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

Lorena Zardo Trindade

Tim Goedemé

Herman Deleeck Centre for Social Policy

University of Antwerp

University
of Antwerp

Outline

Motivation

What is MetaSILC 2015?

Survey

Database

Preliminary analysis

What did we find for 2010?

Recommendations

Next steps

Motivation

EU-SILC is currently the most important **comparative** survey on income statistics in Europe.

- Essential information source for analysing the distributive effects of the welfare state

Growing interest in the distributive effects of particular benefits

- Social Investment Package
- Country-specific recommendations
- Children / Adequacy & sustainability of pension systems
- …

> **It is crucial to cross-country comparative analyses that total household incomes and the various target variables related to more specific income components are well harmonized.**

Motivation

Knowledge of content (aggregation) and comparability of income variables is key

Description of target variables in Doc065 & Quality reports not sufficiently detailed for

- Identifying exact classification of all income components in all countries
- Evaluating level of cross-country comparability ('correct' classification) depends also on question)

Researchers often have difficulties in finding out which income components are being compared

- potential biased the results
- weakened validity
- interpretation of the survey estimates becomes complicated

➢ **Researchers cannot easily evaluate whether target variables are comparable for their research purposes**

What is MetaSILC 2015?

Survey among national statistical institutes/ EU-SILC data producers

- cross-sectional 2015 wave
- covering as many EU-SILC countries as possible

Report and database describing how national income components are aggregated into the EU-SILC target variables

- information on the composition of variables on **total income before and after transfers**; income from **benefits, work and capital**; **social contributions and taxes**
- outlier detection and data processing

Focus on undermining factors for cross-national comparability

- the compliance with Eurostat guidelines
- misclassifications and omitted income sources

➢ **Identify possibilities for improving the definition of target variables and recommendations for updating and improve our new database in the future**

Survey

Online questionnaire in 2 rounds

Questions on all 34 income variables
- Income from benefits with information on "mixed components" or "contributory and means-tested, contributory and non-means-tested, non-contributory and means-tested, non-contributory and non-means-tested components"

For each of the income components
- official name (national language) and the equivalent name in English
- the target variable code and name
- the source of the income information used (register data, questionnaire, imputation)
- information on gross-net conversion
- changes between wave 2010 and wave 2015
- changes planned for future waves
- additional questions on data processing of specific variables (HY030, PY050, PY021)

Database

Information provided
Information not provided, EU-SILC 2015 available
Information not provided, EU-SILC 2015 availability not confirmed

Excel file with information on 25 countries

The exact composition of all income variables of EU-SILC cross-sectional 2015 wave

- 34 variables, over 2000 income components

Latvia, Poland and Sweden

- Income from benefits with information only on "mixed components"

Database

Preliminary analysis

The use of register data in some countries and survey data in other countries may affect comparability across countries

Changes in 2010-2015

- (11 countries) Belgium, Cyprus, Czech Republic, Denmark, Estonia, Latvia, Poland, Portugal, Sweden, Slovenia, United Kingdom
- Most variables, except for PY020, PY021, PY080

Future changes

- (11 countries) Austria, Belgium, Czech Republic, Germany, Estonia, France, Malta, Netherlands, Poland, Spain, Slovenia
- Most variables

Total income before and after transfers (HY010, HY020, HY022, HY023)

Overview of the countries in compliance with the standard definition according to MetaSILC 2010 and 2015

Country	2010				2015			
	HY010	HY020	HY022	HY023	HY010	HY020	HY022	HY023
EU Member States								
Austria								
Belgium								
Bulgaria								
Croatia								
Cyprus								
Czech Republic								
Denmark								
Estonia								
Finland					-	-	-	-
France								
Germany	-	-	-	-				
Greece								
Hungary								
Italy	-	-	-	-				
Latvia								
Luxembourg								
Malta	-	-	-	-				
Netherlands								
Poland								
Portugal	-	-	-	-				
Slovakia								
Slovenia	-	-	-	-				
Spain								
Sweden								
United Kingdom								
Other countries								
Republic of Serbia	-	-	-	-				

Source: MetaSILC 2010 and 2015 Database.

