

Task force on the quality of the Labour Force Survey

Final report

2009 edition

Task force on the quality of the Labour Force Survey

Final report

2009 edition

How to obtain EU publications

Our priced publications are available from EU Bookshop (<http://bookshop.europa.eu>), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

Europe Direct is a service to help you find answers to your questions about the European Union

Freephone number (*):

00 800 6 7 8 9 10 11

(* Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.)

More information on the European Union is available on the Internet (<http://europa.eu>).

Luxembourg: Publications Office of the European Union, 2009

ISBN 978-92-79-14061-7

ISSN 1977-0383

Cat. No. KS-RA-09-020-EN-c

Theme: Population and social conditions

Collection: Methodologies and working papers

© European Union, 2009

Eurostat is the Statistical Office of the European Communities. Its mission is to provide the European Union with high-quality statistical information. For that purpose, it gathers and analyses figures from the national statistical offices across Europe and provides comparable and harmonised data for the European Union to use in the definition, implementation and analysis of Community policies. Its statistical products and services are also of great value to Europe's business community, professional organisations, academics, librarians, NGOs, the media and citizens.

Eurostat's publications programme consists of several collections:

- **News releases** provide recent information on the Euro-Indicators and on social, economic, regional, agricultural or environmental topics.
- **Statistical books** are larger A4 publications with statistical data and analysis.
- **Pocketbooks** are free of charge publications aiming to give users a set of basic figures on a specific topic.
- **Statistics in focus** provides updated summaries of the main results of surveys, studies and statistical analysis.
- **Data in focus** present the most recent statistics with methodological notes.
- **Methodologies and working papers** are technical publications for statistical experts working in a particular field.

Eurostat publications can be ordered via the EU Bookshop at <http://bookshop.europa.eu>.

All publications are also downloadable free of charge in PDF format from the Eurostat website <http://ec.europa.eu/eurostat>. Furthermore, Eurostat's databases are freely available there, as are tables with the most frequently used and demanded short- and long-term indicators.

Eurostat has set up with the members of the 'European statistical system' (ESS) a network of user support centres which exist in nearly all Member States as well as in some EFTA countries. Their mission is to provide help and guidance to Internet users of European statistical data. Contact details for this support network can be found on Eurostat Internet site.

Table of contents

Executive summary	4
Recommendations from the Task Force on the quality of the Labour Force Survey	7
1. Introduction.....	11
1.1. A clear commitment to the quality of the Labour Force Survey.....	12
1.2. The Task Force on the Quality of the Labour Force Survey.....	13
2. Different views on the main quality concerns regarding employment and unemployment.....	17
2.1. Member States main quality concerns with regard to the LFS	17
2.1.1. Accuracy	18
2.1.1.1. <i>Frame, weighting and non-response issues</i>	18
2.1.1.2. <i>Other issues of accuracy of measurement</i>	19
2.1.2. Coherence	22
2.1.3. Comparability	22
2.1.4. Timeliness.....	23
2.1.5. Concerns of burden.....	23
2.2. Users' views	23
2.3. Eurostat's views	24
3. Review of quality issues relating to employment and unemployment data.....	27
3.1. Accuracy.....	27
3.1.1. Sampling errors	28
3.1.1.1. <i>Levels</i>	28
3.1.1.2. <i>Trends</i>	29
3.1.1.3. <i>General remarks</i>	29
3.1.2. Non-sampling errors	30
3.1.2.1. <i>Frame and design issues</i>	30
3.1.2.2. <i>Non-response</i>	34
3.1.2.3. <i>Measurement issues</i>	44
3.1.2.4. <i>Processing issues</i>	49
3.2. Coherence between the LFS and National Accounts estimates of employment.....	50
3.3. Comparability of employment and unemployment statistics	54
3.3.1. Comparability over space.....	54
3.3.2. Comparability over time	56
3.4. Relevance of the ILO concepts of employment and unemployment.....	58
3.5. Timeliness and punctuality.....	60
3.5.1. Timeliness.....	60
3.5.2. Punctuality.....	63
Cited references	65
ANNEX 1 – Reconciliation table between LFS and NA estimates of employment.....	69

Executive summary

1. The present document reports the views of the Task Force on the Quality of the Labour Force Survey (hereinafter LFS) as a result of its six meetings between June 2007 and April 2009. It also takes into account the views expressed by some of the main European institutional users, namely the European Commission's Directorate General "Employment, Social Affairs and Equal Opportunities" (DG EMPL) and the European Central Bank (ECB), and the feedback from the Labour Market Statistics (LAMAS) Working Group in September 2007, April and September 2008.
2. The Task Force was set up by the LAMAS Working Group at its March 2007 meeting. It was coordinated by Eurostat and composed of national delegates with substantial expertise on the LFS from nine Member States: France, Germany, Greece, Italy, the Netherlands, Poland, Portugal, Spain and the United Kingdom.
3. This initiative is in line with the continuous work to improve the quality of the LFS. It was conceived to consolidate the gains achieved in recent years and reinforce the status of the LFS – due to its history, sample size and richness of characteristics – as the main statistical source on the labour market.
4. The goal of the Task Force was to review the quality of the LFS along the dimensions of the quality framework for statistical output of the European Statistical System (ESS), detect weaknesses and recommend improvements. The focus of the review was on the estimates of employment and unemployment, as these are the most relevant and largely used indicators produced by the LFS.
5. Following this review, the Task Force formulated forty-three recommendations on sampling design and sampling errors, weighting schemes, non-response, interviewers and fieldwork organization, survey modes and questionnaire, information for users, coherence, comparability of employment and unemployment statistics, relevance of the ILO concept of employment and unemployment, timeliness and punctuality.
6. As concerns sampling design and sampling errors, the sample should be balanced over geographical areas and reference weeks. This would both improve the national quarterly and yearly estimates and increase the relevance of the LFS by enabling the production of good monthly estimates. Moreover, target population, sampling frame and population estimates should be consistent and up to date in order to avoid overcoverage and undercoverage. The importance of harmonized rotation patterns was also highlighted, to allow comparable longitudinal analysis at European level. Finally, the need was recognized for a clarification of the wording of the precision requirements in Council Regulation 577/98 and for an agreed method to assess compliance with the Regulation.
7. Non-response in the EU, EFTA and candidate countries is rather high (about 20% on average). It is usually selective with respect to employment and unemployment, thus affecting the accuracy of their estimation. Recommendations cover studying, preventing and correcting for non-response. Information on the characteristics of non-respondents should be regularly collected to assess and adjust for non-response bias and to improve fieldwork strategies.

Suitable tools to reassure respondents (such as free-toll numbers or presentation letters) should be introduced, with a special view to increase the participation of non-nationals. The use of the wave approach and of dependent interviewing should be considered to reduce response burden. Finally, weighting schemes should take into account specific characteristics of non-respondents to correct for non-response bias.

8. The role of interviewers is crucial for the accuracy of the survey results. Several recommendations and good practices concerning interviewers' contractual features, training, monitoring and in general on the field-work organization were identified with a view to common guidance, as national arrangements concerning these features tend to vary. In particular, in order to boost motivation and minimize turnover, permanent professional interviewers should be used and their remuneration should be adequate to their crucial role for the quality of the survey. Interviewers' training should cover not only the survey content but also how to conduct the interview and to prevent non-response. Periodic debriefing and focus groups should be organized to review and tackle issues. Interviews should be carried out as close as possible to the reference period, to avoid recall problems and support timely production of results.
9. The LFS should always be carried out by computer-assisted questionnaires, given that the traditional paper-interviewing mode is no longer suitable to cope with the complexity of the survey. However, the impact of self-administered electronic data collection, including web-based modes, on the measurement of ILO labour status should be carefully investigated. The use of mixed modes should be considered, in the light of possible gains relating to response rates, burden and costs along with likely the mode effects. In any case, any changes to modes, questionnaires and other explanatory survey material should be carefully tested and their impact assessed before introduction.
10. Lack of coherence between LFS and national accounts employment estimates is a major concern, as it may harm the credibility of statistics. In this regard, distinguishing between differences in coverage, scope and definitions from inconsistencies that can be ascribed to the accuracy of the different statistics is of the utmost importance. For this purpose the Task Force recommended the use of reconciliation tables between LFS and National Accounts estimates. The value of appropriate communication to users on the nature of incoherence and the need to provide guidance on which source fits which purpose were also recognised.
11. The idea of moving towards higher input harmonization is considered too difficult for the moment because of national specificities and needs. Council Regulation no. 577/1998 together with the 12 principles for the formulation of the questions on labour status laid down in Commission Regulation 1897/2000 remain therefore the basis at European level for comparable statistics on employment and unemployment. However, the principles should be reviewed in order to clarify particular ambiguous points. Such clarifications should not necessarily imply changes in the regulation (necessarily via a new legal act), but should instead be provided as much as possible in working documents such as the explanatory notes.

12. Care should be taken when introducing innovations, as these can negatively impact on comparability of statistics over time. National statistical institutes should always adequately plan and monitor all changes initiated either by Eurostat or by countries in order to assess the statistical effect on time series. Consistent time-series should be produced and disseminated, at least for the headline indicators. For its part, Eurostat should group together innovations it proposes in order to limit the number of potential breaks in time series.
13. The relevance of the ILO labour force concept was confirmed, although the need for supplementary indicators for the ILO unemployment rate, both capturing a wider extent of the labour reserve and allowing longitudinal analysis, was recognized. The variable "Main Status as perceived by respondents", which offers a complementary view to the ILO economic activity status, should be mandatory in the EU-LFS.
14. The timeliness of the EU-LFS can be significantly improved. This would further enhance its relevance for short-term economic analysis. Establishing a release calendar would be similarly helpful. For this purpose it is essential that the twelve-week deadline in the Regulation as the one for final, not first, data transmission is respected.
15. All recommended practices are effective for improving the quality of the LFS and are feasible, as they are already in use in at least one country. Most of the recommendations apply to national statistical institutes, whereas several apply to Eurostat and a few to both.
16. The full list of recommendations, grouped by subject, is provided at the end of this summary. Page numbers in brackets at the end of each recommendation refer to the point in the text where they are discussed.

Recommendations from the Task Force on the quality of the Labour Force Survey

Sampling design and sampling errors

- R1. Address ambiguities in the wording of the precision requirements in Council Regulation 577/98 (art. 3.1 and 3.2) and issues in the assessment of compliance with the Regulation. Investigate ways of improving the comparability of variance estimation. Eurostat should create a group of experts with the mandate to provide recommendations on the above. *(p. 30)*
- R2. Achieve consistency between target population, sampling frame and population estimates *(p. 31)*
- R3. Balance samples over geographical areas and reference weeks, in order to both improve the national quarterly and yearly estimates, and to enable the production of good European monthly estimates. *(p. 31)*
- R4. Regularly review the efficiency of the sample design and adapt it when necessary. *(p. 31)*
- R5. Fully apply the concept of resident population as defined by art. 2(d) of EP and Council Regulation (EC) No 763/2008. *(p. 33)*
- R6. Use sample designs with intra-annual rotation patterns to allow calculation of quarter-to-quarter labour market flows estimates. *(p. 60)*
- R7. Examine options for further harmonization of rotational patterns to enhance comparable longitudinal analysis at European level. *(p. 60)*

Weighting schemes

- R8. Treat collective households separately from private households, for instance when applying weights. This is needed to ensure cross-country comparability and produce meaningful EU estimates, as not all Member States cover collective households. *(p. 33)*
- R9. Apply weighting schemes which ensure the coherence between household and individual estimates. *(p. 34)*
- R10. Optimise weighting schemes to correct for non-response bias, taking into account specific characteristics of non-respondents. *(p. 43)*

Non-response

- R11. Regularly collect information on the characteristics of non-respondents (e.g. by retrieving structural data from the sampling frame or from low-level geographical databases that can be linked to it, or by conducting non-response surveys) to assess and adjust for non-response bias and to improve fieldwork strategies. Specific non-response surveys should be used primarily for improving fieldwork strategies, and care must be taken when using them to adjust the results of the LFS. *(p. 37)*

- R12. Make participation in the LFS compulsory, wherever legally feasible, in order to reduce non-response. *(p. 38)*
- R13. Make use of tools to inform or reassure respondents, such as presentation letters, free-toll numbers and visible identification cards for interviewers. *(p. 40)*
- R14. Arrange for special tools to increase participation of non-nationals, such as translated questionnaires and presentation letters, and free-toll numbers for non-nationals where information is provided in other languages. *(p. 40)*
- R15. Consider implementing the wave approach, as it helps to reduce response burden. *(p. 41)*
- R16. Make use of dependent interviewing to reduce response burden and thus increase participation, but only for variables that tend to be stable from quarter to quarter. Dependent interviewing must not adversely impact on the measurement of the ILO labour status. In particular, the question on having worked for at least one hour during the reference week should be asked by default. *(p. 42)*
- R17. Develop a harmonised approach for the calculation of non-response indicators. Eurostat should initiate work on this issue. *(p. 43)*

Interviewers and fieldwork organization

- R18. Provide specific training to interviewers to avoid refusals by respondents *(p. 39)* and to limit non-contacts. In particular, contact attempts should be carried out at different times of the day and interviewing time should also cover periods after working hours (evenings and weekends), to reduce cases of non-response. Collecting preferred contact time during the first interview may facilitate contacts at further waves. *(p. 38)*
- R19. Make use of permanent professional interviewers. Minimise turnover of interviewers to ensure continuity *(p. 39)*. Interviewer remuneration should be adequate with respect to the interviewers' crucial role for the quality of the survey. *(p. 46)*
- R20. Provide interviewers with appropriate training on how to conduct the interview, covering both communication skills and survey content. Training should be provided on a continuous basis. *(p. 40)*
- R21. Organise periodical focus groups and debriefing for interviewers to review and tackle critical issues. *(p. 40)*
- R22. Instruct the interviewers in general not to adapt questions to the respondent but just to read them out as they are written in the questionnaire, as personal interpretations may lead to biased results. *(p. 46)*
- R23. Continuously monitor and systematically assess the quality of the fieldwork. In the case of telephone interviewing, centralized fieldwork supervision should be used in a telephone facility. In the case of face-to-face interviewing, examples of suitable monitoring methods are regular quality surveys but also other methods like the systematic analysis of interview protocols and behaviour coding of audio recorded interviews. *(p. 47)*

- R24. Keep interviewing periods as close as possible to the reference period, to avoid recall problems and to support timely production of results. Exceptions may be made in particular periods such as holiday seasons, when shorter interviewing periods would result in a low response rate. (*p. 49*)
- R25. Stress, both in the presentation letter and from the interviewers side, that the information collected from individual respondents is treated as confidential and will be used for statistical purposes only. (*p. 49*)

Survey modes and questionnaire

- R26. Consider the use of mixed modes, in the light of possible gains in response rates, burden and costs versus the mode effects (*p. 46*). The impact of self-administered electronic data collection, including web-based modes, on the measurement of ILO labour status should be carefully investigated. (*p. 46*)
- R27. Carefully test any changes to modes, questionnaires and other explanatory survey material and assess their impact before introduction. (*p. 46*)
- R28. Always carry out the LFS by computer-assisted questionnaires, given that the traditional paper interviewing mode is no longer suitable to cope with the complexity of the survey. (*p. 46*)

Information for users

- R29. Accompany published estimates by information about their accuracy, covering in principle both sampling and non-sampling errors. (*p. 30*)
- R30. Make available specific documents to explain the differences between LFS and National Accounts employment estimates and provide guidance to users on which source fits which purpose. The aforementioned documents should be easily accessible to domestic and international users. (*p. 53*)

Coherence

- R31. Produce, at least annually, reconciliation tables between National Accounts and LFS estimates of employed persons to enable targeted improvement measures for the LFS and National Accounts and for communication to users. A suitable template should be developed. (*p. 51*)
- R32. Foster close communication between LFS experts and National Accounts labour market experts, to enhance cooperation and increase the mutual understanding of the methods used in each domain and their respective strengths and weaknesses. (*p. 51*)
- R33. Conduct further research in the areas of marginal employment, employment in black economy activities, employment in private households, illegal immigrants, the influence of non-response and proxy interviews, as these were identified as areas with potential accuracy problems in the LFS that could result in inconsistency¹ with National Accounts estimates. (*p. 52*)

¹ Here inconsistency refers to differences due to accuracy-related errors of statistics. They do not include differences in concepts, scope and definitions.

Comparability of employment and unemployment statistics

- R34. Review the 12 principles for the formulation of the questions on labour status laid down in Commission Regulation 1897/2000. Eurostat should set up a group of experts for this review. Clarifications, if necessary, should be provided as much as possible in working documents such as the explanatory notes; amendments to the wording of the principles (necessarily via a new legal act) should be made only when strictly needed. (p. 56)
- R35. Adequately plan and monitor, at national level, all changes initiated either by Eurostat or by countries in order to assess the statistical effect on time series. (p. 56)
- R36. Eurostat should group together innovations which it proposes in order to limit the number of potential breaks in time series. (p. 56)
- R37. Produce and disseminate consistent time series at least for the headline employment and unemployment indicators. (p. 57)

Relevance of ILO concepts of employment and unemployment

- R38. Define a set of common indicators supplementing the employment and unemployment statistics based on the ILO concept. Eurostat should initiate a Task Force for this purpose². (p. 60)
- R39. Remove the optional status, in the LFS Regulation³, of the variable "Main Status as perceived by respondents" in the EU-LFS. This variable offers a complementary view to the ILO economic activity status and was agreed to be a core variable in all social surveys of the European Statistical System. (p. 60)

Timeliness and Punctuality

- R40. Improve significantly the timeliness of the EU-LFS in order to further enhance its relevance for short-term economic analysis. (p. 63)
- R41. Establish a release calendar for the EU-LFS. (p. 64)
- R42. Comply with the twelve-week deadline in the Regulation as the one for final, not first data transmission. (p. 64)
- R43. Make available to NSIs the Eurostat validation programmes used to process and control national datasets in order to allow a complete check prior to the data delivery and avoid time-consuming retransmissions. (p. 64)

² The Task Force on supplementary indicators was established in September 2008 and started work in February 2009.

³ Council of the European Union, 1998.

1

Introduction

1. Introduction

1.1. A clear commitment to the quality of the Labour Force Survey

Generally speaking, quality is essential for statistics to be trusted, accepted and used.

The quality of labour statistics in general and the Labour Force Survey in particular has been a major concern of the Labour Market Statistics (hereinafter LAMAS) Working Group and its predecessor, the Employment Statistics Working Group, from the very outset. This has been expressed in Regulations, reflected in harmonised definitions, discussed at workshops, seminars and at Working Group meetings.

Concerning the Labour Force Survey, major milestones in the improvement of its quality have been the adoption of Council Regulation (EC) No 577/98 of 9 March 1998 on the organisation of a continuous, quarterly sample survey in the Community; the adoption of Commission Regulation (EC) No 1897/2000 of 7 September 2000 concerning the operational definition of unemployment and the 12 principles for formulating questions on labour status; the adoption of Regulation (EC) No 1991/2002 of the European Parliament and of the Council of 8 October 2002 making the continuous survey mandatory from 2003 onwards⁴; the adoption of Regulation (EC) No 2257/2003 of the European Parliament and of the Council of 25 November 2003 extending the survey characteristics and introducing the distinction between structural and quarterly variables.

Eurostat and the Member States have continuously worked also on a voluntary basis to improve the quality of the survey. Annual quality reports were introduced in 2002 and quarterly accuracy reports were introduced in 2004. Standards and rules for preparing *ad hoc* modules were adopted in 2004. At the initiative of Member States, a programme of annual LFS workshops was started in 2005. A thorough revision of the explanatory notes has been carried out from 2005 to 2007 and a monitoring of the implementation of these explanatory notes is ongoing.

In addition to this, Eurostat has introduced various measures that, while applying to all statistical products, have also concerned labour market statistics. These are the standard definition of quality, transparency in documentation and the adoption of the SDDS standard for presentation of metadata and the European Statistics Code of Practice for the national and Community statistical authorities⁵.

