
- **Accuracy:** the proximity of statistical figures to the true values they estimate.
- **Timeliness and punctuality:** timeliness is the proximity between the release date of statistics and their reference period, while punctuality is the conformance of the statistical authority to the scheduled release dates.
- **Accessibility and clarity:** accessibility refers to the physical conditions in which users can access statistics while clarity is the degree of the users’ facility to interpret the statistics.
- **Comparability:** the ability to compare statistics about the same characteristic between different points in time, geographical areas or statistical domains.
- **Coherence:** the ability to combine, in various ways and for different uses, similar statistics from different sources.

This document is a quality report which describes the different quality aspects of the 4th EWCS in 31 European countries based on reports dispatched from the coordinating organisation to the Foundation, on site audits and the data set itself.

The primary aim of this document is to describe in some detail these dimensions for the 4th EWC survey data which so that users can make the best use of data for their purposes.

The remainder of this document is organised as follows. Each section corresponds to one of the six quality dimensions and contains an introduction to the corresponding quality concept and a description of the related characteristics of the EB surveys.

# Relevance

Relevance is the degree to which statistics meet current and potential users' needs. It refers to whether all statistics that are needed are produced and the extent to which concepts used (definitions, classifications etc.) reflect user needs.

## EWCS users

**European Foundation Stakeholders.** The European Foundation for the Improvement of Living and Working Conditions is one of the first Community institutions established and its stakeholders (represented in the administrative board) include employers, EU policy makers, Governments, and trade unions. The main function of the Foundation is to provide with information and knowledge that assists policy makers and social partners in social dialogue and policy making.

**International organisations** (ILO, OECD, etc). International organisations are using EWCS data in combination with data from other regions into studies, assessments and data products with a wider geographical coverage.

**European organisations** (Eurostat, European Agency for Health and Safety at Work, European Parliament other European bodies). EWCS output (data and publications) are being used in a number of policy documents (evidence is collected by the Foundation). National Governments, evidence of some use of EWCS data for policy making especially for comparison and benchmarking.

**Non Governmental Organisations** (NGOs) also evidence of some use EWCS data to set up priorities in their agenda and support their arguments and advance their goals.

**Academia, Researchers** use EWCS data for carrying out independent research in the area. Researchers and students use the findings of the survey and they are usually interested in detailed data and metadata. They use the microdata that is made publicly available after some embargo time. Their use of data may lead to publications in refereed journals.

**Media, the general public** use the findings as an input to public dialogue and the democratic process. International or national media – specialised or for the general public – are interested both in figures and analyses or comments. The media are the main channels of statistics to the general public. EWCS data have fostered public dialog and debate on working conditions.

## User needs

The needs of EWCS users are included in the process of designing each implementation of the EWC survey during a series of workshops and expert group meetings and through the presence of stakeholders in the Administrative board of the Foundation.

This process to a large extent defines the content of the survey so that it is relevant to current and future needs. The main requirements of users in terms of methodology are:

- Provide results of high accuracy appropriate for policy making purposes at a European Level.
- Achieve sufficient accuracy for detailed figures (by country, occupation, economic activity etc.) and for gender disaggregating so that sizeable differences and trends can be identified. Achieve sufficient accuracy for the analysis of relationships between job characteristics, personal characteristics and outcomes so as to be able to identify substantial differences.

- Produce highly comparable data based on harmonised methodology across member states of the EU, candidate countries and other European countries.
- Adapt the content of the survey to the evolving needs of the Foundations stakeholders and other users while maintaining a core set of variables unchanged so that trends can be identified and estimated.
- Distribute anonymised datasets to interested researchers after a certain period of time needed to protect the confidentiality of respondents (embargo).
- Provide extensive documentation on the survey methodology and implementation including quality assessment so that users can make the best use of data for their purposes.
- Meet the aforementioned user needs with a specified budget.

## Examples of uses

### European Commission

Organisation/author	Report	Date	Analysis
Eurostat	Work and Health in the EU: A statistical portrait, 1994-2002	2004	Extensive analysis (especially by sector) based on 2000 WC survey data (EU-15) on ..  - Violence, intimidation and harassment in the workplace  - Information about workplace risks  - Risk factors for musculoskeletal disorders (MSDs)
European Commission DG Employment and Social Affairs	Social Situation in the EU, 2004	2004	Analysis of data from 2000/2001 surveys on workers reporting health at risk due to job

### International organisations.

Organisation/author	Report	Date	Analysis	Sample findings
OECD	<a href="#">Employment Outlook 2003</a>	2003	Analysis of WC survey data on life quality consequences of long hours / intense work	Increase in number of stress-related symptoms (eg. overall fatigue, stress, sleeping problems...) as longer hours worked.
OECD	<a href="#">Employment Outlook 2003</a>	2004	Multivariate analysis of WC survey data 2000/1 on the relationship between working hours and work-life balance	Work-family life conflict more acute for those working non-standard hours or with little control over work schedules

### Scientific publications

Organisation/author	Report	Date / Journal	Analysis	Sample findings
Kevin Daniels, Loughborough University (UK)	<b>Perceived risk from occupational stress, a survey of 15 European Countries</b>	<i>Occupational and Environmental Medicine</i> 2004;61:467-470	Analysis of WC survey data on stress factors	Cultural variability of stress: "Occupational stress is to some extent at least socially constructed"
Brendan Burchell, Colette Fagan	<b>Gender and the intensification of work:</b> Evidence from the European WC surveys	<i>Eastern Economic Journal</i> , Fall 2004	Analysis of gender dimension of the work intensification trend	Intensity of work has been rising faster for women than for men.  Intensity of work has a negative effect on health and work-life balance
D.Gimeno, F.G Benavides, J.Benach, B.C.Amick III	<b>Distribution of sickness absence in the EU countries</b>	<i>Occupational and Environmental Medicine</i> 2004;61:867-869	Analysis of work absence data from 2000 survey	Sickness absence percentages in Southern European countries were lower than in Central and Northern European countries.
P.Smolders (TNO, the Netherlands)	<b>Work in 27 European countries: Testing the North-South hypothesis</b>	<i>Tijdschrift voor Arbeidsvraagstukken</i> 2004-20, nr3	Cluster analysis of 2000/1 data to show work environment country groupings	Work pressure highest in the service-oriented northern and western European countries (eg. SV, NL, Fin).

### User Consultation

The main instrument through which the Foundation receives users' needs is through expert group meetings in which experts from various user organisations are providing comments and suggestions on the content and the design of the survey. These suggestions are taken into account in the survey design and the questionnaire drafting. This demonstrates a very responsive attitude towards users and their needs and to a large extent assures the continuing relevance of EWCS data. The composition of the expert group includes, members of the Administrative Board of the Foundation, academic experts and experts from international organisations dealing with working conditions issues as well as representatives from national statistical institutes or research organisations dealing with national working conditions surveys. During the meetings both methodological and content aspects of the questionnaire are discussed based on proposals from the EWCS team of the Foundation. The group then provides recommendations on issues where consensus has been reached.

# Accuracy

Accuracy in the general statistical sense denotes the closeness of computations or estimates to the exact or true values.

A word which is immediately associated with accuracy is "error" since accuracy is used to mean "the inverse of the total error, including bias and variance". The larger the error, the lower the accuracy. However, the definition of "error" is wide: it can encompass deficiencies, mistakes, bias, sampling variation etc. In other words, error is the result of all phenomena that can distort the accuracy of the final statistics.

The following typology of errors is commonly adopted nowadays in statistics :

- Sampling errors;
- Non sampling errors, which in turn cover:
  - Coverage errors;
  - Measurement errors;
  - Processing errors;
  - Non response errors; and
  - Model assumption errors.

These various types of errors are analysed in detail in this chapter, as all of them do not have the same impact on accuracy.

Before proceeding to the assessment of each type of error, it would be useful to describe in brief the sampling procedure followed in the 4th EWCS along with an evaluation of the sample based on a comparison with the LFS results.

## Sampling

### Sample Description

The statistical population includes all persons aged 15 or over whose usual place of residence is in the territory of the Member States of the European Union and Acceding, Candidate and EEA countries, and who are in employment during the reference period. Some countries apply a different lower age limit (16 in ES, UK, NO) and some other use an upper age limit (74 in DK, EE, LV, HU, FI, SE, NO) in their LFS population. The exactly same approach was followed in the EWCS.

A person is considered as being in employment if he or she did any work for pay or profit during the reference week for at least one hour. This is the same definition as in the LFS, and the same inclusion and exclusion rules apply.

Usual place of residence is the address where a person usually resides. It may also be his/her legal residence. This may be different from the place where he/she actually is at the time of the survey. For the purposes of the survey, a person's usual residence is the place where they normally sleep. All special cases were treated in exactly the same way as in the European labour force survey.

The reference week was the week that preceded the beginning of data collection in each country.



Multistage sample was used; in the first stage PSUs were selected using stratified random sampling. Area sampling was employed for the selection of the sample. Each country was divided into strata defined by region (at NUTS level 2 or equivalent) and degree of urbanization (metropolitan, urban, rural). The sample was allocated to the strata proportionately to their total (general aged 15 and over) population. Within each sampling point, an address was chosen at random, and formed the first member of a cluster of sampled addresses. The rest of the cluster was chosen with a systematic scheme, i.e. every 3rd address after the first one along a random route.

In four countries, Belgium, Sweden, Netherlands and Switzerland, phone screening was used for selecting the sample. A stratification of the interviews by region and degree of urbanization was nevertheless applied (in BE, SE and CH this stratification could be immediately done because of the use of phone registers, whereas in NL – where random dialling was used – the numbers were assigned to each strata a posteriori). The phone screening in these four countries therefore mainly affected the final stage of selection of interviewers within PSUs: instead of selecting them by the random walk procedure, they were randomly selected from a directory (BE, SE and CH) or from a Random Pulsing Procedure (NL). For issues of coverage and comparability, see the section on "Non Sampling Errors" below.

Clusters have a maximum size of 21 households. One person belonging in the target population was selected from each household, following the last birthday method (selecting the worker who will have the last birthday in the household).

Table 1: *Number of clusters and cluster size in each country*

	<b>Planned N of clusters</b>	<b>Sampae N of clusters</b>	<b>Minimum size</b>	<b>Maximum size</b>	<b>Average cluster size</b>
EWCS 2005	4771	4819	1	21	6.2
Belgium	150	150	5	11	6.7
Czech Republic	150	151	1	15	6.8
Denmark	120	120	2	15	8.4
Germany	258	258	1	5	3.9
Estonia	167	108	1	12	5.6
Greece	125	124	2	16	8.1
Spain	140	143	2	12	7.1
France	250	250	1	12	4.3
Ireland	180	176	2	9	5.7
Italy	133	133	1	11	7.6
Cyprus	60	60	10	10	10.0
Latvia	120	136	2	18	7.4
Lithuania	110	110	6	14	9.2
Luxembourg	55	55	9	12	10.9
Hungary	180	180	3	8	5.6
Malta	60	137	1	10	4.4
Netherlands	120	117	3	12	8.8
Austria	170	125	1	20	8.1
Poland	220	220	3	5	4.5
Portugal	120	120	2	13	8.3
Slovenia	85	106	1	8	5.7
Slovakia	150	151	1	8	6.8

Tabla 1: Number of clusters and cluster size in each country (cont'd)

	Planned N of clusters	Sampae N of clusters	Minimum size	Maximum size	Average cluster size
Finland	200	228	1	15	4.6
Sweden <sup>1</sup>	100	86	1	(155)	12.3
United Kingdom	250	248	1	7	4.3
Bulgaria	150	150	4	8	7.6
Croatia	146	147	2	11	6.9
Romania	200	201	4	7	5.2
Turkey	220	227	2	8	4.5
Norway	276	296	1	6	3.4
Switzerland	106	106	3	21	9.8

The target number of interviews was 1000 in each country with the exception of smaller countries (Estonia, Cyprus, Slovenia, Malta, and Luxembourg) where the target was 600 interviews. Table 2 summarises the achieved number of cases before and after quality control.

Table 2: Number of interviews carried out and retained after quality control

	Number of interviews			Number of interviews	
	before quality control	after quality control		before quality control	after quality control
<b>Belgium</b>	1003	1003	<b>Austria</b>	1009	1009
<b>Czech Republic</b>	1029	1027	<b>Poland</b>	1019	1000
<b>Denmark</b>	1036	1006	<b>Portugal</b>	1092	1000
<b>Germany</b>	1022	1018	<b>Slovenia</b>	615	600
<b>Estonia</b>	602	602	<b>Slovakia</b>	1030	1024
<b>Greece</b>	1002	1001	<b>Finland</b>	1088	1059
<b>Spain</b>	1035	1017	<b>Sweden</b>	1066	1059
<b>France</b>	1164	1083	<b>United Kingdom</b>	1058	1058
<b>Ireland</b>	1046	1009			
<b>Italy</b>	1020	1005			
<b>Cyprus</b>	600	600	<b>Bulgaria</b>	1139	1135
<b>Latvia</b>	1014	1003	<b>Croatia</b>	1045	1011
<b>Lithuania</b>	1017	1017	<b>Romania</b>	1053	1053
<b>Luxembourg</b>	600	600	<b>Turkey</b>	1041	1015
<b>Hungary</b>	1017	1001			
<b>Malta</b>	600	600	<b>Norway</b>	1000	1000
<b>Netherlands</b>	1025	1025	<b>Switzerland</b>	1040	1040

<sup>1</sup> In Sweden the sample selection was based on telephone recruiting and PSUs in large cities that are still one NUTS 5 (settlement) region have been amalgamated into one PSU, hence the large maximum cluster size.

### **Sample evaluation**

In order to evaluate the sample of the 4th EWCS, the distribution of the sample population totals among main demographic variables was compared with that of the Labour Force Survey (AGExSEX, Region, Occupation, and Economic Activity of employer) in each country. The two distributions were compared using a chi-squared statistic in which the expected value was replaced by the LFS estimate assuming the same sample size. We can therefore test whether the sample may be regarded as randomly selected from the universe (the LFS) and if not, which cells represent the main differences for each country and classification.

Significant<sup>2</sup> differences indicate that coverage problems and non response make the sample so different from the universe that substantial biases may exist and weighting is imperative. It also indicates that weights will likely be large resulting to increased variance.

