Reconciliation of EU statistics on income and living conditions (EU-SILC) data and national accounts

VELI-MATTI TÖRMÄLEHTO

2019 edition

STATISTICAL
WORKING PAPERS

eurostat

Reconciliation of EU statistics on income and living conditions (EU-SILC) data and national accounts

VELI-MATTI TÖRMÄLEHTO

2019 edition

Manuscript completed in June 2019

Luxembourg: Publications Office of the European Union, 2019

Theme: Population and social conditions
Collection: Statistical working papers

PDF ISBN 978-92-76-08674-1 ISSN 2315-0807 doi: 10.2785/489634 KS-TC-19-004-EN-N

Abstract

The coherence of household survey data with national accounts has been studied extensively in recent years, following the “Beyond GDP” initiatives. This paper compares income aggregates in EU-SILC and national accounts, adjusts for the main conceptual differences, and discusses factors that could influence the observed discrepancies. Following a proposal by Atkinson, Guio and Marlier (2017), sensitivity of key social indicators to the micro/macro-discrepancies is then examined by adjusting the micro data totals to match the reconciled macro aggregates. Three adjustment methods are tested (simple proportional scaling, calibration to margins, Pareto imputation), and their impact on the measures of income inequality and at risk of poverty compared. In line with other studies, the micro/macro gaps are found to vary significantly across countries, and are more substantial in property and self-employment income compared to wages and salaries and transfers received. The observed gaps are likely to be mostly due to measurement errors and conceptual differences. Adjusting the micro data with the gaps results in significant increases in inequality and median income levels, but more subdued changes in at risk of poverty rates. The results are sensitive to the adjustment methods as well as proper assessment of the micro/macro gaps. Caution is warranted if distributional indicators are computed from macro-adjusted micro data.

Table of contents

1 Introduction([1])

National accounts (NA) includes indicators or data that are useful as such for monitoring households' economic well-being ([2]), but there is a well-recognised need to complement aggregate indicators with distributional information. This information is available only from micro statistics. Recently, OECD and Eurostat have worked towards "distributional national accounts" (DINA), i.e. distributional information fully consistent with household sector income and consumption aggregates. Consequently, the coherence of national accounts and household survey data has been studied extensively in recent years (e.g. Mattonetti, 2013; Fesseau et. al, 2013; Zwijnenburg, 2016; Gregorini et al, 2016; Eurostat, 2018)([3]).

Atkinson, Guio, and Marlier (2017) show that understanding and reconciling the differences in micro and macro sources is important even when the aim is not distributional national accounts per se. They examine the evolution of real incomes from both national accounts and EU-SILC, and identify the main reasons for discrepancies (Atkinson, Guio and Marlier, 2017). Moreover, they discuss the implications of micro/macro coherence for social indicators, such as at risk of poverty rates.

The current paper seeks to address the following recommendation given by Atkinson, Guio and Marlier (ibid, p. 77):

Recommendation 3: The EU-SILC coverage of income by components exercise should be re-done, with a baseline appropriate for the calculation of social indicators.

They also suggest examining the sensitivity of conclusions to data deficiencies: "…the obvious question to ask is how far the AROP (at risk of poverty) and other indicators are affected by proportionate adjustments to different income categories." (ibid). This, alongside other adjustment methods, is the objective of this paper. Accordingly, we experiment with sensitivity analyses of social indicators to the micro data adjustments based on the gap between EU-SILC and reconciled national accounts' estimates. The baseline is the EU-SILC income concept and social indicators.

([1]) The author is from Statistics Finland. The author wishes to thank Sigita Grundiza, Anne-Catherine Guio, Tarja Hatakka, Tara Junes, Pierre Lamarche, Eric Marlier, Francesca Tartamella and Philippe Van Kerm for valuable comments and suggestions. All errors remain strictly the author's responsibility. This work has been supported by the third network for the analysis of EU-SILC (Net-SILC3), funded by Eurostat. The European Commission bears no responsibility for the analyses and conclusions, which are solely those of the authors. E-mail address for correspondence: veli-matti.tormalehto@stat.fi.

([2]) These include disposable income of household sector, which is also adjusted to take into account social transfers in kind and better reflect different social protection regimes. The Social Protection Committee has adopted the growth rate in real unadjusted gross household disposable income ("unadjusted" here refers to the indicator not taking into account social transfers in kind) as the national accounts based indicator; this indicator is now part of the EU portfolio of social indicators (Social Protection Committee, 2015).

([3]) The ECB has focused on reconciling macro and micro data on households' assets and liabilities in the expert group on linking macro and micro data for the household sector (EG LMM).

The paper is structured as follows. First, we discuss the potential reasons for the discrepancies, including generic and survey-specific issues. In the second part, we focus on the main conceptual differences between EU-SILC disposable income and national accounts' gross household disposable income (GHDI). We construct a modified measure of GHDI, which is more comparable with EU-SILC, and compute coverage rates for the main income components accounting for conceptual differences. In the third part, we adjust the EU-SILC data with different methods (simple rescaling, calibration of weights, Pareto imputation) at micro level, and look at sensitivity of the income-based indicators to the adjustments.

We use the National Statistical Institutes' (NSI) version of the EU-SILC cross-sectional database (version spring 2017) from income reference years 2010-2014. The income reference year is the calendar year prior to the survey year (2011-2015), except for the UK and Ireland. We examine only a subset of countries. First, we have no access to German EU-SILC data because it is not included in the data disseminated to NSIs. Second, we do not consider countries who still only have combined data of non-profit institutions serving households (S15 NPISH) and household sector (S14) available in national accounts on gross household disposable income (B6G). These countries were, at the time of writing this paper, Luxembourg, the UK, Ireland and Malta, in addition to Germany. Finally, our choice of countries is further restricted by the availability of macro data on certain transactions needed in the adjustments. The national accounts data are from Eurostat, annual sector accounts, table nasa_10_nf_tr, retrieved in September 2017.

1.1. The starting point: coverage rates before reconciliation

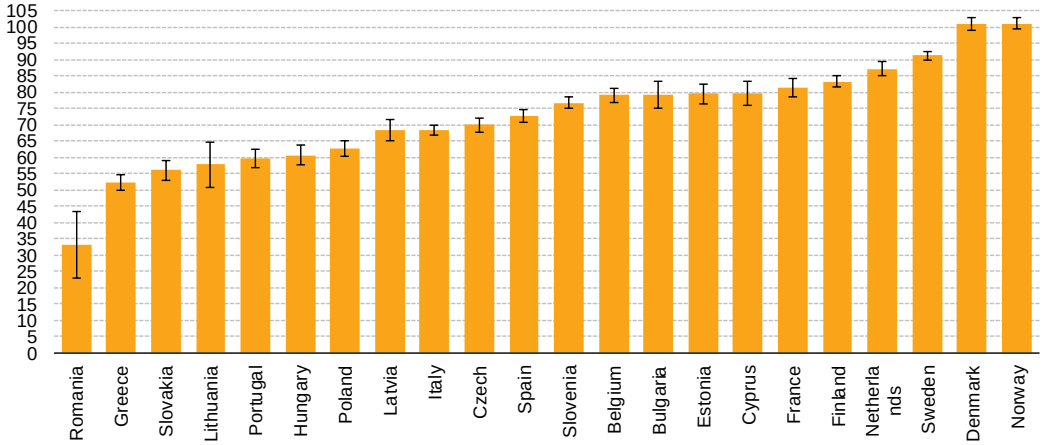
Comparing surveys and national accounts is one way to assess the accuracy of survey estimates, which are subject to various kinds of measurement and estimation errors. A direct comparison does not indicate accuracy because of differences in income concepts, population coverage, and methods. The importance of these vary across the countries, as discussed later in this paper. Nevertheless, the starting point is the coverage of the totals of disposable income before any modifications, shown in Figure 1. The coverage rates range from around one third to more than 100 percent. The variation and the levels are both worrying.

EU-SILC estimates are based on samples, but looking at the estimated confidence limits indicated by the error bars([4]), it is evident that sampling variance is not the root cause of the differences. The unadjusted coverage rates already suggest that the use of administrative data in surveys improves coherence, but also among the traditional EU-SILC "register countries" (Nordic countries, the Netherlands, Slovenia) there is variation that calls for further analysis.

Recent studies of the coverage of unadjusted EU-SILC and national accounts by income component include Gregorini, Grundiza, and Lamarche (2016) and Eurostat (2018). These authors also studied stability of the coverage rates since year 2006. This paper does not aim to repeat the extremely valuable work already done, but attempts to contribute by examining the discrepancies and their causes based on modified, reconciled aggregates that take into account the major conceptual differences.

([4]) Confidence limits are indicative only as they do not take into account design features such as clustering, stratification, or calibration, i.e. assume simple random sampling. Proper variance estimation is not feasible with the information available in the EU-SILC UDB, although this can be partially circumvented using pseudo-design information of Goedemé (2013).

Figure 1: Unadjusted coverage rates of EU-SILC disposable income to national accounts gross disposable income, income reference year 2014 (EU-SILC survey year 2015)
(Sum of EU-SILC HY020, % of NA GHDI)



Reading note: In Romania, the estimated total sum of EU-SILC disposable income was 33 % (+/-10 %) of national accounts gross household sector disposable income, before any modifications.

Source: Eurostat, EU-SILC UDB 2015 (variable HY020) and annual sector accounts tables (transaction B6G S14).

1.2. Identifying the gaps: the OECD methodology and distributional national accounts

Aside from sampling variance, the gap could be due to generic reasons as well as various factors specific to surveys. The OECD Expert Group on Disparities in National Accounts (EG DNA) developed methodologies to produce distributional results for household income, consumption and saving consistent with national accounts concepts using micro data (Zwijnenburg, 2016; OECD,2017). An important part of the work was identification of the reasons for the gaps, and allocation of these to household subgroups. The OECD expert group identified several reasons for the observed gaps:

1. Population differences
2. Conceptual and classification differences
3. Underground economy and illegal activities
4. Measurement and estimation errors in micro data
5. Quality of national accounts data

In the EU-SILC context, point (4) includes challenges such as sampling bias and unit non-response, non-reporting and under-reporting, imputation bias and variance, and errors induced by net-to-gross conversion of incomes. We look in more detail the conceptual differences of micro and macro income concepts in section 2. Before that we discuss selected issues related to sampling, measurement and target populations and the role these might play in the observed discrepancies.

1.3. Estimation of total amounts in national accounts and EU-SILC

In the system of national accounts (SNA), the household sector income aggregates are derived from various data sources, such as tax administration and social security data, pension providers, business register, and other counterpart data. The transactions in the household sector account are not compiled independently from the other sector accounts, and possibly use indirect methods. There may be issues with the allocation of the economy-wide amounts to different institutional sectors, although main items such as wages and salaries or social benefits are received only by households. Some bias may have to be allowed in the household sector to minimize bias and statistical discrepancies in the total economy. Transactions in the household sector may even be estimated as a residual of the other sectors.

In the compilation process, the data sources are confronted (cross-checked), and the estimates are completed (balanced) in a coherent framework (SNA/ESA). Because of data confrontation and completion, the errors and inconsistencies of the primary sources can be corrected. Therefore, the national account estimates of household sector aggregates should have less bias than any single source statistic. Moreover, the sole concern is accurate estimation of the total amount of each transaction, without consideration of its distribution within an institutional sector.

The task is more complicated in a sample survey, where the focus is on distributions([5]). For any given income component Y observed for the responding sample, we can write the estimator of the population total income in a sample survey as

$$(1)\quad Y = \sum W_i Y_i, \; = \sum (\pi_i \rho_i g_i) f(X, \varepsilon)_i,$$

i.e. as the sum of estimation weights W and measured incomes Y. The latter can contain measurement error and is also expressed as a function of the “true” income X and an unknown error term ε. When incomes are observed only for the responding part of the sample, the design weights, computed as inverses of inclusion probabilities, need to be adjusted to account for non-response and ineligible units (over-coverage). To improve the estimates, the weights are usually further adjusted to auxiliary external information. In the formula (1), π_i is the inverse of the inclusion probability (design weight), ρ_i is the inverse of the estimated response propensity (non-response weight), and g_i is the adjustment due to some known population totals (g-weight in the standard calibration approach). Multiplying these gives the estimation weight, i.e. calibrated and non-response adjusted inverse of the inclusion probability.

The crucial issue is what the properties of the measurement error term are, and whether it is correlated with the “true” incomes. Measurement errors of income in the EU-SILC context have been assessed by simultaneous measurement of income from interviews and registers (Jäntti et. al, 2013; Nordberg et. al, 2004) ([6]). Recently, such measurements have been conducted at least for Austria and Spain (Statistics Austria, 2015; Méndez-Martin, 2015). Both countries have implemented the transition from interviewed income data to mostly register data in EU-SILC, and this provides an opportunity to compare how the EU-SILC total income amounts change due to transition.