The introduction of a continuous quarterly labour force sample survey in replacement of the annual spring survey has inevitably changed the statistical landscape. In many Member States policy makers, the market and the general public used to rely on various other sources for assessing the situation in the labour market, not least because of their infra-annual data availability. Foremost of these are the estimates of the National Accounts on the employment levels and monthly statistics on registered unemployment. While the introduction of the continuous Labour Force Survey has significantly improved labour market statistics in the EU as a whole, this has also caused uncertainty among the users, not the least when the levels, trends or the patterns of the LFS statistics differ from that of the National Accounts or the register based unemployment statistics.

⁴ Except Italy from 2004 and Germany from 2005.

⁵ Eurostat, 2005.

The very success of the European Union Labour Force Survey can thus paradoxically be said to have caused concerns with regard to its quality, comparability and coherence. Inevitable questions about reliability and validity follow. In order to consolidate the gains from having adopted an EU-LFS and re-establish full trust by the users, a twofold strategy is envisaged. On the one hand, the quality of the survey must be further improved whenever some weaknesses are detected. The question of coherence with other statistics appears to be of concern in many Member States. Other questions have also been raised, such as the question of the effects of different survey designs (incl. data collection methods, rotation patterns, non-response, proxy answers, questionnaires) on the comparability of the survey results. On the other hand, communication aspects play a major role in reinforcing the credibility of the survey and reassure the users, both by making improvements visible and by highlighting the good quality standards already in place when no need for change is found.

1.2. The Task Force on the Quality of the Labour Force Survey

To address these issues, the Task Force on the quality of the Labour Force Survey was created by the LAMAS Working Group at its March 2007 meeting.

The Task Force was coordinated by Eurostat's Unit F2 "Labour Market Statistics". Nine Member States were represented: Portugal, Poland, Italy, Germany, Netherlands, United Kingdom, France, Spain and Greece⁶. Other Units from Eurostat participated whenever a topic of interest for them was discussed. It was the case of Unit B1 "Quality, classifications", Unit C2 "National accounts – production" and Unit E4 "Regional indicators and geographical information". During the first meeting DG EMPL and the European Central Bank were also invited to express their view as users of the LFS.

Whereas the Labour Force Survey covers over 80 characteristics, the focus of the Task Force was on the measurement of employment and unemployment. This restricted focus was due to several reasons. First, employment and unemployment are the most relevant statistics derived from the LFS. They are used for economic as well as social analysis at international as well as at national level, underpin policy making, business decisions, influence stock exchange quotations. Second, the statistics on employment and unemployment are those which have raised higher concerns, due to the fact that other sources providing different results are often available⁷. Such a situation often confounds users who tend to lean on the most consolidated source. Third, the precision of the LFS results in Council regulation 577/98 is expressed with a view to the variable *unemployment*. Fourth, the effort to assess in detail the full extent of the information derived from the LFS would not be feasible in the relatively short working time the Task Force has at its disposal. Hence, employment

⁶ The Task Force was coordinated by Omar Hardarson until October 2007 and by Nicola Massarelli afterwards. Permanent members of the Task Force were Sonia Torres, Agnieszka Zgierska, Antonio Discenza, Thomas Körner, Johan van der Valk, Debra Prestwood, Sébastien Durier, Miguel Angel García Martínez, Stelios Zachariou, Joachim Recktenwald, Remko Hijman, Frank Espelage and Arturo de la Fuente.

⁷ The main alternative source for employment both at national and international level are the estimates from the national accounts, although at national level other sources such as registers (business registers, social security registers, tax registers) or business surveys may be available. As concerns unemployment, there are no alternative sources at international level, whereas at national level the LFS statistics may be confronted with registered unemployment.

and unemployment might serve as barometers for the quality concerns about the LFS in general. The in-depth review and assessment of the whole LFS process will benefit all the statistics derived from it. Finally, the decision to focus on employment and unemployment served as a criterion whenever there is a trade-off affecting the quality of different statistics. For instance, this may occur for the sample design, which is usually set up in order to minimize costs constrained to precision targets for specific indicators. As a result, the accuracy of other statistics may be negatively influenced.

The quality review of the Labour Force Survey followed the dimensions of the quality framework for statistical output of the European Statistical System (ESS), as laid down in Article 12 of the Regulation on European Statistics⁸: Coherence, Accuracy, Comparability, Relevance, Timeliness and punctuality, and Accessibility and clarity. Some dimensions (notably coherence, accuracy and comparability) were deemed to be of particular importance and were assessed more in-depth than the others (relevance, timeliness and punctuality, and accessibility and clarity).

The work of the Task Force was mainly guided by a Working Programme (Doc.: Eurostat/F2/EMPL/14/07) and an Outline of the final report (Doc.:Eurostat/F2/EMPL/13/07), agreed during its first meeting and adapted over time to take into account the actual pace of the discussion or the reorganization of the topics whenever it was deemed convenient by the Task Force and the LAMAS working group.

The main tasks of the Task Force were to survey the quality of the LFS estimates of employment and unemployment and to make recommendations for improvements, either for immediate actions to be taken or longer-term solutions. The involvement of Member States not represented in the Task Force and the consultation of users was also foreseen whenever considered necessary, although having in mind to keep the burden on them to a minimum.

The Outline of the final report provided a guideline both for the compilation of the documents to be presented to the LAMAS (the interim and the final report) working group and for the discussion.

This final report, as a consequence, follows the structure outlined in that document. After the present introduction, chapter 2 introduces the different views from Member States, users and Eurostat on what are the main quality concerns regarding LFS based employment and unemployment. Chapter 3 deals with the review and (possible) quantification of quality issues relating to the production of the statistics of employment and unemployment. Here one section is devoted to each quality dimension and all the elements affecting it are listed as sub-sections. Recommendations emerging from the different sections or sub-sections are in bold in the text.

The final report also covers the relevant results from the Workshop on LFS quality assurance. This was a second exercise launched by Eurostat with the endorsement of the LAMAS working group⁹ complementing the work of the Task Force on the Quality of the Labour Force Survey. The event

⁸ Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European Statistics

⁹ In its March 2007 meeting the LAMAS Working Group recognised the importance of the quality of quality statistics and agreed on the need to review the current scheme for monitoring and evaluating the quality of the LFS and of statistics derived from it.

was organized by Eurostat in cooperation with the National Statistical Service of Greece and was held in Athens on 23-24 October 2008.

The Workshop had several objectives. Firstly, it aimed at answering to the need to review and reorganise the framework for the quality assessment of the EU-LFS. A second goal was to promote the use of a quality assurance approach at national level. Thirdly, it was conceived as a forum for discussion on how to enhance the harmonization of the main quality indicators of the EU-LFS. Last but not least, it aimed at sharing and discussing national practices in the field of quality assurance and at stimulating the creation of a network of LFS experts across Europe.

Eight different topics were discussed: Quality assurance frameworks; LFS quality report forms; calculation of coefficients of variation; coherence among different sources: monitoring and communication issues; calculation of non-response rates; interviewers and field work monitoring; documentation of processes and communication to users; check plans. Documents and presentations discussed at the Workshop as well as the minutes are available on Circa at the address:

http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/3_workshops/documents&vrm=detailed&sb=Title.

2

Different views on the main quality concerns regarding employment and unemployment

2. Different views on the main quality concerns regarding employment and unemployment

Quality in statistics is a relative concept which depends on the perspective of users and on their cognitive needs. The assessment of quality dimension such as *relevance*, *comparability*, *clarity* or *timeliness* is strictly related to those needs and may vary from user to user. Some may be more interested in certain statistics while other users may judge different indicators more relevant for their purposes. Users interested only in single countries or exclusively conducting cross-section studies may not care about geographical or time comparability as much as those who look at the international situation or make use of time-series analysis. The interest in metadata also largely varies from user to user. Timeliness may be a major concern for users performing short-term economic analysis but less relevant for those interested in structural, long-term assessment. Even the assessment of *accuracy* may be subjective. For instance, in the case of multiple target variables there may be a trade-off between the precision of some of them. Even more striking is the case of the level of detail of the information needed. The same statistic may thus be deemed fit for purpose to those interested in big aggregates but less appropriate to users concerned by territorial or sectional detail.

In addition to this, there is often a trade-off between different quality dimensions. For instance, some statistics are relevant only if they are timely, but more timeliness often implies lower accuracy¹⁰.

Producers have a wider view on the quality of the statistics. While users are ultimately worried about the quality of the *outputs*, producers are in the position to also look at the productive *process*. They are the only ones who can go to the root of problems once quality concerns have emerged and find suitable solutions for improvements. But producers are also aware of constraints they face, so that not every issue may have a (feasible) solution.

In the following, quality concerns are presented from the different perspectives of Member States, users and Eurostat.

2.1. Member States main quality concerns with regard to the LFS

In order to obtain the views of Member States about their main quality concerns on the LFS, the Task Force did not carry out a survey, but pooled the experiences from the nine Member States represented, and then asked the other countries to comment on the outcome, offer suggestions, add solutions or subscribe to the summary of the Task Force. This chapter is the outcome of that process.

The main concerns can be classified under five main headings: questions of coherence, accuracy, comparability, timeliness and increasing burden. Of course not all concerns have the same relevance in all countries. Nonetheless, the following statements can be considered representative of the general situation in Member States.

¹⁰ A typical example concerns flash estimates. The closer to the reference period they are produced and disseminated, the less accurate they are, the bigger will be the revisions of the provisional estimates.

2.1.1. Accuracy

Generally speaking, accuracy-related issues are of great concern, first of all because of the impact of inaccuracies on the reliability of the survey results. Inaccuracies also have a negative impact on other quality dimensions such as comparability or coherence.

2.1.1.1. Frame, weighting and non-response issues

- *Sampling frames*

The sampling frames and population estimates are mostly provided from other departments within the National Statistical Institutes (NSI) and are thus not within the purview of labour market statisticians. Frame and errors of population figures generally affect the estimates of totals, but to a lesser extent the rates and averages. Badly skewed population estimates of certain sub-groups may, however, affect the rates and averages.

In many cases the sampling frames are based on the last Census, with the population estimates also anchored to the last Population Census, which in turn make estimates towards the end of the intercensal period less reliable and more prone to bigger revisions once the new Census estimates are available. In Germany in particular, there is reason to believe that the population estimates may be considerably off, given that the last Census was taken in 1987.

One observed consequence of out-of-date sampling frames is the increased overcoverage, due to vacant or no longer existing housing, as it is the case in Latvia, with 3.2% overcoverage in 2006.

There are issues of comparability with regard to the frames, since these are generally not defined in a similar way across countries. In Poland, e.g., the Census and the LFS data refer to the *de facto* population, whereas in all other countries the usual or main residence is the defining criteria. Even criteria to define the resident population differ among countries.

Covering the *de facto* population, including foreign students and short-term migrants and temporary workers, is nevertheless of increasing interest and relevance for the analysis of the labour market in many countries. In view of the increasing (temporary) emigration from the new Member States, this problem can also lead to an underestimation of employment in the host country. In the UK, the ONS is currently studying ways to cover short-term migration as well.

The sampling frames in many countries are of less quality when it comes to covering certain populations, especially migrants, or they exclude certain groups of people, such as those in institutional households.

In a continuous survey, especially if the focus is on short-term developments, the timing of the actual sampling may be of importance. In some countries the sample is only drawn once a year for all the four quarters of the year, which causes discrepancies between the sample frame population and the actual population later in the year.

- *Weighting procedures*

Council Regulation (EC) No 577/98 does not provide strict guidelines with regard to the calculation of the weighting factors. Most Member States, however, apply the minimum (sex, 5 year age groups, NUTS 2 level regions) in addition to the sampling probabilities. In some weighting

schemes, however, the more detailed age breakdown is ignored, which may affect the estimates on employment and unemployment.

The explanatory notes for the EU-LFS stipulate that in order to achieve consistency between the household data and the individual data, the weighting schemes should assign a single weighting factor to all members of the same household. Not all of the countries providing household data do this.

It could, nevertheless, be asked if the consistency between household data and individual data is desirable. Not all individual behaviour is determined by the household environment, and it could be considered that there is always some (minor) loss of efficiency of the weights when enforcing this constraint.

Since re-weighting back data is very resource intensive, when new population estimates are available after a Census, the new weighting factors may cause apparent breaks in time series until the microdata can be reprocessed.

In short-term statistics the question of coherence of the monthly weights and the quarterly weights may also be of concern. Unless the survey is designed as a monthly survey, fully consistent weights may not be available for each month until after the quarter has been completed.

– *Non-response issues*

Non-response is a quality concern to the extent that it biases the estimation of the characteristics to be measured. Non-response is likely to be selective for categories of persons that in turn are related to the variables. Measuring the impact needs to be at the variable level. An unequal distribution of non-response by age group and work status may therefore affect the estimates.

In many countries, but not all, the response rates are falling, which causes concern, not the least because it is not certain that the assumptions underlying the non-response adjustments continue to hold true. Reasons for non-response differ - one of them is that people refuse to participate in surveys in large cities as it is the case in Ireland, Lithuania and Poland. Rising concerns about data protection issues as in Denmark may add to the problem of non-response. Closely related to this is the panel attrition, which leads to unjustified differences of measurement between survey waves and is often difficult to distinguish from the effects due to different modes of data collection (face to face vs. telephone interviewing). The panel attrition may also cause bias in the trend estimates.

2.1.1.2. Other issues of accuracy of measurement

The issues of accuracy cover a wide range of topics. Some of them reflect random effects, while the most serious can lead to biased estimates. Apart from sampling errors, there are four main sources for measurement errors: the question is not measuring what it was intended to measure, the interviewers may not perform as planned, the respondents can deliberately or accidentally mislead, and the mode for interviewing may give rise to erroneous results. In addition, the professional staff, coders and statisticians may contribute to the errors of measurement, by erroneous design, badly organised field work, errors in classification and bias due to imputation and/or editing.

– *Sampling errors*

It is of concern that in some countries the regional sample size is not sufficiently big for producing reliable estimates. This was stressed in the report of the Commission to the Council and European Parliament¹¹.

In many cases the samples are also not sufficiently large to produce stable results for small domains. That is the reason why countries like Denmark increased the sample size in recent years. Similar concerns arise when using the LFS to produce monthly results with regard to employment and unemployment. It is not only the insufficient sample size that contributes to these difficulties but also design effects, including non-optimal stratification schemes.

– *Respondents*

The respondents may not always recall correctly their situation, give wrong answers (intentionally or unintentionally) or refuse answering some questions. In many household surveys, a sizeable part of the responses are by proxy. In those cases the actual respondent may not know exactly the situation of his or her household members or replace the actual situation with the normal situation. Proxy rates often exceed 40%, topping to 61.2% in Slovakia in 2006. What makes this problem more serious is the increasing share of proxies in the last years, as reported by Estonia.

For some variables, the respondent is not the best source for the characteristic being measured. Industry classification in the LFS is based on the self-declaration of the respondent, which is often not sufficiently detailed or the unit that the respondent refers to may not be the same as the local unit required by the LFS.

It is recognised that some of the technical definitions of employment or unemployment, such as the one-hour criterion or what is defined as a job search may not correspond to the common understanding of either the interviewer or the respondent. Small, peripheral employment or the search for such employment may thus not always be detected in the interview. The Finnish experience when moving from mail inquiry and paper questionnaires to computer-assisted questionnaires confirmed that this may happen. In computer-assisted interviews, tailored questionnaire design and routings of the interview are possible and the interviews are much better controlled. That way, errors can be reduced significantly.

– *The questionnaire*

The questionnaire may also be at the source of certain undesirable outcomes. If it is unbalanced and focussing on one particular area, such as employment, the next time a respondent is asked, either in the next wave or for the next person in the household, he or she will try to avoid mentioning small employment in order to avoid the lengthy questioning. Length, design of the questionnaire, formulation of questions may have an impact which is yet difficult to assess.

The wording of questions is also of importance as it may influence the answer. In some statistics, such as those on hours of work, a single question on the actual number of hours in the reference week, as envisaged by the Regulation may not suffice and lead to serious measurement errors. It may, in particular, not be appropriate to ask the self-employed and the employees the same

¹¹ See Eurostat, 2006 (a).

question or series of questions with regard to hours worked. Altering certain questions depending on which sub-group the respondent belong to, is an option especially in computer-assisted interviewing.

– *The mode*

The mode, especially the difference between face-to-face and telephone interviewing, has also discernable effects on the measurement. The mode effects can be exacerbated by the fact that if telephone interviews are only used in subsequent interviews the length of the questionnaire is also different. In many countries, the complexity of the LFS questionnaire also makes it difficult to render it properly on paper.

– *The interviewers*

The role of the interviewer is vital in the data collection, but it is recognised that in face-to-face interviews the interviewers cannot be monitored or supervised to the same degrees as in call centres. Some countries tackle the problem by re-contacting a sub-sample of households by phone or post, yet in many countries too little is known of how interviewers perform.

Training of interviewers is therefore of concern. At present there are great differences between Member States with regard to the amount and kind of training the interviewers receive.

– *Organisation of the data collection*

The recall problems are also exacerbated by allowing interviews to take place too long after the reference week. Ideally the interview should take place in the week after the reference week. In many countries a significant number of the interviews takes place later, with Council Regulation No 577/98 allowing an interview to be carried out up to five weeks after the reference week. On the other side, longer field periods allow increasing the response rate, which makes their net effect on data accuracy uncertain.

The countries differ with regard to the allocation of reference weeks. Most have a fixed week and a given window of opportunity for the interview to take place subsequently. Others let the logistics of the field work and interviewer work load or travelling routes determine the reference week. The latter may lead to uneven distribution of reference weeks over the year and in particular may cause undesirable cut-offs at the end of the quarter/year.

– *Classification and imputation*

Even when the textual information about the industry, occupation or field of study is sufficient, there could also be differences of interpretation between coders, which lead to different classification outcomes. This could be overcome, however, if the individual information can be linked to business registers or a corresponding administrative source whenever the register information is considered of sufficient quality.

In some countries, missing answers to some numerical variables, such as hours of work or income are imputed. While imputation is intended to correct for the bias that the item non-response causes, the imputation methods may themselves be a source of estimation errors.

2.1.2. Coherence

The Labour Force Survey (LFS) is in most countries situated between social statistics, economic statistics and population statistics. Issues of consistency between these statistics are therefore of utmost importance.

Lack of coherence between the LFS estimates of employment and jobs and job estimates of business surveys or employment estimates of National Accounts have often been of concern for users. The problem of incoherence regards both levels and trends.

In Germany, e.g., the level of employment according to the LFS is about 1.3 million (3.3%) lower than the estimate from the employment accounts. A lower LFS figure can also be found for example in the Czech Republic where the level of employment published by National Accounts exceeds the LFS figures by approx. 5%. In Estonia, in contrast, the employment from National Accounts (domestic concept) was 1.5% smaller than LFS employment in 2006. Sometimes, incoherence only occurs in specific sectors. This is, for example, the case for employment in agriculture (Bulgaria, Poland) or the public sector (Ireland). Different coverage of the shadow economy also matters (Poland).

In the Netherlands, France and Germany the short-term trends are different, which worries the users¹². It appears that if the LFS is not used directly or only peripherally in the production of the employment estimates for National Accounts purposes, the differences tend to be bigger.

Some of the incoherence between National Accounts and LFS is due to different concepts. In order to inform users and facilitate understanding, Switzerland produces schemes presenting the connection between estimates from the LFS and estimates from other surveys.

It is rare that coherence of the unemployment statistics emerges as an issue, as most users are aware of the large conceptual and methodological differences between ILO based unemployment and the registered unemployment. Nevertheless, when register data are used for estimating monthly trends of LFS unemployment within a reference quarter or reference year, any disruption of the assumption of stable relationship between the two will cause perceived problems of coherence.

There are also other fields of statistics where the LFS is visibly different from administrative or other sources, such as population statistics, statistics on self-employed from the tax registers and education statistics. This is for example the case in the Irish LFS which underestimates non-nationals in comparison with the Census (the issue is currently subject to revision). Differences in employment figures for self-employed compared with other sources are reported for Cyprus and Turkey, differences in educational data for Bulgaria and Denmark.

