



9th Workshop on LFS methodology, Rome, 15-16 May 2014

The LFS review in the context of the Eurostat programme for modernising social micro data collections

Anne Clémenceau (Eurostat)

1. Introduction

In line with the orientations endorsed by DGINS in the Wiesbaden in September 2011 and with Communication COM (2009) 404 on “The production method of EU statistics: a vision for the next decade”, Eurostat has been working since 2009 on the modernisation of social statistics. This modernisation strategy includes three main components:

- Designing an integrated system of social micro-data collections including work on standardisation of social variables and statistical frames,
- Modernising population statistics (including demography, migration and population census)
- Streamlining (mainly) administrative data collections in the fields of labour statistics, health statistics, social protection, education statistics, crime and migration¹.

The strategy does not cover demographic projections (entirely under Eurostat responsibility) and enterprises surveys (Structure of Earnings Survey, the Labour cost Survey, the surveys underlying the labour cost index and the job vacancy statistics and the Continuing Vocational Training Survey) as plan is to integrate them in the future in the similar exercise carried out for business statistics (FRIBS project).

The main objective is to increase efficiency and responsiveness to users’ needs while learning from the past, pushing strongly towards harmonisation of variables and modules, and establishing concrete and realistic scenarios.

1

- Labour statistics (Minimum wages and labour market policy statistics)
- Health statistics (causes of death, health care including SHA – System of Health Accounts, morbidity statistics (in planning), accidents at work (ESAW) and occupational diseases (EODS))
- Social protection (ESSPROS – European System of Social Protection);
- Education statistics (education and training systems, participation of adults in lifelong learning, CVTS and in the future possible Education Accounts)
- Crime (police-reported crime, number of prisoners, new crime types (money laundering, human trafficking,...))
- Migration (asylum, residence permits, enforcement of immigration legislation)

This paper will describe the Eurostat plans for modernising social micro-data collections in general (including ideas about the new legal framework and timing for the exercise) and the link between the modernisation strategy and the LFS review.

2. Background

There are currently eight European social data collections for which micro-data is transmitted to Eurostat:

- Labour Force Survey (LFS),
- Statistics on Income and Living Conditions (SILC),
- Information and Communication Technology (ICT) survey,
- Adult Education Survey (AES),
- European Health Interview Survey (EHIS),
- European Health and Social Integration Survey (EHSIS),
- Household Budget Survey (HBS),
- Time Use Survey (TUS)

These data collections are mainly based on survey information. For each of them, the EU requirements consist of a list of variables to be obtained for each statistical unit of a sample of a reference population. They are consequently all based on output harmonisation where the list of variables to be transmitted to Eurostat is defined. A model questionnaire is however developed for some of the data collections (especially for the LFS and SILC ad-hoc modules). They are accompanied by precision requirements and a set of methodological conditions to be followed by the sampling design or the survey vehicle used at national level to collect the data. Standardised concepts and definitions have also to be followed.

These eight social micro-data collections differ widely in terms of:

- Seniority: while LFS and HBS data collections started in the 1980's, it is only from the middle of 2000's that the others emerged
- Frequency: annual or infra-annual collections for LFS, SILC and ICT, every 4/5 years for AES and EHIS, every 10 years for TUS while only one wave for EHSIS has been implemented so far
- Legal framework: most of them are based on EP/Council regulations; three (HBS, TUS and EHSIS) are however implemented under a Gentlemen's Agreement
- Degree of harmonisation: is clearly more limited when the data collection is based on a Gentlemen's Agreement
- Collection unit: most of them have individuals as collection units but LFS and SILC mainly rely on a household collection unit
- For all these data collections, the main data source is a survey but is combined with administrative information. The use of the administrative data is more extensive in some instruments than in others.

3. The LFS review and the modernisation of social micro-data collections programme: complementary and fully integrated exercises

The eight data collections have been developed independently to answer specific policy needs in the fields of labour market, social exclusion, health or lifelong learning.

For the first time in the context of the modernisation of social statistics, Eurostat made a thorough evaluation of the system for two different angles. First, the content and structure of each individual data collection have been thoroughly reviewed, especially for the two main pillars of the system, LFS and SILC – the LFS and SILC review exercises. But in addition, an assessment of the overall system of social micro-data collections – the modernisation of social statistics exercise - took place. Work focussed on the streamlining and integration of all these data collections with the aim to increase efficiency and responsiveness to users' needs.

These two exercises are complementary and fully integrated and consistent as will be highlighted below in the case of the LFS review.

3.1. The Eurostat modernisation programme

The Eurostat modernisation programme covers all eight social micro-data collections and comprises six main components:

A. Review of the content of the social micro-data collections

The content of all surveys has been scrutinised in view of current and future users' needs: some variables have been dropped; others were introduced in order to answer new users' needs (fulfilling the objective of increasing the responsiveness of the system of surveys). The technical characteristics associated with each (group of variables) have also been rethought including Accuracy/Sample size requirements; Reference period; Frequency; Target population (including filtering); and Crossing needs.

