AMELI

Advanced Methodology for European Laeken Indicators

Deliverable 5.1

Quality of EU-SILC data

Version: 2011

Monique Graf, Ariane Wenger, and Desislava Nedyalkova

Contributors to Deliverable 5.1

Chapter 1: Monique Graf and Ariane Wenger and Desislava Nedyalkova, Swiss Federal Statistical Office.

Chapter 2: Ariane Wenger and Desislava Nedyalkova, Swiss Federal Statistical Office.

Chapter 3: Desislava Nedyalkova, Swiss Federal Statistical Office.

Main responsibility

Monique Graf, Ariane Wenger, and Desislava Nedyalkova, Swiss Federal Statistical Office.

Evaluators

Internal expert: Beat Hulliger, University of Applied Sciences Northwestern Switzerland.

Aim and Objectives of Deliverable 5.1

EU regulations 1177/2003 and 28/2004 describe the quality reporting of EU-SILC and the indicators on poverty and social exclusion as to be performed by the member states within the ESS. Since EU-SILC is newly setup, implementation details of the different EU surveys are needed to be known for building an appropriate methodology which takes into consideration possible peculiarities of data and other concerns from praxis.

The aim of the work package consists of two issues:

1. Collecting metadata and quality information of the EU-SILC Laeken indicators including survey designs of the participating NSIs;

2. Further, possible peculiarities of data and procedures will be investigated as possible input for the simulation study.

The present report gives an overview of selected standard quality reports on SILC. The overview focuses on the AMELI relevant tasks such as the income variables and the sampling designs which are essentially needed as input for the simulation study.

Contents

Chapter 1

Description and Quality of the EU-SILC User Database

1.1   Introduction

The goal of this chapter is to collect metadata and quality information on the EU-SILC Laeken indicators both at the European and at the national level.

Section 1.2 is devoted to the user database description. Section 1.2.1 covers the aim of the survey, the legal basis, the reference population, the sampling plan, the surveys, the weights, the imputation and finally the data sets and the variables. Section 1.2.2 is devoted to the construction of the target variables which are divided in four types: household register, personal register, household data and person variables.

In Section 1.3, we quote two documents about the calculation of the overarching and pensions indicators based on EU-SILC, adopted under the Open Method of Coordination (OMC).

The quality documents of EU-SILC are available in Section 1.4 which is divided in three parts. Section 1.4.1 collects national quality reports from the national statistical institutes participating in the AMELI project, namely Austria, Estonia, Germany, Slovenia and Finland. Section 1.4.2 contains quality documents from a supranational perspective. Next, Section 1.4.3 addresses the problem of weighting.

Section 1.5 is devoted to measurement problems. Section 1.5.1 collects documents about the impact of extreme incomes on the precision of poverty and inequality indicators and Section 1.5.2 about their variance estimation.

1.2   User Database and Target Variables

In this section, we find a general description of the user database EU-SILC from the goal of the project to the construction of the target variables. EU-SILC 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).

Clemenceau and Museux (2007) proposed the first evaluation of the project after almost five years of activity, from its launching up to the Conference on Comparative EU Statistics on income and living conditions: issues and challengesïn Helsinki, from the perspective of the European Commission and from the perspective of each country.
The EU-SILC project has been implemented over the last years on a step by step basis. It has been launched in 2003, on the basis of a 'gentlemen agreement' in seven countries (AT, BE, DK, GR, IE, LU, and NO). The user database (UDB) 2004 was issued by 15 countries (AT, BE, DK, EE, ES, FI, FR, GR, IE, IS, IT, LU, NO, PT and SE). Then 26 countries (Norway, Iceland and the 25 member states except Malta) participated in the EU-SILC UDB 2005 - 2007. Bulgaria (BG) and Romania (RO) joined the survey in 2008. Switzerland (CH) and Turkey (TR) will join the project in the subsequent years.

Two types of data are provided:

- Cross-sectional data pertaining to a given time or a certain time period with variables on income, poverty, social exclusion and other living conditions;

- Longitudinal data pertaining to individual-level changes over time, observed periodically over, typically, a four-years period.

The launching of EU-SILC experiences a transition period till 2007 during which NSI's can adapt their tool to common standard, for instance, imputed rent, employer social contribution, income component at gross level. Release 2004 corresponds to cross-sectional data for a limited set of countries. The first longitudinal data were available by July 2007 for the UDB 2005 for 13 countries: AT, BE, EE, ES, FI, FR, GR, IE, IS, IT, LU, NO, SE. In 2008, the longitudinal data are available for the 26 countries. The first 4 years' individual trajectories will be available by July 2009.

In the sequel, certain documents are quoted twice: in Section 1.2.1 for the general description of the UDB and in Section 1.2.2 for the variables.

1.2.1 Description of the UDB

The documents contained in this subsection are classified by the year of the UDB. The 2004 information was designed to be general and to be reusable for subsequent releases. Nevertheless, besides the ground information, each year is devoted to a particular topic. This framework allows for flexibility and different implementations. Information on current status of the implementation in the different Member States (MS) is provided as an addendum to the different sections and complements the general presentation. The transitional measures valid till 2007 are also underlined whenever relevant.

The topics about the general description of the UDB are: aim of the survey, legal basis, reference population, sampling designs, survey characteristics (survey units, modes of collection, survey duration and time, tracing rules), weights (legal aspects, theoretical

aspects and EU-SILC weights), imputation (missing data in EU-SILC and EU-SILC target variables for imputation) and database (data availability, domains and areas, the files, format, the data sets, variables and list of variables).

1. **UDB 2003 - 2006:**
   A general description of the UDB for years 2003 - 2006 is given in BERNARD (2008). The EU-SILC sampling designs for 2004 are given in Annex 2.

2. **UDB 2007:**
   The document of BERNARD (2009) contains the description of the UDB and the EU-SILC sample designs for 2007 (in Annex 3).

3. **UDB 2008:**
   The document EUROSTAT (2010) contains the description of the UDB, the target variables and the secondary target variables.

4. **UDB 2009:**
   The document EUROSTAT (2009f) contains the description of the UDB, the target variables and the secondary target variables.

1.2.2   Description of the Variables

In this subsection, we are interested in the structure of the database which is divided in two parts: the main target variables, i.e. the variables which are present every year and the secondary target variables which are part of the module of the year considered. Indeed, every year the survey treats a particular topic like intergenerational transmission of poverty (module 2005), social participation (module 2006), housing conditions (module 2007), over-indebtedness and financial exclusion (module 2008) and material deprivation (module 2009)(see above).

The documents about the main variables are built in the following way. For each variable, the following is provided:

- Kind of data:

  A. If the variable is a household variable, it is either:

    1. Basic data (basic household data including degree of urbanisation);
    2. Income (total household income and gross income components at household level);
    3. Social exclusion (non-monetary household deprivation indicators, including problems in making ends meet, extent of debt and enforced lack of basic amenities);
    4. Labour information (child care);
    5. Housing (Dwelling type, tenure status and housing conditions, amenities in dwelling and housing costs).

