2005 EPUNET Conference
Panel studies: current issues and challenges
EU-SILC: the new European instrument on Income and Living conditions

1. PRESENTATION OF EU-SILC

EU-SILC (Community Statistics on Income and Living Conditions) is an instrument aiming at collecting timely and comparable cross sectional and longitudinal multidimensional micro data on income poverty and social exclusion. This instrument is anchored in the European Statistical System (ESS).

EU-SILC is a response of the ESS to a political need to develop statistical instrument to unable the so called open method of coordination in the fields of social inclusion and pension reform by which Member policies are evaluated by using a common set of comparable indicators.

EU-SILC is a characterised by framework aiming at defining common rules and ensuring comparability across MS and over time (target population, common definitions and common variables, minimum sample sizes, common tracing rules). The framework allows for some flexibility (sample design, mode of collection (register/survey, moving/fixed income reference period, net or/and gross collection but reported gross, national questionnaire, imputation and weighting). The framework regulation defines a set a target primary variables to be collected each year at household and individual level. The areas covered are income, social exclusion, housing, activity and work, demography, education and health. Additional topics such as the intergenerational transmission of poverty, social participation, housing inadequacy are covered occasionally by specific modules (secondary target variables to be defined each year). This implementation is called the ex ante output harmonised concept. The flexibility allowed will inevitably raise non comparability issues. Whenever flexibility is allowed, Eurostat has issue a series of recommendations in order to foster harmonisation (rotational design, method for weighting, imputation …). Given most national instruments have been designed for EU-SILC using Eurostat recommendations, the degree of harmonisation reached is high. Measuring the impact of main source of non comparability will received much attention during the launching phase of the survey.

Timeliness of the cross-sectional data is a clear requirement of political users. The cross sectional micro for survey period during year T and income reference year T-1 are due to be gathered at Eurostat by November T+1 and made available (to researchers) by March T+2. Longitudinal components are to be gathered by March T+2 and made available for research by July T+2.

Longitudinal dimension has received less priority in the current instrument as compared to ECHP. The focus has been put on persistent poverty measured over a period of 4 years. The rotational design proposed by Eurostat records individual trajectories over 4 years. MS were however free to implement a "true" panel to meet Eurostat needs (as was done in Luxembourg and Sweden for instance).

EU-SILC was launched in 2003 on a full scale pilot basis in seven MS (BE, NO, GR, LU, AT, DK and IE). Laeken and pension indicators computed on the basis of EU-SILC micro data have been released in May 2005. The warming phase has allowed studying the impact of the change of instrument the indicator series introduced by EU-SILC. A working paper should be released so on on this subject[1]. The "true" EU-SILC operation was launched in 2004 in 13 MS (all except NL, DE, UK and the 10 new MS except EE) + NO and IS. The first cross sectional data with income reference year 2003 (2004 operation) will be available researchers by March 2006. In 2005, EU-SILC will reach its full scale extension with the 25 MS + NO, IS. In addition, EU-SILC is under preparation in TR, RO, BG, CH. Finally, a transition period till 2007 is foreseen during which NSI have to adapt their tool to common standard, for instance, imputed rent, employer social contribution, income component at gross level. The full implementation of EU-SILC will thus be completed in 2007. The first longitudinal data (2003-2004 income evolution) are expected to be available by July 2007. The first 4 years individual trajectories will be available by July 2009.

2. **MAIN DIFFERENCES BETWEEN EU-SILC AND ECHP**

EU-SILC can obviously be seen as the successor of ECHP. However, they are some substantial differences:

2.1. **Legal basis**

While the ECHP was launched on the basis of a gentlemen's agreement, it was decided to introduce a legal act for EU-SILC and a Framework Regulation of the European Parliament and of the Council was published in June 2003. Technical aspects of the instrument are developed by five Commission Implementation Regulations ('Sampling and tracing rules'; 'Definitions'; 'List of primary variables'; 'Fieldwork aspect and imputation procedures'; and 'Intermediate and final quality reports')[1]. Each year, a new Commission Regulation is set up on the list of target secondary variables: two have been issued to date ('Intergenerational transmission of poverty'; 'Social participation') and a third (on Housing inadequacy) is under discussion.