As an example, the review of the content of the LFS, both as regards the list of variables in the core of the instrument and the system of ad-hoc modules, results in the following:

- Selection of candidates for new labour market variables (e.g. on quality of employment); self-perceived main activity status turned into compulsory variable; some variables identified for deletion, e.g. situation one year before.
- Adaptation of educational variables and additional variables focused on a 12-month reference period and coverage of non-formal education and training.
- Possible inclusion of further core variables (household income, health): feasibility tests in 2014.
- Evaluation and review of the system of ad-hoc modules: these modules will be called SAM (supplementary annual modules) in future.

B. Modularisation of the European social surveys

For each survey, the variables have been grouped into modules, i.e. a module comprises a number of variables on the same topic and which have exactly the same characteristics mainly same sample size requirement, same frequency and target population.

For the LFS, preliminary result of the modularisation exercise is shown in table 1. The LFS (core and AHMs) has been partitioned into 22 modules and their periodicity (quarterly, annual, biennial) defined. This list is still work in progress under discussion in the LAMAS.

Table 1

Possible modules for the LFS Version April 2014			
Modules and sub modules	Target group	Freq	Vars
1 Technical items	All		
2 Person characteristics	All	Q	7
3 Household characteristics	All	A	5
4 Socio economic background	Age>14	A	2
5 Employment status	Age>14	Q	4
6 Job characteristics	Main job	Q	7
7 Working hours	Main job	Q	5
8 Occupation	Main job	Q	2
9 Characteristics of the workplace	Main job	Q	5
10 Working time arrangements	Main job	2A	6
11 Start of job	Main job	Q	3
12 Second job	Working	A	4
13 Wished hours	Working	Q	5
14 Looking for other job	Working	A	3
15 Previous work experience	Not working	A	7
16 Looking for work	Not working	Q	11
17 Educational attainment	Age > 14	Q	5
18 Participation in education and training (4 weeks)	Age > 14	Q	2
19 Secondary information on participation	Age 15-34	A	3
20 Participation in education and training (12 months)	Age 18-69	2A	7
21 Health module	All	2A	2
22 Supplementary Annual Module	Variable	A	11

C. Standardisation of variables

Another essential component of the strategy, the standardisation of variables, takes place in parallel to the domain specific revisions.

Currently, standardisation focuses on 16 'core social variables'², i.e. variables to be included in all social micro-data collections. Further work is however required as the list is to be reviewed, a thorough evaluation of the explanatory notes is required and more standardisation is needed given that substantial differences remain in the way they are implemented in the different surveys/countries.

The list of variables to be standardised has been extended and work is in progress. This exercise has started with Eurostat internal discussion which will get soon extended to the working groups, thus involving the Member States. Not only core social variables are covered by the standardisation

² Sex; Age in completed years; Country of birth; Country of citizenship at time of data collection; Legal marital status; De facto marital status; Household composition; Country of residence; Region of residence; Degree of urbanisation; Self-declared labour status; Status in employment; Occupation in employment; Economic sector in employment; Highest education completed; Net monthly income of the household

exercise but also a number of common social variables (variables recommended for inclusion in most, but not all social data collections). Two criteria are used for choosing the subset of surveys: relevance and precision considerations.

D. Sampling frames

Work is also required on the sampling frames used for the survey component of social micro-data collections. The objective is to improve quality of the sampling frame in some countries and to ensure that similar sampling frames are used for all surveys in the same countries.

- Quality of the sampling frame is still an issue in a few countries. Such a situation occurs because a sampling frame of individual does not exist or because the existing frame does not allow for representative probability samples to be drawn. Regular updating of the frame can also be problematic.
- In some countries, the sampling frames used differ from one survey to another, especially when the survey is implemented outside the National Statistical Institute.

E. Multi-mode data collections

The future integrated system of micro-data collections is to take on board recent changes in the data collection modes for survey information with an increased use of web and of multi-mode of data collections (combining face-to-face, phone and web-interviews) and associated comparability issues. Results of the work by the ESSnet on multi-mode data collection led by DESATIS to become available in autumn 2014 will be incorporated into the programme of modernisation of social statistics.

F. The target structure for an integrated system of social micro-data collections

Chart 1 gives an overview of the main changes to be implemented in the future integrated system of micro-data collections. They can be summarised as follows:

- Move from 8 to 7 micro-data collections (with the merge of EHIS and EHSIS)
- LFS and SILC are the key pillars of the system:
 - flexibility and responsiveness are mainly achieved for LFS through sub-sampling and frequency requirements (quarterly/yearly/biennial variables) and for SILC through a substantial reduction of the core/nucleus of the survey combined with the introduction of a double system of modules (repeated on a 3 and 6 year basis)
 - Addition of new topics are considered (a new module on health in SILC and new educational or income information in LFS)
- Other detailed information for health and education would be covered by two specific additional periodic surveys (AES and EHIS type). As a consequence of the introduction of new variables on health and education in SILC and LFS, frequency and number of variables to be included in these periodic surveys will be carefully reviewed
- In-depth thinking is also required on the ICT survey (household component) especially as regards the frequency of the data collection (currently annual) given that the results are structural.
- HBS and TUS, currently implemented on the basis of a Gentlemen's Agreement, will require more harmonisation to be considered "European" products and the requirement for joint distributions of income, consumption and wealth (also dealt with in the Stiglitz report) will need to be considered.