2.1.3. Comparability

Although the definition of unemployment is governed in Europe by Commission Regulation (EC) No 1897/2000, there still seems to be some scope for national interpretations of certain aspects. There are still differences between countries with regard to what should be regarded as specific

¹² However, discrepancies between trends in the Netherlands were visible during a short period only. Retrospectively, this was caused by a change of fieldwork in 2004. Since 2005, no significant differences in trends are observed.

steps to find unemployment, what kind of contact with a public employment agency constitutes an active method, to which reference period the availability for work refers, etc.

Comparability in space is also reduced by the use of different age categories when referring to the working age population.

Comparability over time is hampered by innovations introduced by Eurostat, constant reorganisation of field work, improvements to the questionnaire, introduction of new question sequences or response categories.

2.1.4. Timeliness

In some countries field work organisation does not allow for very fast estimates. In Italy, e.g., the data for a given month are only available five weeks after the end of the month.

2.1.5. Concerns of burden

Because of its big sample and well-established organization, the LFS is often seen as the most suitable vehicle to address a wide range of topics. The underlying hypothesis is that the marginal cost of a few more questions within the LFS is much less than the cost of setting up dedicated brand new surveys. There is a concern in many countries that this practice may be stretched too far, both in asking too much detail on some labour market policy related issues and in identifying and monitoring small populations. This concern is further exacerbated by the fact that a number of member states add further 'national' variables or survey modules to the LFS questionnaire, usually in response to user demands.

Adding too many peripheral questions may divert from the focus of the survey, in the mind of the statisticians as well as the respondents and the interviewers. The problem is reinforced by the international nature of the survey where specific issues may have considerable importance for some but are peripheral for others. For example, in Romania as in other Eastern European Countries, part-timers (especially employees) represent a rather insignificant sub-population.

Increased complexity and elaborate filters also increase the danger of mistakes and errors in the application of the questionnaire. In countries which mainly use paper questionnaires the current complexity of the survey may already have exceeded the limits of feasibility.

The mere length and number of questions may also contribute to fatigue and thus errors during the interview and eventually lead to non-response and panel attrition.

Closely linked to the scale of the survey is the fact that, even as it is, many countries as well as Eurostat do not use and analyze the survey data to their full potential. The longitudinal element in the survey is for example mostly ignored.

2.2. Users' views¹³

For DG EMPL the LFS is the key statistical source to measure and monitor the Lisbon Strategy and the Employment Guidelines. In general DG EMPL is satisfied with the data on employment

¹³ Views expressed during the first meeting of the Task Force in June 2007.

and unemployment from the LFS. There are, however, some issues that should be emphasised, in particular the need for flow estimates, coherence with other statistics, especially employment growth in the National Accounts and population and migration statistics. Furthermore, it is important that the current regulations on the implementation of the LFS are fully observed in all Member States, as otherwise the utility of the data is reduced.

DG EMPL suggests some actions that might enhance the quality of the LFS:

- Detailed monitoring of the implementation of the new explanatory notes from 2008 onwards;
- Consider making the explanatory notes a part of the LFS regulation in order to achieve a stricter adherence to common definitions;
- Consider more input harmonisation;
- Regularly pre-test variables in modules to identify potential problems of comparability and common interpretation in advance.

The ECB in the past did not use LFS data as intensively as desired, mainly due to perceived quality problems, but is increasingly looking at the LFS data in order to monitor labour market developments. The focus of the ECB is on the Euro area aggregates, but also country data are monitored. The LFS is currently the main source for activity, employment and unemployment rates. Employment and unemployment data are analysed in combination with data from other sources. The main uses are with regard to short-term developments, longer time series and structurally detailed data. ECB is mainly concerned with the coherence issue, the timeliness and the properties of the time series. As there are problems of coherence, the different messages with regard to labour market developments are difficult to interpret. The ECB considers National Accounts employment estimates more reliable, although they lack the detailed breakdowns offered by the LFS. National accounts results are also timelier than the LFS data, which are published around 90 days after the end of the quarter. In addition, breaks in the LFS series prevent from having long time series, given that back recalculation is rarely performed. The ECB also points out that the ILO concept of employment of at least 1 hour gives reason to produce (supplementing) estimates of underemployment.

The ECB calls for a systematic use of seasonal adjustment and timely publication of monthly figures on the main aggregates. The use of EU-sampling could be considered for achieving this. Consistency with National Accounts data should be improved, with perhaps LFS more extensively used for filling in gaps of the national account estimates. The ECB also calls for improved metadata information, especially with regard to breaks, as well as improved accessibility of the data on the web site.

2.3. Eurostat's views

As regards the *relevance* of employment and unemployment statistics, efforts should be made to meet the increasing demand for flow statistics, at present largely non-existent. On the side of monthly unemployment statistics, in several countries the demand is not met directly by the LFS but by various national sources, which reduces international comparability.

Accuracy issues concern the non compliance of structural data on regional level with accuracy requirements of approximately 8% for the NUTS 2 regions¹⁴ and the high unit non-response often exceeding 20% (in 11 out of 32 countries). The assessment of non-sampling errors is not always easy as they are insufficiently reported by Member States.

Coherence issues concern discrepancies between employment estimates from the LFS and National Accounts, both in levels and trends. In addition, coherence with national publications on the LFS is not always ensured.

Comparability. Output harmonization does not ensure full comparability among national surveys. The definition of unemployment and the 12 principles for the formulation of questions on labour status are set to achieve a good comparability of employment and unemployment statistics via harmonisation of input elements. However, from a first screening of 2008 national questionnaires it emerges that only three countries fully respect all the 12 principles for formulating the questions, which raises concerns with regard to cross-country comparability. A further issue regards time series breaks, which are spread over the period 1998 to 2005, due to staggered introduction of the continuous survey. What makes things even more difficult to deal with is that breaks are not always reported or quantified.

Timeliness and punctuality. Deadlines for delivery of data (Q+12 weeks) are in general respected, in spite of stragglers in every quarter. However, delays are in most cases rather short; no release schedules are set for Eurostat's LFS publications, although they are being considered; and even if timeliness of quarterly publication has improved after moving from Statistics in Focus (SIF) to Data in Focus (DIF) (in the four quarters of 2007 the delay was around 20 weeks after the end of the reference period), publications should be still closer to the end of the quarter.

Accessibility and clarity. Clarity can be improved, especially with regard to the relationship between different datasets containing labour market data.

*Cost and burden*¹⁵. Average interview time increased from 2004 to 2005 by about 50 seconds. The total cost increased from €90 million to €95 million in 2005.

¹⁴ See Eurostat, 2006 (a).

¹⁵ Very rough and not particularly reliable estimates.

3

**Review of quality issues relating to
employment and unemployment data**

3. Review of quality issues relating to employment and unemployment data

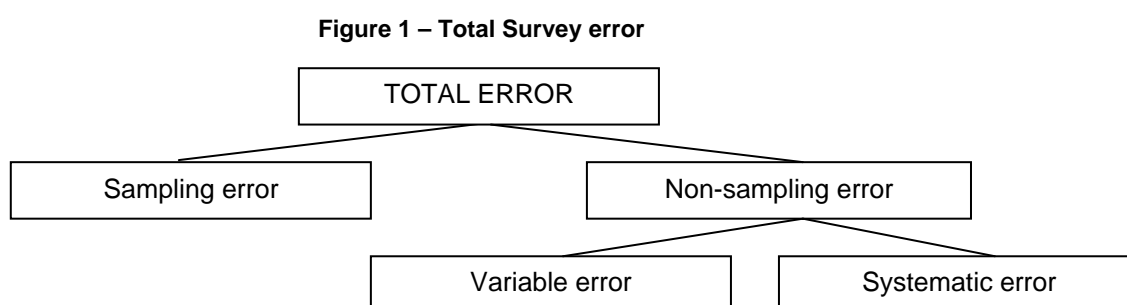
As stated in the introduction, the present review follows the dimensions in the quality framework for statistical output of the European Statistical System (ESS). As a general approach, the Task Force attempted as far as possible to find solid evidence for purported problems in order to remove these from the domain of speculation. However, problems that are difficult to quantify were not disregarded.

The Task Force recognises that not all problems have solutions, but when solutions exist, one can distinguish between solutions that can be applied and solutions that are outside the reach of labour market statisticians. For some problems there are no feasible solutions or even no solutions at all

In the following review topics are presented according to the importance the Task Force attributes to the dimensions in the ESS quality framework for the EU-LFS employment and unemployment statistics, from the most to the less important. The first item is therefore accuracy, followed in turn by coherence, comparability, relevance and timeliness and punctuality.

3.1. Accuracy

Accuracy refers to “the closeness of estimates to the unknown true values” (Regulation 223/2009 on European Statistics, Article 12). When assessing the accuracy of sample surveys such as the Labour Force Survey, sampling errors have to be taken into account in addition to non-sampling errors. Sampling and non-sampling errors in the context of the EU-LFS are reviewed in this chapter. However, larger space is devoted to the latter. In fact, the former exclusively derives from drawing a probability sample rather than surveying the whole target population and can always be controlled (it is always possible to estimate precision indicators such as the variance or coefficient of variation of the estimates knowing the underlying sample design) and improved. The sources for non-sampling errors, instead, are numerous and their impact on the estimates is not homogeneous.



Source: Banda, 2003.

As sketched in figure 1, they may systematically affect the estimates thus introducing a recurrent bias, or may result in variable errors, whose effect in term of both size and sign changes at each repetition of the survey. Non-sampling errors are often unpredictable and are related to the frame and the survey design, to the measurements process, to data processing. While sampling errors can be reduced by an enlargement of the sample or a more efficient sample design, non-sampling errors may even be boosted by an increase in sample size. Overall, if not properly controlled the effect of non-sampling errors on the estimates may be large to the point to undermine their use.

3.1.1. Sampling errors

Council Regulation 577/98 provides certain benchmarks for the accuracy of annual estimates of unemployment at NUTS 2 level (art. 3.1), as well as for the accuracy of the estimate of change between two successive quarters (art. 3.2). According to article 3.1, precision at NUTS 2 level for the annual estimates must not exceed 8% for regions with a population of more than 300,000. Article 3.2 states that, at national level, for the estimate of changes between two successive quarters the coefficient of variation (CV) must not exceed 2% for countries with population of more than twenty million and 3% for countries with less than twenty million population but more than one million. In order to provide a standard, both estimates refer to a level of 5% of the working age population.

Such benchmarks set minimum standard requirements, letting Member States freely choose the sample design they deem most suitable according to their national situation.

3.1.1.1. Levels

As a general assessment, according to the 2006 Commission Report to the Council¹⁶ requirements set in art. 3.1 are broadly respected, although the benchmark of 8% was exceeded for all of the years 2003-2005 in 31 (out of 281) regions in France, Poland and the United Kingdom, Bulgaria and Romania. Poland is currently investigating ways to comply with the regulation in this regard, while France is carrying out a revision of the sampling design which also foresees an increase of the sample.

The Task Force underlines the importance of the precision requirements for the annual estimates as stipulated by Council Regulation 577/98, as far as the relevance of statistics at NUTS 2 level holds. The requirement for certain standards with regard to NUTS 2 regions may, however, prevent the countries to apply the most efficient sampling designs. For this reason, the relevance of such a requirement should be monitored from time to time because it may change since this was originally conceived.

Urged by the Task Force, Eurostat carried out such a review. Sharing DG-Regio's view, Eurostat's Regional indicators and geographical information unit highlighted that regional unemployment rates are a key element in monitoring cohesion policy. In addition, regional unemployment rates were an important element for financial allocations of cohesion policy funds and it can be realistically expected that in the future they will again be used for this purpose. As a conclusion, the precision requirements for NUTS 2 estimates of unemployment are still relevant and should not be dropped from the Regulation.

3.1.1.2. Trends

According to the aforementioned Commission Report to the Council (see footnote 17), eight Member States were not in compliance with regard to the precision for the estimates of changes between two successive quarters. Poland is studying changes to the survey design, while Denmark, Estonia, Latvia and Lithuania have already made design changes, including significant increases in sample size in Denmark, Latvia and Lithuania.

The Task Force points out that the stipulation of article 3.2 on the precision requirement for estimates of changes between two consecutive quarters in the Council Regulation 577/98 leaves room for interpretation. Conscious of that, Eurostat had produced a paper for clarification for the task force on the explanatory notes in 2006¹⁷.

As suggested by the Task Force, issues concerning the wording of LFS precision requirements and Eurostat's assessment of Member States' compliance were discussed at the Athens Workshop on LFS quality assurance.

As regards the first point, two alternative approaches were proposed and discussed in Athens. Following a first one, interpretation of the current wording is possible, so that no change to the Regulation is needed. According to an alternative view, it would be better to change the current precision requirements in the legal act to make them clearer.

Regarding the second point, from the discussion in Athens it emerged that the method proposed by Eurostat is not effective as it requires information at present not available (i.e. the design effect in the theoretical situation of the precision requirements and the correlation between employment statuses at different intervals). Approximations for such information may lead to misleading conclusions about Member States' compliance. As a consequence, an alternative method to assess Member States' compliance with the precision requirements, agreed by Eurostat and Member States, has to be found.

3.1.1.3. General remarks

The Task Force members agree that the estimation of sampling errors should be carried out in a comparable way across Member States. The current situation leads to non-comparable variance estimates. The Task Force recalls that variance estimations in the LFS had been addressed in 2002 by an outside expert (Quantos, 2002), who also suggested the most appropriate variance estimation techniques for each survey design. Again following a Task Force suggestion, the issue of the harmonization of variance estimation was also addressed at the Workshop on LFS quality assurance. Different national practices were presented and discussed, although in the end no indication on which method may be adopted was given. The Task Force, however, highlights that good practice in variance estimation is to err on the side of caution, so as not to give the users false information about the significance of changes between quarters or differences between two groups.

Endorsing the suggestion from the Workshop, **the Task Force recommends to address ambiguities in the wording of the precision requirements in Council Regulation 577/98 (art. 3.1 and 3.2) and issues in the assessment of compliance with the Regulation. Ways of**

¹⁶ See Eurostat, 2006 (a).

¹⁷ See Eurostat, 2006 (b).

improving the comparability of variance estimation should also be investigated. Eurostat should create a group of experts with the mandate to provide recommendations on the above. (R1)

The Task Force also highlights the importance of providing users with information about the accuracy of the estimates, and recommends that their publication is accompanied by information covering in principle both sampling and non-sampling errors (R29). With regard to the publication of sampling errors, practices of Member States greatly vary. Italy, for instance, provides the user with a function or formula, so that the approximate sampling error can be calculated for whichever sub-group and selected number of characteristics. Although an exchange of experience in this field among national statistical institutes would be useful, a standardization of communication of sampling and non-sampling errors is not envisaged.

3.1.2. Non-sampling errors

The clustering of non-sampling errors in four broad groups as done in the following – Frame and design issues, Non-response issues, Measurement issues and Processing issues – is only meant as guidance to the discussion and does not reflect a specific theoretical framework. On the contrary, it is worthwhile underlining that the different issues are often strictly related, so that the choice of the umbrella under which to place them is rather arbitrary. For some elements of the production process it is also possible that conflicting conclusions on how to deal with them may emerge when looking at them from different perspectives. As an example, although different modes of data collection may produce different estimates (the so called *mode effect*), using mixed modes is a way to reduce costs and to reach a larger number of respondents. In cases like these, recommendations should follow a careful evaluation of pros and cons of each situation.

Non-response issues deal with the analysis and quantification of non-response, how to limit it and how to correct for non-response bias once non-response has occurred. Measurement issues cover instruments (questionnaire, explanatory notes, dependent interviewing, mode of data collection), interviewers (training, control, management) and response errors (respondents, recall problems, proxies). Processing issues include editing errors, coding errors, data entry errors, programming errors. Their relevance with regard to the statistics on employment and unemployment from the LFS is however limited.

3.1.2.1. Frame and design issues¹⁸

– *Sampling frame*

Ideally, the sampling frame and the target population perfectly match, so that the sample drawn from the frame is properly representative of the target population. In practice however, all lists of sampling units suffer from the four basic frame imperfections discussed by Kish (1965): a) missing elements or under-coverage, b) clusters of elements, c) blanks, foreign elements or over-coverage, and d) duplicate listings. In the EU-LFS there are three main types of final sampling units, dwellings or addresses (18 countries), households (10 countries) and individuals (4 countries). The frames from which these units are drawn are also different and have different problems.

¹⁸ This chapter draws from Discenza, 2008.

Information in the quality reports about frame imperfections is rather limited. The problem of under-coverage is the most difficult to detect and quantify. The Task Force suspects that such errors could be more of a concern when the sampling frame is based on lists or registers of individuals or households, while lists of addresses or dwellings may be easier to maintain. Maintenance of lists or sampling frames is more difficult in areas of high migration, as it may be more difficult to maintain the correspondence of the frame population to the target population. Migration is, however, more of a problem for sampling frames based on population or household registers, whereas area and dwelling frames are more sensitive to not picking up new buildings. This correspondence can also deteriorate if there is a long period from the drawing of the sample from a given frame until the actual interviewing. In a two stage sampling the updating of the primary sampling unit for the purpose of selecting the ultimate sampling units is usually done in the field by interviewers but only in the first wave. In France this is done in the first and sixth interview, i.e., whenever the interviewers conduct face-to-face interviews. It is to be expected that the deterioration of the sampling frame in between the survey waves is small and the updating of the frame small compared to the updating that occurs in the first wave.

The Task Force considers however that, in case the frame is not frequently updated, the size and composition of the sample should be constantly monitored and adapted to reflect the actual situation whenever it differs from that resulting by the sampling frame. **The Task Force, therefore, recommends to achieve consistency between target population, sampling frame and population estimates (R2).**

– *Sample allocation over time and space*

An even distribution of the sample units across all the weeks of the year, as prescribed by Article 1 of Council Regulation 577/98, is a precondition for avoiding bias in the quarterly estimates due to seasonal effects, which would arise if peaks or lows in the sample actually achieved coincided with specific seasons. In addition, as a by-product from which countries may take advantage, a better representativeness of the sample on a monthly basis is a prerequisite for the production of monthly LFS estimates. At present, some Member States still have difficulties to cope with a strict uniform distribution of the sample across the year. In order to improve representativeness, the monthly (or even weekly) sample should also be balanced over space, as seasonal effects, which can largely vary between geographical areas, can heavily influence the results. **The Task Force therefore recommends to balance samples over geographical areas and reference weeks, in order to both improve the national quarterly and yearly estimates, and to enable the production of good European monthly estimates (R3).**

– *Stratification*

Stratification criteria as well as strata for the Primary Sampling Units (PSU) are usually fixed once, when the survey is setup, and are usually chosen in order to achieve a target precision level for the estimates at minimum cost. But as the population, the economy, the labour market and the territory evolve, a periodical review of the stratification would be necessary. The Task Force therefore suggests that stratification is regularly monitored and adapted when needed. In general, **the Task Force recommends to regularly review the efficiency of the sample design and adapt it when necessary (R4).**

- *Rotating panels*

There are various rotation schemes in Europe, the most popular being the 2-(2)-2, in-for-5 and in-for-6. Denmark and Latvia have already moved from an atypical rotation scheme to the 2-(2)-2 format, with Malta adopting this scheme from 2008.

The aim of rotating panels is double: it allows firstly increasing the precision of estimates of change between two different points in time and secondly producing flow estimates, thus allowing the calculation of important indicators for the analysis of labour market dynamics. A further advantage is the possibility to make use of dependent interviewing to reduce non-response burden. However rotating panels present the typical drawbacks of panels, although these problems are less critical in the light of the short panel duration. Possible drawbacks include panel attrition, panel conditioning and misreporting. Furthermore, because population evolves in time, the longer panels remain in the sample the more they diverge from the actual population's structure. Overlap between quarters may also cause some inefficiency in annual estimates. Besides, rotation can be the underlying cause of other problems, such as non-response and measurement inconsistencies between subsequent survey waves. Comparability of longitudinal data could also be of concern with different rotation schemes. Overall, however, the advantages of rotation patterns outweigh their disadvantages.

- *Target population and population estimates*

The ILO definition of economically active population¹⁹ is not straightforward when it comes to measurement. As a consequence, it leaves room for interpretation which in turn may lead to different coverage. This regards in particular the treatment of collective households and the choice between the resident and the *de facto* population.