---

[1] The continuity of indicators during the transition between ECHP and EU-SILC

This aspect has practical consequences: the instrument anchored in the National Statistical System will be going on even without Commission funding (Actually, Commission is contributing to the funding the first 4 years of the survey at the level of 2/3 of the costs). The instrument will deliver each year micro date according to a well define time schedule according to comparable definitions. The drawback is the relative lack of flexibility of the whole instrument: a new regulation has to be issued when changes have to be brought to the instrument.

2.2.   Survey design

The ECHP survey design was based on a pure panel. It means that the sample of people selected for the first year of the survey (sample persons) were followed-up throughout the subsequent duration of the survey (8 years), wherever they may have moved. Children born to sample women were included as sample persons and followed-up.

The cross-sectional and longitudinal data came from the same survey and were collected and processed at the same time.

Since improving timeliness has been one of the new tool's core objectives, and because it is recognised that the longitudinal dimension takes more time in data production, EU-SILC will provide two types of annual data: cross-sectional and longitudinal data that will be treated according to different timetables and processes.

In this way, the cross-sectional and longitudinal data can conceivably come from separate sources, i.e., the longitudinal dataset does not need to be "linkable" with the cross-sectional dataset at the micro-level. Nevertheless, an integrated design 'the rotational design[2]' was recommended by Eurostat for those countries planning to launch a new operation. This design aims to be the most cost effective and efficient for satisfying both the cross-sectional and the longitudinal requirements.

Under this design the minimum panel duration was reduced from 8 years to 4 years (the number of years of observations necessary for building the longitudinal common EU indicators). Consequently, the impact of cumulative attrition will be lower but history spell much reduced.

The use of a rotational panel allows introduction in the sample of new population sub-groups each year (eg. immigrants), and as a result the cross-sectional data derived from this design will have a wider representativity than data derived from a pure panel. Pooling trajectories over time may help to achieved good transition analysis.

2.3.   Variables

The number of EU-SILC target primary variables is much less than the variables recorded in ECHP (although countries are of course free to include additional variables in their national surveys). In particular, subject measure has been

---

[2] Rotational design refers to the sample selection based on a number of subsamples or replications, each of them similar in size and design and representative of the whole population. From one year to the next, some replications are retained, while others are dropped and replaced by new replications

withdrawn. Nevertheless, the variables necessary to build the indicators adopted under the open method of coordination have generally been retained (with some exceptions)[3]. However, the instrument has been designed to allow targeted modules on less rapidly changing variables to complement sporadically or with a lower periodicity the primary target variables collected every year.

2.4. Definitions

Although the majority of EU-SILC variables were defined in the same way as the corresponding ECHP variables, the total household gross and disposable income and the different income components were redefined to follow as closely as possible the international recommendations of the UN 'Canberra Manual'.

2.5. Income

A key objective of EU-SILC is to deliver robust and comparable data on total disposable household income, total disposable household income before transfers (except old age and survivor's benefits; including old age and survivor's benefits), total gross income and gross income at component level.

This objective will be reached in two steps, insofar as Member States will be allowed to postpone the delivery of gross income at component level and of total household gross income data until after the first year of their operations.

Apart from this, the only data for which delivery will not be compulsory as from the first year of the operation are as follows:

- non-monetary components of employee (with the exception of company cars that is to be calculated as from the first year of the operation) and self-employed income, imputed rent and interest payments that shall be optional from the first year of the operation and compulsory from 2007;

- gross employers' social insurance contributions shall only be included from 2007 if results of feasibility studies are positive.

As it is mentioned above, in EU-SILC the income at component level is recorded 'gross'. In the ECHP, the income components were recorded net.

New components of disposable income have been introduced in EU-SILC:

- Transfers paid to other households (only transfers received from other households were taken into account in the ECHP);

- Tax adjustment (only taxes paid at source were collected in ECHP)

- Taxes on wealth

---

[3] ILO activity status is not an EU-SILC target primary variable. In this way, indicators based on this variable will henceforth be built with the 'current activity status' or 'most frequent activity status'. Furthermore, the distinction between employees and self-employed is no more possible in the calendar of activities, and therefore it may only be possible to compute the working poverty indicator for all people at work, regrouped together.

- Interest paid on mortgage loans

- Imputed rent

- Non-cash employee income ('income-in-kind')

- Value of goods produced for own consumption ('income-in-kind')

- Employers social insurance contributions[4]

- Furthermore, EU-SILC takes into account negative values of self-employment income, which were previously set to 0 in the ECHP.