B. If the variable is a person variable, it is either:

1. Basic data (basic personal data and demographic data);

2. Education;

3. Labour information (basic labour information on current activity status and on current main job, including information on last main job for unemployed, basic information on activity status during income reference period, total number of hours worked on current second/third ... jobs, detailed labour information, activity history and calendar of activities);

4. Health (health, including health status and chronic illness or condition, and access to health care);

5. Income (gross personal income, total and components at personal level).

- Type of variable: cross-sectional or longitudinal;

- Reference period (constant, current, income reference period, last twelve months, since last year, working life and childcare reference period);

- Unit (household, household member, former household member, selected respondent and household members aged 16 and over);

- Mode of collection (constructed, frame, register, interviewer, household respondent, personal interview and proxy);

- Values (range);

- Flags.

Like before the documents are classified by year of the UDB.

Finally, we have:

1. **UDB 2003 - 2006:**
   The description of the target variables is given in EUROSTAT (2008e).

2. **UDB 2005:**
   The secondary target variables used in the module 2005, *Intergenerational transmission of poverty*, are described in EUROSTAT (2008c).

3. **UDB 2006:**
   The secondary target variables used in the module 2006, *Social participation*, are described in EUROSTAT (2008d).

4. **UDB 2007:**
   The description of the target variables is given in EUROSTAT (2009e). The secondary target variables used in the module 2007, *housing conditions*, are described in EUROSTAT (2008b).

5. **UDB 2008:**
   The description of the target variables is given in Eurostat (2010). A description of the secondary target variables in module 2008, *Over-indebtedness and financial exclusion*, is given in Annexe 8.

6. **UDB 2009:**
   The description of target variables is given in Eurostat (2009f). A description of the secondary target variables in module 2009, *Material deprivation*, is given in Annex 8.

Comparison between recommended and provided equivalised disposable income on the basis of EU-SILC UDB 2004-2006 is provided in the next chapter.

1.2.3   Guidelines in Progress

The subject of Eurostat (2008f) is the harmonisation of SILC and EHIS[1]. Part of SILC questionnaire is in fact about health. So it is recommended to coordinate activities at the national level between SILC and EHIS counterparts in order to have EHIS standards implemented in the SILC questions on health according to the new SILC revised guidelines from 2008 data collection onwards Eurostat (2010).

1.3   Indicators

The portfolio of Overarching Indicators is available in Eurostat (2009a). The objectives, definition and computation of these indicators are presented.

Eurostat (2009b) contains a background about the development of Pensions indicators under the Open Method of Coordination. Further, we can find a portfolio of Pension Indicators calculated from SILC with detailed methodological notes about the calculations.

1.4   Quality

This Section contains different documents about quality. In Section 1.4.1 the Quality Reports follow a special structure that is mandatory by the Commission, assuring a continuous documentation of quality. The documents in Section 1.4.2 develop a more prospective point of view about the assessment of statistical quality in EU-SILC. Finally, in Section 1.4.3, we have collected several documents about the weighting procedure in EU-SILC.

---

[1]EHIS: European Health Interview Survey

1.4.1 Quality Reports

All these documents follow the structure outlined in the Commission Regulation No. 28/2004 (EUROSTAT (2004a)). This regulation defines four chapters. The first chapter provides the common cross-sectional indicators and other indicators of interest computed on the basis of EU-SILC. The second chapter deals with accuracy, meaning that all factors that affect the closeness of estimations and results to the exact or true value should be described. The third chapter reports on comparability and describes all differences between the standard EU definitions and the definitions applied by the countries. The fourth and last chapter, reporting on coherence, presents the comparisons of the EU-SILC data with external sources.

The comparative final quality report for 2005, EUROSTAT (2008a) gathers and summarizes all the information contained in the 2005 national final quality reports that the member states have sent to Eurostat. The objective here is to evaluate the quality of the instrument from the European point of view, i.e. by establishing between-country comparisons of some of its key quality dimensions. Quality reports for 2006 and 2007 are also available (EUROSTAT (2009c,d)).

In the following subsection, we find the national final quality reports of those NSI's that are involved in AMELI for different years.

Austria Quality Report

Austria uses a simple random sample of addresses. Sampling units are dwelling units registered in the Central Residence Register (ZMR, Zentrales Melderegister). At the end, all households and individuals living at the eligible addresses were interviewed.

All the responding households in 2004 have been kept in the sample for the 2005 survey. In addition, a new sample of addresses was selected in 2005 by simple random sampling. This sample was supplemented by another sample of addresses, which was drawn likewise. At the end, all the households and the individuals living at the eligible addresses were interviewed.

In 2006 the sample consists of two quarters that entered the survey in 2004, one quarter that started in 2005 and a further quarter that entered the survey in 2006. More details for the UDB 2003, UDB 2004, UDB 2005 and UDB 2006 are available in these quality reports (STATISTICS AUSTRIA (2004)), (STATISTICS AUSTRIA (2005)), (STATISTICS AUSTRIA (2007)), (STATISTICS AUSTRIA (2008a)) and (STATISTICS AUSTRIA (2008b)).

Estonia Quality Report

There are three quality reports, one for the UDB 2004 (STATISTICS ESTONIA (2006)), one for the UDB 2005(STATISTICS ESTONIA (2007a)) and the last one for the UDB 2006 (STATISTICS ESTONIA (2007b)).

The design used is one-stage stratified unequal probability sampling of households, with a household selected with probability proportional to the number of persons aged 14 and over. The Estonian EU-SILC sample is selected according to the following sampling procedure:

- Stratification by county level into three strata by population size: *big* counties, *small* counties and the Hiiu County, which forms a separate stratum as the smallest county in terms of population size.

- A sample of persons aged 14 and over is selected with equal probabilities within strata.

All the households of the selected persons are identified and all eligible persons in the household are interviewed.

As 2006 operation was the third round of EU-SILC in Estonia, the sample comprised of three parts: 1. The first part consists of households selected for the survey in 2004 and followed up in 2005 and 2006. Initially, this part consisted of 4 rotational groups, one of which was to be dropped after the 2004 operation and another after the 2005 operation. However, due to the smaller than expected response rates, it was decided to keep all rotational groups in the sample up to 2006. 2. The second part consists of households selected for the survey in 2005 and followed-up in 2006. 3. Sample of new households introduced into the survey in 2006.

Weighting scheme was generally in line with documents Osier et al. (2006) and Verma (2006), with some peculiarities due to modified rotational scheme.

The sample of year 2006 consists of three subsamples to be weighted independently and combined thereafter:

- s_3: households started in 2004 and their split-offs, participate for the third time (= 4 rotational groups)

- s_2: households started in 2005 and their split-offs, participate for the second time (= one rotational group)

- s_1: households started in 2006, participate for the first time (= one rotational group)

First, we need to calculate base weights of year 2006 for each subsample. Since weighting procedure was different in 2005, base weights of 2005 also need to be recalculated. In the following years, no recalculation will be needed, besides correction for attrition.

More details are available in the above quality reports.

Germany Quality Report

For Germany, we have two quality reports.
For the UDB 2005, FEDERAL STATISTICAL OFFICE GERMANY (2007) and for the UDB 2006, FEDERAL STATISTICAL OFFICE GERMANY (2008).

The German SILC survey is organised as a rotational panel containing 4 subsamples. In 2005 the survey started with 3 quota samples and 1 random sample. In 2006 one quota sample was replaced by a further random sample, i.e. in 2006 there were 2 quota subsamples and 2 random subsamples. And so on, until from 2008 where there are only four random samples.
The sampling frame for the random subsamples is the permanent sample (DSP), a sampling frame recruited among former participants of the German Microcensus (access panel).

