Title:

Validity and cross-country comparability of the EU-SILC income variables

Authors:
Tim Goedemé* and Lorena Zardo Trindade

Herman Deleeck Centre for Social Policy – University of Antwerp
St. Jacobstraat 2
2000 Antwerp
Belgium

* Author for correspondence: tim.goedeme@uantwerpen.be

Abstract

The European Union Statistics on Income and Living Conditions (EU-SILC) are currently the most important comparative survey on income statistics in Europe. The survey is regularly used for studying income patterns, poverty and income inequality in all EU Member States and other countries. Apart from analysing total household income, there is much interest in the distribution of particular income components. For this reason, it is essential to cross-country comparative analyses that not only total household incomes are well harmonised, but also the various target variables related to more specific income components. Given the absence of a database which describes how national income components are aggregated into the target variables for all EU-SILC countries, researchers cannot easily evaluate whether target variables are comparable for their research purposes. As a result, even though general and country-specific descriptions of income target variables are available in the EU-SILC methodological guidelines (Doc065) and in the national quality reports, it is often not clear how each of the national income components should be classified and aggregated into a target variable. Therefore, when analysing a particular income variable in EU-SILC in a cross-national context, researchers might have difficulties in finding out which income components are being compared, potentially biasing the results, undermining their validity, and complicating the interpretation of the survey estimates.

In this paper, we study how individual income components are aggregated into the EU-SILC target variables. In particular, we look at compliance with Eurostat guidelines, misclassifications and omitted income sources, all undermining cross-national comparability. In addition, we look into the setup of the new, more disaggregated income variables, and check whether these are constructed in a more consistent way. On the basis of a survey among national statistical institutes, an analysis of the national quality reports, the comparative quality reports, and an analysis of the EU-SILC data, we are compiling a database which documents the exact classification of income components into the EU-SILC target variables. The focus of the database is on EU-SILC 2015, covering as many EU-SILC countries as possible. The database will contain information on the composition of variables on total income before and after transfers; income from benefits, work and capital; social contributions and taxes. Special attention is given to self-employment income, imputed rent and income from

production for own consumption, as well as outlier detection and data error correction. The database is constructed in the context of the Net-SILC 3 project, and will be made freely available.

For each of the income components, the database contains the official name (national language) and code in the national EU-SILC survey; the equivalent name in English; the target variable code and name; the source of the income information used (register data, questionnaire, imputation); the level of aggregation when it was collected; information on gross-net conversion; whether there have been important changes between wave 2010 and wave 2015 or if there will be important changes planned for future waves. The database allows researchers to identify which income components are covered in EU-SILC and how they have been classified into the EU-SILC target variables.

On the basis of this exercise, we will draw some general conclusions with regard to (1) possibilities for improving the definition of target variables; (2) the most appropriate classification of national income components (e.g. particular benefits) that can be considered 'borderline cases' which are currently classified inconsistently across countries; (3) the (unjustified) exclusion of some income components from EU-SILC target variables; (4) cross-national deviations with regard to the calculation of the EU-SILC total income variables; (5) recommendations for Eurostat to update and improve our new database in the future. To conclude, we are strongly convinced this new database will be helpful for many researchers using EU-SILC and will contribute to increasing the transparency and validity of cross-country comparisons.