All tests are shown in Annex 1. Detailed figures are included in the sample evaluation report to which the interested reader is referred. Here, the main findings for each variable used for the comparison are presented:

#### *Region*

The distribution of the sample across NUTS 2 regions is usually consistent with the universe. This is to be expected since the selection of the sample was based on population totals for each region (albeit general population and not working population). What is notable is that in three countries Norway, Netherlands and Greece there is a large discrepancy between sample and universe. This is due to large discrepancies in the regions Sør-Østlandet and Agder og Rogaland in Norway, Flevoland + Zeeland and Utrecht+Limburg in Netherlands (in this case the NUTS classification was not used in the stratification) and to Attiki and to a lesser extent Ipeiros in Greece. There are also other countries where differences are found significant at the 0.001 level Finland (mainly regions Itä-Suomi + Åland and Pohjois-Suomi) and France (the problem is almost exclusively in the region of Languedoc-Roussillon which had double frequency in the sample than in the general population). Finally, some further countries show differences that are found significant at the 0.01 level: Italy (mainly regions Lombardia and Sicily+Sardinia), UK (mainly South Western Scotland, Gloucestershire, Wiltshire and North Somerset, and Berkshire, Bucks and Oxfordshire), Bulgaria (mainly Severozapaden and Yugoapaden), in Turkey problems are concentrated in the Zonguldak + Kastamonu, a region. produced by collapsing administrative regions.

#### *Age & Gender*

For the rest of the variables there was no stratification and therefore the discrepancies are larger and more frequent. Comparison was performed using the Age by Gender cross classification. In most countries there were significant differences. Only in Germany, Greece and Luxembourg there appears no discrepancy while in Cyprus and Romania the test was significant at the 0.01 level. A main observation is that there is a general trend of overrepresentation of females over males (reversed in Turkey where male workers tended to participate to the survey more often than female ones) and of older workers over young.

#### *Occupation*

In all countries appear there are discrepancies, in several cases quite large, for one or more category of occupation. The only country where discrepancies were small yet still significant was Greece. The large number of ISCO classes at the 1st digit resulting in small number of values for each cell is one cause of the problem. Classification problems have been

---

<sup>2</sup> The common significance level of 0.05 is not used here as the number of tests (over 100) guarantees that at least some will show significant difference when in fact there is no difference. Therefore the 0.001 level is used to positively identify differences from Universe values.

reported by implementing organizations and may have contributed to the large discrepancies. There is no clear pattern however and specific ISCO classes were over represented in some countries while under represented in others. The most substantial trends are an overrepresentation (related to the LFS) in the sample of ISCO categories 1,3,4, 9 and under representation of ISCO categories 6 and 8. In the rest the situation is mixed.

#### *Activity of employer (based on NACE)*

In several countries there are no discrepancies regarding the distribution of people in NACE categories according to the economic activity of the employer (EE, CY, MT, TR, NO, CH). In the rest there appear to be over representation or under representation for one or more categories of employer activity. In general, categories A-F (Primary sector and manufacturing) seem to be under represented while M-Q (Services) over represented in the sample as compared with the composition of the population as described in the LFS.

## **Variance**

Sampling errors arise from the fact that not all units of the frame population, but only a sample of them, are enumerated. The statistics produced from a sample survey will differ from the values, which would be computed if exactly the same survey operations were applied to the whole frame population. The difference is the sampling error.

Sampling errors can be assessed via variance, standard error (the square root of the variance), relative standard error (coefficient of variation – CV, the ratio of standard error over estimate) and the extent of confidence intervals (usually a multiple of standard error).

Variance is computed based on a probabilistic model, which for most of the EWCS estimates is the binomial distribution. The complex design that has been used (stratified, multistage/clustered) is affecting the variance of the estimates (stratification generally decreases the variance and clustering increases it). In addition the adjustment to the universe totals in terms of age, gender, occupation, economic activity of the employer as well as for the working population in each country (the latter affecting total estimates only) that was implemented in order to remove biases due to undercoverage and non-response has also an adverse effect on variance.

The design effect of a sample design is defined as the ratio of the actual estimate's variance to the sampling variance of simple random sampling of the same sample size. This, however, is different for each variable. In order to get an "overall" design effect, a list of questions from the survey were selected and the corresponding standard errors of the estimates (proportions for categorical variables and averages for continuous ones) calculated using SUDAAN, a specialised software for the analysis of data from complex surveys. The set of questions was the same across all countries. Variables involved in estimations (in parenthesis the number of levels without DK, NA, and Refusal): q3a (4), q5(5), q10a-j(7), q11a-m(7), q12(4), q15a(2), q18(4), q23a-f(2), q25a-m(5), q27(3), q28c-e(2), q29a-i(2), q32(2), q33(2), q35(3), q36(4), q37a-f(5), ef3(3), ef6a(2), ef6c(2), ef6g(2), ef6h(2), q6(8), q2d, q7, q8a-b, q13, q34b, ef1 (continuous).

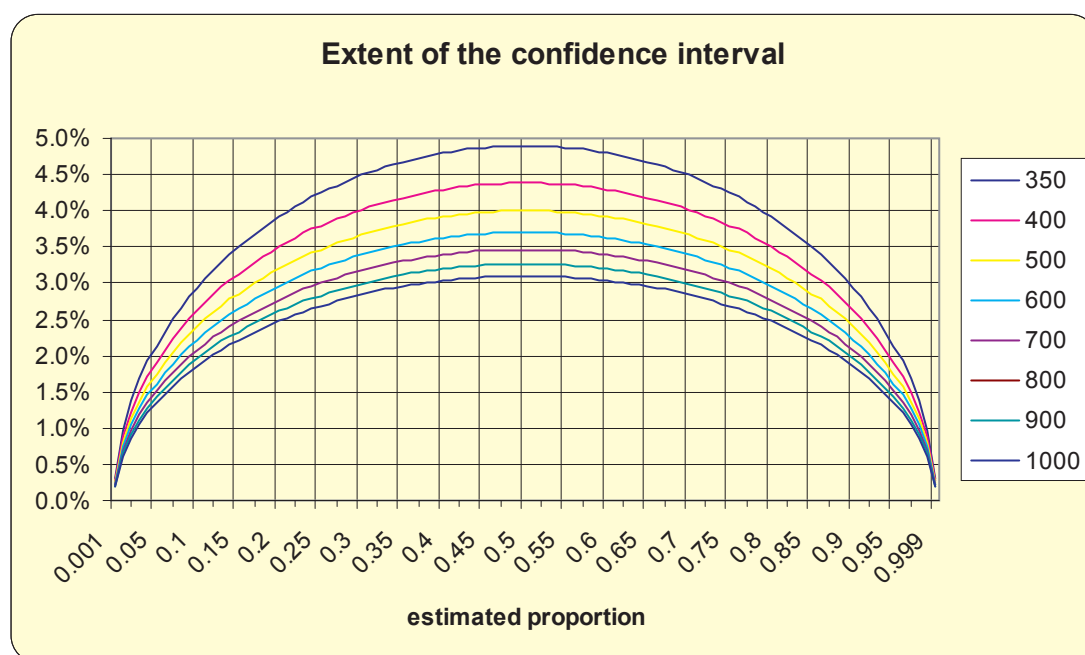
After obtaining the estimated design effects for each estimate, the median was computed and selected as the "overall design effect". The fourth column is just the ratio of total sample size to the estimate design effect.

	Sample size	Median design effect	Effective sample size		Sample size	Median design effect	Effective sample size
<b>Belgium</b>	1003	1.63785	612	<b>Netherlands</b>	1025	1.8193	563
<b>Czech Republic</b>	1027	1.86109	552	<b>Austria</b>	1009	1.83853	549
<b>Denmark</b>	1006	1.59105	632	<b>Poland</b>	1000	1.69645	589
<b>Germany</b>	1018	1.73086	588	<b>Portugal</b>	1000	1.78989	559
<b>Estonia</b>	602	1.47595	408	<b>Slovenia</b>	600	1.4242	421
<b>Greece</b>	1001	1.56024	642	<b>Slovakia</b>	1024	1.80553	567
<b>Spain</b>	1017	2.02836	501	<b>Finland</b>	1059	1.39332	760
<b>France</b>	1083	1.50694	719	<b>Sweden</b>	1059	1.59156	665
<b>Ireland</b>	1009	1.5949	633	<b>United Kingdom</b>	1058	1.3878	762
<b>Italy</b>	1005	1.98831	505	<b>Bulgaria</b>	1135	2.08974	543
<b>Cyprus</b>	600	1.69819	353	<b>Croatia</b>	1011	2.12021	477
<b>Latvia</b>	1003	1.66488	602	<b>Romania</b>	1053	1.83678	573
<b>Lithuania</b>	1017	1.97482	515	<b>Turkey</b>	1015	1.66679	609
<b>Luxembourg</b>	600	1.46452	410	<b>Norway</b>	1000	1.52687	655
<b>Hungary</b>	1001	1.58182	633	<b>Switzerland</b>	1040	1.53947	676
<b>Malta</b>	600	1.72416	348				

It should be noted that the design effect of different variables varies considerably and the median is only indicative.

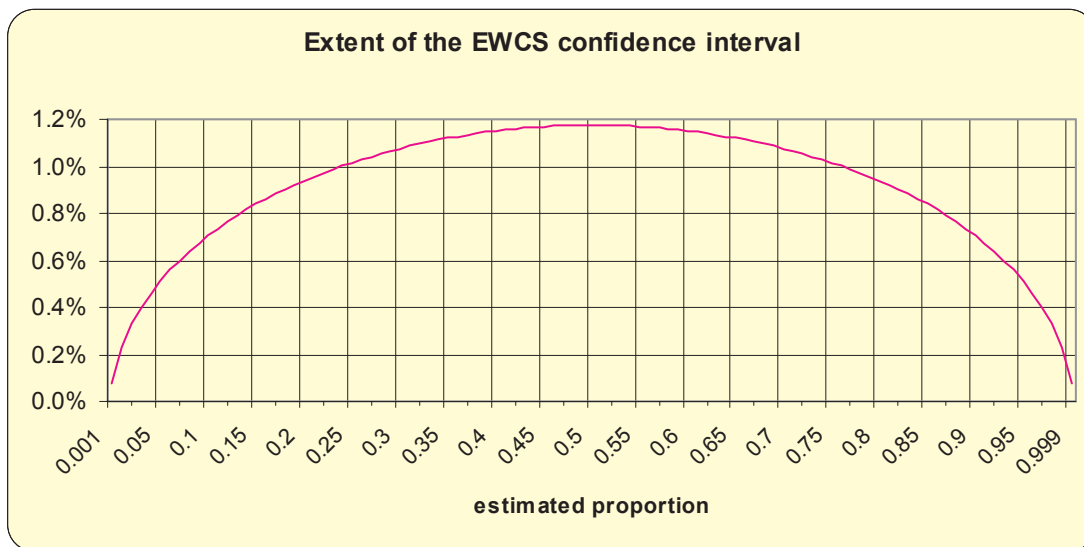
In order to compute the precision of estimates using the median design effects and corresponding effective sample sizes the following graph presenting the (half) extent of the confidence interval based on the effective sample size might be useful. In general the estimates for countries with actual sample size of 1000, the extent of the confidence interval is between  $\pm 3\%$  and  $\pm 4\%$  around the point estimate.

Figure 1: The extent of the confidence interval as a function of the value of the estimated proportion for various sample sizes.



Estimates for the whole area covered by the EWCS (EU-27, candidate countries, Norway and Switzerland) have much larger design effect, as the allocation of the sample was almost uniform among countries resulting in large variations in weights. The median design effect is 4.26 and the corresponding effective sample size is 6968. The extent of the confidence interval is shown in Figure 2.

Figure 2: *The extent of the confidence interval for estimates of the total EWCS covered area as a function of the estimated proportion.*



## Non Sampling errors

### Coverage errors

The main sampling selection method (used in 27 out of 31 countries) for reporting units was the random route procedure. Under-coverage of the random route process is usually small although its actual extent was not measured. Most of it refers to institutionalised persons (e.g. prisoners) and/or those not having a place of residence (i.e. homeless, travellers etc.). The omission of the former group of persons is a design choice while the omission of the latter is a common feature of all area sample surveys.

Other types of under-coverage that can occur refer to errors in the application of the random route procedure. For example, if a door to a housing unit is not obvious the unit might not be counted along the route. These coverage problems depend to a large extent on the experience and training of interviewers in each country. The interviewer's instruction manual, which was prepared and used in instructing the interviewers, gave them precise and clear directions for the implementation of the random route procedure, which aimed at minimising coverage problems from the random route application.

In four countries telephone screening was used:

- Sweden. The Swedish Population and Address Register (SPAR) was used to generate sampling addresses and telephone numbers for telephone screening. The register receives information from the official Swedish notification register and is a common source for sample frames providing a very good coverage of Swedish households.

- Belgium. The major part of the sample was obtained with telephone screening. The method used was to generate telephone numbers based on existing numbers (from telephone directories of fixed and mobile numbers) to which a digit from 0–9 was added thus reaching unlisted numbers.
- Netherlands. A pulsed Random Dialing Digits (RDD) system was used to generate fixed line numbers (pulsing was used to screen fax, unused and other non-eligible numbers).
- Switzerland. The sample was based on telephone directories of Switzerland.

The telephone screening process is expected to have introduced some undercoverage mainly due to mobile-only households (households without a fixed line) or out of grid households. This undercoverage is expected to be minimal in Sweden and Netherlands: in a recent Eurobarometer survey<sup>3</sup> the percentage of households with a fixed telephone line is 100% and 96% respectively. The number of households without a fixed telephone line in Belgium is substantial (27%) however the use of mobile numbers as well as fixed line ones allows reaching also the mobile only population (24%).

In addition to households without telephone, households that are included in do-not-call-me lists are excluded from any telephone-based survey including the EWCS in Sweden and Switzerland.

Bias introduced by coverage problems is alleviated to some extent by weighting which adjusts the distribution of the sample among main variables (age, gender, occupation, economic activity and region) according to a universe description (the EU-Labour Force Survey).

### **Measurement errors**

Measurement errors are errors that occur during data collection and cause the recorded values of variables to be different than the true ones. They are distinguished in three types of errors: errors due to the questionnaire, the respondent or the interviewer.

Errors due to the questionnaire used for data collection. Such errors occur if questions are ambiguously formulated, the wording is inappropriate, the structure of the questionnaire is illogical or if not all possible replies to closed questions are given. Indications of questionnaire errors are included in the reports of implementing organisations based on their experience in implementing the questionnaire. Main observations-problems include

- A difficulty in marking some variables for self-employed (economic activity of employer for instance).
- A difficulty in relating the questions to the particulars of occasional or casual workers.
- Reluctance in giving information about income (EF5) in several countries.
- Questions Q2b,c and in fewer cases have created some problems as they had to be calculated by the respondent. In some cases responses were illogical (Age when stopped full time education+years of work after education greater than respondent's age).