For both random and quota part, the stratification criteria within the German Microcensus were in 2006 (as in 2005): Federal state (Bundesland); Household type; Social status of the main income earner; Household net income; Farm household (separate stratum for each federal state). All the individuals living at the selected addresses were eligible for interview.

Slovenia Quality Report

For Slovenia, we have the quality report of the UDB 2005, STATISTICAL OFFICE OF THE REPUBLIC OF SLOVENIA (2007a) and of the UDB 2006, STATISTICAL OFFICE OF THE REPUBLIC OF SLOVENIA (2007b).

The sample for the Slovenian EU-SILC was selected according to a stratified two-stage design.

- The strata were defined according to the size of the settlement (clusters of enumeration areas defined according to the population size) and its proportion of agricultural households and sorted according to the statistical regions, see STATISTICAL OFFICE OF THE REPUBLIC OF SLOVENIA (2007a) for more details.

- In each stratum Primary Sampling Units (PSU), i.e. enumeration areas, were firstly systematically selected.

- In the second stage seven persons aged 16 and over were selected in each PSU.

Finally, the entire household, a selected person belongs to, was eligible for contact.

Since in 2005 a lower interviewing rate than expected was experienced, the sample for 2006 was enlarged; otherwise the sample size after four years would be too small for longitudinal analysis.

In 2006 the fourth wave from 2005 would be dropped out, but it was kept and divided into three parts, which were reallocated to the remaining three waves from 2005. Therefore, all households which responded in 2005 were interviewed again in 2006.

Since this was decided before data processing of the 2005 survey, selected sampling units were initially renumbered so that instead of four, there are three rotational groups. None of the rotational groups were dropped out in 2005. In 2006 only one new rotational group was added, so there are four rotational groups in 2006.

Finland Quality Report

For Finland, we have three quality reports: UDB 2004: STATISTICS FINLAND (2005); UDB 2005: STATISTICS FINLAND (2008); UDB 2006: STATISTICS FINLAND (2007).

The sampling design of the Finnish EU-SILC survey 2006 is a two-phase sampling design. The Finnish cross-sectional SILC data collection 2006 contains two groups based on the Income Distribution Survey (IDS): one is a new rotation group and the other is a set of responding households of the IDS of the previous year. Both of them had been selected according to the following two-phase sampling design:

- A systematic sampling of persons aged 16 years and more is carried out in the Population Register in order to get the basis for a Master Sample. Then, all the dwellings with at least one selected person are included in that Master Sample.

- The Master Sample is stratified according to socio-economic criteria.

- A simple random sample without replacement of dwellings is selected in each stratum.

Finally, all the households and the individuals living in the selected dwellings were eligible for interview.

There are some changes in the used questionnaires.
The 2005 questionnaire was partly rebuilt. The section on labour market activity was totally reconstructed, since the three tests of the 2004 questionnaire carried out by the Cognitive Laboratory showed major problems in the interviewing process. However, the trouble seemed to come from the ordering of the questions, not from wordings. The ordering of the questions was changed thoroughly in 2005.

The changes in the questionnaire for 2006, which were not so many, were introduced to the interviewers in a separate written report and, of course, in the instructions book. The instructions book is rewritten every year and it is also under constant development. No interviewers' training courses were arranged in 2006 for the interviewers who had been trained in earlier years. The training material was available in the form of the computer-assisted telephone interview (CATI) questionnaire and interviewer instructions and the interviewers were paid to study the material.

1.4.2   Quality Documents

The documents in this section investigate the assessment and possible improvement of the statistical quality in EU-SILC.

Verma (2007) develops and discusses a framework for the assessment of statistical quality in EU-SILC, with focus on comparability as a central dimension of quality. The data quality is viewed as a multidimensional concept, covering not only statistical accuracy but also the relevance, timeliness, comprehensiveness, etc., of the data. There is a broad agreement on what dimensions make up the overall quality of statistical data, and these dimensions are briefly reviewed, noting some relationships between them.
This paper proceeds to identify a number of specific aspects where problems of comparability are likely to arise in EU-SILC, and elaborates methodologies for the study of some of the more important ones among them.

Figari et al. (2007) explore the prospects for using the EU-SILC as the underlying micro-database for policy simulation across the EU. In particular it considers the issues to be addressed, and the advantages arising, from building a database from the EU-SILC for the EU tax-benefit model, EUROMOD. In order to identify the issues and illustrate their importance, a trial database for Spain has been constructed. It is used within EUROMOD to calculate some selected social indicators as well as indicators of work incentives and the effects of fiscal drag in Spain between 2003 and 2006. They conclude that, although transforming the EU-SILC into a database for EUROMOD would require a significant amount of effort, this is likely to be worthwhile because of the consequent improvements in comparability across countries, efficiency in developing and maintaining the model for many countries and simplification of access arrangements. Therefore some suggestions for how to improve the User Database for this purpose are offered.

1.4.3   Weighting

The weighting methodology is an important topic in the EU-SILC survey and several documents address this topic. Indeed, although the calculation of weights is a country's responsibility, Eurostat has provided guidelines on how to calculate weights within the two documents: Eurostat (2003b) and Eurostat (2003c).
They furnish information about the construction of a sample of households and persons, design weights, cross-sectional weights and longitudinal weights.

The objective of Eurostat (2004b) is to further the concept of sample weights and to provide a systematic procedure to obtain them. The weighting procedures are described for the units of the subsamples that appear for the first time in the EU-SILC sample, only in situations where a rotational design is used.

Eurostat (2005) proposes a procedure to compute the cross-sectional weights from second year onwards, when a rotational design with four panels, as recommended by

Eurostat, is supposed to have been applied.

In the first part of the document, the basic principle of the rotational design is recalled and a panel terminology (panel person, sample household and co-resident), appearing throughout the document, is compiled in a second part. The third part provides guidelines for a satisfying weight computation, by distinguishing whether a sample of households or a sample of persons has been drawn.

The proposal of Eurostat (2004c) is to adjust the distribution of children for each year of age. This involves the adjustment of personal cross-sectional weights so as to make the distribution, according to age characteristics, of the children covered in the sample agree with the same information from some more reliable external source (age distribution of children aged 0 to 12 in private households).

Apart from Eurostat's documents, here is a list of papers on the weighting scheme.

Osier et al. (2006) and Verma et al. (2007) build on Verma (2006) and provide a comprehensive description of an integrated system of cross-sectional and longitudinal weighting for rotational household panel surveys. The two papers provide a systematic description of the weighting procedures, also introducing some refinements so as to enhance the consistency and completeness of original recommendations.

The papers begin with a summary of the main features of EU-SILC and an overview of the integrated weighting system for the different types of data coming out of the rotational panel annually. They describe a step-by-step procedure for construction of initial weights to be applied to each new sample as it is introduced into the survey. The objective is to propose a unified structure for the whole weighting procedure for the standard integrated EU-SILC design, covering the initial sample, and its cross-sectional as well as longitudinal development.

The document is outlined as follows: 1. Weighting for the first year of each subsample (panel); 2. Computation of base weights; 3. Cross-sectional weights, year 2 onwards; 4. Longitudinal weights.