Most common issues **reported** for 2015

- Belgium (HY020, HY022, HY023) - regular taxes on wealth (HY120G) is not collected because tax on wealth does not exist in the country

- France - net income of social contributions is retained (HY020, HY022, HY023)

- Netherlands (HY022, HY023) – Eurostat formula is used, but tax on income and social contributions (HY140G) is corrected. Taxes paid on income components that are excluded from HY022 and HY023 are deducted from HY140G

Total income before and after transfers (HY010, HY020, HY022, HY023)

Country	2010				2015			
Overview of the countries in compliance with the standard definition according to MetaSILC 2010 and 2015	HY010	HY020	HY022	HY023	HY010	HY020	HY022	HY023
EU Member States								
Austria								
Belgium								
Bulgaria								
Croatia								
Cyprus								
Czech Republic								
Denmark								
Estonia								
Finland					-	-	-	-
France								
Germany	-	-	-	-				
Greece								
Hungary								
Italy	-	-	-	-				
Latvia								
Luxembourg								
Malta	-	-	-	-				
Netherlands								
Poland								
Portugal	-	-	-	-				
Slovakia								
Slovenia	-	-	-	-				
Spain								
Sweden								
United Kingdom								
Other countries								
Republic of Serbia	-	-	-	-				

Source: MetaSILC 2010 and 2015 Database.

Most common issues **reported** for 2015 *(continuing)*

- Denmark (HY010, HY023) - pension from individual private plans (PY080) is part of PY100 (old age benefits). It does not affect values for HY010, but affects for HY023

- Republic of Serbia (HY010, HY022, HY023) - Company car is (PY021) imputed as missing

- Spain (HY022, HY023) and Slovenia (HY020, HY022, HY023) – used formula based on net amounts instead of gross (Definitions Regulation (No 1980/2003))

Sickness benefits (PY120, PY121, PY122, PY123, PY124)

18 seem to fully comply with the Eurostat definition for Sickness benefits

- countries still include inappropriate components as maternity benefits (should be included in HY050) and other benefits in the computation of Sickness benefits
- Bulgaria, Denmark, France, Greece, Latvia, Poland and Portugal do not fully comply to the Eurostat definition

The components reported seem to have different level of aggregation across countries

- Czech Republic, Estonia, Hungary, Netherlands and Slovenia seem to collect information at a more aggregated level than the other countries consulted.

Eight countries have reported changes in the type of data collection and inclusion and exclusion of components, which may affect comparability across time
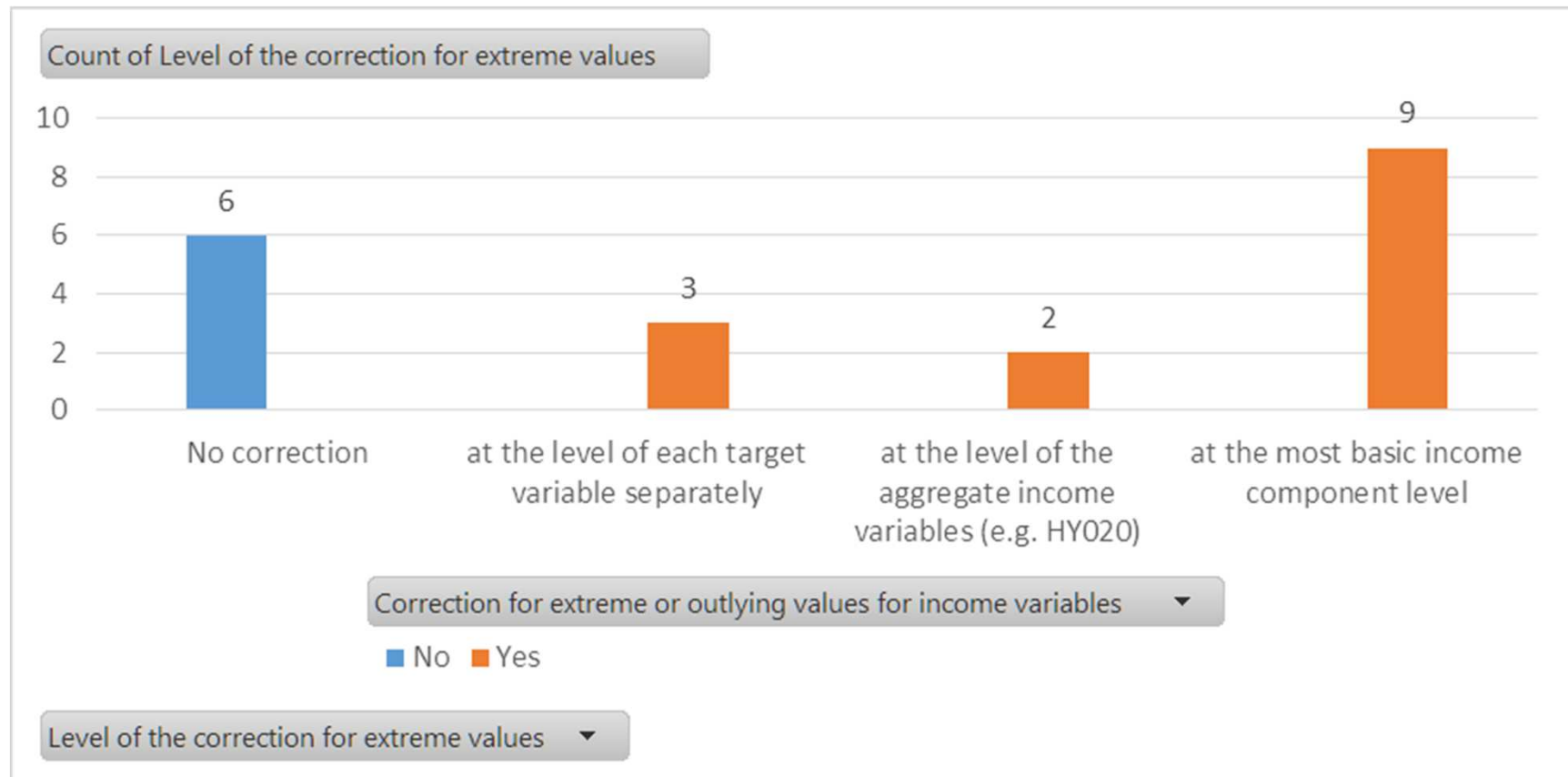
Two countries reported future changes on the computation of sickness benefits