In Austria, the switch to mostly register data (on wages and salaries and transfers) increased the total amount of disposable income only by 2.3 %, but increased in income inequality and at risk of poverty. In Spain, the transition to a mixture of register and interview data increased the total amount

([5]) It should be noted that in four Nordic countries (Denmark, Finland, Norway, Sweden), household income statistics nowadays are based on national entirely register-based sources that cover the whole population. These have neither sampling error nor non-response bias.

([6]) An example of a non-register based approach is Pissarides and Weber (1989), who estimate the extent of unreported self-employment income based on food expenditure equations.

by 14 %, but had much more mitigated impact on inequality than in Austria. (Statistics Austria, 2015; Méndez-Martin 2015, Törmälehto, Jäntti and Marlier, 2017)([7]). More in accordance with the Spanish data, the repeated measurement based on ECHP data in Finland from 1995 and 1999 suggested overall under-reporting of survey data with regard to register data to be around 5-10 %, but with uneven patterns. Those in the bottom decile over-reported their incomes by 10-15 % whilst those in the top under-reported their incomes by 15-20 % (Nordberg et. al, 2004, p. 68).

The empirical evidence on discrepancies between survey and register incomes seems somewhat country-specific. It would require quite some effort to build a generic model of measurement errors applicable to all EU-SILC countries. As a rule of thumb, in this paper we consider coverage rates higher than around 80 % (for wages and salaries and transfers) or even 50 % (for property income) to be acceptable, considering measurement errors related to interview data as well as differences in target populations.

1.4. Sampling design considerations

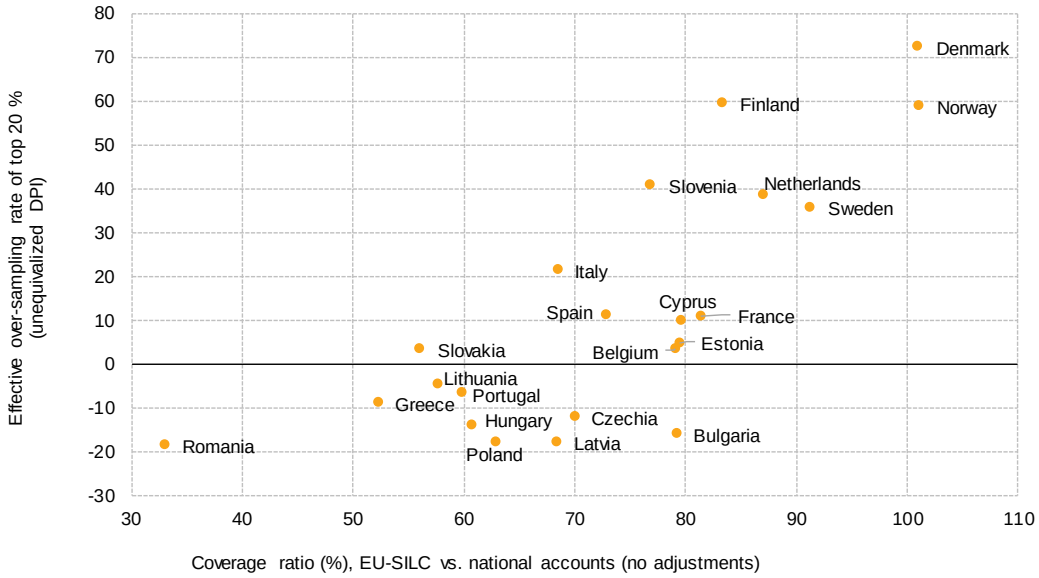
The EU-SILC implementations are based on samples, and sampling designs may affect the precision and bias of the estimation of population totals. We next briefly discuss two features that could have an impact on the estimated total amounts, namely stratification of the sample and calibration of survey weights.

Appropriate stratification of the sample may increase accuracy of the estimated totals. This is very relevant for skewed distributions, and in wealth surveys over-sampling the rich is a highly recommended strategy. The most common sampling design in EU-SILC is stratified multi-stage sampling. Simple random sampling without explicit stratification is used by Denmark, Malta, Iceland, and Norway. Other countries use stratified sampling designs.

On average, close to 40 percent of total disposable income is concentrated to the top quintile. Consequently, having relatively more of the sample in higher income groups should improve estimates of total amounts, which could result in better coverage rates with respect to National Accounts. To this end, Figure 2 maps the effective over-sampling rates in the weighted top quintile versus the unadjusted coverage rates of disposable income in EU-SILC.

([7]) The total amounts are based on author's computations from two versions of the EU-SILC UDB 2011 before/after the transition.

Figure 2: Effective over-sampling rates of the top quintile of non-equivalised disposable income versus coverage rates before any adjustments, 2014 (EU-SILC survey year 2015)



Note: The effective over-sampling rates are calculated as (s20-20)/20, where s20 is the unweighted proportion of households in the sample in the top income quintile, and income quintiles are defined using final estimation weights and household disposable income (no equivalisation, no person weighting).

Reading note: In Romania, 16.3 % (1206/7391) of the EU-SILC 2015 sample was in the weighted top 20 % of total non-equivalised disposable income, and the effective over-sampling rate therefore (16.3-20)/20=-18.5. The coverage rate in Romania was 33 %. In Denmark, the corresponding shares were 34.5 % (2076/6025) and 101 % and the effective over-sampling rate 72.5 %.

Source: EU-SILC UDB 2015 (variable HY020) and Eurostat, annual sector accounts tables (transaction B6G S14).

Although there appears to be a positive correlation between the coverage and "effective over-sampling" rates, this may be spurious. The effective over-sampling rates reflect not only sampling strategies, such as stratification and clustering, but also non-response patterns. With simple random sampling and 100 percent response rate, we would expect the top quintile to have 20 percent of the sample. Differential non-response would change this, as would stratification of the sample and other design features. For instance, the group of "register countries" have more of the sample in the higher income groups. A factor in this could also be sampling of individuals proportional to household size (pps-sampling), which implies that larger households have higher inclusion probabilities.

Differential unit non-response patterns could potentially affect the survey coverage rates with respect to national accounts, for instance if the sample is biased towards middle part of the distribution at the expense of the bottom and top parts. Weighting, and in particular calibration of the sampling weights to auxiliary information may increase accuracy of the sample estimates. In particular, if income data are taken from registers, it can be used in the calibration, and this can improve accuracy of the estimates of total amounts([8]). Some countries are indeed calibrating weights to external sources on household income, including Austria, Denmark, Finland and the Netherlands. The way income variables are used varies, and this method in general is available only for the "register countries", who already may have less measurement errors on income data and therefore better coverage rates.

([8]) The calibration variables available for the sample need to be conceptually equivalent and measured the same way as the available auxiliary information, which is the case with exactly matched register data. Using interview-based income data in calibration is questionable, in particular if there is evidence of substantial measurement error. This could have unexpected effects on the calibrated sample weights.

To gain insight into the effect of income calibration, we use the Finnish EU-SILC, where the total sums of main income components derived from registers are used directly as calibration variables. Table 1 gives an example how this affects the total sums in the Finnish SILC, by comparing two calibration models: one that includes only demographic variables (age/gender/region/household size) and the complete model, which adds several income variables. The point estimate of total actually reduces slightly, whilst the confidence limit is narrower. With or without calibration to income totals, the coverage rate remains at the level, which would not much alter Finland's rank in cross-national comparisons.

Table 1: The effect of calibration on the estimated total amount of disposable income and its standard error, Finland, income reference year 2014

Total disposable income (HY020)		Estimated total amount	Confidence limit		Coverage rate, interval estimate, %	
Weight	Variance estimation method	EUR million	EUR million	%	Lower	Upper
Full calibration	Bootstrap, stratification and calibration	99 450.0	492.0	0.5 %	82.8	83.8
Full calibration	Linearization, stratification	99 450.0	1 306.0	1.3 %	82.0	84.6
Demographic calibration	Linearization, stratification	100 951.0	1 274.0	1.3 %	82.0	84.6

Note: Method 'bootstrap' refers to variance estimation using the rescaling bootstrap method (Rao, Wu and Yue, 1992), with 1000 replicates, taking into account the stratification of the sample as well calibration of the bootstrap weights to the original margins.

Source: Statistics Finland, SILC database 2014 (EU-SILC 2015).

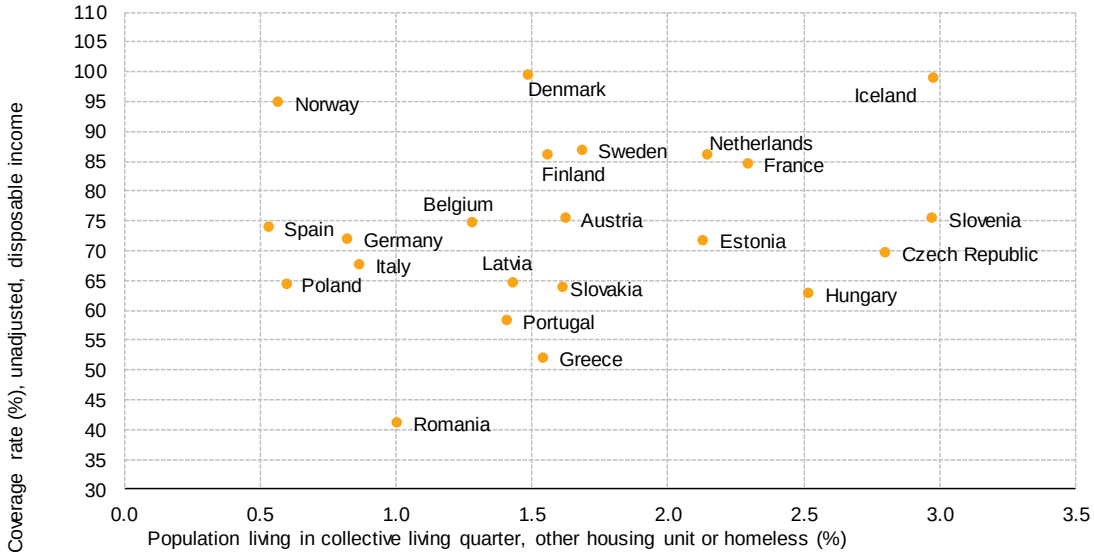
1.5. Population differences

In EU-SILC, the reference population is those living in private households at the end of the year or during the fieldwork period. The reference population excludes certain groups who are included in the national accounts concept of resident population, and therefore their incomes. These include persons in collective households and institutions, such as prisons, hospitals and nursing homes, or retirement homes, as well as persons who died or emigrated during the year.

The definition of a household in national accounts is very similar to that of EU-SILC, and mentions pooling of income and share of consumption of essentials of living (Eurostat, ESA 2010, paragraph 2.118). In practice, the definition of a household is not important, because households are treated as an institutional sector, which includes all resident population.

An estimate of the share of population in non-private households is available from Census. The horizontal axis in Figure 3 shows the data from year 2011. The shares range between 0.5 % to around 3 % of population, although different definitions give slightly different number for certain countries. The figure also indicates that the share of population in non-private households does not correspond to lower unadjusted coverage rates; on the contrary, if anything there is a slightly positive correlation between the two. This would indicate that the population differences account only a small part of the observed variation in the unadjusted EU-SILC/NA coverage rates.

Figure 3: Share of population in non-private households (%) versus ratio of EU-SILC total disposable income to NA gross disposable income in 2011 (EU-SILC survey year 2012)



Note: Non-private households refers to occupants in collective living quarters, other housing units or homeless. *Source* of data: Eurostat, Census Hub 2011.

Usually, the share of income received by non-private households is not known and has to be estimated. The population share is sometimes used to approximate the share of income going to out-scope-population. More refined approaches have been used, e.g. by Eurostat (2013)([9]). In this paper, we do not adjust for population differences on the grounds that these do not explain the observed variation between countries in micro/macro differences.

Some actual insight on the importance of out-of-scope population can be gained from Finnish register data, which contains their detailed income data. The population([10]) is divided into four groups, depending on residency status and living in private households. Table 2 shows the share of different income components going to the population of non-private households at the end of the year. The shares are higher for current transfers received and property income, which are more age-dependent income components and received relatively more by the elderly.

([9]) In the Eurostat reconciliation process (Eurostat, 2013), NA totals going to out-of-scope households were deducted from NA totals, using data from the 2001 Census on 'person living in an institutional household' and 'person living in a private household' (13) by age class, demography data for 2008 for total population by age class, and per capita value of EU-SILC income variables calculated under specific assumptions by age class.

([10]) The total population consists of all those having registered income during a year in any of the Finnish income registers. The size of the population in private households at the end of the year 2015 was 5,4 million. Around 1.5 percent or around 80,000 people were registered in non-private household. The total population at the end of the year was 5.48 million. Moreover, there were 766 thousand persons who had at least some registered income in the Finnish income registers but who were not in resident population at the end of the year.