Ardilly and Lavallée (2007) discusses the longitudinal and cross-sectional weighting for the French SILC survey, that are designed to produce approximately unbiased estimators.

Graf (2009) gives a very detailed description of the weighting procedure (cross-sectional and longitudinal) in the case of the Swiss Household Panel. Similar steps are taken up for the Swiss SILC.

1.5   Variability

The precision of the Laeken Indicators depends on the variability of the underlying income variables. It is thus of importance to compute the variance of the equivalised income and to take the presence of extremes into account.

1.5.1 Extreme Income

The micro-data estimates of welfare indices are known to be sensitive to observations from tails of the income distribution. It is therefore usual to make adjustments to extreme data before proceeding with estiamtion of inequality and poverty statistics. KERM (2007a) systematically evaluates the impact of such adjustments on indicators estimated from the EU-SILC by a Pareto distribution.

KERM (2007b) presents the results of a large scale sensitivity analysis considering both simple, classical adjustments and a more sophisticated approach based on modeling parametrically the tails of the income distribution.

1.5.2 Variance

In OSIER and MUSEUX (2006) the principle of the variance estimation approach based on linearization techniques is presented.

LEITEN and TRAAT (2008) contains the estimation formulae for Laeken indicators and then concentrates on their variance estimation by resampling and linearization methods (Jacknife method). Then, the variance formulae of Laeken indicators, such as median, at-risk-of-poverty threshold, at-risk-of-poverty rate and related quantities are developed, for the EU-SILC design of Estonia.

1.6 Remarks

Up-to-date documentation on EU-SILC can be found on EU-SILC Library. All cited documents are available on demand (in pdf or doc format).

Chapter 2

Comparison between Recommended and Provided Equivalised Disposable Income on the basis of EU-SILC UDB 2004-2006

2.1 Introduction

This chapter is intended as a study of the equivalised disposable income, which is the basic concept in the construction of the EU-SILC Laeken Indicators. First, we recall the definition of the equivalised disposable income. Then, we use the R software to reconstruct this income in the EU-SILC 2004-2006 User Database (UDB). From the results, we compare the equivalised disposable income that we have constructed with the one provided in the UDB and conclude.

2.1.1 Equivalised disposable income

The theoretical definition of the equivalised disposable income, called EQ_INC_i, is the following:

For each person i, EQ_INC_i is defined as the household's total disposable income corrected for individual non-response, $TDHI$, divided by the equivalent household size, EQ_SS.

In other terms,

$$EQ\_INC_i = \frac{TDHI}{EQ\_SS}, \tag{2.1}$$

where:

- $TDHI = HY020 \times HY025$, where

- $HY020$ is the total disposable household income[1],

- $HY025$ is a within-household non-response inflation factor[2][3] and is defined by:

$$HY025 = \frac{HY020_c + HY020_i}{HY020_c},\tag{2.2}$$

  where:

  $HY020_c$: collected total disposable income;
  $HY020_i$: imputed total disposable income.

- EQ_SS is the equivalised household size.

2.1.2 Equivalised household size

In order to take account of the differences in the size and the composition of a household, the equivalised household size is defined using the so called modified OECDßcale, which gives a weight of 1.0 to the first adult, a weight of 0.5 to other household members aged 14 or over and a weight 0.3 to other household members aged 13 or less:

$$EQ\_SS = 1 + 0.5(HM_{14+} - 1) + 0.3HM_{13-},\tag{2.3}$$

where:

HM_{14+} = number of household members aged 14 and over (at the end
              of income reference period)
HM_{13-} = number of household members aged 13 or less (at the end
              of income reference period)

The definition of age as the age at the end of the income reference period can be found in EUROSTAT (2008e).

2.1.3 Total disposable household income

The total disposable income of a household is calculated by adding together the personal income components received by each household member plus income components received at household level minus some deductions. It is described in details in EUROSTAT (2008e). In EUROSTAT (2009a), we can find three definitions of the variable $HY020$, based on gross income components, on net income components, and on a mixture of gross and net income components. However, the recommendations for SILC are to use the gross income components. The first and recommended definition of the total disposable income, based on the gross income components, is given in Table (2.1):

---

[1]From 2007, the variable $HY020$ will represent the corrected for individual non-response disposable income.

[2]Factor used to multiply the total gross income, the total disposable income, or the total disposable income before social transfers, to compensate for the non-response in individual questionnaires. It is necessary to correct the effect of non-responding individuals within a household, otherwise the income of individuals interviewed doesn't sum up to the total household income.

[3]From 2007, the variable $HY020$ will represent the corrected for individual non-response disposable income.

Table 2.1: Recommended definition of the variable *total disposable household income (gross).*

HY020: Total disposable household income=	
+	$PY010G$: Gross employee cash or near cash income
+	$PY020G$: Gross non-cash employee income
+	$PY050G$: Gross cash benefits or losses from self-employment
+	$PY090G$: Unemployment benefits
+	$PY100G$: Old-age benefits
+	$PY110G$: Survivor' benefits
+	$PY120G$: Sickness benefits
+	$PY130G$: Disability benefits
+	$PY140G$: Education-related allowances
+	$HY040G$: Income from rental of a property or land
+	$HY050G$: Family/children related allowances
+	$HY060G$: Social exclusion not elsewhere classified
+	$HY070G$: Housing allowances
+	$HY080G$: Regular inter-household cash transfers received
+	$HY090G$: Interests, dividends, profit
+	$HY110G$: Income received by people aged under 16
-	$HY120G$: Regular taxes on wealth
-	$HY130G$: Regular inter-household cash transfer paid
-	$HY140G$: Tax on income and social insurance contributions

The variables $PY030G$, $PY070G$, $HY030G$, $HY100G$ are only mandatory from 2007 and the variable $PY020G$ is only mandatory from 2007, except company car.

According to the description of the UDB (Bernard, 2008) and of the target variables (Eurostat, 2008e), the equivalised disposable income is represented by the variable $HX090$ and the equivalised household size by the variable $HX050$. So, Equation (2.1) can be written as:

$$HX090 = \frac{HY020 \times HY025}{HX050} \qquad (2.4)$$

Finally, we consider these variables in the SILC UDB 2004, 2005 and 2006.

2.2   Calculation of EQ_SS and EQ_INC_i

The EU-SILC 2004 database contains 15 countries[4] and the UDB 2005 and 2006 databases contain the data for 26 countries[5]. For each country, we calculate the equivalised household size and the equivalised household income based on the gross components using the two

---

[4]The countries represented in the UDB 2004 are Austria, Belgium, Denmark, Estonia, Spain, Finland, France, Greece, Ireland, Iceland, Italy, Luxembourg, Norway, Portugal and Slovenia.

[5]For the UDB 2005 and 2006, the additional countries are: Cyprus, Czech Republic, Germany, Hungary, Lithuania, Latvia, Netherlands, Poland, Sweden, Slovakia and United Kingdom.

R functions described in Sections 2.2.1 and 2.2.3 (see the accompanying documents for code and more details on the used R functions). Then, in Sections 2.2.2 and 2.2.4, we compare the results obtained by calculations with the provided EU-SILC variables.