- Belgium (data will be collected using register data) and Poland (apply ESSPROSS classification on all benefits)

Tax on income and social contributions (HY140G/HY140N)

(to be included)

Outlier treatment

Count of Level of the correction for extreme values

	No correction	at the level of each target variable separately	at the level of the aggregate income variables (e.g. HY020)	at the most basic income component level
	6	3	2	9

Correction for extreme or outlying values for income variables

■ No ■ Yes

Level of the correction for extreme values

Outlier treatment

No clear best practice in EU-SILC, strong variation across countries

More emphasis on trying to impute 'correct' value, e.g.:

- Logical deduction / simulation models
- Back to interviewers / data processing
- Comparison with administrative records
- Imputation based on multivariate regression models
- Comparison with previous waves (panel data)

Corrections through calibration

Check on aggregates at individual level (e.g. total income / consumption):

- Still undesirable outliers
- Accumulation of imputations?

Sensitivity checks are key

What did we find for 2010?

Definition of target variables

- Before transfers variables HY022 & HY023
  - Disposable income, minus transfers
  - In quite a few countries some benefits are taxed
  - In these cases taxes are deducted twice
  - A few countries correct for this, but not all
  - Reduces value of variable as indicator of 'pre-transfer counterfactual'

- Split HY050?
  - Compensation for cost of children (child benefits, birth grants, tax credits,…)
  - Compensation for loss of income - Work-life balance (maternity leave, parental leave, social leave)
  - Public & private alimony together; rest of HY050 in separate variable

What did we find for 2010?

Classification of benefits

- Incorrect classification (PY120 - HY050)

- Relatively many borderline cases, often not consistent
  - some housing cost related benefits (HY060 - HY070)
  - vocational training (PY090 - PY140)
  - maternity benefits (HY050 - PY120)

- Some income components not registered?
  - Tax credits
  - Education-related allowances for children under the age of 16
  - Carers allowance if not paid to disabled person

Difference in information between Quality Reports and Survey results on comparability

Recommendation

More disaggregated information in UDB

A yearly database with the classification of all income components, for every year and all countries + changes documented in quality reports

International panel with social policy experts can help decide on borderline cases (cf. Canberra Group)

Next steps

Further detailed analysis of all other income variables'

Validation by national statistical institutes

Discuss recommendations with stakeholders

Final Report

Distribute MetaSILC 2015 database'