---

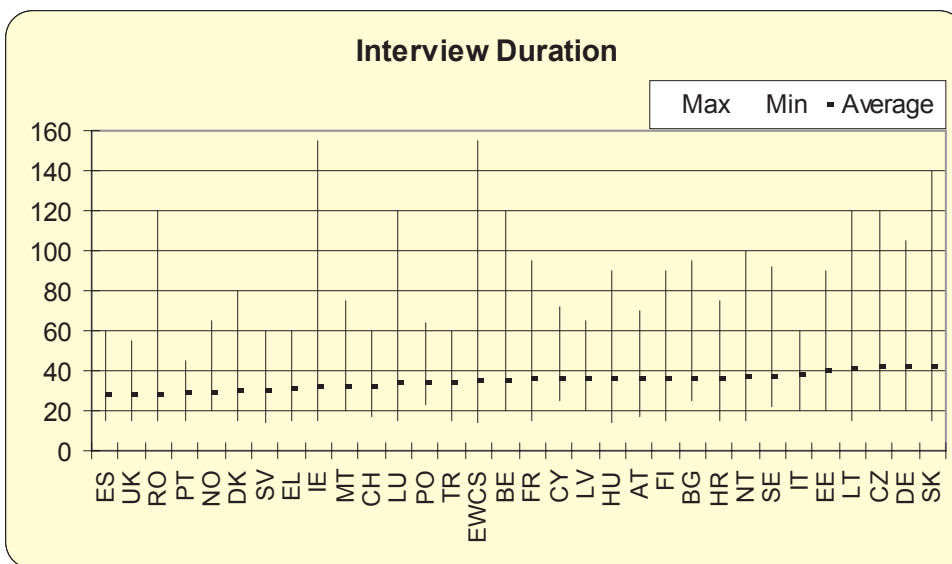
<sup>3</sup> E-Communications household survey. Final report available at [http://ec.europa.eu/information\\_society/policy/ecommm/doc/info\\_centre/studies\\_ext\\_consult/ecommm\\_household\\_study/eb\\_jul06\\_main\\_report\\_en.pdf](http://ec.europa.eu/information_society/policy/ecommm/doc/info_centre/studies_ext_consult/ecommm_household_study/eb_jul06_main_report_en.pdf)

**Errors due to the respondent**, who may consciously or unconsciously give erroneous data. These errors can occur due to mistaken recollection of past events (memory effect), to the tendency to over-report characteristics perceived to be valued by others and not to report characteristics not valued by others (social desirability effects or conditioning), to the lack of the respondent's attention, to the effects of age and education, to fears that sensitive information (e.g. income) may be used for other than statistical purposes, to the tendency to round figures (e.g. age) etc.

The length is one factor that might affect the accuracy of recorded data due to fatigue. On average the EWCS 2005 questionnaire was 35 minutes long. The actual length was a function of a list of factors, most of them are the obvious ones (respondents' talkativeness, different interviewing styles, etc). Also interviews with respondents from larger households had a longer interview time due to the household grid.

The minimum interview length was 14 minutes, while in Ireland we registered a case with 155 minutes duration. The typical length of the interview was between 30 and 35 minutes, and very few were shorter than 20 minutes. The proportion of the latter group is the largest in Spain (reaching 10%). The distribution of interview duration is summarised for each country in Figure 3 and presented in detail in Annex 3.

Figure 3: A three point summarisation of the distribution of interview time in each country



**Errors due to the interviewer**, that occur when interviewers influence the answers given by the respondents. These errors may also be identified and assessed by the qualitative survey that will follow.

*Actions taken to minimise interviewer error*

Measurement errors are controlled if an experienced, well-trained and not overworked field force is employed in the survey.

Interviewers participating in the 4th EWCS survey had at least one year of experience in interviewing.

Training of interviewers was carried out locally by national coordinators, under the guidance of Gallup. Gallup and the European Foundation, prepared an Interviewing Manual, which covered the basic rules of interviewing, sampling techniques in general and about the Random Route Technique, specifically. The manual contained an extensive



explanation of survey concept (e.g. the definition of the labour force, etc.) and questions (EWCS 2006 Questionnaire Glossary). The manuals were translated to each national language in which interviewing was carried out. These local language manuals were used at interviewer training.

Training was provided for those involved in the survey, from interviewers to trainers and supervisors. Group project briefings were carried out, where interviewers could also share their "best practice" techniques of approaching respondents. To cut down costs in most countries training was decentralized and cascaded. For each interviewer on the EWCS team there was at least one opportunity to do a pilot interview before they started the actual data collection. Take-home training materials were provided to interviewers to use as reference material. The content of training included the following topics:

- Administrative issues
- Planning of fieldwork
- Review of all materials provided
- Contacting procedures
- Conducting an interview
- Interview procedures in the field
- Supervision in field and reporting procedures
- Structure of survey team and role of all members of the team

**Interviewers workload.** The amount of interviews carried out by each interviewer may have an effect on the quality of the collected data. It was therefore capped at 25 interviews per interviewer.

Table 3: *Size of field force, distribution of interviews among interviewers*

	<b>Number of interviewers</b>	<b>Minimum number of interviews per interviewer</b>	<b>Maximum number of interviews per interviewer</b>	<b>Average number of interviews per interviewer</b>
EWCS 2005	2745	1	25	10.8
Belgium	59	2	25	17.0
Czech Republic	140	1	22	7.3
Denmark	63	4	25	16.0
Germany	188	1	10	5.5
Estonia	74	1	24	8.1
Greece	66	1	25	15.2
Spain	66	6	25	15.4
France	114	1	25	9.5
Ireland	63	1	25	16.0
Italy	133	2	11	7.6
Cyprus	38	7	15	15.8
Latvia	74	4	24	13.6
Lithuania	101	6	14	10.1
Luxembourg	41	3	23	14.6

Table 3: Size of field force, distribution of interviews among interviewers (cont'd)

	Number of interviewers	Minimum number of interviews per interviewer	Maximum number of interviews per interviewer	Average number of interviews per interviewer
Hungary	129	1	24	7.8
Malta	35	2	25	17.1
Netherlands	94	1	24	10.9
Austria	84	1	24	11.6
Poland	110	1	24	9.1
Portugal	62	2	25	16.1
Slovenia	74	1	23	8.1
Slovakia	134	1	23	7.6
Finland	94	1	25	11.3
Sweden	107	1	25	9.9
United Kingdom	146	1	25	7.3
Bulgaria	105	6	24	10.8
Croatia	80	1	25	12.6
Romania	82	4	25	12.8
Turkey	61	1	25	16.6
Norway	70	1	25	14.3
Switzerland	58	4	25	17.9

In the random route method the sample selection is delegated to a large extent to the interviewers. The extent to which the method was applied correctly is a substantial determinant of the quality of the sample. Interviewers documented their route in two ways: they kept record of each attempted / visited "doors" in the Routing Slip with exact address and dates (as well as the outcome of the visit). They also maintained a document ("Random Route Map") that qualitatively described their route. This "map" is in fact a mixture of actual map / drawings and text information on certain probably ambiguous decisions that might have been taken by the interviewer (i.e. if there was a big factory at the random start and how they reached the actual first "third door").

Supervisors used these two inputs to check the implementation of the random route in a random subsample of the routes used in the survey. Supervisors revisited the route and they double-checked the recorded "doors" (i.e. potentially residential addresses recorded by the interviewer). The supervisors assessed each controlled route with a quality mark:

4 – routes with full compliance

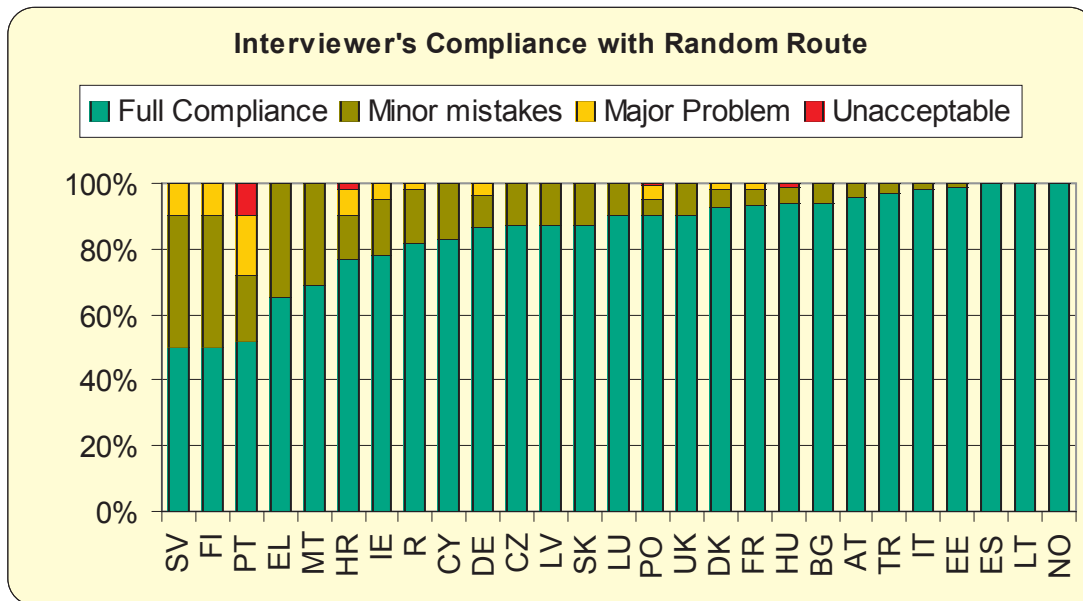
3 – routes with minor mistakes (e.g. one or two missed addresses, did not start at the third address, but the one that was given, etc.)

2 – routes with a major problem, but carried out randomly (e.g. if someone did the random procedure always to the left instead of to the right, but otherwise random rules were followed)

1 – unacceptable routes (e.g. if someone did not observe the random procedure, or route could not be verified at all). Random routes that got a mark of '1' were excluded from the final datafile.

The results are presented in Figure 4 for all participating countries. Generally the procedure was followed without major problems in more than 90% of the routes with the exception of Portugal. In that case the quality control revealed that two interviewers were not implementing the procedure correctly (actually not at all) and all their interviews were excluded from the sample.

Figure 4: Quality control of the sampling selection procedure.



### Processing Errors

Between data collection and the beginning of statistical analysis for the production of statistics, data must undergo a certain processing: verbal responses must be coded (coding), data must be keyed into computers for storage (data entry), data must be checked for inconsistencies and outlying values (data editing), missing and identified-as-wrong values may need to be imputed (imputation), survey statistics have to be adjusted in order to cope with sampling problems, i.e. coverage problems, non-response, unequal probabilities of selection (weighting), the instability of survey statistics estimates should be computed (sampling variance estimation). Errors introduced at these stages are called processing errors.

#### Data coding errors

Two variables, economic activity and occupation were recorded as textual descriptions and were coded afterwards using standard classifications (NACE Rev. 1 and ISCO 88 COM respectively). A small team of coders in each country were trained specifically for this task. They coded the responses in close cooperation with each other based on consensus when there was doubt on the appropriate code of a specific record. National agencies used the translated coding schemes provided by the coordinating organisation along with the instructions on applying the classifications provided by the responsible organisations (Eurostat for NACE rev. 1 and U. of Warwick for ISCO 88 COM). Coding was carried out in a nationally centralised way.

The coding procedure was thoroughly checked, and all problems were reported in a coding report. Ambiguities due to inadequate textual descriptions were resolved by re-contacting the respondents. Difficulty with coding of ISCO were reported by BE, DK, and FR.

Coding was originally planned to be checked by an independent auditor by recoding a random sample of textual descriptions and producing indicators of coding reliability based on intercoder agreement. However, as no tender was received in the related call, this operation was not carried out. Therefore no quality indicators for coding have been produced.

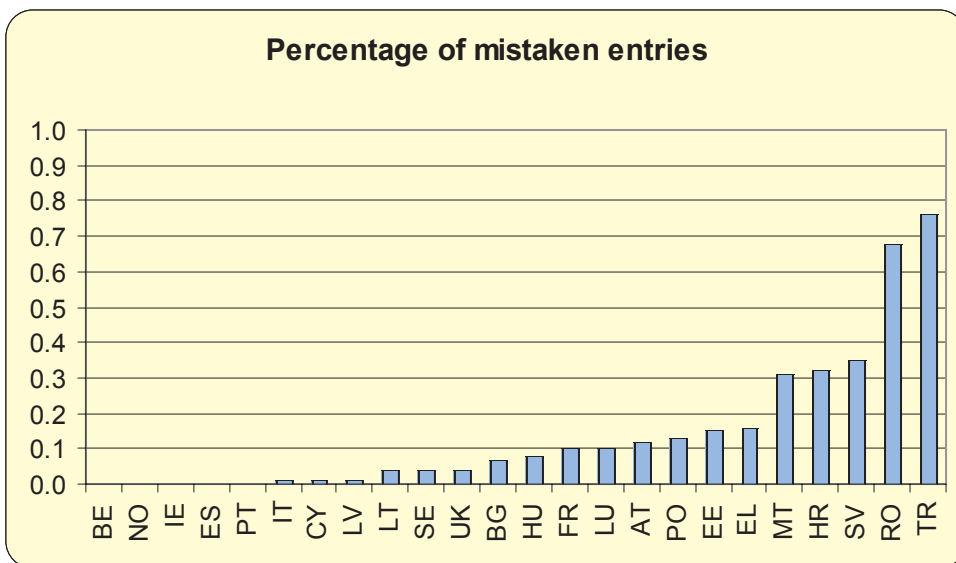
When comparing the sample characteristics to the LFS, we found significant differences in the distribution along occupation categories between the sample and the universe. This must be taken into account when using the occupation variable in the analysis.

**Data entry errors**

In each country where manual entry of the paper questionnaires was performed, a partial data re-entry of 10% of all cases was performed. Mistakes were defined as verified incorrect first entries in each case, where the comparison double entry revealed a difference between the initial and re-entry. The error rates reported are a reasonable estimate of errors remaining in the data set (it includes erroneous entry in both inputs but excludes errors found out during re-entry). The error rate is computed as the number of mistakes discovered divided by all fields (database cells) to be filled in. The highest error rate appeared in Turkey, 0.76% still bellow the threshold of 1% which would have triggered a complete re-entry of the data file.

In each country where manual entry of the paper questionnaires was performed, a partial data re-entry was performed with 10% of all cases. The table below indicates that the error rate remained limited, and data entry was of high quality. Mistakes were defined as verified incorrect first entries in each case, where the comparison double entry revealed a difference between the initial and re-entry. Infrequent cases where the initial and re-entry made the same error, remained therefore undiscovered. We then calculated the rate of mistakes as the number of mistakes discovered divided by all fields (database cells) to be filled in. An initial threshold of 1% error rate was set, which would have triggered a complete re-entry of the complete datafile of the country in case it was reached. The highest mistake rate was 0,76% in Turkey. No complete re-entry was carried out in any country.

Figure 5: The error rates of data entry.