2.2.1 Calculation of EQ_SS

We use the R function `eqSS` created by Andreas Alfons (Vienna University of Technology). This function corresponds to Definition (2.3) of the equivalised household size, when there are no missing values in the variable $RX020$ (see below). The arguments of this function are the household identifier (variable DB030), the year of the interview (variable RB010) and the age of each person in the household. For the age, there are two available variables:

1. $RX010$: age at the end of interview (UDB definition);

2. $RX020$: age at the end of the income reference period (UDB definition).

We have to use the second variable, $RX020$, in the calculation of the equivalised household size. Note that missing values in the variable $RX020$ are systematically imputed by the R function as being a person aged 14 and over. For example, for a person with missing age in a household consisting of 5 persons of age, respectively:

$$\left.\begin{array}{l} P_1\text{: 40 years old} \\ P_3\text{: 38 years old} \\ P_3\text{: 10 years old} \\ P_4\text{: NA} \\ P_5\text{: 3 years old} \end{array}\right\} \Rightarrow HM_{14+} = 3 \text{ and } HM_{13-} = 2$$

we have $EQ\_SS = 1 + 0.5(HM_{14+} - 1) + 0.3 HM_{13-} = 2.6$.

2.2.2 Comparison with $HX050$

We can compare the variable provided by the database, i.e. $HX050$, with the one we just have calculated, i.e. EQ_SS. We note that

$$EQ\_SS = HX050$$

for all persons in all countries except Norway.

Indeed, for the latter we note a few cases where EQ_SS is not equal to $HX050$. There are 11 cases (4 households) in the UDB 2004 and 131 cases (35 households) in the UDB 2006. This is because, in each of these households, one or more persons have the variable $RX020$ (age) missing. And so, we fall in the case reported in the calculation of EQ_SS, where the function EQ_SS attributes the age of these persons as being 14 and over.

The problem comes from the fact that countries do not force the missing age to be included in HM_{14+}. Instead, according to Eurostat (2008e), they use another variable which allows us to know if the person, whose age is missing, is included in HM_{14+} or in HM_{13-}. The variable is $RB245$ (Respondent status). So, if we have missing values in $RX020$, then the variable $RB245$ is used in the following way:

$$\begin{cases} \text{if } RB245 = 1, 2 \text{ or } 3, & \text{then } HM_{14+} = 1 \\ \text{if } RB245 = 4, & \text{then } HM_{13-} = 1 \\ \text{otherwise,} & HM_{14+} = HM_{13-} = 0 \end{cases}$$

2.2.3   Calculation of EQ_INC_i

In order to calculate EQ_INC_i, we use the R function `eqInc`, created by Andreas Alfons. Using as arguments the gross income variables from the recommended definition of $HY020$ (see Section 1.4), the function returns as result the calculated total disposable income, which we will denote $HY020'$, divided by EQ_SS. Then, if we multiply this result by the variable $HY025$, present in the UDB, we obtain the total equivalised disposable income:

$$EQ\_INC_i = \frac{HY020' \times HY025}{EQ\_SS}$$

Note that if one of the arguments (i.e. an income variable) is missing, the function considers it equal to zero. Thus, if all income variables are missing, the obtained result is equal to zero:

$$\frac{HY020'}{EQ\_SS} = 0$$

The obtained equivalised disposable income can have the following different values:

1. real value;

2. 0, when all components are missing and $HY025$ is available;

3. NA, when $HY025$ is missing.

2.2.4   Comparison with $HX090$

We want to compare the so obtained equivalised income EQ_INC_i with the variable $HX090$ present in the UDB. In order to that, we define

$$\Delta = \text{abs}(EQ\_INC_i - HX090) \tag{2.5}$$

The comparison is done using two R functions - `critical` and `f_stat`.

Table 2.2: Comparison between EQ_INC_i and $HX090$ for the SILC UDB 2004-2006 based on the gross variables.

| | UDB 2004 | | | UDB 2005 | | | UDB 2006 | | |
country	C_P	C_H	N_P	C_P	C_H	N_P	C_P	C_H	N_P
AT	0	0	11550	0	0	13043	0	0	14883
BE	0	0	12971	0	0	12757	0	0	14329
CY	-	-	-	0	0	11541	0	0	11069
CZ	-	-	-	8	2	10333	8	2	17830
DE	-	-	-	0	0	31276	0	0	31777
DK	0	0	17290	0	0	15321	7	2	14676
EE	0	0	11665	10	3	11948	10	3	15840
ES	NA	NA	44647	NA	NA	37491	4	1	34694
FI	8	3	29070	5	3	29112	0	0	28039
FR	23092	9227	24987	20187	7971	24245	20768	8240	24940
GR	NA	NA	16849	NA	NA	14878	NA	NA	15190
HU	-	-	-	0	0	17969	0	0	19902
IE	0	0	14272	0	0	15539	0	0	14634
IS	0	0	8870	0	0	8840	0	0	8598
IT	NA	NA	61542	NA	NA	56105	NA	NA	54512
LT	-	-	-	0	0	12117	30	9	12134
LU	2	1	9629	0	0	9661	4	1	10242
LV	-	-	-	NA	NA	9699	NA	NA	10986
NL	-	-	-	0	0	23756	0	0	23096
NO	85	26	15868	12	6	15716	133	36	15479
PL	-	-	-	4442	1135	49044	0	0	45122
PT	14101	4961	14159	12878	4615	12878	12071	4367	12071
SE	829	237	14297	360	106	15319	4069	1027	17149
SI	-	-	-	26605	7966	27679	30128	9138	31276
SK	-	-	-	0	0	15418	8	2	15147
UK	-	-	-	0	0	25504	0	0	23365

C_P, C_H: number of persons, households with $\Delta > 1$
N_P: number of persons in the sample

In Table (2.2), we can see if the equivalised disposable income has the same values as the variable $HX090$. Countries with no UDB 2004 are denoted by „-". Countries with no gross income variables, i.e. Spain (UDB 2004-5), Greece, Italy and Latvia have $C_P = C_H = NA$, because no comparison with $HX090$ is possible.

The following countries use the recommended definition (i.e. based on the gross variables): Austria, Belgium, Cyprus, Germany, Hungary, Ireland, Iceland, Netherlands, Poland (UDB 2006) and United Kingdom. They have:

$$EQ\_INC_i = HX090$$

$$HY020' = HY020$$

The same definition is used by the Czech Republic, Denmark, Estonia, Finland, Lithuania, Luxembourg, Norway and Slovakia except for very few cases where EQ_INC_i is not equal to $HX090$.

For Poland (UDB 2005), the equivalised disposable income was taken from another source (see CENTRAL STATISTICAL OFFICE OF POLAND, 2007, 2008).

Some countries use alternative definitions of $HY020$ to compute the equivalised disposable income (EUROSTAT, 2009a):

Second definition of $HY020$

France, Italy, Spain[6], Greece, Latvia, Slovenia and Sweden used an alternative definition of $HY020$, based on the net variables, see EUROSTAT (2009a).