Table 2: Breakdown of total income received during a calendar year by in and out of scope population in Finland, 2015

Income component	In private households 31.12	In non-private households 31.12	Not in resident population 31.12: deceased	Not in resident population 31.12: emigrated or unknown	Total
Wages and salaries	98.4 %	0.5 %	0.1 %	1.1 %	100 %
Self-employment income	98.7 %	0.4 %	0.3 %	0.5 %	100 %
Property income	97.0 %	0.7 %	0.6 %	1.7 %	100 %
Current transfers received	96.1 %	1.8 %	1.3 %	0.8 %	100 %
Current transfers paid	98.5 %	0.6 %	0.2 %	0.7 %	100 %
Disposable income	97.3 %	1.0 %	0.6 %	1.1 %	100 %

Note: The four groups are a) resident end of year 2015 and living in private household at the end of the year, b) resident end of year 2015 and living in non-private household, c) not resident end of year 2015, deceased during the year 2015, and d) not in resident population end of year 2015 because of emigration or other (unknown) reason.

Source: Statistics Finland, Total statistics on income distribution (TSID).

2 Conceptual differences and coherence of adjusted EU-SILC and National Accounts disposable income

The country variation in the overall micro/macro coverage rates of total sums are not likely to be due only to sampling error, population differences, calibration of the weights or other technical sample design features. We therefore now turn to issues related to concepts and measurement. In this section, we review the relative importance of the main conceptual differences between national accounts and EU-SILC, and compare coverage rates of main income components after adjusting for such differences. Annex 1 contains a more detailed discussion on some of the national accounts' transactions.

The concepts of income and consumption in national accounts contain items that are not actually cash flow for households, and adjustments are needed if the baseline is the micro concept. For instance, Cynamon and Fazzari (2017) consider an adjusted national accounts income concept, which conforms to "market-based cash income under the control of households", and is closer to the micro concept of disposable income.

2.1. Conceptual differences and adjustments

There are important differences between the NA and EU-SILC operational income concepts. The most important of these is gross operating surplus of household sector in national accounts, which is conceptually equivalent to imputed rent in EU-SILC([11]). Net imputed rent would be available in the EU-SILC data for all countries, but it is excluded in this paper because it is not part of the current EU-SILC income definition used for calculating the EU social indicators. Moreover, comparability between countries remains an issue (Törmälehto and Sauli, 2017).

Some transactions in national accounts do not exist in micro sources, or are operationally different although conceptually related. Such transactions include investment income attributed to insurance policy holders (D441) and financial intermediate services indirectly measured (FISIM). The first is an imputed property income flow, and the second affects recording of interest income, but is not relevant at micro level (see annex 1 for further discussion). The standard approach is to ignore these items in the micro/macro comparisons.

([11]) "Net" in EU-SILC refers to imputed rents net of interest repayments on mortgage, but gross of deprecation (consumption of fixed capital). Net operating surplus in national accounts would refer to operating surplus net of consumption of fixed capital, but gross of interest repayments on mortgage.

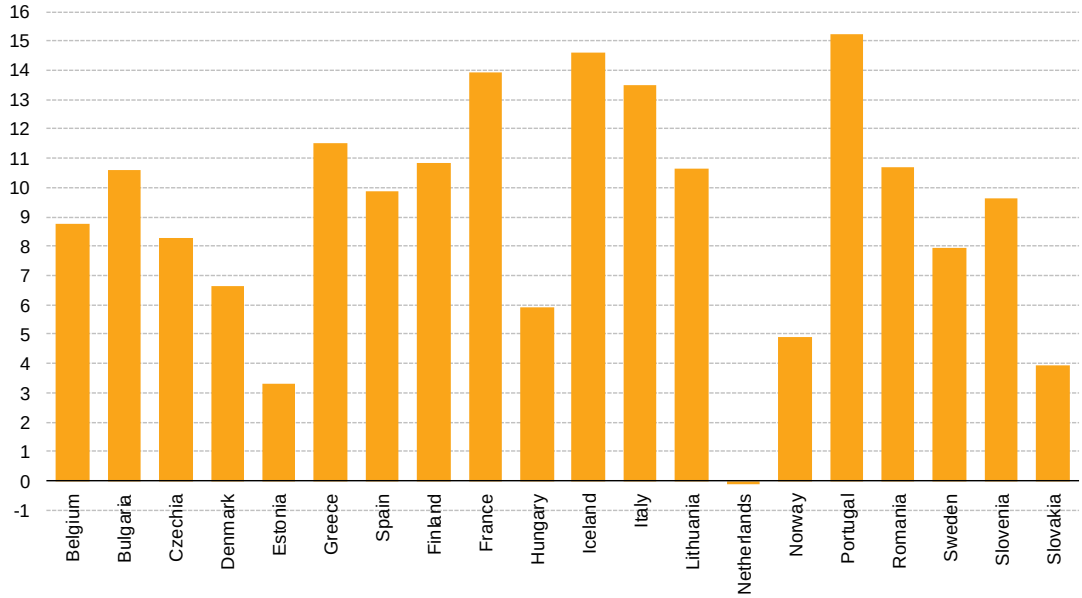
Interest payments are deducted as property income paid or as intermediate consumption (FISIM part) in national accounts([12]). This is not the case in the standard EU-SILC income definition, which excludes net imputed rent. In micro statistics, interest paid on consumer loans is treated as consumption expenditure, while other interest payments are netted off if they relate to loans used to purchase income-generating assets (OECD, 2013, p.52). Mortgage interest payments therefore should be deducted from imputed rents, and interest payments on business loans or investment loans from self-employment income and rental income. In the current EU-SILC concept, there is no counterpart to the national accounts concept of property income paid.

Some items may affect component-by-component comparisons, but net out when considering disposable income. For instance, net non-life insurance premiums are in transfers paid (D71) and claims in transfers received, but the components tend to be almost equal and in practise offset each other. Employers' social contributions received and paid cancel out completely, and are not considered in the adjustments.

Miscellaneous current transfers received (D75) are also excluded. This item includes transfers from non-profit institutions and transfers between households, including money sent by relatives or other households from abroad (within-country inter-household transfers are consolidated). It is relatively quite important in many Eastern European countries, and if not excluded, would affect coverage rates of transfers significantly.

Figure 4 provides an overview of the significance of the macro income components that normally are excluded from comparisons. The share of such items as percentage of NA gross household disposable income varies and ranges from 0 % to almost 15 %. The largest component is usually gross operating surplus of households (see annex 1).

Figure 4. National accounts transactions excluded from comparisons, income reference year 2014 (EU-SILC survey year 2015)
(% of NA gross household disposable income)



Note: Transactions included: investment income of insurance policy holders, gross operating surplus, FISIM interest received, land rents, total actual interest paid. Only countries with all these items available in macro data are included in the figure. See annex 1 for details of the sub-components.

Source: Eurostat, annual sector account tables (nasa_10_nf_tr).

([12]) With the exception of FISIM related to consumption loans, which is consumption expenditure.

The GHDI figures after reconciling the main conceptual differences range from 85 % to 100 % of the GHDI before any adjustments. On average, the conceptual adjustments reduce the benchmark by around 10 percent. For instance, in Italy the total to be used as a EU-SILC benchmark was 86.5 % of the original macro estimate, in Denmark 93.4 %, and in Romania 89.3 %. The adjustments were relatively stable over the years 2003-2016 in most countries. For some transactions and income components, reclassification is necessary to adjust for different in micro and macro statistics, although the components are conceptually comparable. Income from rental of dwellings, other properties or land is recorded in EU-SILC variable HY040. In national accounts, rental income from owned dwellings results from market production of housing services([13]) and is included in mixed income, gross of interest repayments. In EU-SILC, according to guidelines, rental income should be net of interest repayments. To reconcile the information, either mixed income resulting from housing services should be separated from other mixed income national accounts, or rental income should be considered as self-employment income in EU-SILC. In this paper, EU-SILC rental income has been pooled with self-employment income.

National accounts include estimates of hidden economy, which are not shown separately but included in other transactions, such as wages and salaries received. For instance, for Finland in 2014, 0.9% of the total wages and salaries received in national accounts was estimated to be hidden economy wages and salaries, while in self-employment income the share was estimated to be 6.5 %. Without access to national data, specific adjustments are not possible, as discussed later in section 3.4.

To summarise, in this paper the NA gross household disposable income is adjusted as follows to have a more comparable aggregate benchmark:

- Removal of gross operating surplus, non-life insurance, reinvested earnings on direct foreign investment, investment income on insurance policy holders, FISIM part of interest received, net non-life insurance claims minus premiums, and miscellaneous current transfers (D75)
- Adding back interest paid before FISIM allocation (i.e. true interest paid)([14]), rents paid (land rents), and other current transfers paid (D75)

2.2. Coverage of EU-SILC and reconciled national accounts income components

The following sections examine the coverage of the main income categories, reconciling for conceptual differences and aiming for a comparable macro benchmark. Moreover, we only examine countries, which have separate and sufficiently complete S14 household sector accounts. We make no adjustment for non-resident and institutionalised population.

Coverage of wages and salaries

The EU-SILC concept of wages and salaries used here for the comparison include gross employee income (variable PY010G) plus all gross non-cash employee income (variable PY020G). The coverage rates of wages and salaries range substantially between countries, but in general the coverage rates are high, exceeding 90 % in a majority of countries in 2014 (Figure 5). The median of the country coverage rates was 95 % in 2014. The coverage rates are relatively stable within the countries but exceed 100 % in a few countries. A priori one expects the EU-SILC total sum be below the NA total sum. The coverage rates above 100 % may result from sampling variation. However, in

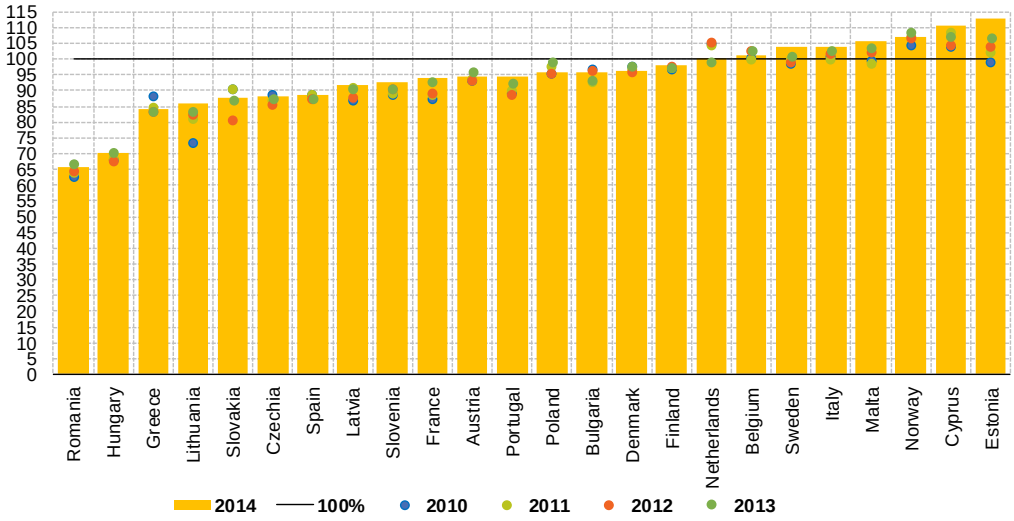
[13] Own account production of housing services results in operating surplus in NA household sector.

[14] Total interest paid before FISIM allocation is split into FISIM part, which is intermediate consumption, and interest paid, which is property income paid. Both parts are added back to NA aggregates in the adjustment, because the baseline is EU-SILC cash and near cash income concept

some cases, also the 95 % interval estimate lies above the macro aggregate, which should not be the case.

Figure 5: Coverage rates of wages and salaries in EU-SILC, income reference years 2010-2014 (EU-SILC survey years 2011-2015)

(% of wages and salaries in national accounts **(PY010G+PY020G) vs. national accounts**)



Note: Only countries with sufficiently complete national accounts household sector data available are included.

Reading note: In the Netherlands, the EU-SILC estimate of total wages and salaries was around 100 % of the national accounts total (in 2013 and 2014 and around 105 % in 2010-2012 (PY010G+PY020G vs. D11R).

Source: EU-SILC UDB 2011-2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

A potential cause of the varying coverage ratios is the use of administrative registers on wages and salaries. Based on the EU-SILC quality documentation, the Nordic countries, the Netherlands Slovenia, France, Austria, Estonia (as of 2013/2014), Switzerland, and use mostly register data on wages and salaries. A mixed method may be used, using data from both registers and interviews, as in Latvia, Italy, Malta, and Spain.. Over time, some observed changes are likely to be result of changing from interview data to register data (e.g. Estonia, partially in 2013 and mostly in 2014).