- Other variables that can take negative values are variables collected under 'property income'.

The content of some variables has changed:

- The social benefits do not contain the income from 'individual pension plans' (this component was included in theory in the ECHP)

- Survivors' and disability benefits paid after the standard retirement age are included in EU-SILC under 'old-age benefits' (and no more in survivors' and disability functions as in ECHP)

- Early-retirement benefits paid for labour market reasons or in case of reduced capacity to work are included respectively under 'Unemployment benefits' or under 'Disability benefits' (and no more in old age benefits as in the ECHP).

- The total household disposable income[5] includes the new components introduced in the project and tax adjustments are taken into account in the calculation of this variable.

The income reference period is more flexible:

- While in the ECHP the income reference period was the previous year, EU-SILC has fewer constraints. In this way, the income reference period may be a fixed 12-month period (such as the previous calendar year or tax year) or a moving 12-month period (such as the 12 months preceding the interview) or be based on a comparable measure.

The impact of the changes between ECHP and EU-SILC have already received some attention from Eurostat. Further studies need to be carried out. It is planned to organize a conference on this theme in the course of 2006.

---

[4] As mentioned above, gross employers' social insurance contributions shall only be included from 2007 if results of feasibility studies are positive.

[5] See in Annex I 'Computation of household disposable income'

3. THE CHALLENGES OF EU-SILC

3.1. EU-SILC as the main source for comparable data on income and living conditions and for the development of analytic expertise in policy evaluation

The instrument has been designed as an answer to a well defined political need. As such it should be used to develop extensive micro economic evaluation of EU and national policies in the field of social inclusion in its relation to employment and growth. There is a need to develop analytical tools based on EU-SILC.

3.2. EU-SILC as the main source for comparable data on income and living conditions for scientific research

3.2.1. Data access

The use of EU-SILC micro data for research is of great importance for Eurostat. First, it ensures the best return for money invested in the EU-SILC project. Eurostat and National publications are de facto limited in scope and extend. They are not fully flexible and do not fill all the user needs. Second, the wide use of EU-SILC data will definitively help to gradually increase the quality of EU-SILC. If adequately organised, researchers and users feedback will allow to point out weaknesses and to suggest improvement of the instrument.

The release to researchers under Regulation 831/2002 is foreseen in the EU-SILC Framework Regulation 1177/2003 where calendars for release as well as constraints for publication have been settled down. Cross-sectional EU-SILC 2004 micro data files should be made available for scientific research by March 2006.

3.2.2. Information content

The release should obviously be compatible with the right of privacy of individuals and the derived principle of statistical confidentiality and the statistical regime developed in Regulation 322/97 and 831/2002. The free access to sensitive micro data with the risk a breach of confidentially would indubitably hamper the delicate and necessary trust individual have in official statistical data compiler.

Wider release such as public release has been evoked and will continue to be discussed. The main priority for Eurostat is to organise the release for scientific research which stipulates that anonymised micro data can be release to researchers under license conditions.

Given the constraint of preservation of privacy right of respondents, the release of micro data is a difficult balance between the information content of the data release and the objective risk of disclosure of private information. In this exercise there is real risk to the protected data are of low value for scientific research given their information content.

Anonymisation of EU-SILC is currently heavily discussed at ESS level. A task force has come up with recommendations of best practices: main disclosure risks have been identified and method for protecting (mainly global recoding/top coding of identifying variables complemented by local suppression) the data has been designed. The TF has recommended benchmarking the proposal against researchers needs in

order to assess the relevance of the choices made. Actually, no one optimal solution exists with respect to anonymisation but a series of equivalent sub optimal solutions which differ by a priori choices. For instance further recoding can often be balanced by limited local suppression on secondary variables. Also, the choice of the combination of identifying variables to be recoded can have a huge impact on the usability of the variable. Anonymisation is all a matter of trade off. One of the objectives of this paper is to initiate dialog with researchers on this issue.

The scope of paper does not allow for a detailed description of the anonymisation methods currently envisaged. Roughly:

- Reference age is provided in year by is top coded (80+)

- The hierarchical structure of the files (household/individual) has been preserved

- The longitudinal structure of the files (linking of individual through waves) has been preserved

- Any direct identifier is obviously removed from the files. Actually these identifiers are neither available at Eurostat. Numeric identifiers have been randomised.