Chart 1: Target structure for an integrated system of social micro-data collections: main changes

Current situation								
LFS	SILC	AES	EHIS		EHSIS	ICT	HBS	TUS
Main changes								
LFS	SILC	AES		EHIS with a disability module		ICT	HBS	TUS
- Change of frequency requirement	- Reduction of the core (nucleus)	- Revised frequency (5 to 6 years)		- Revised frequency (5 to 6 years)		- Reduced frequency to be investigated	- Based on a legal basis	- Based on a legal basis
- Use of sub-sampling	- Longitudinal component (4 to 6 years?)	- Review of the content		- Review of the content			- More harmonisation required	- More harmonisation required
- Compulsory infra-annual rotational pattern	- Introduction of a double system of modules (repeated every 3 and 6 years)							
- Harmonised measurement of flows								
- Operational definitions of employment/unemployment								
Content of each data collection scrutinised in view of users' needs								
Precision requirements defined on the basis of similar criteria								
Core and common variables standardised								

3.2. The LFS review

LFS has, since 2010, embarked on a revision process. The EU LFS review process includes a number of initiatives grouped into five main blocks:

- 1) Review of the content of the LFS, including the list of variables in the core of the instrument and the system of ad-hoc modules (see detailed description in paragraph 3.1 A above)
- 2) Modularisation of the LFS (see detailed description in paragraph 3.1 B above)
- 3) Introduction of methodological changes:
 - Introduction of the infra-annual sample rotation pattern as compulsory feature of the LFS

- Development of the methodology for harmonised measurement of flows with the help of a Task Force
 - Clarifications and improvement of precision requirements
 - Clarifications and improvement of non-response indicators
- 4) Development of an operational definition of employment and unemployment and preparation of a draft model questionnaire. Development of a model questionnaire to better measure working time.
- 5) Improvement of timeliness of data transmissions by a few weeks.

The LFS review is fully integrated and consistent with the Eurostat modernisation programme for social statistics. The first two blocks of the LFS review (content and modularisation) are the direct contribution to the first two components of the modernisation strategy. The last three (methodology, definition/model questionnaire and timeliness) complement (and are consistent with) the general modernisation strategy.

4. Development of a Framework Regulation and associated timing

The ultimate aim is that one single Framework Regulation - FR (European Parliament and Council regulation) for an Integrated System of Social micro-data collections covers the 7 social micro-data collections instead of having eight separate EP/Council regulations. The FR will be an instrument for the modernisation that will facilitate integration, standardisation and efficiency, will help align financial and administrative arrangements across data collections. The FR will cover the basic elements and principles of the system (scope (7 micro-data collections); definitions; principles; standardisation; data sources, quality; strategic planning (rolling planning over a seven year period) and financing). An annex will be included with some key specific information on the different data collections. The FR will be complemented by Commission delegated or implementing acts (“technical packages”).

For the LFS, the situation is as follows:

- Information planned to be included in the annex of the FR is the following:
 - List of EU LFS topics,
 - Precision requirements for EU LFS (including the provision for a mandatory infra-annual rotation pattern),
 - Timeliness of data transmissions.
 - System of programming of SAMs,
 - Quality reporting for EU LFS,
 - Data collection modes,
 - Use of administrative data in the EU LFS.
- Elements which will not be in the Framework Regulation but will probably be dealt with through Commission delegated or implementing acts (“technical packages”) include:
 - Operational definition of employment and unemployment,
 - List of EU LFS modules and variables (consistent with the topics included in the annex of the FR) and their coding.

It is finally to be noted that some elements (methodological solutions for the production of flow estimates, calculation of non-response indicators, explanatory notes for the EU LFS variables) will not be included in legal acts. Finally, further discussion is required on where to include the model questionnaires (on employment and unemployment and on working time).

The planned timetable for FR on social micro-data collections is as follows:

- Early 2016: Discussion by Directors of Social Statistics (DSS) and European Statistical System Committee (ESSC) of the FR (and implementing provisions for the technical packages) after an impact assessment is carried out
- Mid-2016: Adoption of the Commission proposal
- Mid-2016 – End of 2018: Discussion in EP and Council (likely to be a long process)
- 2019: Adoption of the FR and implementing provisions
- 2020: First year of implementation

5. Conclusion

The review of the LFS and modernisation of social statistics are long and challenging projects and the planned timetable for the first framework regulation on an integrated system of social micro-data collections means an adoption of the entire legal package in 2019, a first year of implementation under new legal framework in 2020 and first results available for 2020 quarter 1 due around mid-2020. This paper has highlighted that these two exercises are complementary and fully integrated and consistent.

For the first time, a holistic and coherent approach for the streamlining of all social micro-data collections has been developed where not only individual data collection is thoroughly reviewed but also an overall assessment of the system of social micro-data collections is being made. It results in a target structure for an integrated system of social micro-data collections where increase of efficiency and responsiveness to users' needs are core concerns while making progress in standardisation of variables across data collections and enabling multi-mode data collections to be used.