**Data Editing**

The data editing procedure involves logical screening and code validating embedded in the data collection and data entry, checking for extreme values (outliers) in the continuous variables, qualitative editing, cross-country comparisons for coding consistency and pattern validation and over time comparisons for previously asked questions in order to check coding consistency. Cross country comparisons included the examination of frequency distributions of different countries which may reveal coding problems like reversed scales. Editing was based on a three-tier system of edits.

**Step 1: Field edits**

During this process, all questionnaires in each country/territory were primarily edited by field coordinators, or supervisors who collected the questionnaire from interviewers. They were urged to clarify with the interviewers any inconsistencies that are found in the questionnaire. Subsequently interviewers went back and clarified the parts suspect to error with the respondent. The correction and editing done by the supervisor was noted in each questionnaire, and these were taken consideration in the next phase.

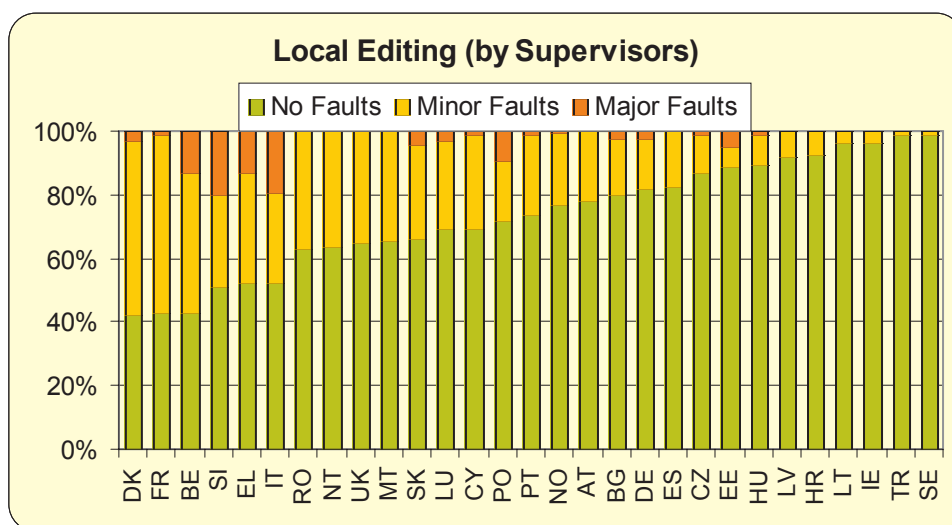
**Step 2: Local edits**

There was an additional quality check done at the level of national institute with a point system to evaluate the questionnaires. The scale of the EWCS quality mark is:

1	Severe and substantial inconsistencies, i.e. cheating or very poor interviewing job overall (these cases should be excluded from the datafile)
2	Severe faults that need/needed to be corrected by contacting the interviewer/respondent
3	Monitor faults than can/could be corrected directly without contacting the interviewer/interviewee
4	Completely filled out questionnaire - no evident faults

The results from this part of editing are shown in Figure 6, in Turkey and Sweden almost no faults were identified by supervisors while about 20% of questionnaires in Slovenia and Italy had major faults which were subsequently corrected by recontacting the respondent.

Figure 6: Results from local editing in percentages among the categories that are included in the datafile.



### Step 3. Logical edits

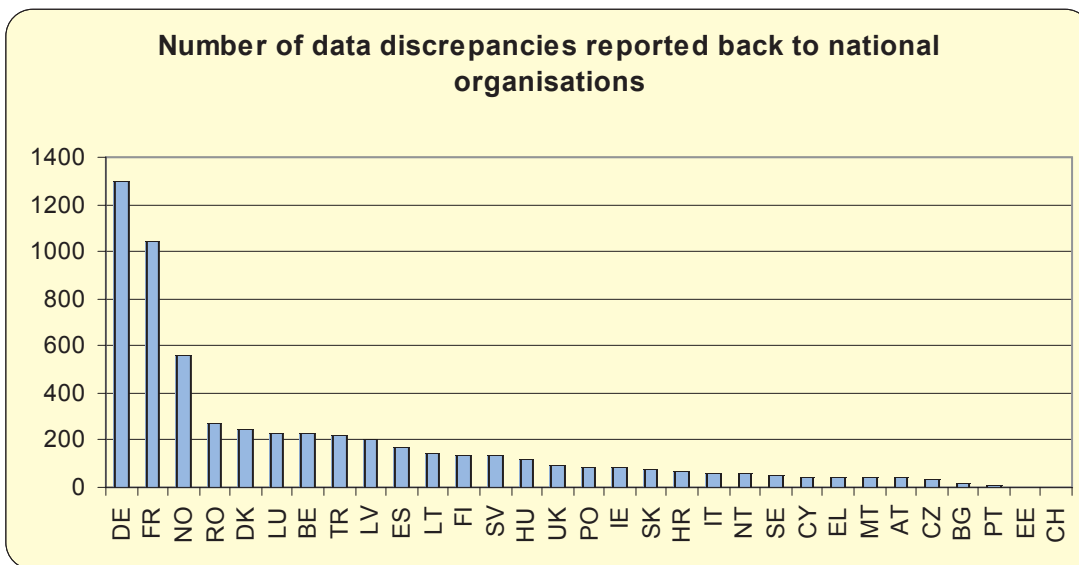
Checks were performed for inconsistent and/or illogical values before and during data entry with the help of appropriate software so as to ensure that records not pose any apparent empirical contradictions.

The definitions of rules involves some judgment and depends on a trade off between false alarms (correct values failing the edit) and failures to detect erroneous values. A set of guidelines was created including 51 rules.

Rules for admissible values were also embedded in the data entry software and all dubious cases were examined in case by case. By keeping the editing as close as possible to the actual questionnaire, the interviewer and the respondent it was possible to control the problem of false alarms (e.g. been able to identify if spouses of the same sex are a mistake or the result of same sex marriage).

In order to ensure that data editing was performed in a uniform way among national organisations the coordinating organisation performed checks in a centralised way as well. Values that failed the edits were referred back to national organisations for re-check. A total of 5817 values were sent back for re-evaluation in a case-by-case basis. The number of values for each country is shown in Figure 7.

Figure 7. Distribution of data discrepancies by country.



The final decision on whether a suspect value is correct or erroneous was made locally and afterwards the final clean data set was dispatched for weighting.

### Weighting

Weights are used in order to compensate for the unequal probability of selection of sample units. Weights are also used in stratification after selection (poststratification) and in adjusting for nonresponse and under-coverage thus removing bias from the sampling and data collection phases of the survey. While weighting is beneficial as it removes known biases from the sample, it also increases the variance of estimates.

Weighting was a three-stage operation and was carried out centrally by the coordinating organisation.

Selection probabilities were computed based on the number of workers in the household and related weights were calculated.

For post-stratification weighting the raking procedure (AKA iterative proportional filtering) was employed. Raking is an iterative optimisation process that essentially performs a ratio adjustment for one class at a time. The method starts with sample weights (based on selection probabilities) and recycles through the classes as necessary (in practice four or fewer times) until a set of calibration weights is effectively found. Trimming was also embedded in the weighting algorithm in order to avoid very large or very small weights that may produce very influential records. The acceptable region for weights was (0.3,3).

The universe data was based on LFS 2004 4th quarter, drawn from Eurostat' NEW CRONOS dissemination database. Where data was missing, the most recent local sources, were used. The variables used for weighting were

- AGExSEX. Figures were obtained for the intercellular probabilities AGExSEX (male under 29, female under 29, male between 30 and 49, female between 30 and 49, male 50 and over and female 50 and over.
- Economic activity. It was based on NACE divisions which were further aggregated in 4 categories,
  - a to f Primary sector and Manufacturing
  - g\_to\_k Services (excluding public administration)
  - l Public administration and defence; compulsory social security.
  - m\_to\_q Other services
- Occupation. The classification used was based on ISCO at first digit level with two further aggregations.
  - isco1 – Legislators, senior officials and managers
  - isco2 – Professionals
  - isco3 – Technicians and associate professionals
  - isco4 – Clerks
  - isco5 & 0 – Service workers, shop and market sales workers and armed forces
  - isco6&8 – Skilled agricultural and fishery workers
  - isco7 – Craft and related trades workers and plant and machine operators and assemblers
  - isco9 – Elementary occupations
- Region. Regions used were based on NUTS regions at level II. When the number of respondents in a cell was small neighbouring regions were combined.

Finally, cross-national weights necessary to produce European figures were computed with multiplying each national weight with the country's proportion in the given larger region (EWCS, EU-25, EU-27, NMS, EURO, EU-12, EFTA). The total number of persons in employment in each EWCS country is as follows:

	<b>Persons in employment</b>		<b>Persons in employment</b>		<b>Persons in employment</b>		<b>Persons in employment</b>
Belgium	4,153	Italy	22,577	Poland	13,999	Bulgaria	2,902
Czech Republic	4,718	Cyprus	339	Portugal	5,052	Croatia	1,532
Denmark	2,740	Latvia	1,017	Slovenia	933	Romania	9,030
Germany	35,981	Lithuania	1,431	Slovakia	2,191	Turkey	21,581
Estonia	596	Luxembourg	186	Finland	2,347	ACC total	35,045
Greece	4,324	Hungary	3,905	Sweden	4,259		
Spain	18,281	Malta	147	UK	28,089	Norway	2,258
France	24,335	Netherlands	8,104	EU-25 total	195,355	Switzerland	3,961
Ireland	1,886	Austria	3,765			EWCS 2005 total	236,619

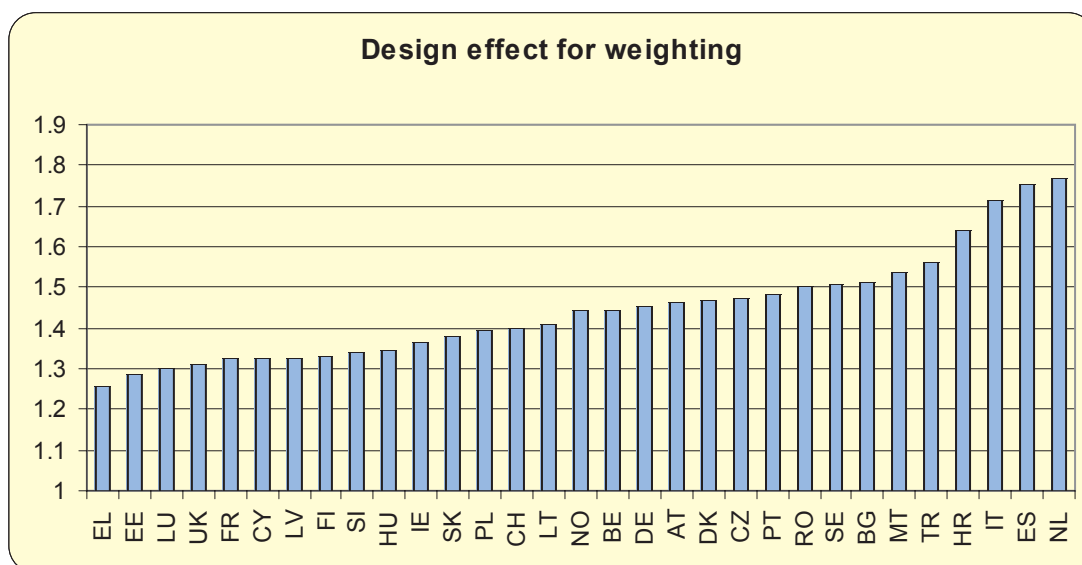
Initially the level of urbanisation (rural, urban, metropolitan) was envisaged to be used in weighting as well. However, relevant figures could not be obtained from the EU-LFS where the related variable is not included in the dataset. It should be noted that the level of urbanisation was used for stratification but the data referred to general population figures (aged 15 and over). It is therefore possible to have some bias due to the disparity between the people in employment and the general population. It should also be noted however that where such disparity exists it will be controlled to some extent by other weighting variables (like region, age, gender).

Weighting, although beneficial for the overall accuracy of results due to the minimisation of bias, increases the variance of estimates. In Figure 8 the design effect (i.e. DEFF the ratio of the true variance of a statistic under the actual design divided by the variance of the unweighted estimates) of weighing is shown for all countries<sup>4</sup>. Values range from 1.25 in Greece to 1.78 in Netherlands.

<sup>4</sup> This is an approximate computation that assumes equal variance among strata. The actual design effect that includes also the effect of clustering, which also has a detrimental effect on the variance, is presented in the sampling errors section.



Figure 8: Approximate design effect for weighting.



### Non Response error

Non-response occurs when the interviewer fails to obtain measurements on sampled units. It is distinguished into unit non-response, which is the failure to obtain a complete response, and item non-response, where the respondent refuses to answer to a particular question.

The unit non-response arises from three reasons:

- Inability to contact the sampled household or person. Non-contacts arise because interviewers cannot gain access to a building, cannot reach anyone at a housing unit or the respondent is away or otherwise unavailable during the interview period.
- Inability of the contacted person to provide responses to the survey (due to illness, disability or language problems).
- Refusal to the interview request

The difference in survey characteristics between respondents and non-respondents creates the non-response error. Non-response error can affect the quality of survey statistics. If, for example, persons with large incomes have a greater tendency not to state their income or not participate in the survey, then average income based on available responses will underestimate the true average income.

The response rate is calculated by dividing the number of completed interviews to the number of eligible persons in the sample. The second number is difficult to have in a random route procedure where non-contact often indicates unknown eligibility. Non-interviews from unknown eligibility arise when it is not known whether a housing unit is involved, housing units that are selected but not attempted or worked, housing units where an address cannot be located and housing units that could not be reached by an interviewer or where in an unsafe area.

Not eligible cases for in-person household surveys include vacant housing units, housing units with no eligible respondents, housing units that are not residences such as businesses, government offices, residential units that are institutions such as prisons and sanatoriums and group quarters such as military barracks and work camps.

Survey outcomes were recorded in the routing slips of the survey using the detailed list of survey outcomes developed by AAPOR<sup>5</sup>. Routing slips were identical in each country, they were created in English and translated in the national languages. A substantial part of the training was dedicated on the application of the routing slip. Large differences in the frequencies of reported codes however indicate that the implementation might not have been harmonised. The main outcomes are presented in Table 4. Rates for telephone and random route sampling are presented separately. In Belgium a portion of the sample was selected via random route while the rest was selected via telephone screening. Outcomes are reported separately for the two parts of the sample.