- All sampling design information (stratum, label of PSU, design weight) and fieldwork information (proxy, length of the interview…) have been removed. Only sampling weights are given. Eurostat is thinking at black box tool based on replication to allow researchers to estimate variance.

- Very disaggregated information related to time has been removed: month of birth, month of move in/ move out.

- The geographical information (NUTS codes and urban/rural classification) is critical information regarding disclosure risk. The proposal of the EU-SILC working group is try to keep NUTS 1 level code and the urban/rural classification. However, it is already knows that for small countries, this level of desegregation is will increased dramatically the risk of identification. Member States are currently assessing this proposal. It is likely that this information can be provided in the more populated countries but not in the other ones.

- Structural variables have been altered according to

Label	Code	Status
SEX	RB090	Unaltered
COUNTRY OF BIRTH	PB210	Local/EU/non EU/world
CITIZENSHIP 1	PB220A	Local/EU/non EU/world
CITIZENSHIP 2	PB220B	Removed
YEAR OF BIRTH or AGE	RB080	Bottom recode (1923 and before)
DWELLING TYPE	HH010	Modality 5 ("Other") put to missing
TENURE STATUS	HH020	Unaltered
NUMBER OF ROOMS	HH030	Top coding (6 and more)
BATH OR SHOWER IN DWELLING	HH080	Unaltered
DO YOU HAVE A CAR?	HS110	Unaltered
MARITAL STATUS	PB190	Unaltered
CONSENSUAL UNION	PB200	Unaltered

EDUCATION (ISCED)	PE040	ISCED 5 and 6 regrouped
ECONOMIC STATUS	PL030	Unaltered
STATUS IN EMPLOYMENT	PL040	Unaltered
OCCUPATION (ISCO-88 (COM))	PL050	Unaltered
NACE	PL110	Grouped at 1 one letter (19 levels)
HOUSEHOLD TYPE	Derived	Unaltered
HOUSEHOLD SIZE	Derived	Unaltered

- Additional specific but limited regrouping can be proposed by Member State to adapt the general scheme to national specificities.

- A priori local suppression on identifying variables should be avoided as much as possible. Only if the affect secondary variables or already non response altered variables, they can be envisaged as a good trade to avoid further global recoding

The here above proposal is the common base to the anonymisation of the national datasets. Fine tuning is the responsibility of MS in cooperation with Eurostat that will seek harmonisation of practices and the respect of best practices.

Further details can be found in two independent papers annexed (to be posted on the conference site) describing anonymisation proposal in more details and a project of structure of the future anonymised UDB.

There is still a theoretical possibility to access the full data at Eurostat premises in a so called secure environment. This possibility has to be strongly motivated because of the human resources implied by this arrangement.

Eurostat is about to launch a wide consultation of researchers on the here above proposal and to collect their feedback in order to tune the anonymisation. Voluntary feedback based on the information provided at the EPUNET conference is very welcome at author e-mail (jean-marc.museux@cec.eu.int) before beginning of September 2005.

3.3.   EU-SILC as a dynamic process seeking improvement of quality

The third challenge will be to organise communication and feedback with scientific community considered as one of the main stakeholder of EU process together with the Commission services in charge of the evaluation of MS policies. Eurostat will organise conference around EU-SILC but there is a need to streamline feedback. The BHPS user group and EPUNET structure are definitely examples of good practice. One has to think at the possibility to draw from these experiences.

As early as 2006, a continuous review of the instrument should be launched in order to improve and adapt regulations for the forthcoming years. Benefiting from scientific expertise would certainly help Eurostat to improve the instrument.

---

[i] Framework Regulation (EC) N° 1177/2003 of the European Parliament and the Council of 16th June 2003 concerning Community statistics on income and living conditions

(http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_165/l_16520030703en00010009.pdf).
Implementing Commission regulations:
N° 1980/2003 (definitions)
(http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_298/l_29820031117en00010022.pdf),
N° 1981/2003 (fieldwork aspect and imputation procedures)
(http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_298/l_29820031117en00230028.pdf),
N° 1982/2003 (sampling and tracing rules)
(http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_298/l_29820031117en00290033.pdf),
N° 1983/2003 (list of target primary variables)
(http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_298/l_29820031117en00340085.pdf),
N° 28/2004 (content of intermediate and final quality reports)
(http://europa.eu.int/eur-lex/pri/en/oj/dat/2004/l_005/l_00520040109en00420056.pdf