Wages and salaries are based on interview data in Romania, Hungary, Croatia, Lithuania, Greece, Czech Republic, Slovakia, Portugal, Poland, Bulgaria, Belgium, and Cyprus. Administrative data may be linked to data and used for checking purposes, for instance in Lithuania and Bulgaria([15]). In Romania, the information is collected only in the survey via paper questionnaires. In Greece, both CAPI and PAPI interviews are used.

Overall, we find no clear pattern in the coverage with respect to the mode of collection or data source. Register countries all have acceptable coverage rates, but also unacceptable in a sense of exceeding 100 %. Certain survey countries have lower rates, but there is no consistent pattern. For instance, in Romania, the coverage rate was 66 % while in Bulgaria it was 96 %, and these were stable over the years 2010-2014. In both Belgium and the Netherlands, the coverage was around 100 %, but only the latter was using register data.

([15]) In Lithuania, administrative data were used for making the survey income data more accurate or for supplementing them.

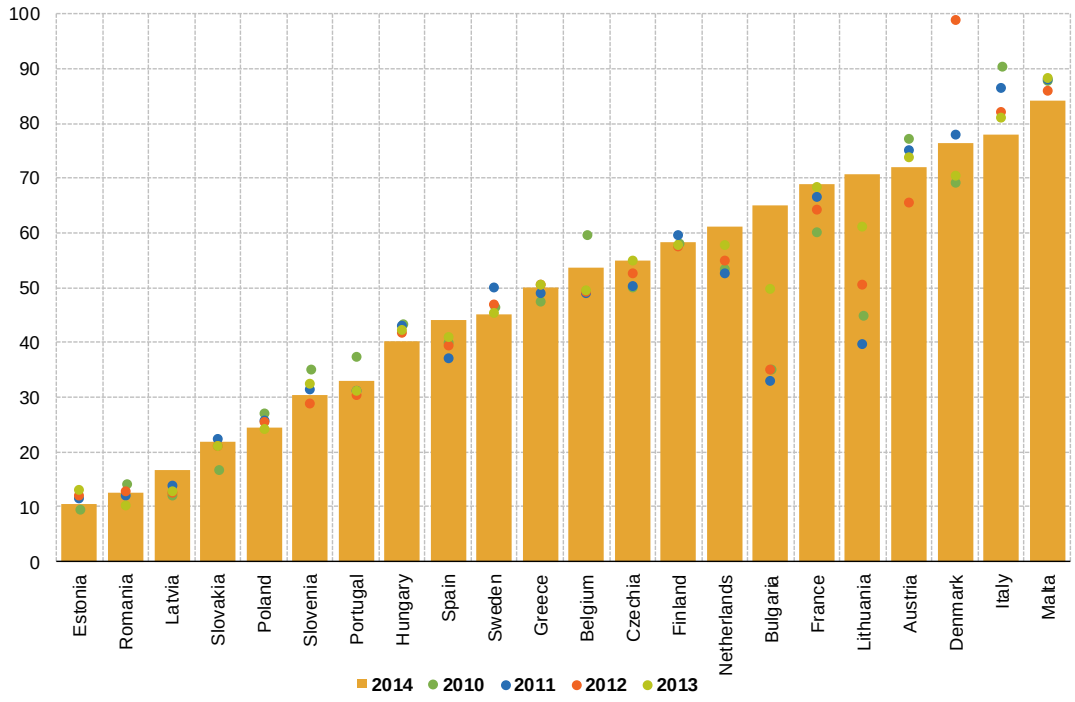
Coverage of self-employment income

The counterpart of EU-SILC self-employment income in national accounts is gross mixed income (B3G), which corresponds to remuneration for work carried out by the owner and members of his family and including his profits as entrepreneur. It is usually derived using administrative records and/or business surveys, and includes often large adjustments for unreported income (OECD, 2013). In the Finnish macro data used in this paper, the estimate of unreported income was 6.5 % of mixed income in 2014. Fesseau et. al (2013) report that in France hidden economy accounted for 14 % and in Sweden as much as 34 % of mixed income.

On average, EU-SILC profits and losses variable (PY050G) covers around half of the NA gross mixed income, but with very wide variation between the countries (Figure 6). The coverage is extremely low in some eastern European countries, whilst it is reasonably high for instance in Italy and Malta. There is significant variation among the "register countries", for instance between Slovenia (~30 %), Sweden (~45 %), Finland and the Netherlands (~60 %), and Denmark (over 70 %).

Figure 6: Coverage rates of EU-SILC profits and losses (PY050G), income reference years 2010-2014 (EU-SILC survey years 2011-2015)

(% of gross mixed income, PY050G /B3G, no adjustments)



Note: Only countries with sufficiently complete national accounts household sector data available are included. Norway is excluded as an outlier.

Source: EU-SILC UDB 2011-2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

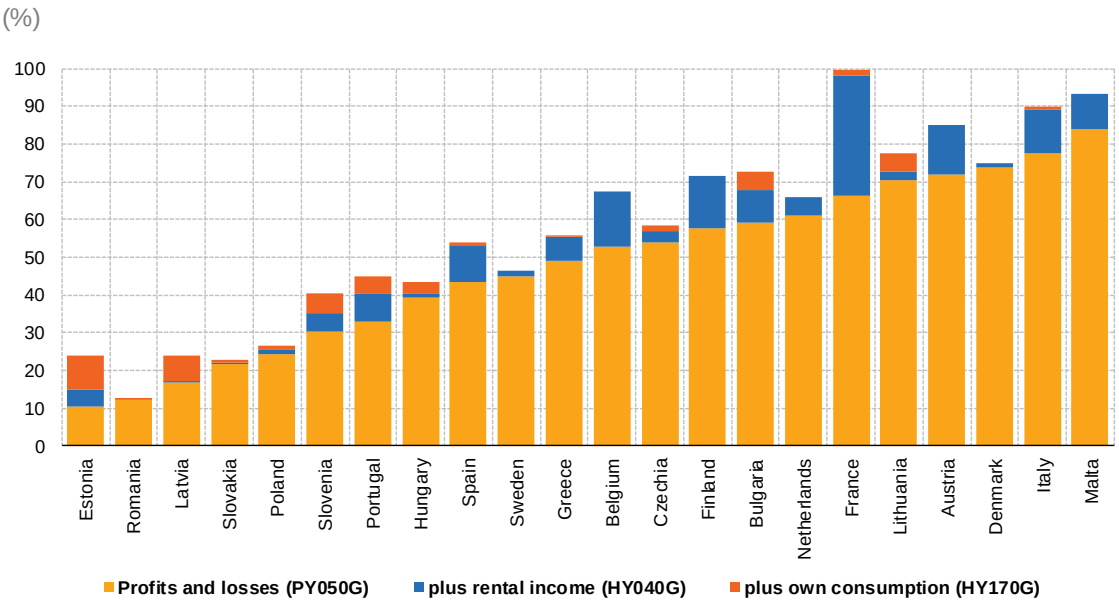
To be more comparable, the concept of self-employment income needs to be modified, at least to account for rental income and possibly also for the value of goods produced for own consumption. Regarding the latter, it refers to the value of food and beverages produced and also consumed within the same household. It is measured (variable HY170G) by several countries but not all. The measurement depends on how important own consumption is for certain population subgroups as well as on the aggregate. Own consumption of goods is not included in the EU-SILC income aggregates; however, conceptually it should be included when comparing the amounts to national

accounts' mixed income (or deducted from the NA amounts, but this is not possible with the available data).

Rental income from property or land is a separate variable in EU-SILC (HY040G), and a component of property income. In national accounts, land rents are property income and other rental income (of market producers) is included in mixed income. Therefore, to be more comparable, EU-SILC rental income from property should be added to self-employment income, and rents from land should be separated to property income. It is likely that rents from residential property is the main component in the EU-SILC variable, so if the partition cannot be done, it might be better to include all rents in self-employment income in the comparisons.

We adjust for these factors in the EU-SILC data by extending the concept of self-employment income to cover value of goods produced for own consumption (HY170G) and rental income (HY040G). We thus end up with adjusted coverage ratios that are higher than the comparison of profits and losses only (Figure 7).

Figure 7: Coverage rates of self-employment income and gross mixed income according to different definitions, income reference year 2014 (EU-SILC survey year 2015)



Note: Only countries with sufficiently complete national accounts household sector data available are included. Norway is excluded as an outlier due to an apparent anomaly in national accounts data.

Source: EU-SILC UDB 2011-2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

Adding self-consumption and rents increases the coverage rates, significantly in some countries such as Estonia (due to own consumption) and France (due to rental income), but the rates generally are still low and remain below 90 % in all countries except Malta, Italy, and France.

If we were to use estimated net mixed income rather than gross mixed income in the comparison, the coverage rates would be much higher (see discussion in Annex 1). Given the uncertainties related to the calculation of net mixed income, and conceptual differences in general, gross mixed income is the benchmark used in this paper.

The method of data collection on self-employment incomes vary. The preferred method to measure self-employment income is to collect or ask data on profit or loss based on accounting books or tax accounts, and if these are not available, then ask about money drawn out of business for non-business purposes. Some countries use administrative registers, essentially following business or tax account rules.

It is notable that two countries with the highest coverage rates of profits and losses, Italy and Malta, use a combination of register and interview data for self-employment incomes. For instance, in Italy the income from self-employment is set equal to the maximum value between: (i) the (net) self-employment income resulting from the Tax Report and (ii) the (net) self-employment income reported by the interviewee. The method is adopted in order to minimise either tax avoidance in the administrative data or under-reporting in the survey data, depending on which of the two is greater. The Italian Quality report notes that this procedure should lead to more comparable data, under the assumption that other countries' self-employment incomes are not underestimated. However, in Italy the coverage rates are very high compared to most other countries, including the "register countries".

Self-employment incomes in EU-SILC can be negative due to losses. In most countries, the share of individuals with negative self-employment income is small, but in Denmark, Sweden, Norway, the Netherlands, Estonia, Spain, and in particular Hungary the shares are high. In terms of total amounts, the share of negatives is significant in Denmark, Estonia, and Sweden. The different treatment/measurement of negative self-employment incomes may distort the comparison, and lead to lower coverage rates in these countries.

Coverage of property income

Coverage rates of property income received are often alarmingly low (see e.g. Gregorini et. al. 2016). Excluding rents, actual and imputed, property income received in EU-SILC consists of only two variables. The first is interest, dividends, and profit from capital investments in unincorporated businesses (HY090G) and the second is pensions from individual private pensions (PY080G). The latter has no direct counterpart in national accounts, and is excluded here from the reconciliations (see annex 1). EU-SILC rental income was already reclassified to self-employment income in this paper.

The macro benchmark for EU-SILC interest and dividends is the sum of interest received and distributed of income of corporations, but it needs to be adjusted for FISIM (see annex 1). Actual interest received (D41G) is available for many countries in national accounts database, but not for all. It should be used instead of interest received (D41), and only countries that have this information available are included in the comparisons. Actual interest is usually below the interest including FISIM, so this adjustment tends to increase the coverage rates.

Distributed income of corporations (D42) has two components: dividends (D421) and withdrawals from income from quasi-corporations (D422). The subcomponents are not available for all countries. Therefore, the whole of D42 has to be used in the comparison. From total property income received, we do not include investment income attributed to policy holders (D44) and reinvested direct foreign investment (D43) because of conceptual differences. Land rents (D45) are also excluded because they were reclassified to self-employment income.

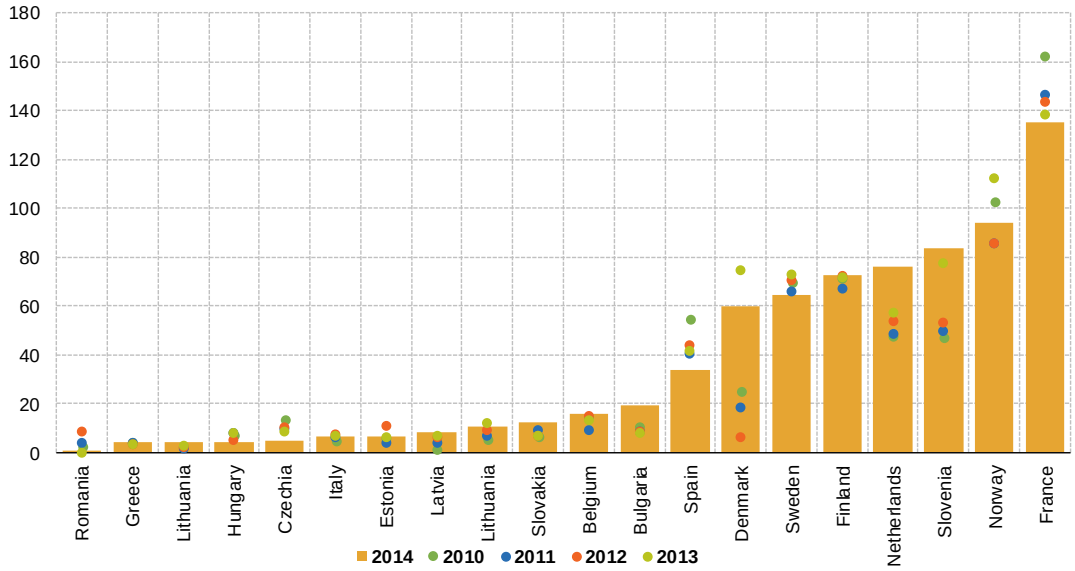
Even after the adjustments, the level and variation in the EU-SILC coverage rates is quite striking (Figure 8). The adjusted coverage rates range from 1 % to 135 %. France is an anomaly, with EU-SILC estimate far exceeding the national accounts value([16]). France combines register data, survey data, and imputations to construct the target variable, which may contribute to the over-estimation.