As regards **collective households**, the Task Force recognizes that, from a conceptual point of view, these should be covered by the LFS, in order to have a more complete picture of the labour market. In addition, this would increase the coherence between the LFS and National Accounts. A study conducted in the UK by the ONS²⁰ shows the feasibility of a social survey focussing on the labour market for people living in communal establishments (prisons excluded). However, it also highlights that the coverage of people living in collective households in UK would have almost no impact on the LFS results, because they account for 1% of the total population. Moreover, almost half of them are 75 or older and 85% are inactive. Recently, ONS conducted a pilot study of communal establishments; the results of the pilot will determine whether a full survey will be conducted.

At present, only a minority of Member States include collective households within the sample. The reasons for those who do not cover them are of technical nature: either there is no frame available to select them or no reliable estimates to weigh them, or it is difficult or impossible to reach and interview the institutional population. A few countries gross up sampling data to the total population, i.e. also including institutional population, although the sampling frame, does not relate to the total population. This causes a mismatch between the frame and the target population.

¹⁹ The economically active population comprises all persons of either sex who furnish the supply of labour for the production of economic goods and services as defined by the United Nations systems of national accounts and balances during a specified time-reference period (ILO, 1982, aforementioned).

The Task Force suggests that technical aspects of surveying collective households be further investigated. However, to countries which already cover them **the Task Force recommends to treat collective households separately from private households, for instance when applying weights. This is needed to ensure cross-country comparability and produce meaningful EU estimates (R8).**

Whether the **resident** or the **de facto population** should be covered is controversial. On the one hand the Task Force has recognized a growing interest to cover foreign students and short-term migrants, which are included in the *de facto* population. Although the *de facto* approach may give the LFS added value by differentiating it from other sources such as employment registers, it would, on the other hand, raise coherence problems with respect to the Census which sticks to the resident population. Furthermore, in many countries frames as well as population estimates for weighting do not allow to properly cover all persons living and working in the country, resident or not. Given the present situation, the Task Force recognizes that EU aggregates can be obtained only with reference to the resident population.

Even when considering the resident population, further harmonization is needed. The necessary time elapsed in a given country to be considered resident is not the same across Member States²¹, although according to the rules of demographic statistics a person belongs to the resident population of a given country if s/he is staying, or intends to stay, on the economic territory of that country for a period of one year or more²². Moreover, migrants inside the EU are likely to be resident in both their home and host countries at the same time, thus being counted twice when aggregating national figures to get EU estimates. At this regard, **the Task Force recommends to fully apply the concept of resident population as defined by demographic statistics²³ (R5).**

The Task Force expresses concern about the updating of **population estimates**, especially when they are tied to the Population Census and the distance with the LFS' reference period increases²⁴. However, although stressing the importance of keeping population estimates up to date, this issue is deemed out of reach of labour market statisticians by the Task Force.

– *Weighting*

The main reason for weighting is to adjust the sample so that it corresponds to the target population on those characteristics, although it is also a way to adjust for unit non-response (see below, page 42). Council Regulation 577/98 states that such a correspondence should be ensured by sex, age (five-year age groups) and region (NUTS 2 level). Most countries make use of more

²⁰ Gatward R., Lound C. and Bowman J., 2002

²¹ E.g. it is 2 months for Poland, 12 for the UK.

²² UNECE and Eurostat, 2006.

²³ See European Parliament and Council of the European Union, 2008, art. 2(d).

²⁴ For instance, in Germany the population projections may be seriously skewed as no Census has been taken since 1987.

information than the one required by the regulation, either to increase the adherence with the target population or to increase the coherence with alternative sources²⁵.

Although the methods of calculating grossing factors differ considerably among countries, two main ones may be identified depending on the detail of the external information and whether or not this external information can be cross-tabulated: 1) inverse of the selection probabilities adjusted *a posteriori* to the population distribution by sex, age groups and other external (administrative) sources, and 2) different variations of adjusting to marginal totals, including generalised calibration and generalised regression.

Concerning the use of external information in the weighting procedure, it should always be verified that it has a good correspondence with the measured variable, with regard to both concept and measurement. Otherwise, using such variables in the weighting procedure can introduce bias in the estimates. As an example, it is not necessarily for the better to weigh observations on the actual region of residence to the registered region of residence. The Task Force also underlines the importance of good population estimates as previously mentioned.

The Task Force also points out that, in a continuous survey, the merit of taking the time dimension into account, such as ensuring that the quarterly estimates are equally distributed over the months of the quarter. Inner consistency with regard to the wave structure of the sample is also of value.

The Task Force also points to the increasing interest in the analysis of the participation in the labour market from the household perspective, both at national as well as international level. More and more often the household is an analysis unit in itself or the household characteristics are used to explain individual outcomes on the labour market. In this regard, main areas of interest for the European Commission are the characteristics of jobless households and of their members and the reconciliation between work and family life. Coherence of household data with individual estimates is a prerequisite for their use in combination and can be achieved by using calibration estimators. Although the introduction of household constraints slightly reduces the efficiency of estimators, **the Task Force** deems overwhelming the advantage given by the enlarged possibilities of labour market analysis and **recommends to apply weighting schemes which ensure the coherence between household and individual estimates (R9)**.

Finally, in order to produce monthly estimates directly from the LFS, the use of monthly constraints ensuring consistency between monthly and quarterly figures may be envisaged.

3.1.2.2. Non-response

Non-response is the failure of a survey to collect data on all survey variables, from all the population units designated for data collection in a sample or complete enumeration. The difference between the statistics computed from the collected data and those that would be computed if there were no missing values is the non-response error.

²⁵ For instance, seventeen countries have a more detailed regional classification (NUTS 3 or even NUTS 4) than the one required by the regulation, while several Nordic countries use register statistics on employment/unemployment directly for weighting. In other countries, different external distributions or sources are frequently used both for weighting and stratification, such as urban/rural distinction, nationality, ethnicity, and size classes of regions or local areas.

There are two main sources of non-response: refusals and non contacts. The former happens when the sampled unit is contacted by the interviewer but refuses to cooperate, while in the latter case no contact at all takes place between the interviewer and the presumed respondent.

According to the Task Force non-response is a major issue for the quality of the statistics on employment and unemployment from the Labour Force Survey. A minor concern is related to the fact that non-response reduces the actual sample size, thus boosting the variance of estimates. What is more worrying is the non-response bias which arises when the characteristics of non respondents are not random and are instead correlated with the variables concerned²⁶, which is very likely with (un)employment.

The size as well as the composition of non-response by reason of non-response largely varies among EU countries²⁷. Apart from the peak of 68% for Luxembourg, the non-response rate ranges from 34.2% for Denmark to 5% for Germany. It is below 15% in only ten out of 32 countries, while exceeds 20% in eleven countries. Non-contacts are the main reason for non-response in fourteen countries, refusals in eleven (table 1).

Table 1 – Non-response in the LFS by reason (Annual average 2007)

Country	Non-response rate (%)	Incidence of non-response reasons on total non-response			Country	Non-response rate (%)	Incidence of non-response reasons on total non-response		
		Refusals (%)	Non-contacts (%)	Other reasons (%)			Refusals (%)	Non-contacts (%)	Other reasons (%)
LU	68.0	26.5	25.0	48.5	NL	17.8	47.9	27.4	24.3
DK	34.2	-	-	-	IS ²⁾	17.5	56.5	31.4	12.2
LV	34.0	31.1	45.5	23.3	MT	17.3	10.7	89.3	-
EE	31.3	34.2	56.6	9.3	FR	16.8	22.8	55.0	22.1
UK	30.4	72.2	27.8	-	PT	15.9	18.3	65.4	16.3
PL	24.3	59.3	33.7	7.0	IE	15.4	24.8	30.2	45.0 ³⁾
BE	23.6	11.5	44.9	43.5	TR	14.5	1.9	68.0	30.1
CH	20.9	31.1	56.9	12.0	NO	13.6	-	-	-
LT	20.2	39.4	52.6	7.6	HU	13.1	30.7	61.0	8.2
FI	20.2	65.0	34.2	0.8	IT	11.6	33.2	49.9	16.9
CZ	20.2	73.5	26.1	0.6	EL ²⁾	10.5	34.0	-	65.6
SI	19.0	64.2	7.1	28.6	AT	8.0	1.9	97.8	-
ES ¹⁾	18.9	40.7	59.3	-	SK	6.9	40.2	39.6	22.2
BG	18.9	23.3	72.1	4.6	RO	5.0	22.7	38.4	38.9
SE	18.4	50.1	46.8	3.0	CY	3.5	82.8	8.5	8.8
HR	18.0	49.1	23.2	27.7	DE	5.0	-	-	-

1) First interviews only. 2) Data refer to 2006. 3) Eurostat calculation

Source: 2007 yearly quality reports. (Germany: 2006)

²⁶ Following Bethlehem (2002), the bias for an estimator of the population mean is a direct function of the covariance between response probabilities and the values of the target variable. The estimator is unbiased if there is no correlation, while the bias becomes larger the stronger the correlation.

²⁷ Non-response rates are provided by the countries and are not fully comparable. The issue of comparability of non-response assessment is addressed below in this chapter.

Dealing with non-response is not straightforward, as the reasons for non-response are not homogeneous and its impact on the quality of the survey results may be different in different situations. In addition, the literature shows several cases in which pursuing higher response rates at any cost ends up with increasing, rather than reducing, the bias of estimates²⁸. This happens when late respondents are more similar to early respondents than to final non-respondents, which further shrinks the representativeness of the sample. Any decision on how to deal with non-response should therefore be preceded by careful analysis of non-respondents, having in mind that there are no universal recipes fitting all the situations and that each survey is different from the others²⁹.

Dealing with non-response means two things: trying to maximize participation and / or correcting for non-response *ex post* as best as possible. For both purposes, different approaches have to be followed for refusals and non-contacts. The following paragraphs reflect the findings of the Task Force on the practices to quantify and analyze non-response bias, to prevent or limit non-participation and to correct for it.

A further issue for the EU-LFS concerns the comparability of non-response indicators, as different figures may reflect different national practices in the calculation of non-response rates rather than different levels of survey participation. Cross-country comparability of non-response assessment will be addressed in the last paragraph of this chapter.

– *Analyzing and quantifying non-response bias*

Approaches to study non-response go from the use of information retrieved from the sampling frame, the use of *paradata*³⁰, comparison with other sources such as the Population Census, the basic question approach to *ad hoc* follow-up surveys³¹.

Experiences with follow-up surveys seem to be rather limited among the countries in the Task Force. At present only Insee is carrying out a Non-response Survey (hereinafter NRS) for the French LFS³². Its main features consist in a 2 week-extension of the interviewing period, during which non-responding units are asked to answer to a light version of the LFS questionnaire³³. With respect to the main LFS a different mode – a self-administrated paper questionnaire – is used. A positive side-effect of introducing this NRS is a better organization of the field work in general.

²⁸ See Merkle and Edelman, 2002; Stoop, 2005.

²⁹ It should not be overlooked that pursuing higher participation often requires considerable financial efforts. It is therefore crucial that actions are focused to the real issues and are effective to overcome them. Investing in the monitoring and assessment of non-response may therefore prove being a more efficient allocation of funds than just blindly pursue higher response rates.

³⁰ “Process data, or *paradata* as they are also called, are data that are fed back from the process they inform about. For example, paradata may inform about whether a certain transaction was successfully processed or not, and if not, what kind of error occurred. Paradata may be used to signal problems in the design or operation of a public information system, and they may be used as a point of departure for improvements” (Sundgren, 2005). In the specific case of the household surveys they consist in the information on the field work process and may also include reasons for refusals or interviewers’ observations on the dwelling of the non responding sample unit or the neighbouring area.

³¹ For a review of different approaches to the study of non-response see Stoop, 2005 (aforementioned)

³² A more detailed description of the French LFS is reported in Durier and Thélot, 2008.

³³ Such approach is known in the literature as the basic question approach (see Kersten and Bethlehem, 1984).

Although response to the NRS is low (less than 25%) its results are taken into account for the LFS estimates. Integrating the NRS results within the main Labour Force Survey has not straightforward effects. If for instance the focus is on employment, taking into account the results of the NRS would in general result in a reduction of non-response bias only if the employment rates of respondents to the NRS and final non-respondents (i.e. those not participating either in the LFS or in the NRS) are both higher or lower than the employment rate of the LFS respondents³⁴. However, such hypotheses are difficult to test and literature shows cases where higher efforts to increase response rates end up in larger bias. Another problem is that deviating estimates on (un)employment from this NRS could be a result of using a different mode.

Because of the above mentioned quality issues, the Task Force, with the exception of France, would not recommend other Member States to carry out similar NRS, as doubts remain about its effectiveness as a tool to assess non-response bias and even more about the opportunity to integrate its results into the LFS (not to mention the impact on timeliness). Indeed, response to non-response surveys (as the French case shows) is low and the information on final non-respondents is not sufficient to check its effects on non-response bias. Concerns remain also about the use of different survey mode and about its cost.

The Task Force, however, recognises the importance of studying non-response. In particular, given its potential high impact, further research should be carried out into the effect of non-response on the estimation of the labour status. **The Task Force therefore recommends to regularly collect information on the characteristics of non-respondents (e.g. by retrieving structural data from the sampling frame or from low-level geographical databases that can be linked to it, or by conducting non-response surveys) to assess and adjust for non-response bias and to improve fieldwork strategies. However, specific non-response surveys should be used primarily for improving fieldwork strategies, and care must be taken when using them to adjust the results of the LFS (R11).** The Task Force also suggests devoting a seminar to this subject.

– *Increasing contacts*

Prior to any interview, a contact between the interviewer and the interviewee has to be established. Contacting units in the sample may not be an easy task. Two main reasons may be an obstacle for the interviewer to reaching respondents: either respondents are not at home or there are some impediments preventing to get through (households only having mobile phones is an example of such impediments for CATI interviewing. Phone numbers not publicly available, e.g. not in the phonebook, is a further impediment for countries only using CATI). As concerns PAPI or CAPI, obstacles to reach respondents are the presence of alarm systems or watchdogs. In addition, the spreading of numeric codes instead of names on entry phones is becoming an increasing issue. As these situations become more and more common, contacting sampled units is becoming more and more difficult. Research and experience show that the probability of contact is not uniform across the population. It is lower in urban areas, for (mainly young) single persons and for those who have

³⁴ The case in which the employment rate of respondents of the NRS is in between the employment rates of the LFS respondents and final non-respondents is known in literature as *continuum of resistance model* (see Lin and Schaeffer, 1995). This model assumes that response propensity is correlated with the variable of interest.

some activity outside their home such as workers or students³⁵. Such evidence is of major concern for the Labour Force Survey, as all these features are not neutral with respect to the participation to the labour market. As a consequence, great attention must be devoted to reasons for non-contacts in order to limit them as much as possible.

At this regard, the Task Force recommends to provide specific training to interviewers to limit non-contacts. In particular, contact attempts should be carried out at different times of the day and interviewing time should also cover periods after working hours (evenings and weekends). Collecting preferred contact time during the first interview may facilitate contacts at further waves (R18).

Longer interviewing periods, that is the time span after the end of the reference period when the interviews take place, also increase contact probability. However, the marginal gain in terms of higher contact rates rapidly decreases with time. On the other hand, drawbacks such as recall problems soon become predominant. Consequently, the Task Force in general would recommend short interviewing periods (see R24). The bulk of the sample should be interviewed in the two weeks following the reference week. Exceptions may only be justified in particular periods such as holiday seasons when normal interviewing span would risk to give a low response rate.

How to deal with impediments is less obvious. Interviews on mobile phones may also be used when this is the only way to reach respondents. Another possible way out is to try a contact by using a different mode (e.g. a self-administered mail questionnaire). Because of the influence that different modes have on the respondents³⁶ (the so called *mode effect* – see chapter 3.2.2.3), the Task Force suggests that tests are performed before following such practice.

– *Preventing and limiting refusals*

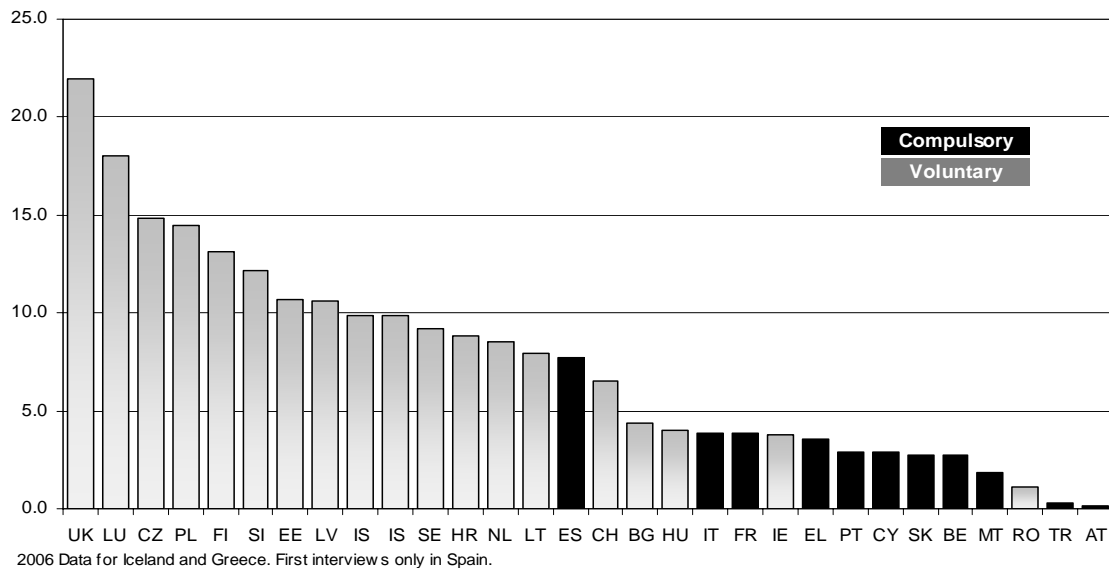
Refusals concern the unwillingness of respondents to answer the questions or their inability to participate. Refusal however presumes that a contact between the interviewer and the selected household or individual has previously happened. Actually, various reasons for refusals may have different impact on the results. Unavailability (e.g. due to lack of time or illness) may be random and have no or little effect, while mistrust in the institutions or lack of interest in the subject may be selective and hence introduce serious bias. When trying to boost participation or to convert refusals, the different reasons for non-participation should be kept into account, in order to adopt the strategy which best suits each situation.

Refusal rates in the EU, EFTA and candidate countries range between 21.9% in the UK and 0.2% in Austria. It can be noticed from figure 2 that generally speaking response rates are higher in countries where answering the LFS is compulsory. The Task Force recognizes that the legal status of the survey depends on national traditions and attitudes towards surveys and is therefore out of reach of LFS managers. In spite of this, **it recommends to make participation in the LFS compulsory, wherever legally feasible, in order to reduce non-response (R12).**

³⁵ See Stoop, 2005, aforementioned; Campanelli, Sturgis and Purdon, 1997; Groves and Cooper, 1998.

³⁶ For a more extensive description of the mode effect see subsection 3.2.2.3.

Figure 2 – Refusal rate by legal status of the LFS (Annual average 2007)



Source: 2007 yearly quality reports.

It is the role of **interviewers** that is crucial to limit refusals. Actually, it is the interviewer who gets through the household or person in the sample, and the way they interact will result in cooperation or in a refusal. Experience shows that expertise is a key factor in obtaining cooperation. More experienced interviewers can usually better *tailor* the approach to the respondents, finding the best way to motivate them or raise their curiosity about the survey. However, not all tailoring features necessarily require previous experience. Practice has shown that respondents usually feel more comfortable with interviewers that are similar to them, who are dressed like them and speak their own dialect or with the local accent (Stoop, 2005, aforementioned, page 150). Establishing a contact with the interviewees and gaining their trust is not always straightforward. Having the same interviewer for all survey waves would therefore avert the need to repeat this crucial step each time. However, this may not always be feasible, especially if different modes are used for different waves. A further element is the motivation of interviewers. Those who are more confident about getting the interview usually show higher response rates. Recommendations from the Task Force in this regard are:

- **Make use of permanent, professional interviewers. Minimize turnover of interviewers to ensure continuity (R19)**, as a high interviewers' turnover or lack of motivation may be a serious threat for the quality of the survey³⁷.
- **Provide specific interviewer training on how to approach households to avoid refusals (R18)**, especially for inexperienced interviewers.