Table 4: Survey outcomes of the 4th EWCS survey

	(1) Interview	(2.10) Refusals and break-offs	(2.20) Non-contact	(2.30) Other eligible, non-interview	(3.10) Unknown if HH	(3.20) Unknown if eligible R	(4) Not Eligible	(no worker)	estimated ineligible ratio
AAPOR code:	I+P	R	NC	O	UH	UO			e <sup>6</sup>
EWCS 2005	29,766	15,287	5,872	3,430	1,332	11,460	25,650	18,993	0.65
F2F sampling, overall	25,764	11,784	4,623	2,942	646	7,671	19,233	15,146	0.66
Telephone sampling overall	4,002	3,503	1,249	488	686	3,789	6,417	3,847	0.58
1a Belgium telephone	876	886	222	81	332	631	1,408	969	0.53
1b Belgium face-to-face	127	49	188	53	14	18	70	56	0.87
2 Czech Republic.	1,027	198	32	29	55	165	230	154	0.88
3 Denmark	1,006	514	448	165	35	328	541	481	0.77
4 Germany	1,022	142	18	107	9	474	301	244	0.81
5 Estonia	601	229	143	103	15	43	312	280	0.74
6 Greece	1,002	586	323	109	34	19	926	692	0.66
7 Spain	1,035	480	28	3	36	0	423	399	0.74
8 France	1,094	386	174	55	23	218	510	422	0.75
9 Ireland	1,009	238	425	177	29	158	682	488	0.74
10 Italy	1,020	563	13	47	0	632	563	455	0.72
11 Cyprus	600	81	56	98	24	308	523	307	0.63
12 Latvia	1,003	304	69	58	8	212	853	753	0.47
13 Lithuania	1,017	202	94	122	14	187	396	331	0.77
14 Luxembourg	600	866	162	193	38	91	1,131	896	0.51
15 Hungary	1,001	641	146	136	6	82	1,623	1,322	0.31
16 Malta	600	396	102	107	7	111	974	470	0.61
17 Netherlands	1,027	1,294	130	126	134	2,020	2,691	1,229	0.52
18 Austria	1,009	412	138	45	6	57	300	240	0.85

<sup>5</sup> The American Association for Public Opinion Research, 2004, Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for surveys, 3rd Edition, Lenexa, Kansas

<sup>6</sup> The estimated ineligible ration is computed according to the proportion of non-workers among all contacts with known employment status.

Table 4: Survey outcomes of the 4th EWCS survey (cont'd)

	(1) Interview	(2.10) Refusals and break-offs	(2.20) Non-contact	(2.30) Other eligible, non-interview	(3.10) Unknown if HH	(3.20) Unknown if eligible R	(4) Not Eligible	(no worker)	estimated ineligible ratio
AAPOR code:	I+P	R	NC	O	UH	UO			e <sup>6</sup>
EWCS 2005	29,766	15,287	5,872	3,430	1,332	11,460	25,650	18,993	0.65
F2F sampling, overall	25,764	11,784	4,623	2,942	646	7,671	19,233	15,146	0.66
Telephone sampling overall	4,002	3,503	1,249	488	686	3,789	6,417	3,847	0.58
19 Poland	1,000	888	299	87	84	829	1,089	858	0.62
20 Portugal	1,000	361	28	50	5	50	195	164	0.89
21 Slovenia	612	557	206	100	10	217	415	354	0.76
22 Slovakia	1,024	247	142	125	28	343	822	561	0.64
23 Finland	1,059	625	4567	368	47	1,029	1,346	1,273	0.49
24 Sweden	1,059	614	216	107	7	359	1,245	715	0.64
25 UK	1,058	1,127	298	197	19	683	1,175	1,102	0.59
26 Bulgaria	1,135	369	47	92	1	234	1,092	831	0.49
27 Croatia	1,014	650	116	75	14	646	999	72	0.59
28 Romania	1,053	180	164	103	14	179	1,219	904	0.40
29 Turkey	1,036	145	110	76	65	200	215	103	0.92
30 Norway	1,000	348	198	62	6	158	308	254	0.84
31 Switzerland	1,040	709	681	174	213	779	1,073	934	0.64

The main indicators, computed based on survey outcomes are:

**Cooperation rate** is the ratio of the number of cases interviewed to the number of all eligible units ever contacted. There are both household-level and respondent-level cooperation rates. The rates computed here are household-level rates. They are based on contact with households, including respondents, rather than contacts with respondents only.

$$\text{COOP3} = \frac{I}{(I + P) + R}$$

Cooperation Rate 3 (COOP3) defines those unable to do an interview as also incapable of cooperating and they are excluded from the base.

<sup>6</sup> The estimated ineligible ration is computed according to the proportion of non-workers among all contacts with known employment status.

**Contact rate** measures the proportion of all cases in which some responsible member of the housing unit was reached by the survey.

$$\text{CON2} = \frac{(I + P) + R + O}{(I + P) + R + O + \text{NC} + e(\text{UH} + \text{UO})}$$

Contact Rate 2 (CON2) includes in the base only the estimated eligible cases among the undetermined cases.

**Refusal rate** is the proportion of all cases in which a housing unit or respondent refuses to do an interview, or breaks-off an interview of all potentially eligible cases.

$$\text{REF2} = \frac{R}{(I + P) + (R + \text{NC} + O) + e(\text{UH} + \text{UO})}$$

Refusal Rate 2 (REF2) includes estimated eligible cases among the unknown cases similar to RR3 and RR4 above.

**Response rate** is the ratio of completed interviews divided by all eligible units including the proportion of units of unknown eligibility that are estimated to be eligible.

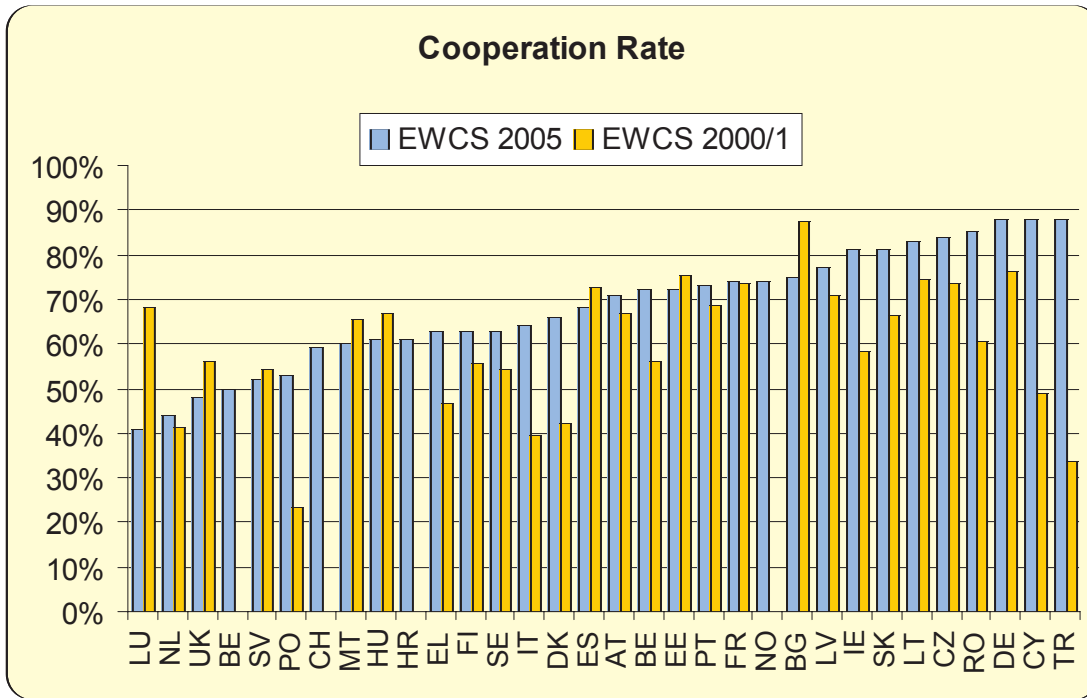
$$\text{RR2} = \frac{I}{(I + P) + (R + \text{NC} + O) + e(\text{UH} + \text{UO})}$$

The values of the main indicators for the whole sample are:

<b>cooperation rate:</b>	COOP3	<b>66%</b>
<b>contact rate:</b>	CON2	<b>77%</b>
<b>refusal rate:</b>	REF2	<b>24%</b>
<b>response rate:</b>	RR3	<b>47%</b>

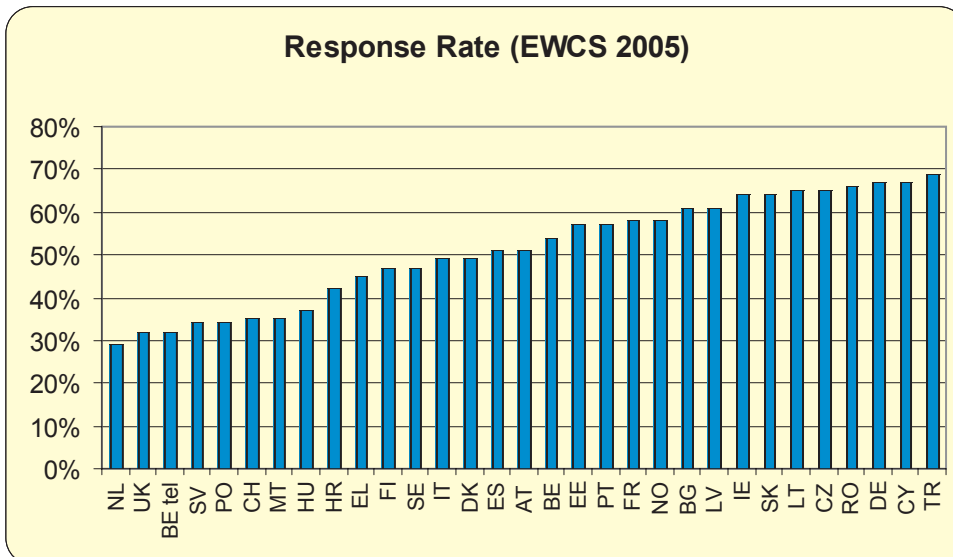
The cooperation rate is presented in Figure 9 in comparison with the same indicator of the 3rd EWCS. The rate ranges from 41% in Luxembourg to 88% in Turkey, Cyprus and Germany. In all but 7 countries (LU, UK, MT, HU, ES, EE and BG) there is an improvement in the contact rate.

Figure 9: Cooperation rate of the last two EWCSs



The response rate is presented in Figure 10. The response rate ranges from 28% in Netherlands to 67% in Turkey. The overall response rate was 47%. No comparable rates were calculated during the 3rd EWCS.

Figure 10: Response rates (RR3) of the EWCS survey by country. The overall response rate was 47%.



The values of the main indicators from survey outcomes is presented in Table 5.

Table 5: *Main indicators from survey outcomes*

	<b>cooperation rate:</b>	<b>contact rate:</b>	<b>refusal rate:</b>	<b>response rate:</b>
AAPOR code:	COOP3	CON2	REF2	RR3
EWCS 2005	0.66	0.77	0.24	0.47
F2F sampling, overall	0.69	0.80	0.23	0.51
Telephone sampling overall	0.53	0.67	0.30	0.34
1a Belgium telephone	0.50	0.72	0.34	0.34
1b Belgium face-to-face	0.72	0.51	0.11	0.29
2 Czech Republic	0.84	0.85	0.13	0.69
3 Denmark	0.66	0.70	0.21	0.42
4 Germany	0.88	0.76	0.08	0.61
5 Estonia	0.72	0.83	0.20	0.54
6 Greece	0.63	0.83	0.29	0.49
7 Spain	0.68	0.97	0.31	0.66
8 France	0.74	0.81	0.20	0.58
9 Ireland	0.81	0.72	0.12	0.51
10 Italy	0.64	0.78	0.27	0.49
11 Cyprus	0.88	0.75	0.08	0.57
12 Latvia	0.77	0.89	0.20	0.65
13 Lithuania	0.83	0.84	0.13	0.64
14 Luxembourg	0.41	0.88	0.46	0.32
15 Hungary	0.61	0.91	0.33	0.51
16 Malta	0.60	0.86	0.31	0.47
17 Netherlands	0.44	0.66	0.35	0.28
18 Austria	0.71	0.88	0.25	0.61
19 Poland	0.53	0.69	0.31	0.35
20 Portugal	0.73	0.95	0.24	0.67
21 Slovenia	0.52	0.77	0.34	0.37
22 Slovakia	0.81	0.79	0.14	0.58
23 Finland	0.63	0.68	0.21	0.35
24 Sweden	0.63	0.80	0.28	0.47
25 UK	0.48	0.77	0.36	0.34
26 Bulgaria	0.75	0.91	0.21	0.65
27 Croatia	0.61	0.77	0.29	0.45
28 Romania	0.85	0.85	0.11	0.67
29 Turkey	0.88	0.78	0.09	0.64
30 Norway	0.74	0.81	0.20	0.57
31 Switzerland	0.59	0.59	0.22	0.32

# Comparability

Comparability aims at measuring the impact of differences in applied statistical concepts and definitions on the comparison of statistics between geographical areas, non-geographical domains, or over time. We can say that comparability is the extent to which differences between statistics are attributed to differences between the true values of the statistical characteristics.

The factors that may cause lack of comparability between two statistical figures can be grouped into two major categories: those related to the concepts of the survey and those related to the measurement and estimation processes.

## Geographical Comparability

The statistical concepts we are going to examine include the definitions of the variables and the questions asked in order to measure them, the statistical unit the data refer to, the target population and the reference period of the surveys.

### Variables

Harmonisation is a main priority in the EWCS and therefore the variables have been implemented in a harmonised way based on a common master questionnaire and an elaborate process of translation.

Two web-based tools have been used to facilitate the process: a questionnaire database and a database-driven, semi-automated translation assistance system systems.

The questionnaire database included all elements of the master questionnaire including filters, instructions, question text, answer categories, etc.

In the beginning of the process the unchanged questions along with their translated versions were included in the database and their translations were checked by national experts in each language.

Afterwards, two forward translations were developed by well trained survey specialists with an excellent command of English. A third, synthesized version was agreed upon by the two translators and the local research director.

Translated questions then were translated back to English by an independent translator with an excellent knowledge of the local language and of English (but not necessarily survey research experience), and absolutely no familiarity with the original source questionnaire. Typically these translations were prepared by professional translators commissioned by the coordinating organisation. The new English version (produced by the back-translation) was then compared with the original version thus identifying any weaknesses in the translation.

It should be noted that in the evaluation of the back-translation, emphasis was placed on conceptual and functional rather than linguistic equivalence. Discrepancies were discussed until a satisfactory version is reached. A short list of discrepancies that remained unresolved after this review, were clarified in cooperation with the European Foundation.

National adaptations were then developed (e.g English for Ireland and Malta, French for Belgium, Luxembourg and Switzerland etc.) following a similar process (two forward translations and compromise and a back-translation)

Once all questionnaires were finalised and adapted within the network, the European Foundation circulated the versions among its network of country experts. They were invited to comment and to propose changes on the questionnaires and based on this last input the questionnaires were finalised.

### Statistical Population

The statistical population was persons in employment according to the Eurostat Labour Force Survey criteria (having worked for pay or profit at least one hour in the week preceding the survey). Any special cases (students, apprentices etc.) were treated in exactly the same way as in the LFS using the LFS manual and contacts with Eurostat's LFS team. In order to be consistent with the LFS output there were some small differences in the age bands of the statistical population. These are included in Table 6.