Aside from France, the coverage rates are above 50 % only in Denmark, Sweden, Finland, the Netherlands, Slovenia, and Norway – all "register" countries. In Spain, which also uses register data, the coverage is better than in the rest of the countries. The coverage rates are very low, even below 10 %, in most "survey" countries and also in countries other than Spain that combine interview and register data (e.g. Estonia and Latvia).

([16]) For France, the macro data includes all components, actual and FISIM interest, dividends, and withdrawals. The FISIM correction is small in France, and does not explain the difference. Withdrawals and dividends are roughly of equal size. Excluding withdrawals and using actual interest would even further increase the coverage rate.

Figure 8: Coverage rates of EU-SILC interest, dividends and profit sharing, income reference years 2010-2014 (EU-SILC survey years 2011-2015).

(% of actual interest received, dividends and withdrawals in national account - HY090G / (D41G+D42))



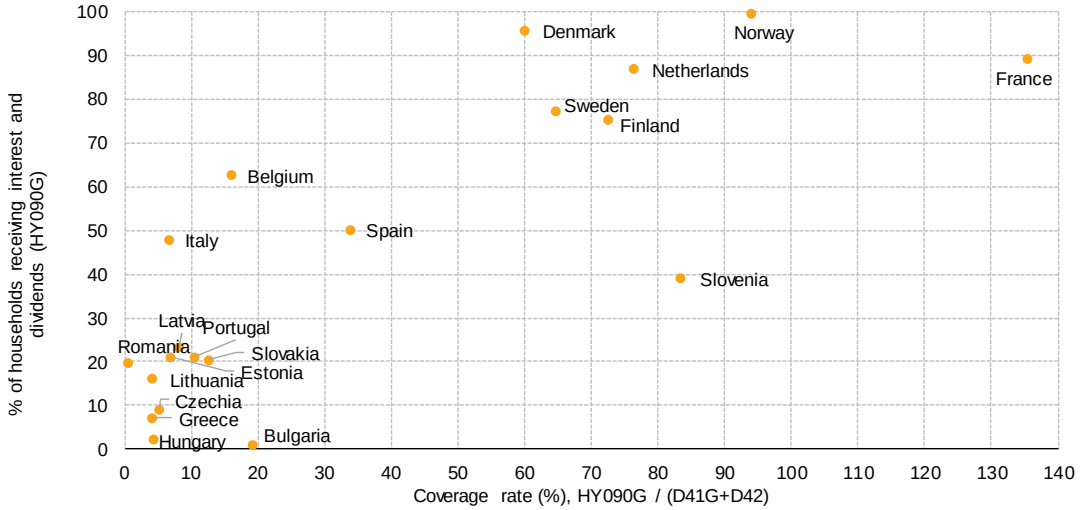
Note: Only countries with sufficiently complete national accounts household sector data available are included.

Source: EU-SILC UDB 2011-2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

The share of households with non-zero interest, dividends, or profit sharing also varies greatly among the countries (Figure 9). In Norway, nearly all households had non-zero values whilst in Bulgaria, less than 1 % had reported a value. The coverage rates of property income are positively correlated with the share of households receiving property income. If we were to correct the EU-SILC micro data using national accounts figures as a benchmark, this would affect very different populations in different countries.

As discussed in annex 1, imputed investment income of insurance policy holders has no direct counterpart in EU-SILC, and is excluded from the comparison. It is a significant component of total national accounts' property income received in many countries. For instance, in France it comprises around 40 % of total property income received. Its exclusion here from the comparison is justified but not a trivial choice, and the measurement of this item warrants further discussion.

Figure 9: Share of households receiving interest, dividends and profit sharing (%) vs. coverage rates EU-SILC/NA, income reference year 2014 (survey year 2015)



Reading note: In Spain, 50 % of households received interest or dividends, which covered 34 % of national accounts' total value of distributed income of corporations and actual interest received.

Source: EU-SILC UDB 2015.

Coverage of transfers received

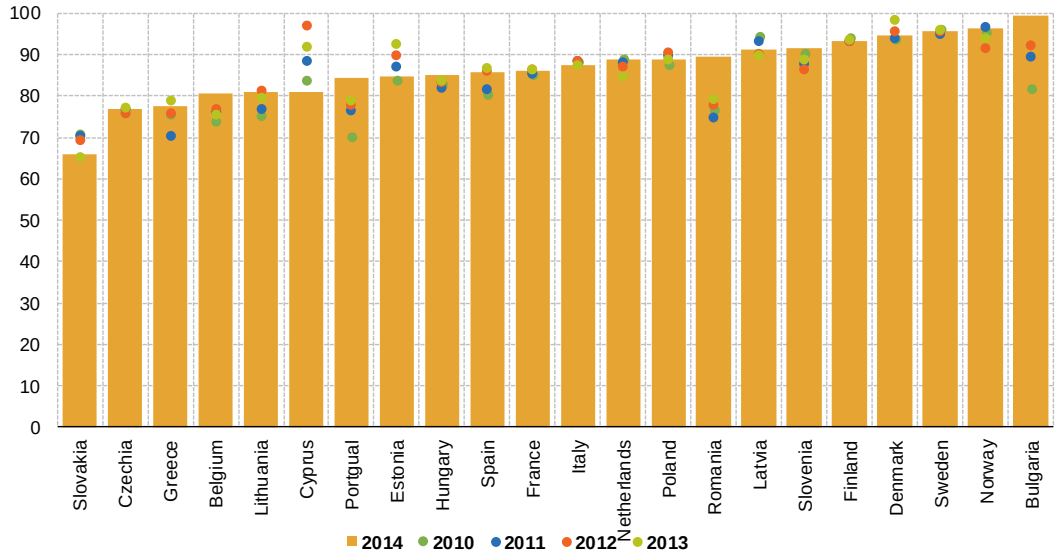
On transfers received, the EU-SILC definition used in the comparison covers current transfers excluding inter-household transfers received (HY080G). This can be compared with national accounts' social benefits other than social transfers in kind (D62). Thus, the equivalency links the NA transaction D62R social benefits other than social transfers in kind to EU-SILC total social transfers.

The coverage rates are relatively high in a number of countries, and consistently above 80 % in a majority of countries (Figure 10). The median of the country coverage rates is around 87 %. The highest coverage rate in income year 2014 was in Bulgaria, which according to the documentation is not using register data. It is followed by the Nordic countries, Slovenia and Latvia, all using administrative data. France, Italy, Spain and Estonia also use register data on transfers, supplementing it with interview data. In addition to Bulgaria, also Poland consistently achieves fairly good coverage rate without using register data.

Current transfers are likely to be received more by the institutionalised population, and thus the expected value of the coverage rate should be a couple of percentage points below 100 %. For instance, as shown earlier for Finland the in-scope population received 96.1 % of total current transfers measured in administrative registers, while the coverage rate with respect national accounts is 93.3 %.

Figure 10: Coverage rates of EU-SILC current transfers received (excl. inter-household transfers), income reference years 2010-2014 (EU-SILC survey years 2011-2015)

(% of national accounts' social benefits other than social transfers in kind (D62R))



Note: Only countries with sufficiently complete national accounts household sector data available are included.

Source: EU-SILC UDB 2011-2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

Coverage of transfers paid

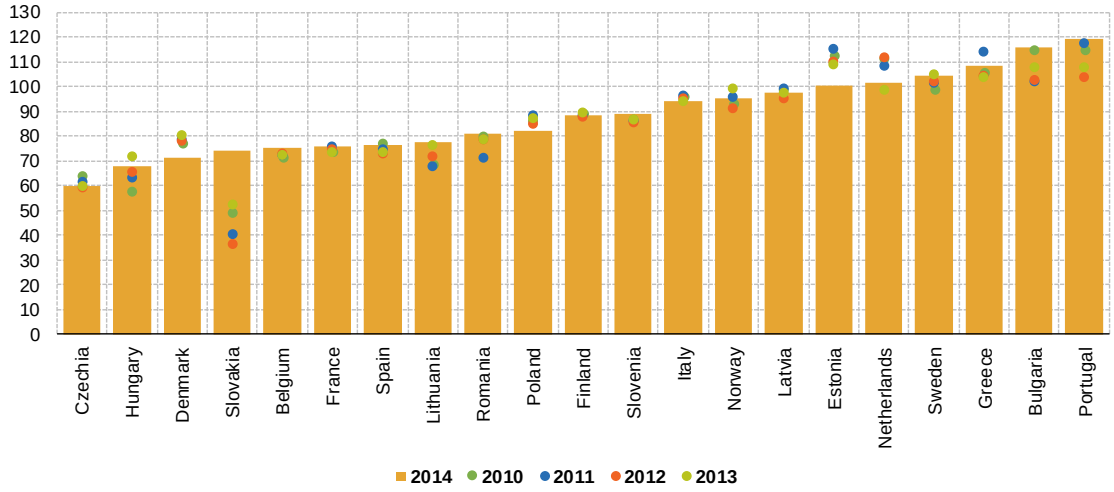
On transfers paid, the reconciled national accounts estimate covers taxes on income (D51), other current taxes (D59), and households' actual social contributions (D613). The main exclusion is employers' actual and imputed social contributions (D611+D612). These cancel out in GHDI, because it also included in compensation of employees. The EU-SILC estimate that we use in the comparison includes taxes on income and social contributions (HY140) and regular taxes on wealth (HY120). Regular inter-household cash transfers paid is not included, because it for the most part is a within-sector transaction in national accounts([17]).

The median coverage rate of the taxes and social contributions is close to 90 %, but there again is quite some variation among the countries (Figure 11). The Czech Republic, Hungary, Slovakia and Denmark stand out as countries with the lowest coverage rates. In contrast, in Portugal, Bulgaria, Greece, Sweden, the Netherland and Estonia the coverage rates even exceed 100 %.

([17]) We also exclude households' social contribution supplements (D614) from the comparison. Thus, the equivalency is that national accounts transactions D51+D59+D613 link to EU-SILC target variables HY120G+ HY140G.

Figure 11: Coverage rates of taxes on income and actual employees' social contributions, income reference years 2010-2014 (EU-SILC survey years 2011-2015).

(%)



Note: Only countries with sufficiently complete national accounts household sector data available included. Cyprus is excluded as an extreme outlier.

Source: EU-SILC UDB 2011-2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

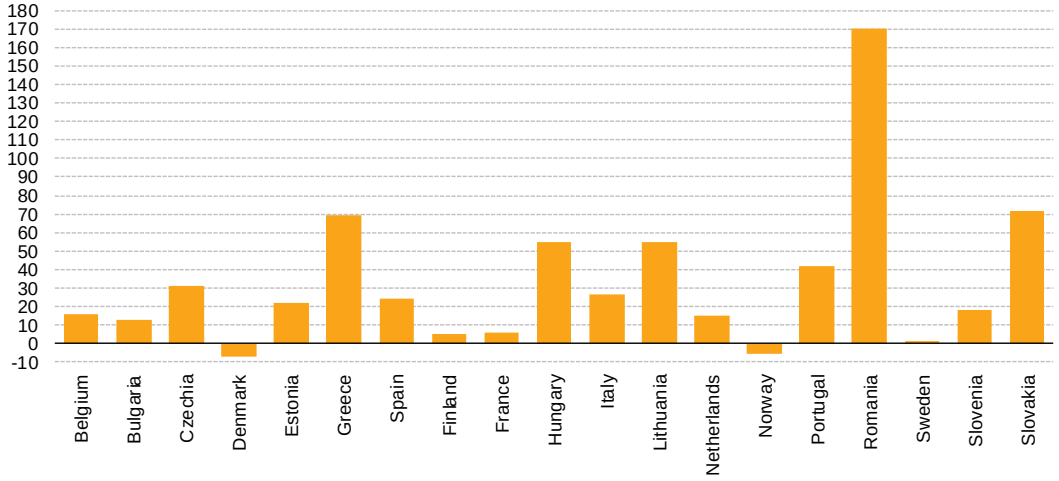
There is more variation between the country coverage rates in taxes than in wages and salaries. One would expect that properly measured coverage rates of taxes would correlate highly with those of wages and salaries. There are cases that suggest some comparability issues with the coverage rates of taxes. For instance, the low coverage rate in Denmark compared to the other register countries calls for an explanation; likewise, the difference in coverage rates in Portugal and Spain require further examination, which however is beyond the scope of this paper.