³⁷ It should be noted different deployment strategies for the interviewers may be adopted for different modes (personal or telephone interviewing).

- **Provide interviewers with appropriate training on how to conduct the interview, covering both communication skills and survey content. Training should be provided on a continuous basis (R20).** The use of mentors to help in difficult situations would also help.
- **Organize periodical focus groups and debriefing for interviewers to review and tackle issues (R21).**

Offering the respondents different **modes** is another way for tailoring the interview to their preferences. Some people may not trust letting in a stranger but would answer to the telephone or to a self-administered questionnaire. Web questionnaires on internet may appeal youngsters. However, the same general concerns about the mode effect expressed in the previous paragraph suggest being cautious.

The Task Force recommends instead to make use of several tools to inform or reassure respondents, such as presentation letters to sampled units, free-toll numbers to answer their questions or reassure them and identification cards to be visibly worn by interviewers (R13). Presentation letters are by now common practice among National Statistical Institutes. A good presentation letter should introduce the survey, highlighting its importance for the country and for the respondent itself. It should also present the NSI, explain why the household or the person has been chosen among thousands or millions, underline the confidential use of the retrieved information exclusively for statistical purposes, provide references for further information.

A specific issue concerns **non-nationals**, which are often underrepresented in the actual samples. One reason is the communication issue, which arises when they do not speak the national language. In order to increase their participation the **Task Force recommends to arrange for special tools such as translated questionnaires and presentation letters, and non-nationals-dedicated free-toll numbers where information is provided in other languages (R14).** On the other hand, the Task Force also recognises that refusals from non-nationals are not only due to language problems but also to cultural differences. *Ad hoc* strategies to reach non-nationals should therefore be envisaged.

Further tools often mentioned as effective to get cooperation are **incentives**. The ONS tested them on the UK LFS³⁸. As an overall assessment, the use of incentives seems to have increased response rates. Monetary unconditional incentives have a higher effect on boosting participation than other kind of gifts. Incentives also seem to have a positive motivational effect on interviewers, but this effect fades away when these become usual practice. In addition, the use of incentives on a continuous basis is rather expensive. For these reasons, incentives may be used only when it is deemed necessary (e.g. holidays, for household entering the sample for the first time, etc.). Furthermore, incentives may have highly selective effects, thus increasing rather than reducing non-response bias. Indeed, in none of the two countries represented in the Task Force where incentives have been tested (the other one being the Netherlands) they have been permanently implemented. To what extent and in what circumstances incentives are effective in boosting participation remains doubtful in the case of the LFS. The issue would need to be further investigated. For this reason

³⁸ More details on this experience are provided in Prestwood, 2008.

the Task Force deems that at this stage no clear recommendation can be envisaged. However, the Task Force agrees that, for reasons of cost-effectiveness, incentives should not be used on a continuous basis.

Response burden is often a reason to escape a survey. Especially for subsequent waves, remembering the length of the previous interview sampled units may decide to quit. Keeping response burden low is therefore a key feature to enhance participation. Two practices can help in this regard: the *wave approach*³⁹ and *dependent interviewing*.

At the moment only a few countries make use of the wave approach, namely the Netherlands, Spain⁴⁰, Germany, France, Norway, Finland, the United Kingdom and Bulgaria. However, only some of them make a full use of it, while others apply it only for some of the structural EU-LFS variables. All other countries do not use it at all. The Dutch experience with the use of the **wave approach** for the LFS shows a reduction by 20 to 29% in the duration of the interview (Kösters, 2008). Shorter interviewing time has further advantages. The most relevant is that it is cost-effective, for instance because shorter interviews are more suitable for CATI. Financial resources becoming available may be better allocated to improve other aspects of the LFS. On the other hand, the wave approach is not free of drawbacks. Some disadvantages are of a technical nature: different weights have to be computed for the structural subsample; different questionnaires have to be set up, managed and maintained at the same time. Other apprehensions concern the results. The use of a subsample gives rise to consistency problems between the quarterly and yearly estimates. By including constraints in the weighting scheme of the yearly dataset consistency may be ensured for the main variables, but not for all. At the same time the wave approach shrinks the analytical potential of the LFS, as some variables can no longer be used on a quarterly basis. In addition, for variables which concern small population groups only or with a large number of response items, the reduction in the number of available cases would limit the potential for analysis by regional or even basic socio-demographic variables. The choice of the variables to be surveyed using the wave approach is therefore heavily depending on the priorities determined by users and stakeholders.

At present, concerns about the drawbacks have made most countries refrain from using the wave approach. Regarding the technical obstacles, many concerns seem to be based on a priori evaluation. Most of the technical issues in fact are related to the setup of the system, but are no longer a burden once it is in place. On the other hand, the experience of the countries which have implemented the wave approach shows that an extensive use of it allows reducing response burden and costs. Overall, these advantages outweigh drawbacks. Therefore, **the Task Force recommends to consider implementing the wave approach (R15).**

³⁹ The *wave approach* is used for the *structural variables* which need be surveyed on a yearly basis, differently from *core variables* for which quarterly figures are required. Generally speaking, in order to get for yearly structural variables the same reliability as for quarterly core variables, questions concerning the formers can be asked each quarter to a subsample of the full quarterly sample, e.g. the first wave. Thus, when rotating patterns with four waves are used, yearly results for the structural variables turn out to be yearly averages collected over a sample that on yearly basis has approximately the same dimension of a full quarterly sample and which is evenly spread all over the year. In case of rotating patterns with more than four waves, the use of more than one wave may be needed in order to attain sufficient precision of the estimates.

⁴⁰ The Spanish experience in the implementation of the wave approach is described in García Martínez, 2008 (a).

Generally speaking, **dependent interviewing** contributes to reduce response burden at subsequent waves thus helping to limit refusals. Experimental research (Jackle and Lynn, 2007) has also shown that it is an effective tool to prevent temporal inconsistencies among waves. However, because respondents tend to confirm the information from the previous wave, variations between waves are likely to be underestimated. For this reason, dependent interviewing is suitable for variables which are rather stable in time (e.g. sex, education level, usual situations at work such as usual working hours, etc.) but not for those which are more prone to changes. Although Regulation 1897/2000 allows to shortly verify the status of employed or inactive (but not of unemployed!) with reference to the previous wave when they seem to be stable, dependent interviewing could nevertheless impact on their measurement.

In general the Task Force recognizes dependent interviewing as an effective tool to reduce response burden and thus increase participation, and **recommends to make use of it for variables that tend to be stable from quarter to quarter. However, dependent interviewing must not adversely impact on the measurement of the ILO labour status. In particular, the question on having worked for at least one hour during the reference week should be asked by default (R16).** This question is crucial to identify the ILO labour status of respondents. Indeed, the spreading of marginal employment and of temporary, often short-term contracts make being employed a less and less stable situation. However, even working situations apparently stable should always be verified. For instance, a permanent employee in the public sector may have started a long-term absence for which he/she should be considered as inactive. This would likely be overlooked if only confirmation of the job stated at the previous interview is asked.

– *Correcting for non-response*

How to deal with non-response once the data collection is over and non-response has occurred? If no bias can be detected, that is missing units are *completely at random* (Bethlehem, Cobben and Schouten, 2007), there is no need to correct for non-response and the only drawback will be a larger variance. If instead non-response produces biased results, ad hoc constraints can be introduced into the **weighting** scheme. In principle, this can be done when some characteristics of non-respondents are known (e.g. nationality, household dimension or composition, characteristics of dwellings, earning bands, etc.) and reliable information from external sources is available. The underlying hypothesis is that the target and auxiliary variables are correlated (such missing data mechanism is known as *missing at random*). If this is the case, missing units can be represented by those with the same characteristics who participate. However, participation to the survey may be related more to the unknown characteristics we want to observe, i.e. labour status, rather than to any structural feature (this mechanism is called *not missing at random*). If this is the case, special constraints are useless to correct for non-response. A drawback of including more constraints, although rather limited in size, is a reduction in the efficiency of the estimator.

Following the hypothesis that auxiliary variables are correlated with the employment status, simply applying the minimum set of constraints foreseen by Council Regulation 577/98 already helps reducing the non-response bias. Indeed, most countries include a larger set of constraints in the weighting scheme. National practices at this regards are largely different from one another with respect both to the number and the kind of auxiliary variables used. **The Task Force** deems that agreeing on a common set of constraints is neither feasible nor advisable. Nonetheless, it

recommends to optimize weighting schemes to correct for non-response bias, taking into account specific characteristics of non-respondents (R10).

Substituting non-responding units is not a widespread practice among Member States. It is the case in Spain, Italy and Cyprus. Actually, its effects are controversial. While it helps keeping the actual sample size close to the theoretical one, its impact on sampling errors could in principle be high. Furthermore, in case of selective non-response, non-response bias may be increased rather than reduced. Italy reported that from their experience random sampling of substitute units does not help getting a more representative actual sample. On the other hand, selective substitution for specific subpopulations such as non nationals may be the only way to have enough of them within the sample, in order to get estimates of reasonable precision.

A further way to correct for non-response is imputation. **Copying** information for non-responding units from previous survey waves if available is a sort of such practice. It relies on the assumption that correlation between labour status in the two periods is high (actually equal to 1). Nonetheless, if this hypothesis in general holds for employment, it is less certain that it fits unemployment. As an example, for Italy such correlation at a distance of three months is 0.4.

– *Comparability of non-response assessment*

The Task Force agrees on the need for homogeneous calculations of non-response indicators across Member States. At present, comparability is hindered because several elements are not necessarily treated in a homogeneous way: a) the codification of final dispositions and the treatment of temporal dispositions to define final dispositions; b) substitutions. Countries which substitute non-respondent units should calculate non-response rates on the basis of the participation of the original sample and not on the basis of the achieved one. In the last two situations the non-response is thus kept at an artificially low level compared to countries that accrue it over survey waves or do not substitute missing units; c) the treatment of non-response in secondary survey waves. Some countries, especially those who do not attempt refusal conversion, do not consider refusals in the first wave as part of the refusals in the subsequent waves; d) the different unit for calculation, which is the household for most of the countries, except for Denmark, Finland, Sweden, Iceland, Norway and Switzerland which calculate non-response on the basis of individuals; e) the treatment of units of unknown eligibility; f) the treatment of Primary Sampling Units non-response; g) the formula for non-response indicators. **The Task Force, therefore, recommends to develop a harmonised approach for the calculation of non-response indicators. Eurostat should initiate work on this issue (R17).**

In order to allow Eurostat to have comparable indicators two alternatives were proposed: a) Member States make calculations on the basis of an agreed method and provide Eurostat with the required indicators; b) Member States send a table of final disposition codes letting Eurostat calculate non-response indicators according to agreed methods. The Task Force favours the first option, as this is the only way to ensure that the same information is disseminated nationally and at EU level.

Eurostat will submit to the LAMAS working group a proposal for homogeneous calculation of non-response indicators. As suggested by the Task Force, the proposal will take into account the

existing literature and in particular the AAPOR standard definitions of final dispositions of codes and outcome rates⁴¹.

A further element of discussion concerns which breakdowns should be used for the calculation of non-response indicators. In Athens it was suggested to include the breakdowns by quarter, wave, non-response reasons, survey mode, NUTS 2 and NUTS 3 (for countries providing NUTS 3 data from the LFS). However, some countries may not be able to distinguish between different non-response reasons. As a preliminary action, Eurostat should carry out a review of information Member States actually collect. For the rest, the level of detail at which the non-response information should be provided is deemed of no concern. Nonetheless, only information that is in fact used should be collected.

3.1.2.3. Measurement issues

Measurement issues cover instruments (questionnaire, explanatory notes, dependent interviewing, mode of data collection), interviewers (training, control, management) and response errors (respondents, recall problems, proxies). Conclusions on these topics are presented in this chapter, except for dependent interviewing, which is covered above in the section on non-response.

– *Questionnaire*

The questionnaire has a huge impact on the final results, as even slight differences in the wording of questions, in the order they are presented to respondents or in filters being applied may influence the answer.

All these elements can easily lead to biased results which do not correspond to the target concepts. The set up phase of a questionnaire is of course crucial for the accuracy of the estimates, but also modifications imply delicate steps. From the slightest amendment of a single question to a complete revision or setup of the form, accurate tests should be performed in order to assess different approaches and select the most appropriate ones. Testing techniques stretch from laboratory testing, to cognitive tests, to pilot surveys which reproduce on a smaller scale the full survey conditions. In the European Statistical System (ESS) it is considered recommended practice to carry out questionnaire testing in a large number of different situations:⁴² For instance, systematic questionnaire testing becomes necessary when new questions are added to the questionnaire, existing questions are being modified (even if apparently minor changes are made), the data collection instrument is being changed (e.g. change to CATI questionnaire), or poor data quality has been indicated by a review of non-response rates and biases, validation against other surveys or re-interview studies, deficiencies in internal consistency or other evidence.

Because the smallest deviation may influence response, the interview must be as standardized as possible. As a consequence, the questionnaire must be precise, ruling out any possibility of misinterpretation. At the same time, the wording should allow a natural conversation situation. Because any external intervention is subjective, in an ideal situation the need for explanations by the interviewers should be zero. At this regard, the language used should follow the common one, avoiding the use of technical or legal words. Even the introductions to subsets of questions or

⁴¹ See AAPOR, 2008.

⁴² Eurostat, 2006 (c), p. II.

shifts between different topics should be foreseen in a standardized way whenever needed, in order to prevent different approaches from different interviewers. Reference dates should be clearly specified, in order to provide clear guidance to the respondents. When a long set of questions refers to the same reference period, it is good practice to mention it again from time to time, in order to keep respondents focused on the right period. As an example, if separate questions are asked for the 13 job search methods in the Commission Regulation 430/2005, the reference period should be reminded every three or four questions, in order to avoid that respondents unconsciously switch to different concepts (e.g. the current situation). Nevertheless, an experienced interviewer should be able to recognise if the answer is in effect in line with the goal of the question and intervene if it's not the case.

A review of national questionnaires has been carried out by the Task Force, especially to assess their compliance with the definition of unemployment and the twelve principles to survey labour status as required by the Commission Regulation 1897/2000. Apart from these guidelines, the EU-LFS is not input harmonized. Hence, national questionnaires largely diverge. This situation, however, raises concerns about cross-country comparability rather than accuracy. The issue is therefore addressed in the related chapter.

– *Survey modes*

The use of mixed modes is quite common among Member States. This practice has several advantages: it allows reaching different groups of population and reducing costs and response burden when telephone interviewing is used in subsequent waves. Pressure for cost reduction and the spreading of internet use among the European population are raising the interest of national statistical institutes for web data collection also for household or individual surveys. However, the mode has an influence on the answers provided by respondents. A survey of the possible reasons for such influence is provided by Thomas (2008). Time pressure, usually higher in telephone interviewing, may affect recall or judgement process. In face-to-face interviews such aspects may be enhanced also by the use of cards or other visual stimuli. The impact of the mode is particularly relevant for open ended questions, which tend to return more detailed information in face-to-face interviews, and for questions with many answer categories, where the order in which they are presented matters. Especially in telephone interviews where no cards can be shown and all items have to be read by the interviewer, respondents tend to concentrate on the latest items and to forget the first ones (*recency effect*). If needed, it is easier for interviewers to provide guidance within a face to face rather than a telephone interview. When social desirability is somehow involved the interaction by telephone is more neutral, whereas physically in front of an interviewer social acceptable answers tend to be provided. Finally, different results from the various modes may arise from different response rates and characteristics of the units that can be reached with each of them⁴³.

Research on the effects of different modes on the Labour Force Survey in the UK (Thomas, 2008, Elliot, Rainford and Eldridge, 2006) and in Spain (García Martínez, 2008 (b)) found some mode

⁴³ For instance, persons who don't have a fix line cannot be reached by telephone if interviewing through mobile phones is not allowed. Likely they are not randomly spread among the population, but their characteristics are correlated with the survey variables.

effects on some variables but not on the labour status, which was the target of the Task Force review. Such findings suggest that, if only employment and unemployment figures are taken into account, making use of different modes may improve survey participation and reduce costs and response burden at the same time. **The Task Force thus recommends to consider the use of mixed modes, in the light of possible gains in response rates, burden and costs versus the mode effects (R26).** However, such conclusion should not bring to the straightforward conclusion that modes may be indifferently used. As a recommended practice, **changes to modes (as well as also to the questionnaire and other explanatory material) should always be carefully tested and their impact be assessed before being introduced (R27).** As a further recommendation, **the LFS should always be carried out by computer-assisted questionnaires, given that the traditional paper interviewing mode is no longer suitable to cope with the complexity of the LFS (R28).** The Task Force suggests the possibility to enforce the use of computer-assisted modes by regulation. Concerns also include the use of self-administered questionnaires. To what extent the increasing use of web interviewing is an alternative to the traditional self-administered paper and pencil questionnaire is still to be verified. Therefore, **the impact of self-administered electronic data collection, including web-based modes, on the measurement of ILO labour status should be carefully investigated (R26).**

– Interviewers

In addition to ensuring participation, interviewers play a key role in retrieving the right information from the interviewees. Instructions to interviewers, their working conditions (including remuneration), how their work is controlled and how field work in general is organized have a large influence on that.

The Task Force agrees that **training to interviewers** should be centralized, as the information deteriorates when passing through different levels of trainers. However, when interviewers are spread all over a country, gathering them all together in a same place at the same time may be expensive. Therefore, when only minor changes occur centralized training may concern only regional supervisors which in turn would train interviewers in their region. The use of self-training instruments or e-learning applications may also be used in such cases. Final tests for supervisors and interviewers are deemed as good practice. **In particular, the Task Force recommends to instruct the interviewers in general not to adapt questions to the respondent but to just read them as they are written in the questionnaire, as personal interpretations unavoidably differ from one another, thus introducing interviewer effects as a relevant source of bias (R22).** Computer-assisted modes instead permit to foresee different question wording for different kind of respondents, so that interviewers always have a homogeneous approach in each situation.

As directly experienced in the Italian LFS, professional interviewers perform better than occasional ones. Besides, expertise increases with experience. For this reason, interviewers' turnover should be limited as much as possible. In this regard, as a relevant motivational feature, **interviewers' remuneration should be adequate with respect to the interviewers' crucial role for the quality of the survey (R19).**

A related issue concerns whether interviewers should be dedicated only to the LFS. There is no evidence that interviewers carrying out several surveys at the same time perform worse than those dedicated to a single survey. In addition, this may be feasible – but not necessarily convenient from

an economical or organizational point of view – when interviewers are directly working for NSIs, but not when they work for third parties, which is very common especially for CATI interviews.

Several practices to **control** interviewers are followed by the countries in the Task Force. In the field of CATI interviews, central supervisors monitoring the interviews in the telephone facility allow a direct and cost-effective control of the interviews and should be standard practice for this data collection mode. Examples are provided by INE Spain⁴⁴, Statistics Austria, and others. In contrast, for face-to-face interviewing no direct control method is available. In this case, a number of complementary approaches should be used in order to guarantee an equivalent level of monitoring as in telephone interviews. This can be achieved by the use of regular quality surveys as practised for example at INE Portugal and the US Census Bureau (see Torres, 2008, Sezer, 2008 and Bureau of Labour Statistics, 2002). The quality survey carried out at INE Portugal is considered a good practice example. Carried out by senior interviewers, LFS central staff or regional supervisors in the two weeks following the LFS interview, it allows to assess survey bias and variability as well as to detect eventual fraudulent behaviours of interviewers. Further methods providing less systematic monitoring, but also requiring fewer resources than regular quality surveys, include:

- the use of a set of indicators on the field work to evaluate interviewers' performance or to discover possible systematic cheating (e.g. if shorter paths along the questionnaire are more frequent than they should be);
- the analysis of the protocols of computer assisted interviews (allowing, e.g., to look at the time devoted to the individual questions; see Stadler B., 2008);
- the analysis of audio recorded interviews using behaviour coding (a method for instance suitable to monitor the actual degree of standardisation of interviews; see Furubjelke/Rackner, 2008 and Pascale/McGee, 2008).

