

# Introduction to the European Union Statistics on Income and Living Conditions (EU-SILC)

Dr Alvaro Martinez-Perez  
ICOSS Research Associate

1. EU-SILC data
2. Data Quality Issues
3. Issues working with EU-SILC
4. Access to the data

" Euro SILC (EU-SILC) Conditions (EU-SILC), launched in 2003 is the first ex-post harmonised micro-data set to provide comprehensive data on incomes and a large number of other social and economic domains across all 28 member states of the enlarged EU (it also covers a number of other countries)



## " Two modules

Cross-sectional module (2004 onwards): Intergenerational transmission of poverty, Social participation, Housing conditions, Over-indebtedness and financial exclusion, Material deprivation, Intra-household sharing of resources,

## " Target population

All private households and their current members residing in the territory of the countries at the time of data collection. All household members are surveyed, but only those aged 16 and more are interviewed

## " Sample size (minimum size of the overall population surveyed every year):

XS: 130,000 HH and 270,000 individuals aged 16 and more are interviewed

L: ABOUT 100,000 HH and 200,000 individuals aged 16 and more interviewed

- " The XS and L files are released separately in the EU-SILC data. Member states are allowed to use different survey instruments to collect XS and L data, and there is no requirement that these data sets be linkable. Indeed, they are not because even when cross-file identifiers to allow linkage are provided by the National Statistical Offices these are changed by Eurostat prior to the public release of the data



''





" Exceptions to the four-years rule:

France: nine-year panel

Norway: eight-year panel

VE, owl y€|qDz €|o, z kxov

" Exceptions to the rule of household surveys:

Finland, the Netherlands, Norway, Slovenia, and Sweden have based data collection on administrative registers, using registers to collect several variables, and obtaining other info via interviews

f sr k ·|oz |o}ox -k-\$ o zo|}yx, x -r o r y€}or yv



## " Sampling and data collection:

DE: Data collected using a mix of quota and random samples. Quota samples should not be used to infer info about the population. No indicator of sample type in the data set

ES & IE use substitutes for non-respondents. Non-respondent substitution undermines the probability nature of the sample. Ideally, substitutes should not be used but the data sets do not contain an indicator to identify them

RECOMMENDATION: Data from these countries should not be used in statistical analyses with the aim of inferring info about national populations

" Sampling and data collection:

EU-SILC does not report the year in which each sample member was selected. It appears that some sample members of the same rotational group were first interviewed in a different year than others but not possible to identify them

" Complex sample design indicators:

In order to obtain correct estimates sample design needs to be



## " Weighting:

Weights are provided with the EU-SILC data. However, the documentation is unclear on whether these are only design weights (correcting for the probability of selection into the sample) or whether they adjust for non-response, and if so, in what countries.

It appears that some countries correct for non-response but they do so in a non-consistent way across countries.

Different treatment of non-response, and especially correcting for it in some countries but not others, may lead to biased cross-country comparisons

## " Relationships between household members

EU-SILC does not provide a household grid (a series of variables documenting the relationship between each member of the household and each of the other members)

EU-SILC provides only three variables: the personal identifiers of each individual, his or her spouse or partner, and his or her mother and father, where these are resident in the same household (able to identify which people are living as part of a couple, and/or with their children or parents)

This limitation means that the nature of many relationships cannot be established

## " The parent/child relationship

For each individual with one or more parents living in the same household, the relationship is recorded in the household grid.

However no distinction is made between biological parents, adoptive parents, foster parents, and step parents

If a full household grid were to be collected making the proper distinctions, this problem would be solved



## " Issues with the longitudinal module

The EU-SILC documentation states the rule of following up those sample members aged 16 and over should be re-interviewed if they leave the original household and start living somewhere else. However the implementation of this rule has varied widely from country to country, and has not been particularly comprehensive anywhere.

There are two groups of people who may be particularly affected: young adults and people who divorce or separate



## " Issues with the longitudinal module

This means that, although the percentage of individuals and households followed from year to year in the longitudinal sample is not bad, the percentage of individuals followed on leaving their family home is extremely low, making the EU-SILC in its current form unsuitable for analysing transitions for some of the groups of most interest to social scientist: young adults and separating couples





## " Income aggregation

The NSAs gather info about incomes from a large number of sources. However, in the data released by EU-SILC, these incomes are aggregated into a much smaller number of variables

(+) Increases comparability across countries

(-) Decreases the level of detail in the data

(-) As benefit systems differ between countries, the income sources contained in each of the aggregate variables vary between countries, and it is not always clear from the documentation what the components of the aggregate variables are

(-) Some income components are reported at the household level, not allowing the analyst to distinguish which individuals receive them and how much each receives



- " Reference period mismatch between income and non-income information
  - 3. Temporal mismatch between the income reference period and the current reference period. Almost all other data in EU-SILC refers to the moment in which the interview takes place. This affects the reliability of any analysis focusing on the relationship between income and any other variable
- To overcome this problem linking income variables at  $t$  with all other at  $t-1$  would mean to lose one wave which for this short panel is problematic



## " Net *versus* gross income

Information on gross incomes has been recorded since 2007:

1. Quality of info is not uniform across countries: data on income components may be collected either gross or net of taxes-3( )-25(h)4

## Contract

" UoS EU-SILC data received contains the XS (2004-2011) and L (2005-2010)

" Contract signed with Eurostat until 31/12/2015. New releases of data should be made available until then

" After that date the contract needs to be amended to extend its validity





The  
University  
Of  
Sheffield