Table 6: Age limits for inclusion in the statistical population.

	Minimum age	Maximum age		Minimum age	Maximum age
<b>Belgium</b>	15	-	<b>Austria</b>	15	-
<b>Czech Republic</b>	15	-	<b>Poland</b>	15	-
<b>Denmark</b>	15	74	<b>Portugal</b>	15	-
<b>Germany</b>	15	-	<b>Slovenia</b>	15	-
<b>Estonia</b>	15	74	<b>Slovakia</b>	15	-
<b>Greece</b>	15	-	<b>Finland</b>	15	74
<b>Spain</b>	16	-	<b>Sweden</b>	15	74
<b>France</b>	15	-	<b>United Kingdom</b>	16	-
<b>Ireland</b>	15	-			
<b>Italy</b>	15	-			
<b>Cyprus</b>	15	-	<b>Bulgaria</b>	15	-
<b>Latvia</b>	15	74	<b>Croatia</b>	15	-
<b>Lithuania</b>	15	-	<b>Romania</b>	15	-
<b>Luxembourg</b>	15	-	<b>Turkey</b>	15	-
<b>Hungary</b>	15	74			
<b>Malta</b>	15	-	<b>Norway</b>	16	74
<b>Netherlands</b>	15	-	<b>Switzerland</b>	15	-

### Reference period

The reference period was the week (from Monday to Sunday) preceding the week the interview took place. Any differences related to the reference period depend on the survey period in each country. All national surveys started at the same week (19/92005), however in some countries the implementation of interviews took more time to complete (especially LU, BE, UK). The distribution of interviews for whole survey period is shown in Table 7.



Table 7: Interviews completed at the end of each week of the campaign.

Completed Interviews	09-19 - 09-25	09-26 - 10-02	10-03 - 10-09	10-10 - 10-16	10-17 - 10-23	10-24 - 10-30	10-31 - 11-06	11-07 - 11-13	11-14 - 11-20	11-21 - 11-27	11-28 - 11-30
Belgium	12	61	127	130	200	250	420	550	950	1,003	1,003
Czech Republic	98	207	810	930	940	1,027	1,027	1,027	1,027	1,027	1,027
Denmark	40	175	332	502	707	1,006	1,006	1,006	1,006	1,006	1,006
Germany	60	233	358	502	843	950	1,018	1,018	1,018	1,018	1,018
Estonia	11	146	237	470	NA	NA	602	602	602	602	602
Greece	30	201	321	492	752	1,001	1,001	1,001	1,001	1,001	1,001
Spain	7	47	NA	364	679	900	1,017	1,017	1,017	1,017	1,017
France	65	NA	541	857	857	967	NA	981	1,083	1,083	1,083
Ireland	35	92	295	442	678	905	967	1,009	1,009	1,009	1,009
Italy	279	683	805	1,005	1,005	1,005	1,005	1,005	1,005	1,005	1,005
Cyprus	67	195	252	435	557	600	600	600	600	600	600
Latvia	104	214	302	399	748	799	1,003	1,003	1,003	1,003	1,003
Lithuania	35	219	539	974	998	1,017	1,017	1,017	1,017	1,017	1,017
Luxembourg	8	36	65	100	110	131	162	257	413	600	600
Hungary	197	467	650	1,001	1,001	1,001	1,001	1,001	1,001	1,001	1,001
Malta	125	246	324	448	584	600	600	600	600	600	600
Netherlands	16	50	134	337	648	848	1,025	1,025	1,025	1,025	1,025
Austria	30	92	129	150	370	579	838	999	1,009	1,009	1,009
Poland	229	308	468	810	903	1,000	1,000	1,000	1,000	1,000	1,000
Portugal	73	125	165	428	544	613	852	1,000	1,000	1,000	1,000
Slovenia	81	179	263	350	470	560	600	600	600	600	600
Slovakia	400	638	828	900	999	1,024	1,024	1,024	1,024	1,024	1,024
Finland	4	55	172	325	521	642	867	970	1,059	1,059	1,059
Sweden	5	89	325	574	752	1,059	1,059	1,059	1,059	1,059	1,059
United Kingdom	12	22	135	135	229	400	715	821	978	1,034	1,058
Bulgaria	454	650	1,135	1,135	1,135	1,135	1,135	1,135	1,135	1,135	1,135
Croatia	350	435	445	580	610	900	940	965	1,011	1,011	1,011
Romania	224	286	757	972	999	1,053	1,053	1,053	1,053	1,053	1,053
Turkey	165	259	NA	389	637	1,015	1,015	1,015	1,015	1,015	1,015
Norway	28	87	181	318	392	561	807	993	1,000	1,000	1,000
Switzerland	53	190	331	519	705	793	867	1,028	1,040	1,040	1,040

## **Methodological issues**

The methodology used in the different stages of the survey process was quite harmonised.

**Sample Selection.** The method used in most countries was the random route selection. In 4 countries (Belgium, Sweden, Netherlands and Switzerland) the sample was selected by telephone screening. Telephone numbers were selected using a telephone register or random dialling digits and an appointment was arranged for the interview. This choice was made in order to obtain better response to the survey as it was suggested that potential respondents would be more likely to participate in the survey if they are first contacted via telephone. This seemed to create no substantial problems in Sweden and Netherlands as the coverage of population is 100 and 96% respectively (see the section on coverage errors, page 15). In Belgium, due to the large percentage of out-of-grid households the random route was tried at first, however, after the first two weeks of the campaign, the problems were so great that the sample selection procedure changed to telephone screening coupled with a semi-random digit scheme that ensured coverage of out of directory and mobile-only households.

The difference in the sample selection procedure might create incomparabilities in two ways:

1. Due to differences in coverage of the statistical population.
2. Due to differences in the choice of people to participate in the survey or not (assuming that people react different to the survey request if it is made at the doorstep than via telephone).

In both cases incomparabilities will occur if these differences are to some extent correlated with the variables measured. This is something that we can not establish but also can not rule out. It should be noted however that both coverage errors and non-response errors are controlled by weighting so any comparability problems in the unweighted estimates should be reduced in the weighted results.

The random route is a process that delegates to a large extent the sample selection to interviewers. Its implementation is important to be done in a uniform manner, according to the specifications. During the beginning of fieldwork an audit was carried out by the Foundation that discovered that the random route (more specifically the replacement of sampling units) was not carried out in accordance to the specifications in two countries. This was corrected immediately and also guidelines were sent to all participating national organizations emphasizing the proper way to carry out the sample selection in the random route.

**Data collection.** The Uniformity during data collection was ensured with the development of guidelines and training material for all respondents which was subsequently translated in all national languages. In 29 countries a printed questionnaire was used to collect data (PAPI) while in Finland and Switzerland CAPI was used.

**Data processing.** Data entry, coding and editing was done by national organisations while weighting was performed by the coordinating organisation. Data entry was performed manually in most countries, in 6 countries (CZ, DK, DE, NL, SK, NO) the questionnaires were scanned and data was captured by OCR software while in two data was captured during the interview (CAPI). Data coding was performed by competent coders based on uniform instructions. Based on the reported problems and the distribution of occupation categories we infer that it is possible to have inconsistencies in the coding of the occupation variable. Data editing was performed using uniform methodology at different stages of the process and detailed guidelines for specific checks. It was also checked globally from the coordinating organisation. Finally, weighting was done centrally by the coordinating organisation based on LFS data, so no incomparabilities were introduced during weighting other than potential problems that the LFS data might have.

Finally issues that are unrelated to the EWCS may cause lack of comparability to some extent. More specifically cultural differences related to working conditions or differences in the related legal framework between any two countries may also affect the comparability of results. This is of course intertwined with the interpretation of results.

## **Comparability over time**

The 4th EWCS marks 15 years of data collection by the European Foundation. The first survey took place in 1991 and covered 12 member states. The second edition was done in 1996 and covered 15 member states. The third edition was done in 2000 and covered the same 15 member states, as well as Norway. It was complemented in 2001/2002 by a survey in 13 acceding and candidate countries. During these 15 years the survey has changed considerably in both concepts and methods.

### **Survey concepts**

The statistical population was always the same in broad terms. However, in the 4th EWCS the survey population has been defined in more detail so that it is in line with the LFS definition.

The variables measured in each EWCS change following the evolving needs for working conditions information and the feedback from previous survey implementations. Out of a total of 118 variables measured (some multiple questions measure more than one variable) 32 are completely new, 34 were used in the 2000/2001 surveys but in a different, but completely comparable way and 52 have remained unchanged (trend questions). A complete list of variables and the concordance with the variables used in the 2000 and 2001 surveys is included in Annex 5.

The 2000 surveys in EU-15 took place in March-April and the 2001 surveys in accession and Candidate countries in May-July (the 1995 survey in the EU-15 countries took place between 27/11/1995 and 19/1/1996). As the 2005 survey was implemented between September and November differences in working conditions that may be affected by seasonal variations should be interpreted with care.

### **Methodological Issues**

The main aspects of methodology have not changed since the 2000/2001 survey. Both surveys were implemented using a multistage sampling design based on the random route procedure. There were a maximum of 3 revisits, one working person was selected in each household by the next birthday method and the data collection was done using PAPI. A known bias, the absence of selection probability weighting in all past EWCSs, has been corrected for the 2000/2001 surveys by re-weighting the results (Also the final concurrent LFS results were used in the re-weighting process). However detailed documentation of the implementation that formed part of the Quality Assurance Plan of the 2005 survey is not available for the 2000/2001 surveys and therefore a detailed comparison in methodological issues is not possible.

Finally it should also be noted that the databases of all previous EWCS have transformed in the format and metadata structure (classifications, variables) of the 2005 EWCS.

# Coherence

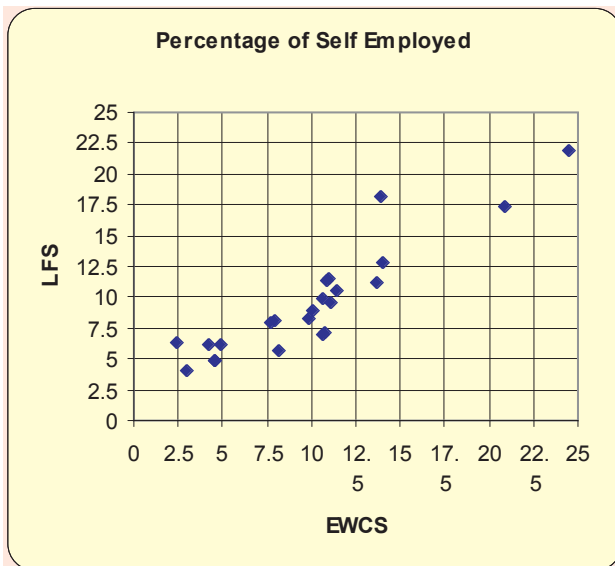
Coherence of statistics is their adequacy to be reliably combined in different ways and for various uses. The basis for assessing the coherence of EWCS with other sources is inconsistencies between results.

As this quality assessment takes place, data from Eurostat's LFS are available for the same quarter (3rd 2005). The overlap between EWCS and LFS is small and includes (besides the variables already used in weighting) 4 variables: employment status, working hours, part-time work and temporary employment.

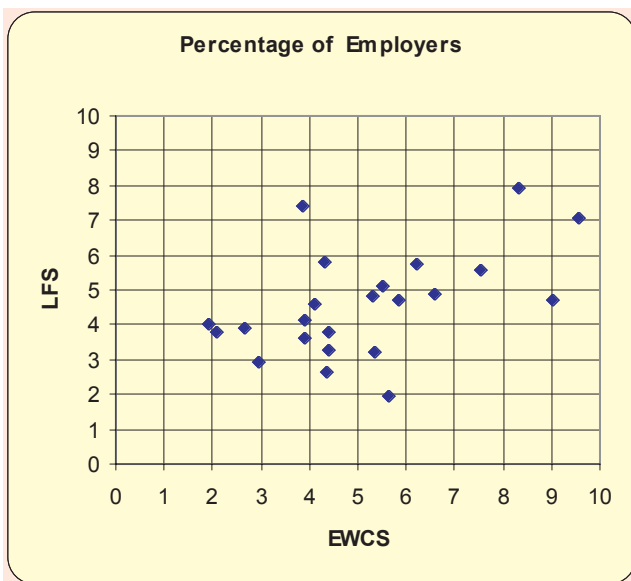
## Employment status

The employment status variable is not identically used in the two surveys. The LFS distinguishes self-employed, employers and employees, as the EWCS does. However instead of the "other" category of the EWCS it uses the family workers category.

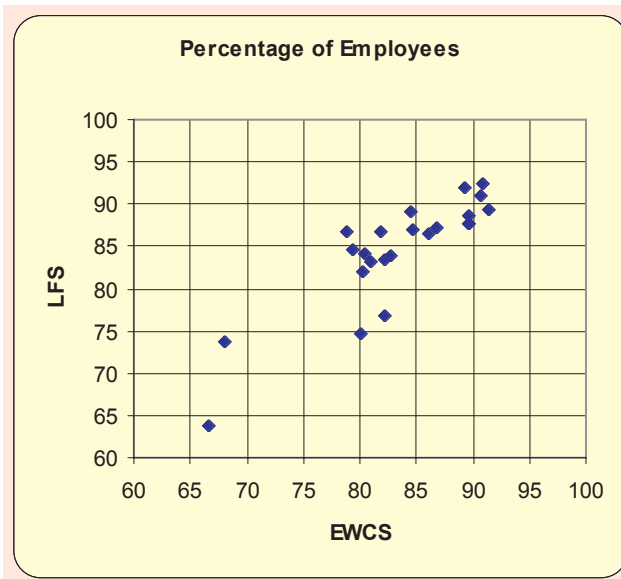
In the following figures we present scatterplots of country estimates from the LFS against the ones from the EWCS for these 4 levels:



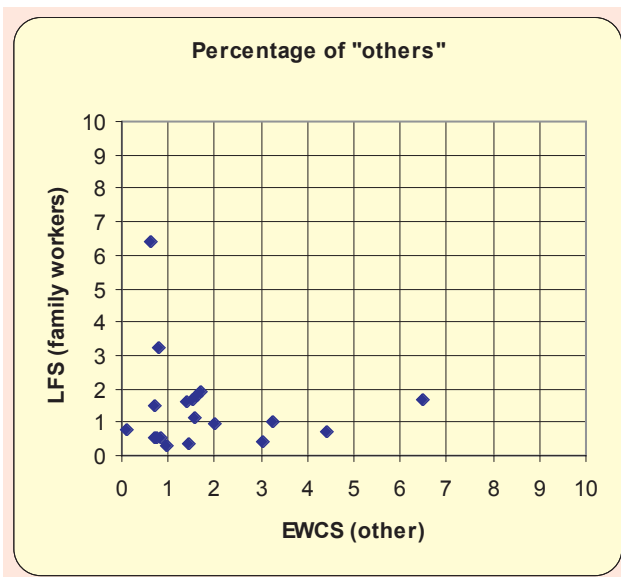
The percentage of self employed seems to be consistent between the two surveys. There is no bias (a t-test fails to find a difference). This is evident from the position of the data cloud with respect to the main diagonal (it is roughly symmetrical around the diagonal). There are however some large differences. In DE the estimates are 2.36% for the EWCS against 6.26% for the LFS.