3 Adjusting EU-SILC survey data with national accounts data

The difference between the national accounts' disposable income and the EU-SILC disposable income is the "gap", which can be interpreted to measure bias in the survey estimates. The gap can then be allocated to households in order to "correct" the distributional estimates derived from a sample survey([18]). The results of such exercises depend crucially on a) how well the gap can be assessed, and b) how it is allocated to households. The adjustment process needs to address the questions of "how much" and "to whom". In this paper, the aim is not for macro consistency, and the gaps are based on modified national accounts aggregates.

Figure 12 shows the overall gaps of disposable income in 2014, relative to EU-SILC total. The gaps ranged from negative to the extreme of 170 % in 2014. This is by how much, on average, we should increase the total disposable income in EU-SILC to be more coherent with a comparable national accounts aggregate. Overall gaps are low in France and the Nordic countries. Note however, that the gaps of the main components may offset each other in such a way that the overall gap appears small.

Figure 12. Gap between EU-SILC and national accounts disposable income (after adjustments), income reference year 2014 (EU-SILC survey year 2015).
(% of EU-SILC disposable income)



Reading note: In Belgium, the national accounts' gross household disposable income was 226.6 billion in 2014. After adjusting for the main conceptual differences (see annex 1), the NA total decreases to 206.7 billion. The difference to EU-SILC total disposable income (179 billion) was 27.6 billion, or 15.4 % of EU-SILC total.

Source: EU-SILC UDB 2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

([18]) In the macro-consistent DINA context, the gaps would need to be allocated in full to households or household groups. Ideally the allocation should be done by making adjustments to the micro data (Zwijnenburg, 2016). This would go towards a micro-founded household sector account, by building a macro-consistent micro data set by record linkage, imputations, statistical matching, and other reconciliations (Coli and Tartamella, 2017).

The allocation of the gaps to households ("to whom") is obviously a major challenge, and crucial for the distributional results. Atkinson, Guio and Marlier (2017, p.78) suggested that sensitivity of social indicators to micro/macro gaps could be examined by proportional scaling of the survey incomes to national accounts by income categories. This is a very straightforward method, which obviously does not reflect the actual distribution of the difference between "true" and measured incomes. However, it is a feasible method and serves as the baseline adjustment in this paper.

Aside from reaching macro-consistency, the gap between national accounts and survey estimates have been used to adjust top of the income, consumption, or wealth distributions. The adjustment method is commonly semi-parametric and relies on Pareto interpolation based on income shares, following Atkinson (2007). The method is useful in particular when tax data is not available for such adjustments (see Jenkins, 2017, for adjustments based on tax data).

Examples of national accounts-based Pareto adjustments include Lakner and Milanovic (2013), and Goda and Sanchez (2017). The latter adjust the top 10 % and top 1 % shares of the Luxembourg Income Study (LIS) data to national accounts totals, distinguishing between labour and capital income. The Pareto adjustment may be combined with other adjustments or be partial, for instance Chandy and Seidel (2017) allocate half of the gap between survey and national accounts estimate to the top decile. A key challenge with the Pareto method is that the choice of the share allocated to the top is quite arbitrary, but it has a large leverage on the results.

Instead of adjusting the incomes, the survey weights can be calibrated so that the estimates of the EU-SILC total amounts match comparable external aggregates. This approach is very useful when exactly corresponding register data can be record-linked to the sample, but it is an option also for the kinds of sensitivity analyses tested in this paper. Consequently, we re-weight the EU-SILC data to reduce the gap between the estimates of total sums.

To summarise, we experiment with three adjustment methods and their combinations in our sensitivity analysis: 1) simple proportionate adjustments, 2) adjustment of the survey weights by calibration to margins, and 3) Pareto-imputations.

3.1. Simple proportionate adjustments by income components

The proportionate adjustments change the levels of each income component, but not their relative distributions. When done by income categories, the income compositions of households change, and consequently also the overall income distribution. The adjustments for regular income components, such as wages and salaries or pensions, are typically much smaller and apply for a larger number of households, in contrast to self-employment or property incomes.

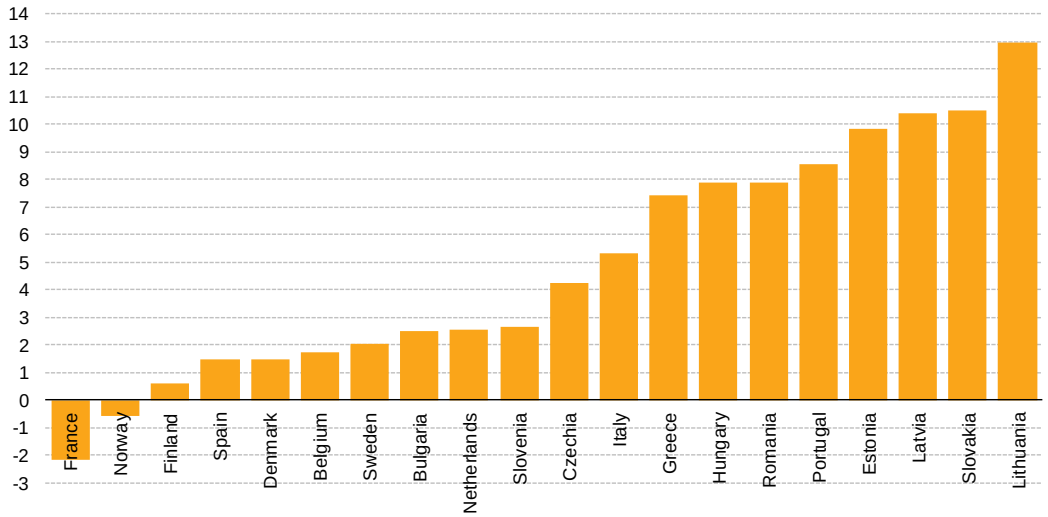
The outcome depends on how many and what type of income components are scaled up. We construct a "scaled-up" version of SILC incomes by adjusting wages and salaries, self-employment income and rents, interest and dividends, transfers received, and taxes paid([19]). All adjustment factors are based on the reconciled income concepts reviewed in section 2.2. That is, the aim is not macro-consistency but adjustment of survey data to assumed data imperfections.

The scaling factors are reported in the annex, and are in a few countries very large, which leads to substantial changes in the income distribution. Figure 13 shows the change in Gini-coefficient. If we use are a rule of thumb of 3 percentage points as a significant change, then in 9 countries the scaling markedly changes the income distribution. The scaling factors are particularly large for property income and self-employment income, which leads to quite large changes in inequality due to increases in the income share of the top decile.

([19]) In principle, gross incomes should be scaled up and taxes then re-computed based on these incomes. Proportional scaling of taxes had only a small impact on the results reported in this section.

Figure 13: Change in Gini-coefficient after simple proportional scaling of main income components to modified national accounts' aggregates, income reference year 2014 (EU-SILC survey year 2015)

(% points)



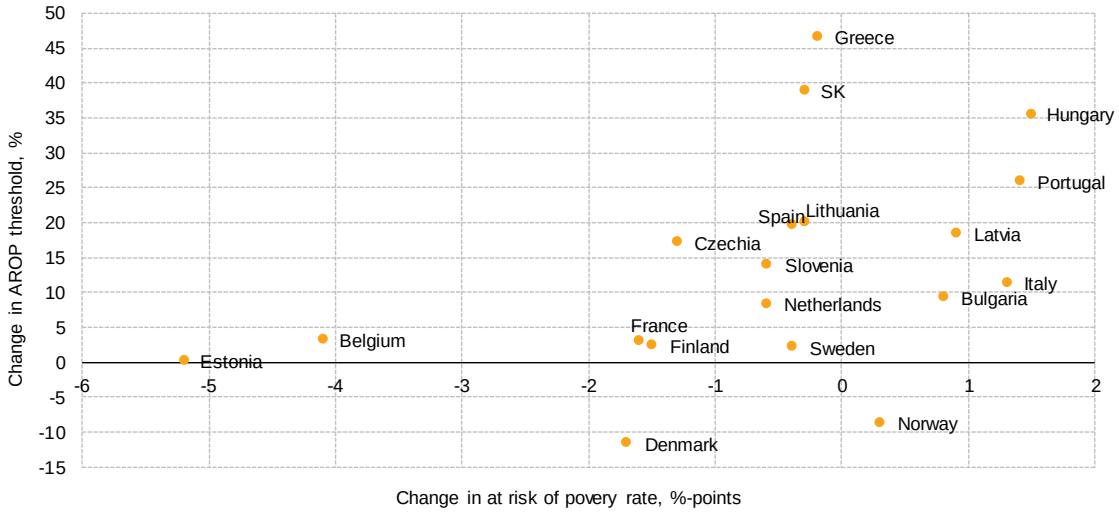
Reading note In Italy, the Gini-coefficient increased by 5.3 percentage points when the EU-SILC main income variables in the 2014 micro data were proportionally scaled up to corresponding national accounts totals.

Source: Author's elaboration from EU-SILC UDB 2015.

Figure 14 plots the changes in at risk of poverty rate and the threshold (i.e. median income) after the scaling up. The effect on at risk of poverty rates is more subdued than in case of inequality, but median income level changes significantly in a few countries. The changes in the AROP rates are within +/- 2 percentage points in most countries. There is a marked decline in Estonia and Belgium, however. The outcome is a result of the scaling factors, their relation, the income structures, and the shape of the original distribution, i.e. a fairly complex process.

The median and therefore the AROP threshold increases significantly in several Southern and Eastern European countries, but very little in Belgium, France, Finland, and Sweden. In Denmark and Norway, it decreases. Although there are very large increases in median incomes, AROP rates can remain almost unaffected (e.g. Greece).

Figure 14: Change in at risk of poverty rate and threshold (60 % of median income) after simple proportional scaling of main income components to modified national accounts' aggregates, income reference year 2014 (EU-SILC survey year 2015)



Note: Romania is excluded as an outlier (increase in threshold by 135 %, AROP-rate by 0.4 %-points).

Reading note: In Italy, at risk of poverty rate increased by 1.3 % and threshold by 11.5 % when the EU-SILC main income variables in the 2014 micro data were proportionally scaled up to corresponding national accounts totals.

Source: Author's elaboration from EU-SILC UDB 2015.

3.2. Re-weighting

As an alternative to simple proportional scaling, the household estimation weights (DB090) can be reweighted by calibrating them to national accounts' income totals. That is, the original sampling weights are changed as little as possible in such a way that the sum of income totals equals the external benchmarks. Deville and Särndal (1992) provide a theoretical description of the calibration technique.

Re-weighting also changes the estimated number of households receiving income (unless these are included in the calibration model), in contrast to simple proportional scaling, where only the income amounts change. The calibration constraints should be orthogonal to each other, which largely is the case when the main income components are used in calibration. However, transfers paid are highly correlated with income received, and cannot be used as a calibration variable. Moreover, very large gaps and/or small number of income recipients may cause non-convergence of the calibration, and indeed this was the case when attempting to use self-employment and property income for countries with very low coverage rates.

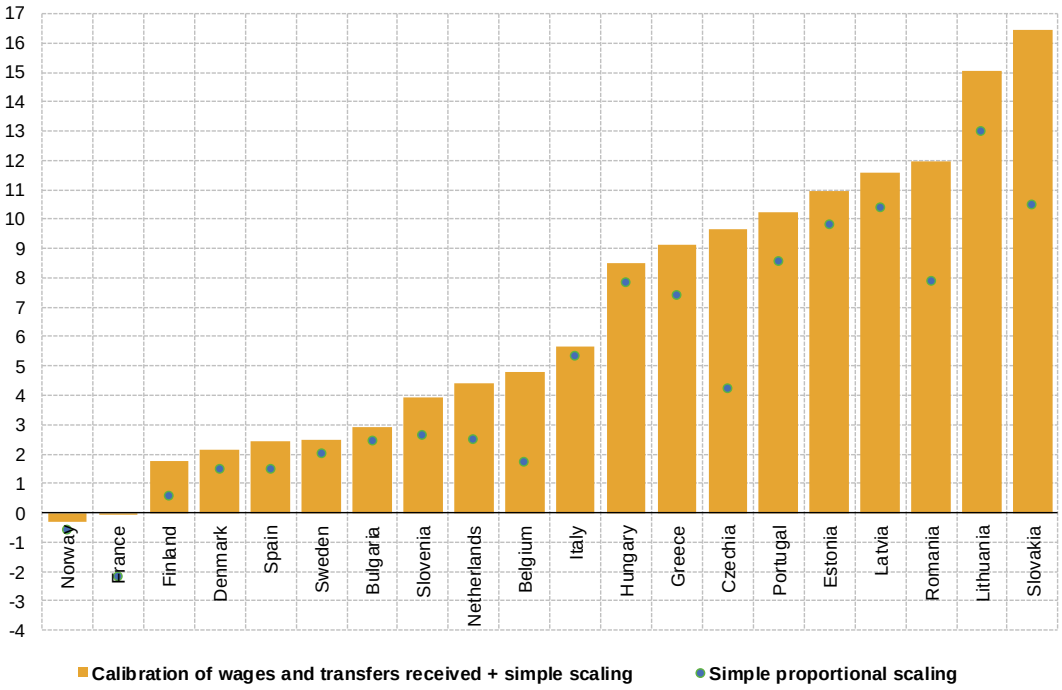
For these reasons, the calibration was restricted to wages and salaries and current transfers received. The calibration totals were the same that were used in the simple proportional scaling. In addition, the calibration constraints include also original estimate of the household size distribution (in four categories), and the number of households and population size, in order to preserve the population structure, and to control for unexpected changes in the weights([20]).