The net variables present in Definition (2.6) have the same meaning as those of Definition (2.1) except the variable $HY145N$ which does not exist as gross. It corresponds to repayments/receipts for tax adjustment.

$$
\begin{aligned}
HY020 = {}& + HY040N & - HY120N & \quad + PY010N \\
& + HY050N & - HY130N & \quad + PY020N \\
& + HY060N & - HY145N & \quad + PY050N \\
& + HY070N & & \quad + PY090N \\
& + HY080N & & \quad + PY100N \\
& + HY090N & & \quad + PY110N \\
& + HY110N & & \quad + PY120N \\
& & & \quad + PY130N \\
& & & \quad + PY140N
\end{aligned}
$$

$$(2.6)$$

Note that for France, the variable $PY020N$ (company car) is included in $PY010N$. For Italy, the variable $PY120N$ is missing because it could not be separated from cash or near cash employee income and is recorded under this variable (i.e. $PY010N$).

When we use as arguments of the function eqInc the net income variables composing the second definition of $HY020$, the function returns as result the calculated total disposable income, which we will denote $HY020''$, divided by EQ_SS. Thus, in this case, we find that France, Spain, Greece, Latvia and Sweden have:

$$
\begin{aligned}
EQ\_INC_i &= HX090 \\
HY020'' &= HY020
\end{aligned}
$$

---

[6]In the UDB 2004-2005, Spain uses the second definition of $HY020$ to compute EQ_INC_i, but in the UDB 2006, it is not clear if Spain used the recommended or the alternative definition of $HY020$.

In Slovenia (UDB2005-6) (see Statistical Office of the Republic of Slovenia, 2007a,b) and Italy (UDB 2004), we have $HY020'' = HY020$, but $EQ\_INC_i \neq HX090$, because other sources are used.

Third definition of $HY020$

The definition used by Portugal corresponds to the third definition given in Eurostat (2009a) (Section *Detailed methodological notes*), i.e. the total disposable household income is computed as *the sum for all household members of personal income components plus income components at household level*,

- *of which some are net (net of income tax, net of social contributions or net of both) and others gross,*

- *or all of them net but some of them net of tax at source, others net of social contributions or net of both, once the tax on income and social insurance contributions (HY140N), the regular taxes on wealth, the regular inter-household cash transfer paid and the employers' social insurance contributions are deducted.*

So Portugal uses the following definition of the total disposable household income:

$$
\begin{aligned}
HY020 = &+ HY040N &&- HY120N &&+ PY010N \\
&+ HY050N &&- HY130N &&+ PY020G \\
&+ HY060N &&&&+ PY050N \\
&+ HY070N &&&&+ PY090N \\
&+ HY080N &&&&+ PY100N \\
&+ HY090N &&&&+ PY110N \\
&+ HY110N &&&&+ PY120N \\
&&&&&+ PY130N \\
&&&&&+ PY140N
\end{aligned}
\tag{2.7}
$$

The variable $HY145N$ is not available and the variable $PY020N$ neither but in this case it is the gross component, $PY020G$, that is used. So, when we use as arguments of the function `eqInc` the income variables from the third definition of $HY020$, the function returns as result the calculated total disposable income, which we will denote $HY020_{PT}$, divided by EQ_SS. Thus, in this case, we find

$$
\frac{HY020_{PT} \times HY025}{HX050} = HX090
$$

where $HY020_{PT} = HY020$.

2.3   Conclusion

This study on equivalised disposable income clarifies the definition of the total disposable income ($HY020$) used by each country in the EU-SILC 2004-2006.

Austria, Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, Germany, Hungary, Ireland, Iceland, Lithuania, Luxembourg, Netherlands, Norway, Poland (UDB 2006), Slovakia and United Kingdom use the recommended definition of $HY020$, defined in (2.1), and have $EQ\_INC_i = HX090$ (see Table (2.2)). For Poland (UDB 2005), the equivalised disposable income was taken from another source. Eurostat recommends to use this definition, but some countries use alternative definitions.

Indeed, Greece, Spain, France, Italy, Latvia, Slovenia and Sweden use the second definition of $HY020$, defined in (2.6), where only the net income variables are used, except for Italy (UDB 2004) and Slovenia (UDB2005-6) who use other sources.

Finally, Portugal uses a mix between net and gross income components to compute the total disposable income $HY020_{PT}$ defined in (2.7) and obtains $EQ\_INC_i = HX090$.

Note that, Greece, Spain, France, Italy and Portugal are supposed to deliver gross income data as from 2007, which is an important step towards the harmonisation of the survey.

Chapter 3

Summary and recommandations

In this chapter, we present a short summary of the previous chapters and give our general remarks and recommendations.

3.1 Summary

In Chapter 1, we revised and summarized the existing documentation on EU-SILC, i.e. description of the user data base, national quality reports, EUROSTAT quality reports, EUROSTAT guidelines, documents about the social inclusion, pensions and healthcare indicators. We presented a detailed bibliographical record.

In Chapter 2, we presented the different manners in which the equivalised disposable income is calculated in the EU-SILC participating countries. Using R functions created at TUWIEN and SFSO, we calculated the equivalised household size and the equivalised total disposable income for each country and we compare the results with the corresponding variables present in the UDB. Peculiarities in the data and differences between countries have been investigated.

3.2 Recommendations

Conforming to the initial aims of Work Package 5, we will therefore present some basic recommendations, which could be helpful for the simulation study. These recommendations include:

- the sampling designs,

- the income variables.

The EU-SILC Commission Regulation on sampling and tracing rules EUROSTAT (2003a) suggests the selection of nationally representative probabilistic samples. A table summarizing the sampling designs by country is given in Eurostat's Intermediate Quality Report

2007 (see EUROSTAT, 2009d). We can note that, in 2007, we have either sampling of addresses, or sampling of households or sampling of individuals. Although, there are still many differences between the countries (sampling frame, sampling designs), we can note that the most used sampling design is the stratified multi-stage sampling design. A summary of all used sampling designs (2004), written by Tobias Schoch (FHNW), is given in the Appendix. More information can be found in the above cited report. In general, it would be rather restrictive to recommend a single sampling design, as each country uses specific designs. We are only able to point out the large use of stratified multi-stage sampling.

In EUROSTAT (2009d), Annex 5, Table 13, we can see whether the EU-SILC countries use the recommended definition for the household income components, i.e. the gross income components. For the variable $HY020$, which is in the core of the calculation of the equivalised disposable income, see Chapter 2, we can note that the definitions used by all countries, in 2007, are either fully or largely comparable to the recommended definition.

For the simulations, it is not really possible to use the gross income variables. What we can recommend is that a sufficient number of the principal income components be used for approximately reproducing the empirical gross total equivalised income.

3.3   Acknowledgements

We would like to acknowledge the contributions of Andreas Alfons (TUW), who has written the R code for calculating european Laeken indicators, given in a supplement to this document, and Tobias Schoch (FHNW) who has written a summary on EU-SILC 2004 samling designs, given in the Appendix.

Bibliography

Ardilly, P. and **Lavallée, P.** (**2007**): *Weighting in rotating samples: The SILC survey in France.* Survey Methodology, Vol. 33 (No. 2), pp. pp. 131–137.
URL http://www.statcan.ca/english/freepub/12-001-XIE/2007002/article/10490-en.pdf

Bernard, B. (**2008**): *EU-SILC User Database Description.* Version 2006-1, European Commission. Directorate F: Social Statistics and Information Society, Luxembourg.