The percentage of employers presents larger discrepancies in 15 countries the EWCS estimate is larger while in 8 the LFS estimate is larger. Although a t-test still fails to find a significant difference the EWCS seems to somewhat overestimate the percentage of employers. The variability is about the same (the largest deviation is in Austria where the EWCS estimate is 9.01% against 4.69% of the LFS) in absolute terms (although the correlation coefficient is smaller).



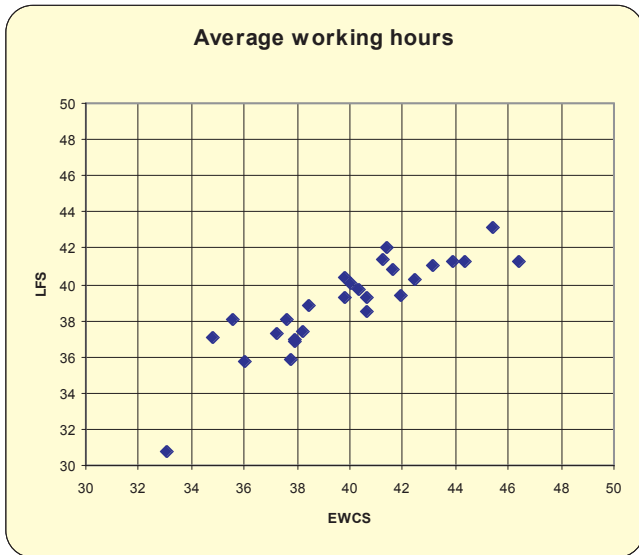
The percentage of employees is showing roughly the same discrepancies although the effect is smaller in relative terms as the numbers are larger. In 16 countries the EWCS estimates are smaller than the LFS ones and in 8 they are larger. Again a test for bias fails to find substantial differences.



The last category of both surveys is clearly inconsistent. The differences are large and the correlation between the values is very small. This inconsistency stems from the different implementation in the questionnaire of the two surveys. In the EWCS an "other" category is included while in the LFS the category refers to family workers. This inconsistency might explain to some extent the differences in the other categories of the type of employment variable.

## Average hours worked

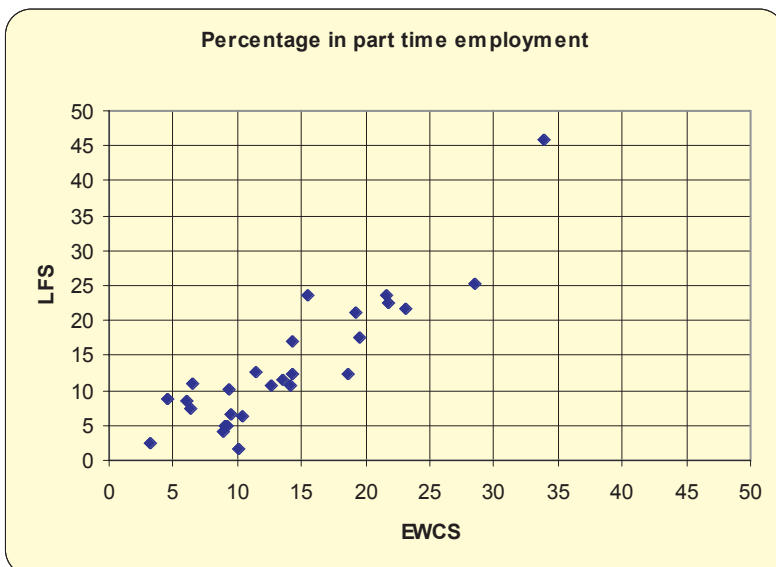
One of the continuous variables measured in the EWCS is the number of hours usually worked by each respondent. The same variable is measured in the LFS and averages are published. By computing averages from the EWCS data we can compare the indicator average hours worked for the two surveys based on a scatterplot as before



The differences found between the two surveys are substantial (a t-test finds significant differences in the means) and in most cases the EWCS estimates are larger than those of the LFS. Also the overestimation is more prevalent in the large values (EWCS averages of over 42 hours per week are always larger than LFS averages). This is a surprising result. One commonly quoted reason for non-response is the busy working lives of potential respondents. It was therefore expected that due to the low response rate of busy people the EWCS would actually underestimate working hours. On the contrary the EWCS seems to overestimate working hours at least over the LFS estimates.

## Part-time employment

Part time employment is also included in the LFS list of variables. The proportion of people employed in part-time contracts is disseminated and can be compared with the relevant EWCS averages.



The picture follows earlier results. There seems to be no systematic bias and the values are highly correlated. Some differences however are quite large: in Netherlands and Germany the LFS estimates are 46% and 23.6% respectively while the EWCS estimates are 33.9% and 15.4% respectively. On the other hand in Ireland and Bulgaria the LFS estimates are 12.4% and 1.7% respectively while the EWCS ones are 18.6 and 10.1 respectively. These differences are very large and cannot be explained by chance only. Especially the Bulgarian EWCS estimate is six times larger than relevant LFS one; this indicates potential problems in the implementation of the specific variable/question in these countries in the EWCS, the LFS or both. It should be noted that while in the LFS variables are specified but questions are not, in the EWCS, the questionnaire was developed once and was translated for the survey in each country. To some extent, the percentage of part-time work in the LFS is constructed in different ways in different countries (in some cases based in the number of hours worked, in others based in the contractual situation reported by the respondent, etc), whereas in the EWCS the percentage of part-time is directly based on the answers of respondents (whether they consider themselves working part-time or full-time).

### Temporary employment

The last variable used in this comparison is the most problematic one. As with the part-time employment it is a variable that is understood, regulated and practiced in a different manner in different member states.



The comparison with the use of the scatterplot reveals a relatively small correlation (the pearson correlation coefficient is 0.628 while in most of the other variables it was close to 0,90) between the two variables with a tendency for overestimation in the part of the EWCS. A reasonable hypothesis is that in some countries (especially in the UK and Ireland where estimates are much larger) people understand their formally indefinite contract as been only temporary (i.e. that perceptions are different than reality).

Michalis Petrakos Agilis - SA

# Annex 1: Sample Evaluation tests

	N	Min weight	Max weight	Age & Sex			Regions (based on NUTS2)			Occupational groups (based on ISCO88)			Activity of employer (based on NACE)		
				X2	DF	p-value	X2	DF	p-value	X2	DF	p-value	X2	DF	p-value
BE	1,003	0.33	3.02	48.40	5	0.00000000	17.86	9	0.03678041	160.63	7	0.00000000	106.8849	3	0.00000000
CZ	1,027	0.41	3.71	46.46	5	0.00000001	9.35	7	0.22825897	181.39	6	0.00000000	22.50459	3	0.00005122
DK	1,006	0.37	3.40	30.25	5	0.00001318	0.00	0		203.42	7	0.00000000	116.5571	3	0.00000000
DE	1,018	0.33	3.01	9.76	5	0.08229594	17.67	12	0.12608565	77.95	7	0.00000000	113.1001	3	0.00000000
EE	602	0.36	3.27	30.88	5	0.00000991	0.00	0		33.51	5	0.00000298	5.834349	4	0.21186544
EL	1,001	0.33	3.01	9.28	5	0.09832142	39.66	10	0.00001941	28.12	8	0.00045232	18.90675	4	0.00081980
ES	1,017	0.33	3.00	112.57	5	0.00000000	17.34	12	0.13727389	299.89	7	0.00000000	169.7942	4	0.00000000
FR	1,084	0.37	3.33	29.37	5	0.00001957	44.10	17	0.00033137	93.96	6	0.00000000	17.10004	3	0.00067404
IE	1,009	0.33	3.01	53.98	5	0.00000000	0.76	1	0.38244707	95.50	7	0.00000000	31.97166	4	0.00000194
IT	1,005	0.33	3.01	176.04	5	0.00000000	25.19	10	0.00499506	267.36	6	0.00000000	56.93146	2	0.00000000
CY	600	0.44	3.96	19.83	5	0.00134208	0.00	0		80.20	5	0.00000000	3.403937	2	0.18232430
LV	1,003	0.38	3.42	60.59	5	0.00000000	0.00	0		45.99	7	0.00000009	53.07037	4	0.00000000
LT	1,017	0.33	3.04	117.31	5	0.00000000	0.00	0		37.79	4	0.00000012	92.00105	4	0.00000000
LU	600	0.34	3.11	10.95	5	0.05246877	0.00	0		45.83	4	0.00000000	10.14981	3	0.01733471
HU	1,001	0.33	3.00	40.37	5	0.00000013	16.01	6	0.01368195	104.07	7	0.00000000	35.41866	4	0.00000038
MT	600	0.33	3.00	31.55	5	0.00000728	0.00	0		162.15	7	0.00000000	1.876085	3	0.59851964
NL	1,025	0.33	3.00	184.85	5	0.00000000	67.28	8	0.00000000	62.84	5	0.00000000	171.97	3	0.00000000
AT	1,009	0.33	3.00	30.34	5	0.00001264	14.68	8	0.06570412	162.05	6	0.00000000	30.38573	4	0.000000408
PL	1,000	0.33	3.03	41.03	5	0.00000009	22.04	13	0.05481590	99.24	5	0.00000000	131.1807	4	0.00000000
PT	1,000	0.33	3.00	34.97	5	0.00000152	2.93	5	0.71058116	323.87	7	0.00000000	88.11653	3	0.00000000
SL	600	0.44	4.02	21.50	5	0.00065262	0.00	0		65.54	6	0.00000000	17.75725	2	0.00013934
SK	1,024	0.35	3.14	134.21	5	0.00000000	7.11	3	0.06858031	60.60	7	0.00000000	31.17253	4	0.00000282
FI	1,059	0.33	3.03	27.87	5	0.00003854	20.03	3	0.00016754	56.60	5	0.00000000	209.3563	4	0.00000000
SE	1,059	0.36	3.31	47.35	5	0.00000000	2.94	7	0.89018689	245.29	7	0.00000000	45.16259	3	0.00000000
UK	1,058	0.39	3.51	25.57	5	0.00010813	46.44	23	0.00263915	46.25	7	0.00000008	8.329402	3	0.03967261
BG	1,135	0.36	3.29	79.39	5	0.00000000	19.59	5	0.00149004	190.45	7	0.00000000	51.57159	4	0.00000000
HR	1,011	0.33	3.01	140.66	5	0.00000000	3.10	3	0.37681194	197.45	5	0.00000000	183.3779	3	0.00000000
RO	1,053	0.34	3.09	15.27	5	0.00926220	3.56	7	0.82926786	231.89	8	0.00000000	150.3582	3	0.00000000
TR	1,015	0.33	3.04	106.44	5	0.00000000	37.57	19	0.00673739	174.61	8	0.00000000	0.08617	1	0.76910352
NO	1,000	0.34	3.13	24.51	5	0.00017315	71.56	6	0.00000000	235.94	7	0.00000000	1.509845	3	0.67999987
CH	1,040	0.38	3.42	68.21	7	0.00000000	9.17	6	0.16431515	108.18	7	0.00000000	2.652467	3	0.44836558



## Annex 2: Error rates as reported by the national institutes

	Re-entry (%)	Number of cases re-entered	Mistakes (N )	Error rate
Belgium	10	101	0	0
Czech Republic	none, scanned	..	..	..
Denmark	none, scanned	NA	0	0
Germany	none, scanned	..	..	..
Estonia	10	60	24	0,15
Greece	10	100	42	0,16
Spain	10	101	0	0,00
France	10	110	30	0,10
Ireland	100	all cases	NA	NA
Italy	10	100	3	0,01
Cyprus	10	60	2	0,01
Latvia	10	100	3	0,01
Lithuania	10	100	10	0,04
Luxembourg	10	60	15	0,10
Hungary	10	100	21	0,08
Malta	10	60	48	0,31
Netherlands	none, scanned	..	..	..
Austria	10	100	32	0,12
Poland	10	100	35	0,13
Portugal	10	100	0	0,00
Slovenia	11	66	60	0,35
Slovakia	none, scanned	..	..	..
Finland	CAPI	..	..	..