[20] The calibration is done with SAS-macro CALMAR using the logit distance function and weight trimming (upper bound 3, lower bound 0.1 set for the g-weights).

Reweighting also changes the income components that are not adjusted. Thus, after the re-weighting the gaps for these income components have to be recalculated. Self-employment and property income and transfers paid were rescaled proportionally to match the national accounts totals. The total sums of disposable income in the calibration exercise and the simple rescaling experiment are therefore almost equal.

For all countries, the calibration experiment results in higher Gini-coefficients compared to the simple proportional scaling (Figure 15). The difference is substantial in some countries, such as Slovakia, Romania, Czech Republic, and Belgium. There is little difference between the two approaches in Sweden, Bulgaria, Italy, and Hungary.

Figure 15: Change in Gini-coefficient after calibration and simple scaling of main income components to modified national accounts' aggregates, income reference year 2014 (EU-SILC survey year 2015)
(% points)

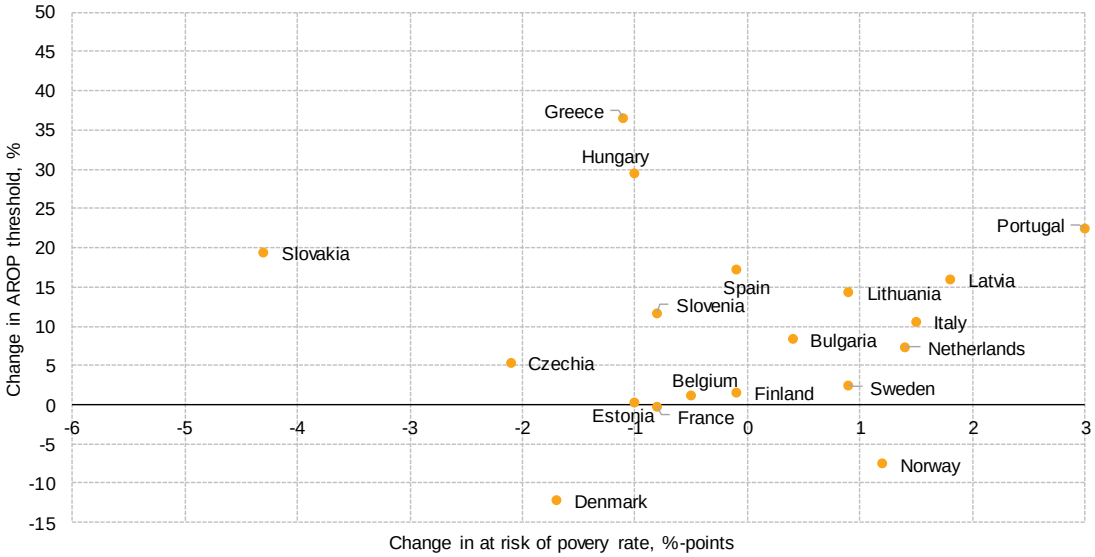


Note: Calibration + simple scaling: wages and salaries and current transfers are reweighted to match modified national accounts' totals, and other income components are proportionally rescaled.

Source: Author's elaboration from EU-SILC UDB 2015.

On at risk of poverty rates, the calibration approach results in somewhat more modest changes than the simple rescaling approach (Figure 16). The changes now range from -4 percentage points in Slovakia to + 3 percentage points in Portugal. The two outliers of the simple proportional scaling (Belgium and Estonia in Figure 14) now have only small changes in the AROP rates. The changes in the AROP thresholds are also more modest in the calibration approach compared to the simple rescaling in a few countries, such as Hungary, Greece, Lithuania, Slovakia, Spain, and Romania.

Figure 16: Change in at risk of poverty rate and threshold (60 % of median income) after calibration and simple scaling of main income components to modified national accounts' aggregates, income reference year 2014 (EU-SILC survey year 2015)



Note: Romania is excluded as an outlier (increase in threshold by 109 %, decrease in AROP-rate by -1.2 %-points).

Reading note: In Italy, at risk of poverty rate increased by 1.5 %-points and threshold by 10.7 % when the EU-SILC main income variables in the 2014 micro data were calibrated and proportionally scaled up to corresponding national accounts totals.

Source: Author's elaboration from EU-SILC UDB 2015.

3.3. Semi-parametric modelling

Instead of scaling incomes of all recipients or changing weights of all households, the gap can be explicitly allocated to a certain segment or segments the distribution. For instance, property income is very concentrated to the top of the distribution, and the gap between survey and national accounts could be entirely or mostly allocated to the very top. This can be done with simple methods, but in this paper, we experiment with semi-parametric Pareto imputation. To preserve the distributional aspect, the values in the top tail are modified or replaced with values, which correspond to a Pareto distribution, while retaining the survey values for the rest. In practice, the top tail is fattened so that the semiparametric estimate of the total sum corresponds to national accounts total.

The distribution of income (and wealth) in the upper tail is often assumed to follow a power law distribution, such as Pareto distribution. The complementary cumulative Pareto distribution function of income y is the following:

(2) $$1 - F(y) = ({}^{k}/_{y})^{\alpha}, \text{ where } \alpha>1, k>0$$

where k is the scale (threshold) parameter above which the power law is assumed to hold, and α is the shape parameter, which measures the heaviness of the right tail. In this paper, we create household-specific adjustment factors by drawing randomly from a Pareto-distribution in such a way, that the tail income total equals the measured values plus the gap to national accounts. This is based on implied tail income of a Pareto distribution (k, α), which is:

(3) $$N * E(x) = N * \frac{\alpha k}{(\alpha-1)}, \alpha>1$$

where N is the sum of EU-SILC weights above income threshold k, and E(x) is the first moment of the distribution, estimated as a weighted sample mean above the threshold k.

The Pareto-replacement strategy is the following:

1. Determine the absolute gap (in EUR) between the survey estimate and the adjusted NA aggregate
2. Determine the fraction of this gap that is to be allocated to the top of the distribution (e.g. 1 or 0.5)
3. Define the threshold k for the top of the distribution, i.e. the segment of the distribution of population size N whose income distribution is assumed to be Pareto distributed, e.g. the 95th percentile of the conditional distribution
4. Compute the empirical Pareto shape parameter α that corresponds to the revised estimate of total income accruing to the top (i.e. the survey estimate + the gap), making use of relationship (2)
5. Draw M random draws of size N from the Pareto distribution (α,k), and use these to adjust the observed values exceeding k in SILC sample of size n; in this paper M is set to 250.

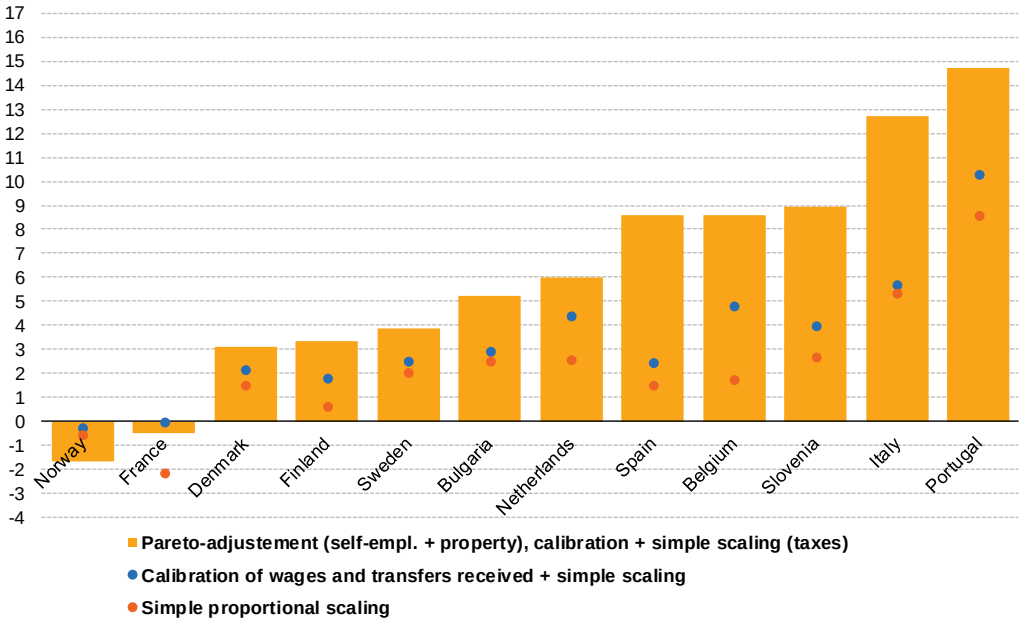
In this paper, we modify the distribution of combined self-employment income and interest and dividend to follow a Pareto distribution([21]). The Pareto assumption is more reasonable for these income components because their distribution tends to be much more skewed than in the case of wages or transfers received. They are combined because the number of households receiving interest and dividends is very low in some countries. For wages and salaries and transfers received, we use the calibration approach, and for transfers paid the simple proportionate scaling.

The threshold is set at the 95th percentile of the conditional distribution of the sum of self-employment income and interest, dividends and profit sharing. For all those above the threshold, the measured values of each household are adjusted so that the tail distribution conforms to a Pareto-distribution. Note that the adjustment factor is household specific, i.e. the Pareto adjustment is done at micro level by changing the observed value of each household in the top. The shape parameter α is determined from the gap between the EU-SILC estimates and the national accounts' sum of gross mixed income, actual interest received, and distributed income of corporations. The EU-SILC estimates and gaps are based on weights that are calibrated to NA wages and salaries and transfers received.

After the Pareto adjustment, the total sums were checked to conform to the benchmarks, and the empirical Pareto index was checked to be close to the value of the theoretical distribution from which the samples were drawn. These checks failed for a few countries, mainly with small sample sizes and large gaps. Figure 17 shows the results for the countries for whom the Pareto imputation could be conducted.

([21]) The assumption that top incomes are Pareto distributed over a certain income threshold is not easy to verify with EU-SILC data, and this would be well beyond the scope of this paper. Such verification is not needed, either, because the hypothesis is that the tail should be Pareto-distributed with certain parameters, but the observed data is not a sample from that population distribution due to measurement errors (under-reporting) and estimation errors (non-participation of the very well off not compensated with re-weighting).

Figure 17: Change in Gini-coefficient after Pareto-replacement of top 5 % of self-employment income and interest and dividends, income reference year 2014 (EU-SILC survey year 2015)
(% points)



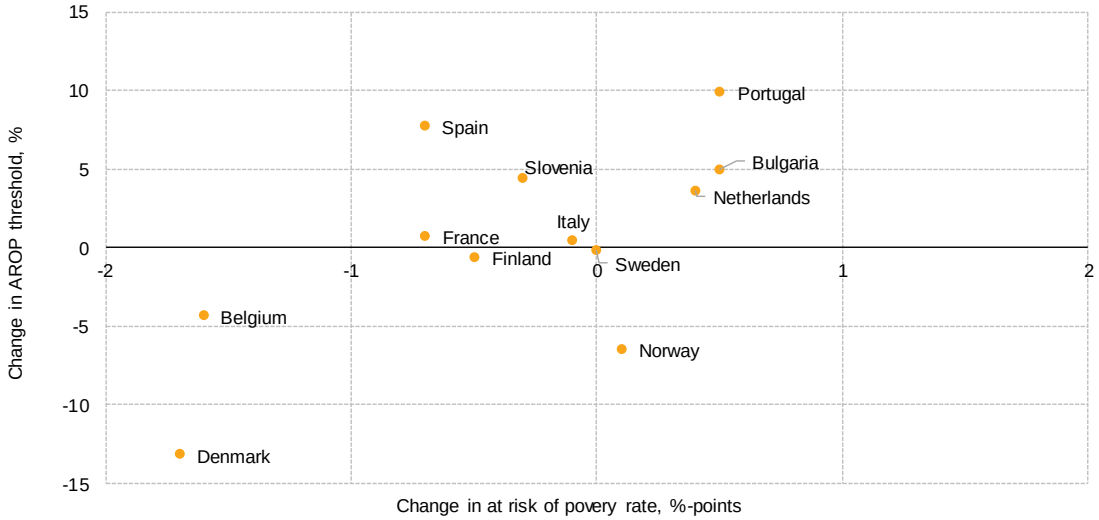
Note: Pareto-replacement of top 5 % of conditional distribution of self-employment income and dividends, calibration of wages and transfers, and simple rescaling of taxes.