For its crucial impact on the accuracy of the final results, **the Task Force recommends to continuously monitor and systematically assess the quality of the fieldwork by using the methods mentioned in this paragraph (R23).**

– *Response errors*

When respondents do not know the required information, or remember it incorrectly or misreport or hide it, either deliberately or not, this can lead to biased results.

Proxy respondents, i.e. persons providing information concerning somebody else selected for the survey whose situation concerning the survey topics they are supposed to know, are at least in theory more prone to inaccuracies. The definition of proxy interview is not always straightforward. A survey of interviewers carried out by Destatis showed that interviewers code similar cases in different ways. Similar evidence comes from France, where interviewers asked for guidance on cases when the respondent is assisted by somebody else. Although the Task Force recognizes that a

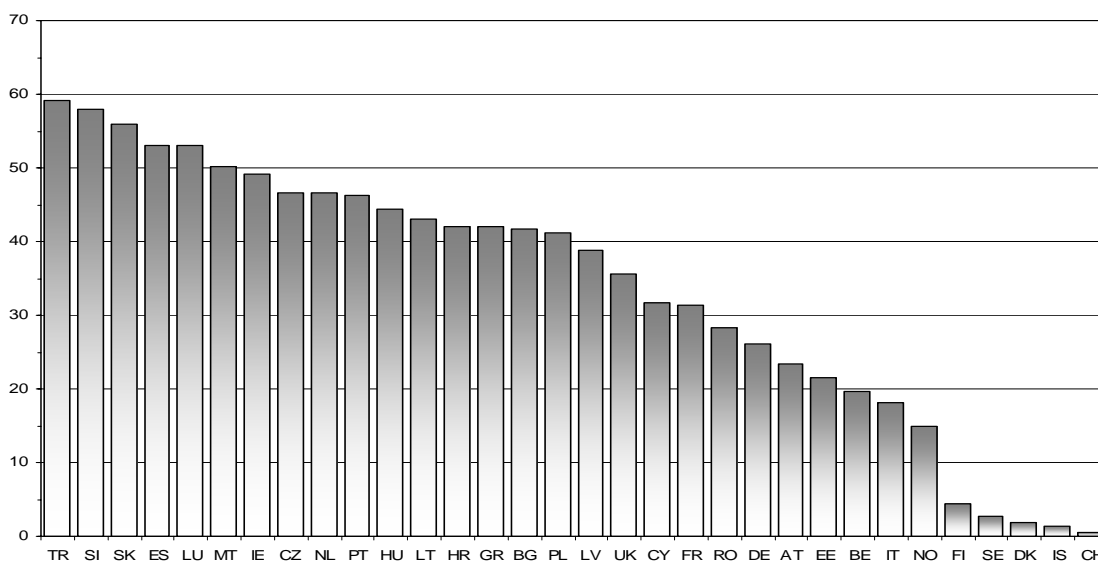
⁴⁴ A systematic monitoring program of the interviewers work is followed in the Spanish LFS since 1971. In 2006 the program was rearranged in order to focus in CAPI interview. The results of quality assessment of the Spanish LFS in 2007 (at the moment available only in Spanish) can be consulted at http://www.ine.es/docutrab/eval_epa/evaluacion_epa07.pdf

definitional issue exists, it is deemed of secondary importance. However, Eurostat may provide guidance at this regards, for instance in the EU-LFS explanatory notes.

Potential bias due to proxy answers is a more serious concern. Evidence shows that proxy is selective with regard to the main structural characteristics, but also that employment and unemployment indicators differ between proxy and direct respondents (although this may be due to the different structural characteristics). However, from this one cannot directly infer that proxy answers are inaccurate or biased. Asked for an overall assessment on the quality of proxy interviews, the German interviewers deemed them quite reliable. A follow up survey carried out in Germany found no significant differences in inconsistencies between proxy and direct answers, although inconsistencies in those two cases may arise from different sources. Similar evidence was found by Istat when checking longitudinal inconsistencies. At the same time, other studies indicate that the use of proxy interviews might lead to a slight underestimation of the number of employed persons (see Kleven/Lagerstrøm/Thomsen 2008). Nevertheless, these studies also suggest that allowing proxy interviews results in a reduced measurement error compared to the higher non-response bias to be faced in a survey without any possibility of proxy answers (see Thomsen/Villund 2008).

The share of proxy interviews varies widely across countries (figure 3). In Italy, for instance, it is smaller than for other countries, especially because in CATI interviewers are instructed to arrange appointments to call back and get a direct answer. Doing the same for face-to-face interviews would be both too expensive and too burdensome for the households, as it would imply visiting them more than once.

Figure 3 – LFS proxy rate (Annual average 2007)



Source: Eurostat, LFS.

Proxy answers may be suitable for some variables and less suitable for others like wages, working hours or marginal employment. Any assessment of potential proxy bias should therefore be carried out for specific variables.

In general, direct interviews should be preferred to proxy interviews. However, at present clear evidence of biased results caused by proxy interviewing is not available. More studies should be carried out to assess potential bias. In addition, proxy interviews also have several advantages: they remain an indispensable tool to increase response, to reduce interviewing times thus increasing timeliness and to minimize recall problems if direct interviews can only be carried out later than proxy ones and are therefore considered unavoidable.

Recall is overall a minor issue within the time span for data collection allowed by the LFS Regulation (5 weeks after the reference week). Evidence from the Portuguese follow-up survey shows high consistency for employment and unemployment, meaning that when asked after some time respondents still provide the same answer. Inconsistencies, however, are large for variables that require recall (as hours actually worked) and variables that require codification (activities, occupation, etc.). In the Netherlands, the European definition of reference periods, especially the ones for job search and availability to start work, is thought not respondent friendly. The current situation is asked for instead. This unavoidably leads to an uneven distribution of the sample by reference week along the year, for instance because in holiday periods less interviews are carried out. This in turn is likely to introduce bias in the estimates. However, the Task Force underlines the importance of recall problems. **In order to minimize them, but also to support timely production of results, it therefore recommends to keep interviewing periods as close as possible to the reference period. Exceptions may be made in particular periods such as holiday seasons, when shorter interviewing periods would result in a low response rate** (see above, chapter 3.2.2.2) (R24).

As concerns the variables employment and unemployment, these may be influenced by **misreporting** of hidden and/or illegal work. Respondents may deliberately or unconsciously provide socially desirable answers. A particular problem concerns respondents' perception that the LFS data may be linked to other administrative sources for cross-checks. For instance, they would tend not to state minor or unofficial jobs if they fear that this could impact on their unemployment benefits or taxes. Asking ID numbers, required to match individual records when the LFS information is supplemented by other retrieved by registers, is a very delicate issue. A possible expedient to minimize misreporting is asking them at the end of the interview.

Even though there is always a risk of misreporting, the way the survey is presented to respondents is crucial in order to minimize it. At this regard **the Task Force recommends to stress, both in the presentation letter and from the interviewers side, that the information collected from individual respondents is treated as confidential and will be used for statistical purposes only** (R25).

3.1.2.4. Processing issues

With the use of computer-assisted interviewing (CAI) techniques, processing issues are judged to be of minor concern for the quality of the LFS, as errors due to editing, coding or data entry are at least in theory largely reduced. The same holds for cross-sectional inconsistencies, as most checks can directly be implemented in the software. However, the precondition for this reduction of errors is that the software used for data entry, coding and editing is well conceived, interviewer-friendly and has been systematically tested prior to its use in the data collection. Main problems concern

systematic errors that not necessarily create inconsistencies in a record⁴⁵. The example of an interviewer coding all employed in the agricultural sector was given. The use of CAI minimizes the need to process data after data collection. However, incompatibility rules should be carefully conceived in order to obtain consistent data without ruling out improbable but possible combinations (“over-editing”).

A specific issue concerns datasets that Eurostat receives from Member States. Processing problems arise with all kinds of datasets, for quarterly, yearly and *ad hoc* module results. Special attention should be paid to the additional conditions in case the wave approach is used for the data collection of structural variables. In this case the consistency of totals of the ILO labour status by sex and age groups between the yearly dataset and the annual average of the quarterly results should be ensured and the coverage of the *ad hoc* module sample by the sub-sample used for the structural variables should be respected. In order to allow countries to detect inconsistencies at national level before transmission to Eurostat, Eurostat should make available to NSIs the check program it uses to process and control national datasets⁴⁶.

3.2. Coherence between the LFS and National Accounts estimates of employment

Lack of coherence is a problem when it comes to the output, simply because users are puzzled in front of different figures referring to identical or similar concepts. However, on the input side incoherence can actually help to identify measurement errors in the sources that can be targeted by specific actions.

The focus in this context is solely on the incoherence between employment in the LFS and in National Accounts (NA) estimates. Indeed, such statistics are produced by all countries in the European Statistical System following common guidelines. The ILO concept of employment which the LFS aims at in theory covers all work performed within the production boundary of the United Nation System of NA (SNA). Discrepancies between employment estimates (with the NA estimates usually higher than the LFS) can give rise to questions from users.

The situation is different for unemployment, for which the European figures exclusively refer to the ILO concept as surveyed by the LFS. At national level alternative figures may be available, mainly data on registered unemployment. Such figures however reflect national situations which may largely differ from one another and whose underlying concepts are very different from the ones in the LFS.

The lack of coherence between NA and LFS employment estimates does not worry users in the same way in all countries represented in the Task Force, as emerges from an *ad hoc* survey on communication issues set up by the Task Force and answered in cooperation between LFS and NA experts in the Task Force countries. While incoherence is seen as a major concern in Germany, the Netherlands, Greece and France, in Spain, Poland, Portugal and Italy this seems not to be the case. Even in countries where users find it hard to understand the gap the situation may be different. Indeed, only in Germany the issue has had a large echo and has triggered a public debate, whereas

⁴⁵ Especially the coding of professions and employment sector require some efforts (e.g. keeping the thesauruses up-to-date or achieving reliability of the coding operations independently of the interviewer).

⁴⁶ As Eurostat's control program is to be reprogrammed in 2009, the Task Force agrees to wait for the new version.

in the other countries the discussion has remained confined to specialists. Questions from the users mainly concern gaps in levels (reported by all four countries), while trends attract attention only in Germany⁴⁷ and the Netherlands. Also the level of concern differs from country to country. Incoherence issues are considered to bear the risk of undermining the credibility of the NSI in Germany and Greece but do not stretch so far in France and the Netherlands.

Acknowledging the differences in the perception of the issue in the different countries, a review of the reasons of the gap between the two sources helps detecting where the LFS might need to be improved. When discrepancies are due to different scope or definitions, suitable communication strategies to inform and guide users have to be set up.

The approach the Task Force followed to address the coherence issue foresaw three stages:

- 1) Understanding and quantifying the lack of coherence, distinguishing differences in concepts, scope and definitions from inconsistency resulting from accuracy problems;
- 2) Singling out targeted measures to decrease the inconsistency by improving the accuracy of the LFS;
- 3) Clear communication of the (remaining) incoherence to users.

A reconciliation table and a questionnaire on communication issues formed the basis of stages 1 and 3 respectively. The table was thought to assess the size and nature of discrepancies between LFS and NA employment figures. The latter aimed at finding out about NSI experience as regards the communication on coherence problems. The forms have been filled out in cooperation between LFS and NA experts.

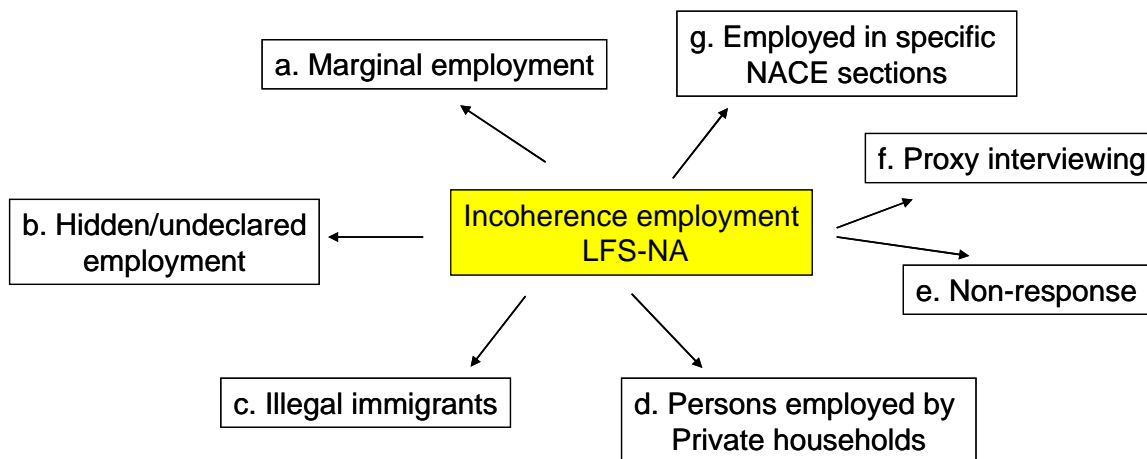
Following this exercise, the Task Force holds the view that the production and subsequent analysis of reconciliation tables is essential for a proper understanding of incoherence between NA and LFS estimates of employed persons. **The production, at least annually, of such reconciliation tables is therefore recommended in order to enable targeted improvement measures for the LFS and National Accounts and for communication to users. A suitable common template should be used in order to allow cross-country comparisons (R31).** As an example, the one used by the Task Force for the aforementioned exercise is presented in Annex 1.

Improved relations between the two statistical sectors, which tend otherwise to ignore each others, could be a positive side effect of the exercise. **The Task Force recommends to foster close communication between LFS experts and National Accounts labour market experts, to enhance cooperation and increase the mutual understanding of the methods used in each domain and their respective strengths and weaknesses (R32).**

The use of reconciliation tables allowed identifying a range of potential sources of incoherence (figure 4), either related to a biased measurement of specific areas of employment, such as marginal employment, employment in black labour market activities, employment in private households, illegal immigrants, or emerging from data collection, as in the case of non-response and proxy interviews.

⁴⁷ This might be due to the fact that the quarterly trends are biased due to a variety of methodological effects in the German case.

Figure 4 – Potential sources of incoherence between LFS and NA estimates of employment



The measurement of specific groups of employment leads to incoherence only if NA estimates include persons employed in these groups on top of those measured by the LFS, or if other more exhaustive sources are available. Non-response and proxy interviews lead to incoherence if corrections for the related bias are carried out by national accountants, or indirectly through the use of other sources.

It can be concluded from the reconciliation tables that the actual sources of incoherence are quite diverse across countries. Even when a potential source of incoherence identified in a country does not emerge as such elsewhere, this may be a signal of an accuracy problem. Improvement in the LFS would then benefit also the accuracy of NA estimates. **The Task Force therefore recommends to conduct further research in the areas identified as affected by potential accuracy problems in the LFS that, on the basis of evidence coming from reconciliation tables, could result in inconsistency with National Accounts estimates: marginal employment, employment in black economy activities, employment in private households, illegal immigrants, the influence of non-response and proxy interviews (R33).**

The first potential source of incoherence identified by the Task Force is the estimation of marginal employment. Generally speaking, the Task Force recognises the difficulties involved in capturing small jobs with a household survey. Although quantified evidence of underestimation of marginal employment is only available in Germany, marginal employment is possibly underestimated in the LFS of other countries as well and could be a source of incoherence if more accurate sources were available and used in NA estimates. Potential measures to improve the situation include improvements to the questionnaire or even additional questions. More evidence and experimental testing is needed.

Further underestimation of employment in the LFS is likely to come from persons employed in the black economy. Some evidence of the underestimation of persons employed by private households, which account for a substantial part of the black economy, comes from an Italian study as well as Greek, German and Dutch estimates. By contrast, a study reveals that the underestimation of Polish black market employment in the LFS is limited. Potential measures to improve the situation

include improvements to the questionnaire or even additional questions. More evidence and experimental testing is needed.

The measurement of illegal immigrants is not done in an identical way across countries. Even though theoretically part of both the sampling frame as well as the reference population, in practice it is very difficult to survey illegal immigrants with the LFS. Besides affecting comparability, it could lead to incoherence with NA estimates. This is however not thought to be significant in most countries, given that working illegal immigrants are unlikely to be fully included by NA on the labour input side. Even when NA take into account illegal immigrants for their estimates, most of the difference with the LFS should be ascribed to differences in coverage and scope, as illegal immigrants often live in collective households or in accommodations such as shacks, caravans, etc. that are excluded from LFS sampling frames.

Non-response has been identified by the Task Force as a possible source of incoherence with a potentially high impact. Quantified evidence is available in France and the UK, indicating an underestimation of employment even after weighting. There is however not enough evidence to conclude on the direction of non-response bias for employment, especially after weighting.

Limited evidence from the UK (1995 study) and Germany (2008 study) indicates a potential underestimation of employed persons by the LFS due to proxy interviews. The scale of the underestimation appears to be limited, however, with both studies showing a very high consistency between proxy interviews and direct answers. None of the countries represented in the Task Force corrects for this effect, making it a potential source for incoherence albeit on a relatively small scale and dependent on the percentage of interviews being proxy.

While the comparison of employment between LFS and NA estimates by NACE section could provide some useful insight for specific sections such as transport and construction, in general such an exercise is unlikely to result in tangible conclusions on the incoherence in level of employed persons because of different ways of obtaining the NACE information in the LFS and other sources used in NA.

Countries in the Task Force follow different approaches also in the communication of incoherence to users. Some countries (Spain, Germany and Greece) provide specific metadata on the differences, while Poland and France only provide separate metadata for the LFS and NA employment estimates and the Netherlands disseminate a reconciliation table. No metadata are available for the Portuguese users.

Metadata are thought for specialized users only in Spain and Greece, while in Germany and Poland they are addressed also to non-professional users. Except for Germany, where specific documents are disseminated, guidance to users is provided on request.

As a general statement, the Task Force underlines the importance of appropriate communication to users to maintain or restore the credibility of statistics. Communication concerns what to communicate, i.e. the content of the information, and how to reach the users. **The Task Force recommends to make available specific documents to explain the differences between LFS and National Accounts employment estimates and provide guidance to users on which source fits which purpose. The aforementioned documents should be easily accessible to domestic and international users (R30).**

3.3. Comparability of employment and unemployment statistics

Comparability is to be ensured over time and over space. Issues concerning comparability over time arise whenever changes in concepts, organization or methodology are introduced in the LFS. The statistical effects thus introduced are often difficult to isolate from the economic ones, introducing breaks in time series which hamper the correct assessment of the evolution of the labour market.

Issues of comparability over space are instead related to the fact that the LFS is output harmonized but not input harmonized. Different survey designs, organization, questionnaires, instructions to interviewers, target populations, weighting schemes are at present in use in Member States. The regulatory framework given by Council and Commission Regulations and gentlemen agreements (as for the explanatory notes) provides a minimum set of common standards.

3.3.1. Comparability over space

Comparability problems emerge from the lack of harmonization of national surveys, from the survey design to the organization of the entire production process, and are often entwined with issues concerning other quality dimensions. Considerations, conclusions and recommendations regarding comparability over space are explicitly mentioned in other chapters of this report, where aspects relevant for other quality dimensions are discussed. It is the case of the chapter 3.1.2.1 on frame and design issues and chapter 3.4 on the relevance of the ILO concepts of employment and unemployment⁴⁸, where actions to enhance comparability have been singled out and recommended. In other fields such as data collection, management of interviewers and fieldwork organization, treatment of non-response, it is much more difficult to improve cross-country comparability, unless significant steps towards input harmonization are made. This option (e.g. a common questionnaire) is however unrealistic, as deemed too restrictive with regard to national specificities and needs. However, the recommendations of good practices to improve for instance accuracy may also result in a higher standardization, which in turn would make national surveys more similar to each other.

In this chapter, geographical comparability is solely addressed with a view to Commission Regulation 1897/2000, as this regulation is the main tool at European level to provide the basis for comparable statistics on employment and unemployment. This regulation covers the ESS implementation of the ILO definition of unemployment and states the 12 principles for the formulation of the questions on labour status.