# Annex 3: Duration of Interviews

		Interviews' duration			Duration categories				
		average	MIN	MAX	<20	20-29	30-39	40-49	50<
	EWCS 2005	35	14	155	1.6	25.7	42.0	20.7	10.0
1.	Belgium	35	20	120	..	27.6	39.8	21.5	11.1
2.	Czech Republic	42	20	120	..	11.6	31.4	28.8	28.2
3.	Denmark	30	15	80	7.7	37.6	38.3	13.1	3.4
4.	Germany	42	20	105	..	9.6	31.1	33.3	25.9
5.	Estonia	40	20	90	..	12.6	41.9	23.8	21.8
6.	Greece	31	15	60	0.5	27.8	56.2	14.4	1.1
7.	Spain	28	15	60	10.4	48.8	24.7	14.7	1.4
8.	France	36	15	95	0.3	9.5	51.8	28.1	10.2
9.	Ireland	32	15	155	4.5	37.9	36.7	15.1	5.9
10.	Italy	38	20	60	..	12.0	41.8	31.5	14.6
11.	Cyprus	36	25	72	..	9.2	56.0	27.2	7.7
12.	Latvia	36	20	65	..	16.6	55.3	23.8	4.3
13.	Lithuania	41	15	120	0.5	19.4	31.9	20.6	27.6
14.	Luxembourg	34	15	120	3.4	32.1	35.9	14.8	13.7
15.	Hungary	36	14	90	0.6	18.1	44.7	27.0	9.7
16.	Malta	32	20	75	..	37.3	45.5	11.8	5.3
17.	Netherlands	37	15	100	0.1	8.8	51.0	28.9	11.2
18.	Austria	36	17	70	0.3	14.1	46.9	27.6	11.2
19.	Poland	34	23	64	..	20.6	54.3	21.2	3.9
20.	Portugal	29	15	45	3.3	44.1	35.4	17.2	
21.	Slovenia	30	14	60	5.0	38.8	41.8	11.7	2.7
22.	Slovakia	42	15	140	0.2	13.5	36.3	24.5	25.5
23.	Finland	36	15	90	0.5	21.2	48.8	20.5	9.0

## Annex 4: Edit rules

Question	Method of correction
Questionnaire ID	Check match with the Routing Slip, in case of difference, clarify. Also check to have unique ID for each case
P1	The date of the interview should be identical with the date written at the given ID on the Routing Slip. The latter is regarded to be the basic data, questionnaire data might be corrected against that, and if the code box is empty, data can be written from it (correctable error)
P3	In case of a too short interview time the interviewee should be recalled, the length of the interview checked. Unexplained length below 10 minutes should be checked,
P6	Check against sample stratum – must match the urbanisation category you used in the PSU allocation table
P7	Check against sample stratum – must match the region category you used in the PSU allocation table
P9	Make sure that you will have the necessary number of clusters
HH1	The minimum number is 1. The written number should be identical with the number of rows in the household table. If there is more rows in the HH table, correct HH1 accordingly.
Household table column “A”	If incomplete at the interviewee, it can be filled in on the basis of the name/sex indicated on the Routing Slip (correctable error). Incomplete code in the next rows can perhaps be corrected, and if this is not possible, then should be checked over the telephone. It can be concluded e.g. if the interviewee is a male, in the incomplete row the spouse is listed (code 01 in column “C”), then the sex is female, and the other way around.
Household table column “B”	Various data are logically traceable in agreement with data in column “C”. E.g. the child of the interviewee should be obviously born later than the interviewee himself, and certain difference is expected to be there in age, e.g. at the age of 10 no one could have a child. The same relation can be expected for parents, and for grandchildren, as well. If the given person is younger than 1 column “D” only number 10 can be written, if the code is different, it should be corrected to 10 (correctable error).  Care should be taken that not the year of birth but age should be written in, if year of birth is coded, then it should be changed to age (2005 – year of birth) – correctable error. Code field should not remain empty. Number below 15 should not be in the first row. Such questionnaire should be checked carefully.
Household table column “C”	In relation to the age in column “B” certain corrections, substitutions can be done, and in cases when this is not unambiguous, then it should be checked with the interviewee. (age and sex should not contradict to relationship status, i.e. a 5 years old cannot be a spouse of anyone, and same sex marriages are more likely to be mistakes than not)
Household table column “D”	If column “B” contains a number below 14, the given person cannot possibly be active.
Q1A	If there is a difference, always the circled code should be accepted, not the one written in the code box.
Q1B	Only to be asked if Q1A=2, if it is 1, the code box should remain empty. If Q1A=1, but Q1B has a code, then Q1A should be changed back to 2. At Q1B, if the interviewer wrote nationality down, because did not know which code to use, then it should be coded during checking (correctable errors)
Q2A	If the description is not detailed enough, further reconciliation with the interviewee should be done. E.g. not acceptable: skilled worker, entrepreneur, intellectual, etc. In case of illegible writing the instructor should have the text rewritten in large printed letters by the interviewer.
Q2B	In case of unrealistic numbers the interviewee should be consulted, e.g. above 40, and in most countries anything below 6 is suspicious as well
Q2C	The sum of Q2B and Q2C cannot be bigger than the age of the respondent in the “B” column of the Household table. If significant difference exists, then the respondent probably misunderstood question Q2B or Q2C, which can only be clarified with him. (non-correctable error)
Q2D	The number usually should not be bigger than Q2C (however in some cases it might happen, because of internships, etc). Pay attention to numbers higher than Q2C and if there are more such numbers than check with interviewer
Q3A	The appropriate skips should be checked. Unnecessary questions can of course be corrected, unasked questions can only be clarified with the respondent. If the response is 1, then Q7 may only have 0, if it is 2, then only numbers bigger than 0. For codes 3 and 4 there is no such correspondence. (correctable error)
Q3B	Number or code can only be here if Q3A=3, otherwise the question should not be asked, if in such a case a code is found, it should be crossed. (correctable error)

<b>Question</b>	<b>Method of correction</b>
Q3C	The number of months can only be less than 12, if it is more, then it should be changed to years, e.g. 14 months is 1 year and 2 months. If the question had to be asked, but code is missing here, this can only be clarified with the respondent, and in lack of that, only data omission can be recorded (incorrigible error)
Q4	If the description is not detailed enough, it should be clarified with the interviewee, in case of illegible writing the instructor should have the text rewritten by the interviewer. Unacceptable e.g. industry, agriculture, etc.
Q5	Problematic cases are might be noted by the interviewer, on the basis of which the correctness of classification can be checked, otherwise the circled code should be accepted.
Q8A	Numbers above 90 (if working five days a week) or 120 (if working seven days a week – see Q8B) are unrealistic, the respondent should be consulted, and can only be accepted if the interviewer casts light on the meaning of the extreme number with some notes. If it cannot be clarified on the basis of the questionnaire, clarification with the respondent should be attempted.
Q8B	Only numbers below 8 may be here. If the number is 6 or 7 then it is certain that in either Q14C or D there is a number different from 0. The correlation is not unequivocal the other way around. In case of logical error either data omission should be recorded, or if possible it should be clarified with the respondent.
Q9A	Q9B is only to be asked if the code is 2 1, if there is no code, data omission should be recorded.
Q9B	Added to Q8A the result cannot be unrealistic (above 90 or 120), if so, it should be checked, consulted with the respondent. If the interviewer wrote a number, but Q9A is not 2, then at Q9a the code should be changed to 2. (corrigible error)
Q10	The location of the beginning of the rotation should be marked, if it is not, then most probably the interviewer did not follow the rules of interviewing, which could result in quality problem. No row may remain empty, 9 should be coded if there is no answer in any of the rows. The responses of the interviewee should be accepted without any correction, even if his employment would suggest the existence of different circumstances. If the responses of the interviewee in the questionnaire continue to be noticeably irrelevant to his way of living, occupation, the head of the management should have the work of the interviewer fully checked. The checking should cover irrelevant answers, as well.
Q11	The same notes are applicable to this table as to Q10. If the interviewer jots down any comment of the respondent, because the answers cannot be unambiguously coded for certain items, these are to be coded by the supervising personnel at the Institute.
Q13	Care should be taken to write the number of minutes. If for example the interviewer writes 2 hours, it should be changed to 120 (corrigible error). In case of unreal numbers data should be clarified with the interviewer/respondent, but if the interviewer himself provided a note to the question justifying the suspicious data, then it can be accepted. The code can also be 0, but this should be in agreement with the other data.
Q14A-B	Only numbers lower than 31 can be represented, in case of bigger numbers the interviewee should be consulted, and if this is not possible, then no data to be circled (non-corrigible error) If this number is around 30-31, it is only possible if the respondent works 7 days a week, and works through the whole Sunday and Saturday. (( Q8B, Q14C, Q14D) If the correlations of the 4 questions are not right, then the given questions can only be corrected in consultation with the respondent.
Q14C-D	4 or less can be here. In case of irrelevant data, those set out at the above question should be followed.
Q14E	If this number is 30, it is only possible if the respondent works 7 days a week, and works through the whole Sunday and Saturday. ( ( Q8B, Q14C, Q14D) If the correlations of the 4 questions are not right, then the given questions can only be corrected in consultation with the respondent.
Q15A	In the case of a part-time worker, it is unlikely that he or she works 40 hours or more a week (Q8A) If data conflict, the appropriate code should be determined.
Q16	Following the circling of positive answers, if the interviewer does not mark the negative answers, 2 should be circled and coded (correctable error)
Q20A	Two negative answers is also possible.
Q25	For omitted items “no data” is to be coded
Q26A	If the respondent does not have a co-worker (Q6=0), then code 1 is excluded, therefore, obviously, such respondents could not have been asked either 26A1 or 26A2.
Q26B,Q26B1	If he does not work together with colleagues, only alone, then Q26B cannot be 1. To be corrected to 2.
Q28	If items A, B are 1, then data should be found at the number of days, and if there are none, then data omission should be marked, or, if possible, it should be consulted with the respondent. Number of days.

Question	Method of correction
Q30-31	Not to be asked from self-employed, if it was asked, the data should be crossed. If the interviewer did not ask from an employee, then, if possible, respondent should be recontacted, perhaps in a follow-up telephone interview. Q31 can only be 7, if the respondent does not have a superior, which can only happen very rarely, as we are talking about employees.
Q33-Q33A	If Q33=2, but Q33A is filled, then code 2 should be corrected to 1 in Q33. If the interviewer only marked the positive answers in Q33A, then negative answers should be marked, but beforehand it should be checked with the interviewer whether the respondent could not or did not want to answer the given item? If there is no mark only in row Q, then 2 should be coded there.
Q34B	Only numbers smaller than 365 and bigger than 000 are allowed. If the interviewer wrote DK or REFUSE on the dotted line, then 888 or 999 is to be coded in accordance with the meaning of the text.
Q34C	In case of no data or a horizontal line by the interviewer, 000 should be coded. Neither C1 and C2, nor their total can exceed the value at Q34B. If no correspondence is found, it should be clarified with the respondent, otherwise no data can be selected as basis.
Q35	Not to be asked if the Age column in the first two rows of the Household table is 60 or more. In such cases the code should be crossed, an obvious interviewer mistake, but can be corrected. Otherwise "no data" is to be coded.
Q37	In case of certain items for those not employees, some are difficult to understand, here the DK or NA response is acceptable.
EF1	Has to be in agreement with Q2A and and Q2B. Cannot be corrected unless it is consulted this with the interviewee. If there are significant discrepancies, (e.g. ISCED 2 level education, high level professional who stopped full time education at the age of 26) than code EF1 as missing.
EF3	If the household is 1 person (HH1=1), then this can only be 1. If this is not the number, it has to be corrected
EF4-EF41	EF4.1 is to be asked only in case of code 1, if other code has a number it has to be crossed, if code 1 does not have a number then data omission should be coded.
EF6	Items G and H have special significance besides the general correction guidelines set out above for the tabular questions. Questions EF6G and H should be asked in case the answer here is 1, if the interviewer unnecessarily asked the questions then the answers are to be crossed, if he did not ask these, then data omission should be recorded, or if possible, it should be clarified with the respondent (non-correctible error). For A-K, any of the items will need to have a positive mention, otherwise the case should be checked.
EF7	Should only be asked from entrepreneurs and self-employed, and not from employees. Any of the A-E items will need to have a positive code, otherwise the case should be checked.

# Annex 5: Questionnaire Concordance

2005	2000	2001	Comments
EU-25	EU-15	NMS-10, RO, BG	
HH1			New
HH2			New
HH3			New
Q1a-b	Q1	Q1a-c	Modified
Q2a	Q2a	Q2a	Trend
Q2b		EF8	Trend (omitted in 2000 survey)
Q2c			New
Q2d	Q3a	Q2c	
Q3a	Q4a	Q35a	Trend
Q3c	Q4b	Q35b	Modified
Q3d	Q4c	Q35d	Trend
Q4	Q5	Q3	Trend
Q5	Q6		Modified (omitted in 2001 survey)
Q6	Q7	Q4	Trend
Q7	Q8	Q5	Trend
Q8a	Q14	Q11	Trend (see 2001 survey questionnaire for full version of Q)
Q8b			New
Q9a	Q9	Q3	Modified
Q9b	Q10	Q7	Modified
Q10.1	Q11.1	Q8.1	Trend
Q10.2	Q11.2	Q8.2	Trend
Q10.3	Q11.3	Q8.3	Trend
Q10.4	Q11.4	Q8.4	Trend
Q10.5	Q11.5	Q8.5	Modified
Q10.6			New
Q10.7	Q11.6	Q8.6	Modified
Q10.8	Q11.7	Q8.7	Trend
Q10.9			New
Q10.10			New
Q11.1	Q12.1	Q9.1	Trend
Q11.2			New
Q11.3	Q12.2	Q9.2	Trend
Q11.4			New
Q11.5	Q12.3	Q9.3	Trend
Q11.6			New
Q11.7	Q12.5	Q9.5	Trend
Q11.8	Q12.6	Q9.6	Trend
Q11.9			New
Q11.10	Q12.7	Q9.7	Trend

2005	2000	2001	Comments
EU-25	EU-15	NMS-10, RO, BG	
Q11.11	Q12.4	Q9.4	Trend
Q11.12			New
Q11.13	Q12.8	Q9.8	Trend
Q12	Q13	Q10	Modified
Q13	Q15	Q12	Trend
Q14a-e	Q16a-e	Q13a-e	Trend
Q15a	Q17a	Q14a	Modified
Q15b	Q17b	Q14b	Modified
Q16a	Q18a	Q15a	Modified ... now incorporates 2000-Q18b shift question
Q16b	Q18c	Q15b	Modified
Q17a	Q19a	Q16a	Modified (2001 version used as model for 2005)
Q17b	Q19b	Q16b	Modified (2001 version used as model for 2005)
Q18	Q20	Q17	Trend
Q19			New
Q20a	Q21a	Q18a	Modified (some answer categories deleted)
Q20b	Q21b	Q18b	Trend
Q21	Q22	Q19	Modified
Q22a	Q23a	Q20a	Trend
Q22b	Q23c	Q20c	Trend
Q23	Q24	Q21	Trend
Q24	Q25	Q22	Trend
Q25.1	Q26.1	Q23.1	Modified (answer categories changed from binary to scale)
Q25.2-4			New
Q25.5	Q26.2	Q23.2	Modified (answer categories changed from binary to scale)
Q25.6	Q26.5	Q23.5	Modified (answer categories changed from binary to scale)
Q25.7	Q26.3	Q23.3	Modified (answer categories changed from binary to scale)
Q25.8-13			New
Q26	Q27b	Q24b	Modified
Q27	Q28	Q25	New (new phrasing of 2000-Q28)
Q28	Q29	Q26	New (extended version of 2000-Q29)
Q29			New
Q30.1			New
Q30.2-11	Q31.1-10	Q28	Trend
Q31	Q33	Q29	Trend
Q32	Q34	Q30	Trend
Q33	Q35	Q31	Trend with some minor modifications, ie. amalgamating the three muscular pain symptoms into one category
Q34a-c	Q36a-c	Q32a-c	New, modified and extended version of 2000-Q36a-c
Q35	Q37	Q33	Trend

2005	2000	2001	Comments
EU-25	EU-15	NMS-10, RO, BG	
Q36	Q38	Q34	Trend
Q37.1-4			New
EF1			New
EF2	EF10	EF10	Trend
EF3		EF11	Trend
EF4	EF19b	EF19	Trend
EF5.1-7	EF20.1-10	EF20.1-10	Modified version of 2000-Q20.1-10 combining some of the categories.
EF5a.1-7			New
EF6	EF21	EF21	Trend
EF7	EF22	EF22	Modified with further sub-answer categories for the financial participation questions
EF8	EF23	EF23	Modified with further sub-answer categories for the financial participation questions
P1-12	P1-12	P1-13	Trend