Bernard, B. (**2009**): *EU-SILC User Database Description.* Version 2007-2, European Commission. Directorate F: Social Statistics and Information Society, Luxembourg.

Central Statistical Office of Poland (**2007**): *EU-SILC 2005 Operation.* Final Quality Report, Central Statistical Office of Poland, Warsaw.

Central Statistical Office of Poland (**2008**): *EU-SILC 2006 Operation.* Final Quality Report, Central Statistical Office of Poland, Warsaw.

Clemenceau, A. and **Museux, J.-M.** (**2007**): *EU-SILC (Community Statistics on Income and Instrument).* Comparative EU Statistics on Income and Living Conditions : Issues and Challenges. Proceedings of the EU-SILC conference (Helsinki, 6-8 November 2006), pp. 11–36.
URL http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-007/EN/KS-RA-07-007-EN.PDF

Eurostat (**2003**a): *Commission Regulation (EC) No 1983/2003.* Report, European Commission. Directorate F: Social Statistics and Information Society.
URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:298:0034:0085:EN:PDF

Eurostat (**2003**b): *First Ideas on Weighting.* Technical report, European Commission. Directorate E: Social statistics. Meeting of the Task Force. Statistics on Income and Living Conditions.

Eurostat (**2003**c): *First Ideas on Weighting.* EU-SILC 123/03 Addendum, European Commission. Directorate E: Social statistics. Meeting of the Task Force. Statistics on Income and Living Conditions, Luxembourg.

Eurostat (**2004**a): *Commission Regulation (EC) No 28/2004.* Report, European Commission. Directorate F: Social Statistics and Information Society.
URL http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:005:0042:0056:EN:PDF

Eurostat (**2004**b): *Cross-Sectional Weighting: First Year of Each Sub-Sample.* EU-SILC 134/04, European Commission. Directorate D: Single Market, Employment and Social Statistics. Meeting of the Working Group, Statistics on Income and Living Conditions Statistics on Income and Living Conditions.

Eurostat (**2004**c): *First Ideas on Weighting for Child Care Data.* Technical report, European Commission. Directorate D: Single Market, Employment and Social statistics. Meeting of the Task Force, Statistics on Income and Living Conditions.

Eurostat (**2005**): *Cross-Sectional Weighting: From second year of the survey onwards.* Technical report, European Commission. Directorate D: Single Market, Employment and Social statistics.

Eurostat (**2008**a): *Comparative Final EU Quality Report 2005 (Version 2).* Report, European Commission. Directorate F: Social Statistics and Information Society.
URL   http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/quality_assessment/comparative_quality_1/comparative_2005/_EN_1.0_&a=d

Eurostat (**2008**b): *Description of Secondary Target Variables: Module 2007.* EU-SILC 065/Annex, European Commission. Directorate F: Social statistics and information society.

Eurostat (**2008**c): *Description of SILC User Database Secondary Target Variables, Module 2005: Inter-generational transmission of poverty. Version 2005.2.* Technical report, European Commission. Directorate E: Social and regional statistics and geographical information system.

Eurostat (**2008**d): *Description of SILC User Database Secondary Target Variables: Module 2006: Social Participation. Version 2006.1.* Technical report, European Commission. Directorate E: Social and regional statistics and geographical information system.

Eurostat (**2008**e): *Description of SILC User Database Variables: Cross-sectional and Longitudinal.* Version 2006.1, European Commission. Directorate F: Social Statistics and Information Society.
URL   http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/data_dissemination/udb_user_database/variables_descriptiondoc/_EN_1.0_&a=d

Eurostat (**2008**f): *Harmonisation of SILC and EHIS Questions on Health. Guidelines for the Health Questions in SILC from 2008 Operation Onwards.* Technical report, European Commission. Directorate F: Social statistics and Information Society. Unit F-3: Living conditions and social protection statistics. Unit F-5: Health and food safety statistics.

Eurostat (**2009**a): *Algorithms to compute Overarching indicators based on EU-SILC and adopted under the Open Method of Coordination (OMC).* Technical report, European Commission. Directorate F: Social Statistics and Information Society.
URL   http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/data_dissemination/42_indicators/methodology_040209/_EN_1.0_&a=d

Eurostat (**2009**b): *Algorithms to compute Pensions Indicators based on EU-SILC and adopted under the Open Method of Coordination.* Doc LC/18/07/EN, European Commission. Directorate F: Social statistics and information society.
URL http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/data_dissemination/42_indicators/pensions_portfolio/_EN_1.0_&a=d

Eurostat (**2009**c): *Comparative Final EU Quality Report 2006 (Version 3).* Report, European Commission. Directorate F: Social Statistics and Information Society.
URL http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/quality_assessment/comparative_quality_1/final_euqr06_version/_EN_1.0_&a=d

Eurostat (**2009**d): *Comparative Intermediate EU Quality Report 2007 (Version 5).* Report, European Commission. Directorate F: Social Statistics and Information Society.
URL http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/quality_assessment/comparative_quality_1/intermediate_euqr07/_EN_1.0_&a=d

Eurostat (**2009**e): *Description of SILC User Database Variables: Cross-sectional and Longitudinal.* Version 2007.2, European Commission. Directorate F: Social Statistics and Information Society.

Eurostat (**2009**f): *Description of Target Variables: cross-sectional and longitudinal.* EU-SILC 065 (2009 operation), European Commission.
URL http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/guidelines_questionnaire/operation_guidelines_1/silc065_highlightedpdf/_EN_1.0_&a=d

Eurostat (**2010**): *Description of Target Variables: cross-sectional and longitudinal.* EU-SILC 065 (2008 operation), European Commission.
URL http://circa.europa.eu/Public/irc/dsis/eusilc/library?l=/guidelines_questionnaire/operation_guidelines/silc065_version/_EN_1.0_&a=d

Federal Statistical Office Germany (**2007**): *Quality Report on EU-SILC 2005.* Final Report, Federal Statistical Office Germany.

Federal Statistical Office Germany (**2008**): *Quality Report on EU-SILC 2006.* Intermediate Report, Federal Statistical Office Germany.

Figari, F., **Levy, H.** and **Sutherland, H.** (**2007**): *Using the EU-SILC for Policy Simulation: Prospects, Some Limitations and some Suggestions Perspective.* Comparative EU Statistics on Income and Living Conditions : Issues and Challenges. Proceedings of the EU-SILC conference (Helsinki, 6-8 November 2006), pp. 347–368.
URL http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-007/EN/KS-RA-07-007-EN.PDF

Graf, E. (**2009**): *Weightings of the Swiss Household Panel SHP_I wave 9, SHP_II wave 4, SHP_I and SHP_II combined.* Methodology Report 338-0058 / ISBN: 978-3-303-00425-8, Swiss Federal Statistical Office.
URL http://www.bfs.admin.ch/bfs/portal/en/index/news/publikationen.Document.126386.pdf

Kerm, P. V. (**2007**a): *Extreme Incomes and the Estimation of Poverty and Inequality Indicators from EU-SILC.* IRISS Working Papers Series, IRISS.
URL http://iriss.ceps.lu/documents/irisswp69.pdf

Kerm, P. V. (**2007**b): *Extreme Incomes and the Estimation of Poverty and Inequality Indicators from EU-SILC.* Comparative EU Statistics on Income and Living Conditions : Issues and Challenges. Proceedings of the EU-SILC conference (Helsinki, 6-8 November 2006), pp. 83–94.
URL http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-07-007/EN/KS-RA-07-007-EN.PDF

Leiten, E. and **Traat, I.** (**2008**): *Variance of Laeken Indicators in Complex Surveys.* Technical report, Statistical Office of Estonia.