The Task Force takes the view that the principles should not be interpreted as rigidly as rules. What really matters is the spirit of the principles, which aims at an accurate measurement of the employed and unemployed. Should evidence emerge that questions formulated in strict accordance with the principles are not easily understood or correctly interpreted by respondents, deviations from the letter of the principles may be accepted as long as the spirit of the principles is respected. In such cases, supported solid evidence is essential. As an example, a follow-up survey carried out by Destatis found that asking the main status in contrast with the ILO status improves the accuracy of

⁴⁸ It should be noted that issues concerning comparability of some quality indicators, i.e. sampling error estimates and non-response rates, are also addressed in this report in chapter 3.2.

employment for special subgroups such as marginally employed people, whereas principle 1 stipulates to completely separate the two sets of questions⁴⁹.

In order to assess the comparability of the way information for the identification of the ILO labour status is collected, a review of national questionnaires and flowcharts⁵⁰ sketching the questions on employment and unemployment was carried out taking the 12 principles as reference. Concerning employment, all countries with few exceptions separate the questions on the ILO labour status from that on the main status. The treatment of people absent from work on the other hand is not homogeneous. Clarifications are needed at European level to define *job attachment*, in particular for groups like family workers, self-employed, people on lay-off, who are not treated exactly the same across countries with respect to the ILO labour status.

Several issues may affect the comparability of unemployment statistics: the order of questions to distinguish the unemployed from the inactive, which changes from country to country; the indication of reference periods for job search and availability to start working, which is not always clearly stated; job search methods, which are asked following several approaches; explicit mentioning that search for any kind of job should be considered, often overlooked. In addition, some of the principles on unemployment need clarification. In particular, principle 9 requires that passive job search methods are asked in addition to active ones, whereas principle 11 requires asking job search methods at least until three are mentioned. These two principles seem at odds with the definition of unemployment, which requires the use of one active method to meet the criterion of active job search. Therefore, the reasons why three active search methods should be collected and why passive methods should be asked need clarification.

One reason for the requirement of three active methods is that this would allow the further analysis of this information (e.g. improve possibilities to analyse job search patterns rather than just concentrating on verifying the ILO unemployment status). But if the aim is to assess which method is the most effective, the order with which they are listed to respondents should be randomized and in principle all methods should be asked. This can be done with computerized questionnaires. The approach followed by INE Spain, which allows the respondent to freely list the methods he/she used mentioning the others only afterwards, in fact produces a similar effect, at the same time potentially collecting the information for the whole set of methods. Generally speaking, the way job search methods are surveyed may influence unemployment estimates by differently stimulating respondents' recall and/or socially desirable answers and should therefore be harmonized. Three main approaches are currently followed in national questionnaires: a first one consists in asking a single multi-response question where all the methods are listed; a second approach foresees asking a separate question for each method; the third is the Spanish one mentioned before. These three approaches are likely to affect the estimates of unemployment differently.

As regards the passive methods, these could possibly be used to build up extended measures of labour supply. Evidence from Italy shows that the probability of transition into employment does not significantly differ between those actively searching and those using passive methods only.

⁴⁹ See Köhne-Finster and Körner, 2009.

⁵⁰ The analysis on questionnaires is presented in Zachariou and Zgierska, 2008. It considered only questionnaires that were available in English. As a consequence, it was not possible to check the situation for Denmark and Luxembourg, whereas 2006 questionnaires were used for Latvia, Finland and Romania.

Further points are also of concern and should be reviewed, such as the correct implementation of the reference period for the availability to start working and in which cases the contact with a Public Employment Office should be considered as an active search action.

Recognizing the central role of the principles for cross-country comparability, the **Task Force recommends that the 12 principles are reviewed in order to clarify specific ambiguous points. Clarifications, if necessary, should be provided as much as possible in working documents such as the explanatory notes. Amendments to the wording of the principles (necessarily via a new legal act) should be made only when strictly needed. Eurostat should set up a group of experts for this purpose (R34).** When reviewing the principles the efforts and experiences made by Member States to get in line with them should be taken into account. However, a major revision of the principles is not deemed necessary.

3.3.2. Comparability over time

Changes introduce breaks in time series, thus reducing possibilities of analysis over time. From the producer point of view, changes are burdensome, because, in addition to the ordinary working activity, they require time and extra efforts to put them in place. For these reasons, changes should be introduced only if they really lead to an improvement of the survey, either in its content or in the methodology.

In recent years several innovations have been introduced in the EU-LFS. In addition, changes have often followed a staggered implementation, as not all countries have implemented them at the same time. This creates special problems of comparability over time for the EU-LFS results, as breaks in each national survey impact on the EU aggregates.

Two issues influence comparability over time: the number and frequency of changes and their implementation patterns by Member States on the one hand, and the approach to time series breaks on the other.

As concerns the first point, a *change policy* bundling the introduction of changes at predetermined points in time would go in the direction of reducing the number of time series breaks. Efforts from Member States to implement innovations all at the same time, changing the widespread practice of derogations currently in place, would go in the same direction, at least for EU statistics. However, derogations, which lead to a staggered introduction of changes, are sometimes vital to get consensus to introduce any change at all and should be allowed if need be. Indeed, changes usually need extra resources and can only be introduced when the budget is available. In addition to this, as accuracy may be more relevant than time-series consistency, it should always be assessed whether improvements should be introduced as early as possible. In any case, the **Task Force recommends that all changes initiated either by Eurostat or by countries are adequately planned and monitored at national level in order to assess the statistical effect on time series⁵¹ (R35).** This information may then be used both to inform users or as input for back-calculation models. **It also recommends that Eurostat groups together innovations which it proposes in order to limit the number of potential breaks in time series (R36).** On his side, Eurostat committed to explore ways to bundle the introduction of changes for EU purposes.

⁵¹ Split samples are an example of assessment method.

NSIs' attitude towards breaks in time series is not homogeneous. In the Netherlands a special unit is in charge of back-recalculations, in order to produce long time-series. In the UK users are very keen on comparability over time, especially for economic analysis. In the present period of financial and economic turmoil, special attention is paid to what happened in previous periods of economic downturn. The ONS transparently communicates changes and revision to users and, for the main changes, consults with the main institutional users. They are also setting up a general model to detect discontinuities in time. In Italy back-recalculations are performed only when major changes are made to the survey, as it was the case in 2004 with the introduction of the continuous labour force survey. In Portugal the issue of time-series breaks is rather a matter of communication and dissemination policies. Rather than providing users with back-recalculated long-term time series, INE Portugal informs them about the changes that occurred. Back-recalculations are produced only when new population estimates become available for an entire decade, based on the new Census results. In Spain the main changes are reported to users through an specific body depending of the national Highest Statistical Council (working group of short-term labour statistics). When it is considered necessary the changes are also explained in media, to specialised economic journalist, etc. That was the case of the methodological changes in 2002 (unemployment definition according to Reg.1987/2000) and in 2005 (new population estimates, implementation of CATI and wave approach).

The Task Force recognises the importance of long homogeneous time series to users. For this reason, **it recommends to produce and disseminate consistent time series at least for the headline employment and unemployment indicators (R37)..**

As concerns back-recalculations, a first best would be to operate them at micro data level, as it is the case in UK for the LFS for data going back to 1992. This would ensure full consistency between micro and macro data for all variables in the datasets and all breakdowns. In Spain, before publishing the results of the first quarter 2005 with the new methodology, new series and microdata files were produced with the new population figures for the period 1996-2004 in order to use the same population basis for weighting the new one for 2005. At the same time, when the results of the first quarter 2005 were disseminated, estimates of the impact of the new methodology were provided (estimates produced with the new and the old method). As this is not always possible, some countries disseminate consistent time series at macro level. This approach raises several issues, such as the choice of indicators and breakdowns to recalculate and the inconsistencies between those and the ones obtained from the non-adjusted micro data. A further option to deal with breaks in time series consists in providing users with a general model for back-recalculation. However, users are not keen on making calculations themselves and are rather interested in data ready to use. Special care should be devoted to the dissemination of back-recalculated data, as these may confound users of the original data. For this reason, NSIs should ensure that only one official series for each aggregate/indicator is disseminated.

3.4. Relevance of the ILO concepts of employment and unemployment⁵²

According to the agreed definition within the European Statistical System, relevance is understood as “the degree to which statistics meet current and potential user needs. It refers to whether all statistics that are needed are produced and to the extent to which concepts (definitions, classifications, etc.) reflect user needs.” (Eurostat, 2003). This definition implies that relevance is always a relative notion. The LFS is a survey with a large variety of users, who at least in part have diverging user needs. Hence an assessment of the relevance of the labour force concept is only possible in relation to certain user needs.

In the field of employment and unemployment, the labour force concept of the ILO⁵³ is the central conceptual backbone of the EU Labour Force Survey. The relevance of this concept is therefore a key factor for the relevance of the LFS results. The ILO labour force concept is essential for international comparisons, for which there is currently no alternative. In addition, the notion of *work* at the basis of the distinction between those “economically active” (or in the labour force), i.e. employed and unemployed, and those who are considered “economically inactive” is in close connection with the System of National Accounts (SNA) production boundary. In this view, employment to be measured is that underlying GDP. A main implication of such a connection is that all persons who provide the supply of labour for the production of goods and services have to be included, even if the input to the production is as small as one hour per week. For the same reason, the one-hour-criterion is a prerequisite for the consistency of the LFS with National Accounts data on production, as input and output of economic production have to be consistent regarding their scope. The labour force concept is therefore highly relevant, and indispensable, for National Accounts and economists who require a complete coverage of even the smallest input of labour. Therefore, there is no need for a revision of the ILO labour force concept when it is looked at from an economic perspective or when it is considered for international comparability.

A minor issue concerning the consistency with the SNA as well as with the European System of Accounts (ESA) is given by the coverage by the ILO definition of employment of persons with a job but not at work in the reference period. Indeed, these do not contribute to GDP. On the other hand, they continue to receive (parts) of their wages which should be reflected as well. If such a period of non-productivity exists for a longer period, one might argue that such persons should rather be considered as inactive rather than employed. It might be further discussed, also in cooperation with National Accounts, whether the definition of the persons with a job but not at work could be improved.

However, there is a point to make concerning the ILO definition of unemployment. It intends to capture only a restricted part of the whole labour reserve, i.e. the one showing a strong attachment to the labour market. It is not meant to measure the entire labour reserve. The full extent of the labour reserve is difficult to quantify and will generally tend to be significantly bigger than the category of the unemployed according to the labour force concept. Persons without work who are not classified as unemployed can still be potential workers. Research has shown that the transition

⁵² The present chapter is based on the document "*Relevance: summary of TF members' views*", prepared by Körner, 2008.

⁵³ See ILO, 1982.

from a state outside the labour force into employment can be substantial⁵⁴. The need for supplementary indicators is therefore recognized⁵⁵.

Another weakness of the ILO definition of unemployment is its limitation as an indicator for developing and evaluating labour market policies. The main problem is that a substantial number of persons who are not classified as unemployed are nevertheless of high interest for labour market policy. Persons without work that are part of the claimant population or those registered at a Public Employment Agency belong to the target population of labour market policies. They are important independently of their availability or search activities. The ironic situation could be that they are not available or not seeking work because they are in a (training) programme in order to get them to work. The restrictiveness of the criteria to be considered unemployed may also turn to be misleading when the ILO concept is used as a reference for funds distribution to disadvantaged areas. As the borderline between unemployment and inactivity is thin, areas where there are many discouraged workers may not meet the requirements if the eligibility criteria are fixed in terms of unemployment rates.

The concept is also less useful for analysing labour market behaviour from a social point of view. The ILO categories only partially reflect individuals' situation on the labour market as they perceive it. This can be seen when contrasting the ILO status with the main status or the own perception of the respondents as to his/her labour status. Indeed, as table 2 shows, a substantial part, almost 3 out of 10, of persons that consider themselves as unemployed are inactive according to the ILO concept.

Table 2 – Distribution of persons aged 15-64 by labour force status for a given main status (%). EU-LFS 2005. EU-25 without Germany, Poland and the United Kingdom

Main status (self declared)	Status according to the ILO Labour Force Concept		
	Employed	Unemployed	Inactive
Employed	99,7	0,1	0,3
Unemployed	4,6	68,0	27,5
Student	9,3	2,0	88,8
Other inactive	6,4	2,0	91,7

Source: Eurostat, 2007, p. 120

In order to capture these nuances the statistics based on the ILO labour force concept should therefore be complemented by a range of statistics, such as full-time/part-time employment, self-declared main status, labour reserve and time related underemployment. Such indicators can be obtained from the information already available in the EU-LFS. **The Task Force recommends**

⁵⁴ See for instance Jones and Riddell, 1999, Brandolini, Cipollone and Viviano, 2004, Garrido and Toharia, 2004, and Bradbury, 2006.

⁵⁵ An ILO proposal for indicators complementing the ILO unemployment rate was recently discussed by the 18th ICLS (Geneva, December 2008). Suggestions from the Conference were to carry on the work on the development of such indicators (ILO, 2008).

therefore to define a set of common indicators supplementing the employment and unemployment statistics based on the ILO concept. Eurostat should initiate a Task Force for this purpose⁵⁶ (R38). Among supplementary indicators for the ILO unemployment, the main status as perceived by respondents themselves provides basic information to supplement ILO concepts. This variable is also one of the so-called Core Social Variables to be implemented in all social surveys of the European Statistical System⁵⁷. **The Task Force recommends therefore to remove the optional status, in the LFS Regulation, of the variable "Main Status as perceived by respondents" in the EU-LFS (R39).**

The Task Force underlines the increasing importance of a description of the dynamics of the labour market, capturing the movements in and out of unemployment, employment and inactivity. The importance of having better information on the flows (in addition to the stocks) has recently been stressed by the United Nations Statistical Commission and in the Cracow Action Plan of the European Statistical System (e.g. ONS UK, 2008, pp. 28-31; Eurostat 2009; UN Statistical Commission 2009). As a consequence, the LFS might need to be adapted to enable the analysis of dynamics on a comparable basis. This, however, is possible within the labour force concept and would rather require a broader use and better harmonisation of rotation panel designs in the EU-LFS. Given its potential for the analysis of labour market dynamics, **the Task Force recommends to use sample designs with intra-annual rotation patterns to allow calculation of quarter-to-quarter labour market flows estimates (R6). The Task Force also recommends to examine options for a further harmonization of the rotational patterns to enhance comparable longitudinal analysis at European level (R7).**

The conclusion of the Task Force is that the ILO definitions of employment and unemployment continue to be appropriate for economic analysis as well as for international comparisons. In addition, the one hour criterion used to define employment provides conceptual consistency between labour market statistics and National Accounts. On the other hand, the ILO concepts are less suitable for developing labour market policy and social labour market analysis. For these purposes additional indicators are needed. In spite of such limitations, the ILO labour force concept can be considered highly relevant for users sharing its economic perspective. However, it is important that this perspective is clearly communicated to users, in order to provide unambiguous guidance for the use of statistics based on this concept and prevent wrong interpretations.

3.5. Timeliness and punctuality

3.5.1. Timeliness

Timeliness is the time lag between the end of the reference period and the dissemination of results. This quality dimension is not considered a major concern by the Task Force. However, its importance is growing. With the dissemination of quarterly results the LFS has more and more become a short-term statistics, with the consequent widening of its audience. At this regard, users interested in short-term economic analysis, as well as internal users from National Accounts, push

⁵⁶ The Task Force on supplementary indicators was established by the LAMAS WG in September 2008 and started work in February 2009.

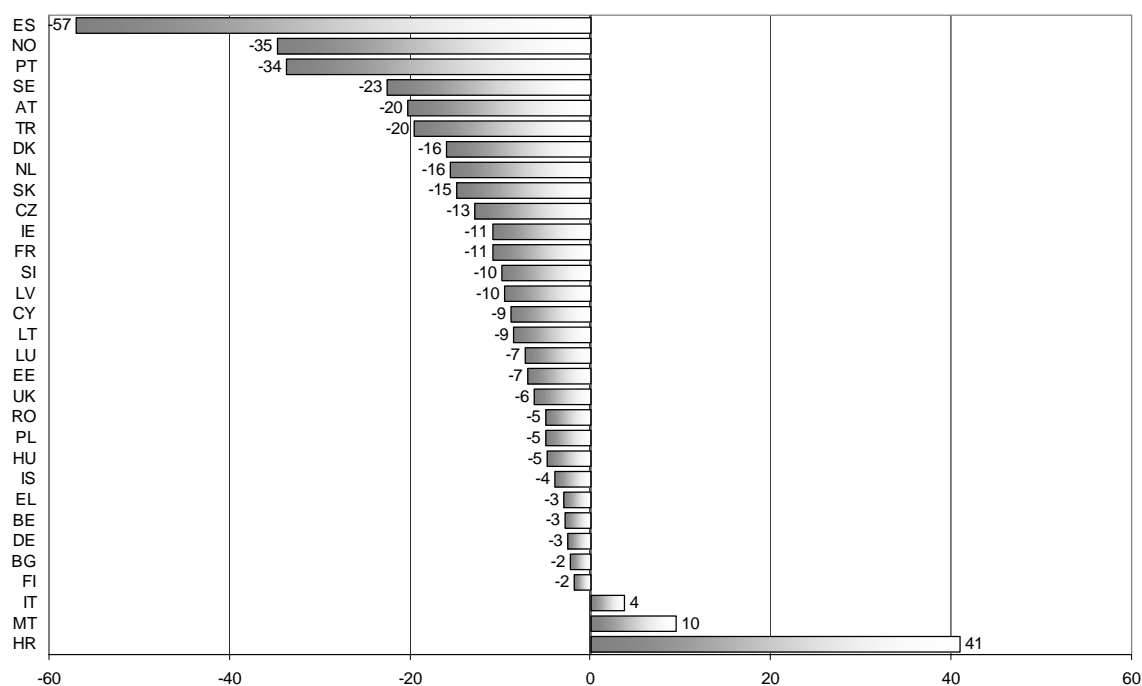
⁵⁷ See Eurostat, 2007.

for a reduction of the dissemination lag measured from the end of the reference periods, especially for headline labour market indicators. In this sense, improving timeliness would also increase the relevance of the LFS.

However, a trade off between timeliness and other quality dimensions, in particular accuracy, is widely recognized. The Task Force therefore agrees that an improvement in timeliness should never be achieved at the expense of a significant loss in data accuracy. At the same time, improvements of timeliness that do not adversely affect accuracy should of course be implemented without delay.

At present, the LFS regulatory framework foresees data transmission to Eurostat within 12 weeks after the end of the reference period. However, a few countries, of which Spain is the most prominent example, regularly send good quality data (and publish them at national level) largely ahead of the deadline (see figure 5). This result is achieved through a good general organization of the survey, from preparatory tasks, to data collection, processing, calculation of weighting factors.

Figure 5 – First data transmission: average days before/after the deadline (Average 2008)



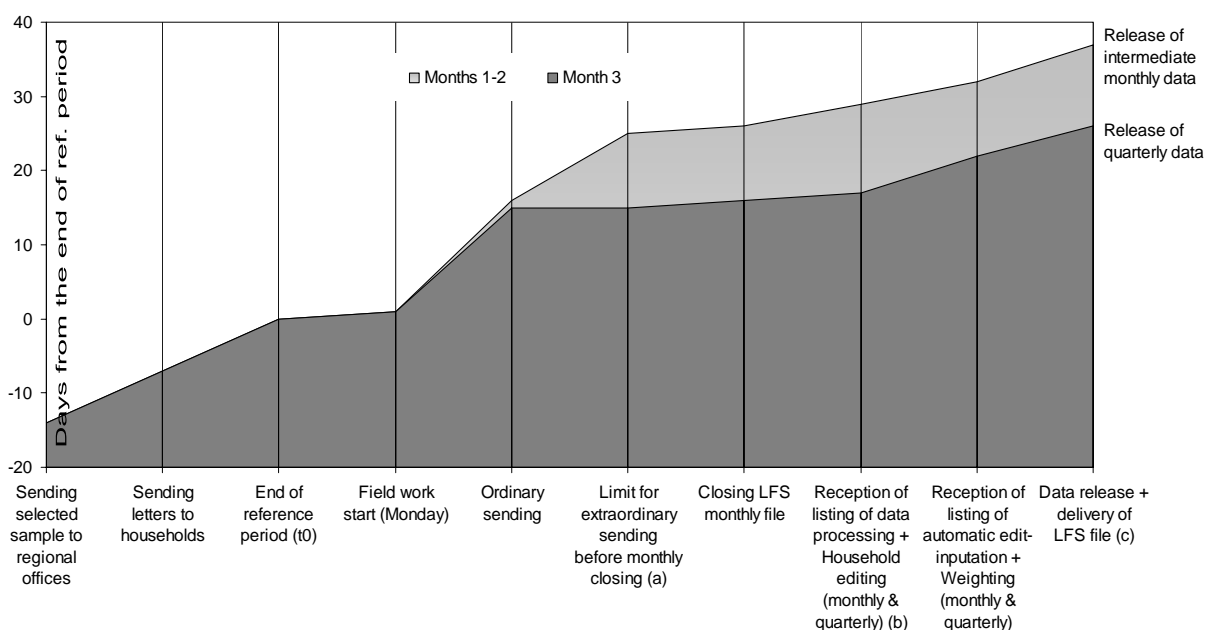
Source: Eurostat, LFS.

Several elements allow the Spanish LFS being so timely: a clear commitment to do it, originating from strong user needs, which gives the LFS a high priority within INE's activities; the organization of the field work (among other features, on average about 74% of interviews are carried out the week after the reference week and 94% within three weeks); the automation of procedures; the coordination with other units providing external input (e.g. population figures for weighting); a release calendar. A minor drawback is that interviewing time is reduced in the two last reference

weeks of each quarter, which sacrifices some response and produces a (slightly) uneven distribution of interviews by reference week. In addition, high automation, essential for timeliness, involves high complexity which makes changes difficult.

Figure 6 sketches the main activities in the Spanish LFS organization. The list of sampled units is sent to regional offices two weeks before the beginning of the field work. During the following week sampled households receive presentation letters from INE. The field work then starts the day after the end of the reference period. Achieved interviews are sent in electronic format to the central office at the end of the second week of field work. Interviews carried out later than two weeks after the end of the reference period are still accepted if sent by 25 days after the end of the reference period for the last week of the month. This "extraordinary sending", however, is not allowed for the third month of the quarter, coinciding with the quarterly release. Data processing and editing as well as the calculation of weighting factors is then performed on a monthly basis, ending 32 and 22 days after the reference period respectively for the first two months of each quarter and the third one. Data are then released around 5 days later in both cases.

Figure 6 – Calendar of activities for the Spanish LFS



Notes: (a) Extraordinary sending allowed for all weeks until the end of the reference month. Time lag refers to the last week. (b) This operation concerns all the weeks of a month. Time lag refers to the last week. (c) From the last reference week of the month. Cumulated moving sample of three months, coinciding with the quarterly results for month 3.

Source: INE Spain.

The Spanish case shows that it is possible to improve timeliness without significantly reducing quality. However, some national specificities may prevent single countries from achieving such a goal. In Italy, for instance, both interviewers' contractual arrangements and the processing after data collection prevent the LFS from being timelier. Istat reports that, as the interviewers are contractually allowed to carry out interviews up to five weeks after the reference week, there is no

means to force them to be faster. In addition, although consistency checks are implemented directly in the software, lengthy *ad hoc* processing is needed each quarter after the end of data collection. What is more, some countries still use paper questionnaires, which reduces timeliness because both data entry has to be performed at the end of data collection and heavier control and processing is needed.

A general feeling is that most countries could streamline their organization in order to improve timeliness with no or negligible loss in accuracy. **In general, the Task Force recommends to improve significantly the timeliness of the EU LFS in order to further enhance its relevance for short term economic analysis (R40).** In particular, the ILO labour status by the main breakdowns (e.g. sex and age) could be produced much earlier than the official deadline of 12 weeks after the end of the reference period, for instance because data editing and processing often concern other variables (mainly classifications or household variables). This would allow improving the timeliness of the headline labour market indicators, thus enlarging the use of the LFS and increasing its relevance. At this regard, the Task Force supports Eurostat intention to explore Member States' possibility to anticipate the delivery to Eurostat at least of the ILO labour status aggregates by the main breakdowns.

3.5.2. Punctuality

The first data transmissions to Eurostat in most (but not all) cases take place by the 12 weeks deadline of the LFS Regulation. However, the first transmissions are only a first step towards data dissemination. As a matter of fact, after receiving national datasets Eurostat processes them and asks NSIs to validate some main results within two weeks. Sometimes several iterations of transmission and processing are needed to get a file ready for release. In practice, it is not possible to know beforehand when each country's data will be ready for publication. In addition, given that Eurostat needs data from all the Member States to calculate and disseminate EU aggregates, with the present organization this takes place at the pace of the slowest NSI.

A further consequence of the present situation is that detailed tables for timelier countries are disseminated by Eurostat earlier than the main results, what contradicts the relative importance of the two data sets. In order to avoid this Eurostat may ask countries to deliver the main variables at an early stage. This has drawbacks, as it would lead to a double data transmission, i.e. a double burden both for NSIs and Eurostat. In addition, at least for some countries this may not be feasible, either because all variables in the dataset become available at the same time or because they are disseminated nationally all together. However it is an option that should be considered. The possibility for countries to send some data to Eurostat under embargo before national publication should also be explored, as it would considerably facilitate Eurostat's work.

The present transmission/validation procedures do not allow Eurostat to have a release calendar, whereas several national Labour Force Surveys have one. The Task Force underlines that a release calendar is standard good practice for short-term statistics and required by the European Statistics Code of Practice (principle 6). This would furthermore increase the EU-LFS visibility and users' satisfaction, by giving them certainty about when they will have the information they need. A release calendar is useful also for data producers. As the moment of data release can have an impact on politics, a release calendar would shield from political influence, which in turn would increase

the credibility of statistics. It would also influence the priority set for data release from NSIs. For these reasons, **the Task Force recommends to establish a release calendar for the EU-LFS (R41)**. At least at a first stage, the release calendar should have conservative dates in order to ensure availability of all Member States' data.

In order to allow an EU-LFS release calendar some adjustments in the validation procedures would be needed beforehand. On its side, Eurostat may distribute its consistency check programmes and the programmes to calculate main indicators to NSIs, so that they can send pre-validated data. It should also work on minimising the time needed to process datasets received from countries and send results back for validation. NSIs, on their side, should not wait for the end of the two weeks for validating data but should do it as soon as possible. Rules allowing Eurostat to publish non-validated data as provisional and/or use them for the calculation of European aggregates should also be set up.

The Task Force reminds that the deadline set in the Regulation refers to the final transmission date, i.e. allowing data dissemination without further retransmissions. **It therefore recommends to comply with the twelve-week deadline in the Regulation as the one for final, not first, data transmission (R42)**. It also recommends that Eurostat make available to NSIs its validation programmes used to process and control national datasets in order to allow a complete check prior to the data delivery and avoid time-consuming retransmissions (R43).

Cited references

- American Association for Public Opinion Research, *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, 5th edition. Lenexa, Kansas: AAPOR, 2008.
- Banda J.P., *Nonsampling errors in surveys*, United Nations Secretariat – Statistical division, Expert Group Meeting to Review the Draft Handbook on Designing of Household Sample Surveys, 3-5, December 2003, at http://unstats.un.org/unsd/demographic/meetings/egm/Sampling_1203/docs/no_7.pdf.
- Bethlehem J., *Weighting nonresponse adjustments based on auxiliary information*, in Robert M. Grooves, Don A. Dillman, John L. Eltinge and Roderick J.A. Little (eds.) *Survey Nonresponse*, New York, Wiley, pp. 265-287, 2002.
- Bethlehem J., Cobben F. and Schouten B., *Nonresponse in Household Surveys*, Statistics Netherlands – Methodology Department, September 2007.
- Bradbury K., *Measurement of Unemployment*, Federal Reserve Bank of Boston Public Policy Briefs No. 06-2, 2006.
- Brandolini A., Cipollone P. and Viviano E., 2004, *Does the ILO definition capture all unemployment?*, in Temi di discussione del Servizio Studi 529, Roma, Banca d'Italia, 2004.
- Bureau of Labour Statistics, *Current Population Survey, Design and Methodology*, Technical Paper 63RV, March 2002.
- Campanelli P., Sturgis P and Purdon S., *Can you hear me knocking: an investigation into the impact of interviewers on survey response rates*, The Survey Method Centre at SCPR, London, 1997.
- Council of the European Union, *Council Regulation (EC) No 577/98 of 9 March 1998 on the organisation of a labour force sample survey in the Community*, Brussels, 9 March 1998.
- Discenza A., *Frame and design issues: summary of TF members' views*, Doc.:Eurostat/F2/EMPL/03/08, March 2008, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2008/quality_meeting_06-07/documents&vm=detailed&sb=Title
- Durier S. and Thélot H., *The non-response survey of the French LFS*, Doc.:Eurostat/F2/EMPL/03/08 Annex_II, March 2008, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2008/quality_meeting_06-07/documents&vm=detailed&sb=Title
- Elliot D., Rainford L. and Eldridge J., *Interview Mode Effects in the UK Local Labour Force Survey*, in Survey Methods Bulletin, 57, pp 50 - 60, Office for National Statistics, 2006.
- European Commission, *Commission Regulation (EC) No 430/2005 of 15 March 2005 implementing Council Regulation (EC) No 577/98 on the organisation of a labour force sample survey in the Community concerning the codification to be used for data transmission from 2006 onwards and the use of a sub-sample for the collection of data on structural variables*, OJ No L 71/36, Brussels, 15 March 2005.
- European Parliament and Council of the European Union, *Regulation (EC) No 763/2008 of 9 July 2008 on population and housing censuses*, Strasbourg, 9 July 2008.
- European Parliament and Council of the European Union, *Regulation (EC) No 223/2009 of 11 March 2009 on European Statistics*, Strasbourg, 11 March 2009.

- Eurostat – Working Group on Quality in Statistics, *Definition of Quality in Statistics*, Document Doc. Eurostat/A4/Quality/03/General/Definition, Luxembourg, October 2003, at <http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/ess%20quality%20definition.pdf>
- Eurostat, *European Statistics Code of Practice for the national and community statistical authorities*, Luxembourg, 2005, at http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/code_practice.pdf
- Eurostat, *Report from the Commission to the Council and the European Parliament on the implementation of Council Regulation (EC) No 577/98*, 2006 (a).
- Eurostat – Task Force on Revision of the LFS explanatory notes, *The precision requirements in the EU-LFS*, Doc.:Eurostat/F2/EMPL/15/2006, Luxembourg, 2006 (b), at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2006/explanatory_19-20062006/documents&vm=detailed&sb=Title
- Eurostat, *Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System*, Luxembourg 2006 (c), at <http://epp.eurostat.ec.europa.eu/portal/page/portal/quality/documents/RPSQDET27062006.pdf>
- Eurostat – Task Force on Core Social Variables, *Final report*, Luxembourg, 2007.
- Eurostat – Task Force on the Quality of the Labour Force Survey, *Outline of the final report – Revised*, Doc.:Eurostat/F2/EMPL/13/07, Luxembourg, October 2007, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2007/quality_18-19_october/documents&vm=detailed&sb=Title
- Eurostat – Task Force on the Quality of the Labour Force Survey, *Draft of the working program – Revised*, Doc.:Eurostat/F2/EMPL/14/07, Luxembourg, October 2007, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2007/quality_18-19_october/documents&vm=detailed&sb=Title
- Eurostat – *Final Report of the Task Force on Statistical Challenges*, Luxembourg 2009.
- Furubjelke, E. and A. D. Rackner, *Monitoring as a Method for Evaluation and Development of Survey Questionnaires*, paper presented at the LFS Workshop on Quality Assessment, Athens 23-24 October 2008.
- García Martínez M.A., *The wave approach in the Spanish LFS*, paper prepared for the Task Force on the Quality of the Labour Force Survey, mimeo, 2008 (a).
- García Martínez M.A., *The interview mode in the LFS. The recent experience from the Spanish LFS*, paper prepared for the Task Force on the Quality of the Labour Force Survey, mimeo, 2008 (b).
- Garrido L. and Toharia L., *What does it take to be (counted as) unemployed? The case of Spain*, in *Labour Economics*, 27, pp. 329-361, 2004.
- Gatward R., Lound C. and Bowman J., *A pilot survey of people living in communal establishments*, in *Labour Market Trends*, pp. 141-149, March 2002.
- Groves R.M. and Cooper M.P., *Nonresponse in household interview survey*, Wiley, New York, 1998.

- International Labour Organization, *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment*, Adopted by the 13th International Conference of Labour Statisticians, Geneva, 1982, at <http://www.ilo.org/public/english/bureau/stat/download/res/ecacpop.pdf>
- International Labour Organization, *Report of the Conference*, 18th International Conference of Labour Statisticians, Geneva, 2008, at http://www.ilo.org/wcmsp5/groups/public/---dgreports/---integration/---stat/documents/meetingdocument/wcms_101467.pdf
- Jackle A. and Lynn P., *Dependent Interviewing and Seam Effects in Work History Data*, in *Journal of Official Statistics*, Vol.23, No. 4, pp. 529-551, 2007.
- Jones S.R.G. and Riddell W.C., *The Measurement of Unemployment: an Empirical Approach*, in *Econometrica*, 67-1, pp. 147-162, 1999.
- Kersten H.M.P. and Bethlehem J.G., *Exploring and reducing the nonresponse bias by asking the basic question*, in *Statistical Journal of the U.N. Economic Commission for Europe* 2, pp. 369-380, 1984.
- Kish, L., *Survey Sampling*, Wiley, New York , 1965.
- Kleven, Ø., Lagerstrøm B.O. and Thomsen, I., *A Simple Model for Studying the Effects of Proxy Interviewing: The Norwegian Labour Force Survey as a Case Study*, *Statistics Norway* 2008, at http://www.ssb.no/emner/06/90/doc_200802/doc_200802.pdf
- Köhne-Finster, S. and Körner, T., *Microcensus Follow-up Survey – A Project of the Federation and the Länder to Measure the ILO Employment Status*, in: *Methods, Approaches, Developments* 1/2009, at <http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/EN/Content/Wissenschaftsforum/MethodsApproaches/Infos/MethodsApproachesDevelopment,templateId=renderPrint.psm1>
- Körner T., *Relevance: summary of TF members' views*, Doc.: Eurostat/F2/EMPL/05/08, March 2008, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2008/quality_meeting_06-07/documents&vm=detailed&sb=Title
- Kösters L., *The wave approach in the LFS: good practice*, 3rd Workshop on LFS Methodology, Stockholm, May 2008, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/workshops_methodology/stockholm_2008/documents&vm=detailed&sb=Title.
- Lin I. and Schaeffer N.C., *Using survey participants to estimate the impact of nonparticipation*, in *Public Opinion Quarterly*, Vol. 59, pp. 236-258, 1995.
- Merkle D.A. and Edelman M., *Nonresponse in exit polls: a comprehensive analysis*, in Robert M. Grooves, Don A. Dillman, John L. Eltinge and Roderick J.A. Little (eds.) *Survey Nonresponse*, New York, Wiley, pp. 243-258, 2002.
- Office for National Statistics (UK), *Report of the Office for National Statistics, United Kingdom on Labour Statistics*, Meeting document for the 39th session of the UN Statistical Commission, 28-31 February 2008.
- Pascale, J. and A. McGee, *Using Behaviour Coding to Evaluate the Effectiveness of Dependent Interviewing*, in *Survey Methodology* 34, 2008, pp. 143-151.
- Prestwood D., *Use of incentives to reduce non-response in UK*, Doc.: Eurostat/F2/EMPL/03/08 Annex_I, March 2008, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2008/quality_meeting_06-07/documents&vm=detailed&sb=Title

- Quantos SARL, *Report on Variance Estimation Methods in EU*, Version 3, 24 July 2002, available from the “Quality in Statistics” CIRCA Interest Group at <http://forum.europa.eu.int/Public/irc/dsis/Home/main>
- Sezer D., *Quality Control Study In Turkish Labour Force Survey*, presentation at the LFS Workshop on Quality Assessment, Athens 23-24 October 2008.
- Stadler, B., *The Analysis of Time Protocols of the Austrian LFS*, presentation at the LFS Workshop on Quality Assessment, Athens 23-24 October 2008.
- Stoop I. A.L., *The hunt for the last respondent – Nonresponse in sample surveys*, Social and Cultural Planning Office of the Netherlands, The Hague, May 2005.
- Sundgren B., *What is a Public information system?*, in *International Journal of Public Information Systems*, vol. 2005:1, 2005.
- Thomas M., *Survey Mode Effects: Sources, Evidence and Questions for Discussion*, 3rd Workshop on LFS Methodology, Stockholm, May 2008, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/workshops_methodology/stoekholm_2008/documents&vm=detailed&sb=Title.
- Thomsen, I. and Villund, O., *Using administrative registers to evaluate the effects of proxy interviews in the Norwegian Labour Force Survey*, paper presented at the IAOS Conference on Reshaping Official Statistics, Shanghai, China, 14-16 October 2008. At http://www.stats.gov.cn/english/specialtopics/iaos/Papers/CS_28_4_Thomsen_et_al.doc
- Torres S., *The control of interviewers in the Portuguese Labour Force Survey by using a Quality Survey*, Doc.:Eurostat/F2/EMPL/12/08, September 2008, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2008/meeting_september_2008&vm=detailed&sb=Title
- UNECE and Eurostat, *Recommendations for the 2010 Censuses of Population and Housing*, Conference of European Statisticians, Geneva 2006. at <http://www.unece.org/stats/documents/ece/ces/ge.41/2006/zip.1.e.pdf>
- UN Statistics Commission, *Report of the International Labour Organization on labour statistics and on the outcomes of the Eighteenth International Conference of Labour Statisticians*, 40th Session of the UN Statistical Commission, 24 - 27 February 2009, at <http://unstats.un.org/unsd/statcom/doc09/2009-12-EmploymentStats-E.pdf>
- Zachariou S., Zgierska A., *Revised summary of national questionnaires showing ILO employment and unemployment status*, Doc.:Eurostat/F2/EMPL/14/08, at http://circa.europa.eu/Members/irc/dsis/employ/library?l=/working_group/meetings/2008/meeting_september_2008&vm=detailed&sb=Title

If the LFS is not the main source for NA employment estimates, try nevertheless to fill in the table by using available sources. Information should refer to the most recent year with final National Accounts data, unless another year is preferred for particular reasons. Please change the prefilled figures as well as the reference year if more up to date figures are available. Information should be provided in thousands, not in percentage.

Please fill in only yellow cells.

ANNUAL RESULTS				Reference Year	
(1)	(2)	(3)	(4)	(5)	(6)
	Sign (a)	Thousands (b)		Source used	Notes / description
A)		18,510	NA employment - domestic concept (in persons)		
B)	+		Residents working outside the economic territory		
C)	-		Non residents working inside the economic territory		
D)	=	18,494	NA employment - national concept (in persons)		
E)	-		Conscripts		
F)	-		Employed persons living in collective households (if not included in LFS figures)		
G)	-		Unpaid trainees and apprentices (if not included in LFS figures)		
H)	-		Farmers only producing for own-consumption (if not included in LFS figures)		
I)	-		Employed persons aged less than 15		
J)	+/-		Other discrepancies in definitions, concepts, coverage (if applicable, please specify)		
K)	=		Difference due to definitions, concepts, coverage $K=\text{SUM}(B:C,E:J)$		
L)	+/-		Difference between source used in NA and LFS for specific economic activities (e.g. agriculture, public administration, etc.) (if applicable, please specify and add lines needed)		
M)	+/-		Adjustments for hidden/undeclared employment <u>not included in L</u> (if applicable)		
N)	+/-		Other adjustments (if applicable, please specify and add lines needed)		
O)	+/-		Residual		
P)	=		Difference due to inconsistencies among different sources $P=\text{SUM}(L:O)$		
Q)	=	17,866	LFS (c) employment (in persons) $Q=A+/-K+/-P$		
General comments on filling in the table					

(a) Please replace +/- with + (-) if, to arrive to LFS estimate, the adjustment must be added (subtracted) to NA

(b) Please fill in the cells in this column with absolute values only. Provide the appropriate sign of each figure in column B on the left

(c) Please indicate in col. (6) if the LFS figures are adjusted to some extent to fit (better) with NA figures