Osier, G. and **Museux, J.-M.** (**2006**): *Variance Estimation for EU-SILC Complex Poverty Indicators using Linearization Techniques.* Technical report, European Conference on Quality in Survey Statistics, Proceedings of Q2006.

Osier, G., **Museux, J.-M.**, **Seoane, P.** and **Verma, V.** (**2006**): *Cross-sectional and Longitudinal Weighted for the EU-SILC Rotational Design.* Working paper, Universitá degli Studi di Siena.
URL   http://www.iser.essex.ac.uk/files/survey/ulsc/methodological-research/mols-2006/scientific-social-programme/papers/Osier.pdf

Särndal, C.-E., **Swensson, B.** and **Wretman, J.** (**1992**): Model Assisted Survey Sampling. New York: Springer.

Statistical Office of the Republic of Slovenia (**2007**a): *Final Quality Report EU-SILC 2005 (draft).* Final Report, Statistical Office of the Republic of Slovenia.

Statistical Office of the Republic of Slovenia (**2007**b): *Intermediate Quality Report: EU-SILC 2006 (draft).* Intermediate Report, Statistical Office of the Republic of Slovenia, rihard Inglic, Marta Arnez, Rudi Seljak, Matija Remec, Stanka Intihar, Martina Stare.

Statistics Austria (**2004**): *Quality Report relating to EU-SILC 2003 Operation.* Technical report, Statistics Austria, Vienna.

Statistics Austria (**2005**): *Quality Report relating to EU-SILC 2004 Operation.* Technical report, Statistics Austria, Vienna.

Statistics Austria (**2007**): *Final Technical and Financial Implementation Report Relating to the EU-SILC 2005 Operation.* Final Report, Statistics Austria, Vienna.

Statistics Austria (**2008**a): *Final Quality Report relating to the EU-SILC Operation 2004-2006.* Final Report, Statistics Austria, Vienna.

Statistics Austria (**2008**b): *Intermediate Quality Report relating to the EU-SILC 2007 Operation.* Intermediate Report, Statistics Austria.

Statistics Estonia (**2006**): *EU-SILC 2004 in Estonia: Final Quality Report.* Final Report, Statistics Estonia.

Statistics Estonia (**2007**a): *EU-SILC 2005 in Estonia: Final Quality Report.* Final Report, Statistics Estonia, Tallinn.

Statistics Estonia (**2007**b): *EU-SILC 2006 in Estonia: Intermediate Quality Report.* Intermediate Report, Statistics Estonia, Tallinn.

Statistics Finland (**2005**): *Intermediate Quality Report of the Finnish EU-SILC 2004.* Intermediate Report, Statistics Finland.

Statistics Finland (**2007**): *Intermediate Quality Report relating to the EU-SILC 2006 Operation in Finland.* Intermediate Report, Statistics Finland, pauli Ollila, Marie Reijo, Hannele Sauli.

Statistics Finland (**2008**): *Final Quality Report relating to the EU-SILC 2005 operation.* Final Report, Statistics Finland, pauli Ollila, Marie Reijo, Hannele Sauli.

Verma, V. (**2006**): *Output II.1(b). EU-SILC weighting procedures: an outline.* Technical report, ISR: International Social Research.

Verma, V. (**2007**): *Issues in Data Quality and Comparability in EU-SILC.* Comparative EU Statistics on Income and Living Conditions : Issues and Challenges. Proceedings of the EU-SILC conference (Helsinki, 6-8 November 2006), pp. 287–309.

Verma, V., **Betti, G.** and **Ghellini, G.** (**2007**): *Cross-Sectional and Longitudinal Weighting in a Rotational Household Panel: Applications to EU-SILC in Statistics.* Statistics in Transition-new series, 8 (1), pp. 5–50.

Appendix A

EU-SILC sampling designs 2004

Table A.1: Summary of the 2004 SILC Sampling Design

Country	(P)SU	$p_1(\cdot)$	SSU	$p_2(\cdot)$
Simpe random cluster samples				
Austria	household address	SI [households] (PPS no. hh members); interview of all household memebers		
Iceland	persons aged 16 and more (register)	SI [persons]; interview of all household members		
Malta	households	SI [households] (PPS no. hh members); interview of all household members		
Stratified random (cluster) samples				
Cyprus	STSI [geographical regions]	SI [households]		
Estonia	ST [county size] SY	SY [persons]; interview of all household members		
Lithuania	STSI [urbanization]	SI [households]		
Slovakia	STSI [NUTS3] and [urbanization]	SI [dwellings]		
Two-stage designs				
Belgium	municipalities	ST [NUTS2], pps [no. household]	households	SY (40 hh per PSU)
Finland	persons from register (Master sample)	SY [persons] (atypical PSU)	dwellings	STSI [socio-econ criteria]
Greece	blocks	ST [NUTS2] and [urbanization], pps [no. dwellings]	addresses	SY
Hungary	1) localities/ counties	ST [election district] and [no. dwellings], pps [no. dwellings]	dwellings	SY
	2) blocks	ST [election district] and [no. dwellings], SY		

Table A.1 (continued)

Country	(P)SU	$p_1(\cdot)$	SSU	$p_2(\cdot)$
Two-stage designs				
Czech Republic	census Enumeration units	ST [NUTS4] and [no. dwellings], pps	dwellings	SI/SIR [dwellings]
Ireland		STSI [county] and [urbanization]	households	SI/SIR
Italy		ST [admin region] and [no. residents], pps [no. residents]	households	SY
Latvia	census areas	ST [urbanization], pps [no. households]	addresses (households)	SI/SIR
The Netherlands	municipalities	ST[geographical criteria], pps [no. addresses]	addresses	SI/SIR
Norway	municipalities	ST [socio-econ criteria], pps [population size]	persons	SY (special: self-weighted design) All the households the selected persons belong to were then interviewed.
Poland	census areas	ST [NUTS2] and [urbanization], pps [no. dwellings]	dwellings	SI/SIR
Portugal	census sections from master sample	SY	households	SI/SIR
Slovenia	enumeration areas	ST [size of settlement] and [proportion of agricultural hh], SY	persons	7 persons per strata. All the households the selected persons belong to were then interviewed.
Spain	census sections	ST [admin region] and [no. dwellings], pps [no. dwellings]	dwellings	SY
United Kingdom	postcode sectors	ST	addresses	(composed from several samples)

Notation: hh: household; SI: simple random sampling without replacement; SIR: simple random sampling with replacement; SY: systematic sampling; ST: stratified sampling; SIC simple random cluster sampling, $STSI$: stratified sampling with SI sampling in each stratum; *pps*: probability-proportional-to-size with replacement; πps: probability-proportional-to-size without replacement (Särndal et al. (1992)).

Reference: Table is based on Bernard (2008)

Notes: The countries Denmark, France, Germany, Luxembourg and Sweden have a special sampling design in the 2004 SILC exercise (not shown).