Source: Author's elaboration from EU-SILC UDB 2015.

The results indicate that Pareto-imputation approach increases inequality compared to the other methods. This is expected since the total gap in this method is allocated to a fairly small number of households in the top of the distribution, whereas in the other approaches the gap is distributed to the whole conditional distribution. The choice of whether the whole gap or some fraction of it should be allocated to the top is arbitrary, but sensitivity analysis regarding this is left for future work.

As expected, the Pareto-adjustment of self-employment and property income instead of proportionate scaling has relatively small impact on the at risk-of-poverty (Figure 18). Overall, the results are quite similar to the calibration approach. The small differences are probably explained by the pooling of self-employment and property income together rather than by the Pareto imputations.

Figure 18: Change in at risk of poverty rate and threshold (60 % of median income) after Pareto-replacement of top 5 % of self-employment income and interest and dividends, income reference year 2014 (EU-SILC survey year 2015)



Note: Only countries with valid Pareto imputation. The scale is different in Figure 18 compared to Figure 16.

Source: Author's elaboration from EU-SILC UDB 2015.

A source of uncertainty in the adjustments comes from reliance on cross-national data in assessing the gaps. Proper assessment of the gap would benefit greatly from having access to detailed data available only in national sources. For instance, in the case of Finland large imputed items such as net growth of forests and own account building should be removed from the NA aggregate. This alone would result in coverage rate of around 80 % instead of 70 % as implied by Eurostat database. The estimate of hidden self-employment income was 6.5 % in NA, and this could also either excluded or allocated to households in a more nuanced way – it certainly should not be allocated entirely to the top of the distribution.

Regarding property income, for Finland the coverage rate of interest, dividends and profit sharing using cross-national data was 73 % in 2014 for Finland. With more detailed national data, we can further adjust the estimated gap, and the coverage rate with comparable items rises up to 93 %. Consequently, the scaling factors for self-employment and property income derived from national data would be markedly lower than those derived from cross-national databases.

4 Conclusions

The present paper aimed to reconcile the main conceptual differences between EU-SILC and national accounts income aggregates, and compared the estimated total sums based on the adjusted concepts, and finally conducted a sensitivity analysis of key inequality and low income indicators. With regards to all these aims, the study revealed significant differences between countries, types of income, and also between different types of indicators.

The discrepancies between EU-SILC estimates of total amounts and national accounts totals, prior to any adjustments can be very large and are not due to sampling variation. Other features specific to sample surveys and population differences may play some role, but differences in measurement and concepts are the most likely candidates for such large gaps as well their variation among the countries. Overall, the use of register income data improves coherence with national accounts, although there are certain inexplicable differences in the sub-components. The magnitude of the conceptual differences, such as imputed rents and property income attributed to insurance policy holders, also vary between the countries. The necessary adjustments decrease the macro benchmark markedly more in certain countries than the others.

The gaps for property income and self-employment income remain very large even after reconciliation of concepts, while the situation is satisfactory with wages and salaries and transfers received. The assessment of the gaps is challenging and laborious and depends on the quality of the national accounts data and level of detail available in the cross-national databases. If the aim is for "correcting" micro data with macro data, it is essential to measure the micro/macro gap accurately so that it reflects the likely bias in the survey estimates rather than conceptual or methodological differences.

The conceptual and measurement issues are particularly complex for self-employment and property income. Treatment of quasi-corporations, and delineation of "producer" households in general, may differ in national accounts between countries, and income received by households from quasi-corporations may be recorded in mixed income, withdrawals from quasi-corporations, or even as wages and salaries. Even if pooled together, the coverage rates of self-employment and property income remain low in many countries. More work is certainly needed on concepts and measurement of self-employment/mixed income and property income, including methods to impute or model property income based on asset values or external information.

Without knowledge of the distribution of the measurement error, the micro data adjustments are essentially arbitrary. We would not recommend using simple proportional scaling and are in favour of more nuanced approaches. Any serious adjustments should be done with a microsimulation model and properly accounting for what is deemed to be correction of measurement error using macro data and what imputation for reaching consistency with macro aggregates. Moreover, rescaling to national accounts totals using cross-national databases should be exercised with caution, and drawing any distributional conclusions from such adjusted data should be accompanied with sensitivity analysis and full transparency with regard to data sources and methods.

References

Atkinson, A. B. (2007) Measuring Top Incomes: Methodological Issues. In: Atkinson, A.B. and Piketty, T. (eds.): *Top Incomes over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*. Oxford: Oxford University Press.

Chandy, L. and Seidel, B. (2017), "How much do we really know about inequality within countries around the world? Adjusting Gini coefficients for missing top incomes". Brookings Institute Op-Ed, February 17 2017.

Coli, A. and Tartamella, F. (2017), "Developing Households' sub-sectors accounts: Pros and cons of the top-down and the bottom-up methods". Statistical Journal of the IAOS 33(2017) 525-535.

Cynamon, Barry, Z. and Fazzari, Steven, M. (2017), "Household Income, Demand, and Saving: Deriving Macro Data With Micro Data Concepts". Review of Income and Wealth, Volume 63, Issue 1.

Deville, J.-C. and Särndal, C-E. (1992), Calibration Estimators in Survey Sampling. Journal of the American Statistical Association, Vol. 87. No. 418, Theory and Methods, 376 - 382.

ECB and Household Finance and Consumption Network (2016), The Household Finance and Consumption Survey: methodological report for the second wave". European Central Bank Statistics Paper Series 17/2016.

Eurostat (2010), European system of accounts ESA 2010. Publications Office of the European Union, Luxembourg.

Fesseau, M., Wolff, F and Mattonetti, M. L (2013), "A Cross-country Comparison of Household Income, Consumption and Wealth Between Micro Sources and National Accounts Aggregates." OECD Statistics Directorate Working Paper No. 52.

Gregorini F., Grundiza S. and Lamarche P. (2016), "Distributional Income Indicators in a Micro-Macro Data Integration Perspective". Paper prepared for the 34th General IARIW Conference, Dresden, Germany, August 21 – 27, 2016.

Goda, T. and Sanchez, S. (2017), "Market and disposable top income shares adjusted by national accounts data". LIS working paper series 175, August 2017.

Goedemé, T. (2013), "How much confidence can we have in EU-SILC? Complex Sample Designs and the Standard Error of the Europe 2020 Poverty Indicators". *Social Indicators Research*, Vol. 99, no 2, pp. 285-299.

Eurostat (2018), "Comparison of household income: European Union Statistics on Income and Living Conditions and National Accounts." Eurostat Methodological note, May 2018.

Jenkins, S.P. (2017), "Pareto Models, Top Incomes and Recent Trends in UK Income Inequality". Economica, Vol. 84, Issue 334.

Méndez Martin, J.M. (2015), “The change of data sources in Spanish SILC”. Paper presented at the EU-SILC best practises workshop, 16-17 September 2015, London.

Nordberg, L., Rendtel, U., Basic E. (2004), “Measurement Error of Survey and Register Income”. In Ehling and Rendtel (eds.), “Harmonisation of Panel Surveys and Data Quality”. Statistisches Bundesamt, Wiesbaden, 2004.

OECD (2013), “OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth”. OECD Publishing.

OECD (2017), “OECD expert group on disparities in national accounts framework – results from the 2015 exercise”. OECD Statistics Working Paper No. 76.

Pissarides, C.A. and Weber, G. (1989), “An expenditure-based estimate of Britain’s black economy”. Journal of Public Economics 39, 17-32. North Holland.

Rao, J.N.K, Wu, C.F.J and Yue, K. (1992), “*Some recent work on resampling methods for complex surveys*”, Survey Methodology, Vol. 18, pp. 209-2017.

Social Protection Committee (2015), *Portfolio of EU Social Indicators for the monitoring of progress towards the EU objectives for social protection and social inclusion 2015 Update*, European Commission, Brussels.

Statistics Austria (2014), “Documentation on the register-based back calculation of EU-SILC 2008-2011 in Austria”. December 2014.

Törmälehto, V-M, Jäntti, M. and Marlier, E. (2017), “The use of registers in the context of EU-SILC”. In Atkinson, Guio and Marlier (eds.): “Monitoring Social Inclusion in Europe”, Eurostat Statistical Books, 2017 edition. Publications Office of the European Union, Luxembourg.

Törmälehto, V-M, and Sauli. H. (2017), “The distributional impact of imputed rent in EU-SILC 2007-2012”. In Atkinson, Guio and Marlier (eds.): “Monitoring Social Inclusion in Europe”, Eurostat Statistical Books, 2017 edition. Publications Office of the European Union, Luxembourg.

Zwijnenburg, J. (2016), “Further enhancing the work on household distributional data: techniques for bridging gaps between micro and macro results and nowcasting methodologies for compiling more timely results.” Paper prepared for the 34th General IARIW Conference, Dresden, Germany, August 21 – 27, 2016

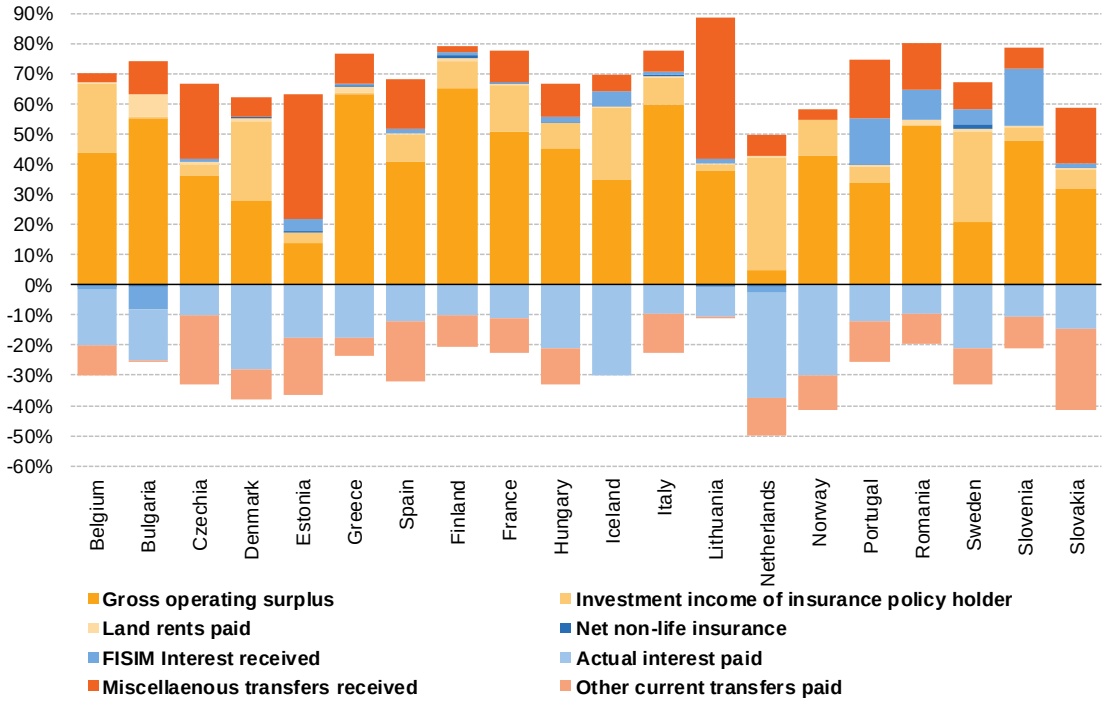
.

Annexes

Annex 1: Review of national accounts items excluded from the comparison

This annex review certain issues related to the conceptual differences between national accounts' gross household disposable income (GHDI) and EU-SILC. Figure A1 shows the relative importance of the items that are excluded from GHDI when computing the overall coverage rates. The most important is gross operating surplus, which is conceptually equivalent to imputed rent in EU-SILC. In the following, we discuss some of the more elusive conceptual differences instead of imputed rent. The figure suggests that there are comparability issues with national accounts data, but these are not explored further.

Figure A1: Main items excluded from the comparison, % of total sum of excluded items. 2014
(% of total sum of excluded items)



Note: Only countries with sufficiently complete national accounts household sector data available are included.

Source: Eurostat, annual sector account tables (nasa_10_nf_tr).

Property income attributed to insurance policy holders

Property income received by insurance policy holders is the imputed flow of income received from the investment of insurance technical reserves, which are assets owned by households but invested by insurance enterprises or pensions funds. Households do not actually receive the interest before the policy matures. It is further broken down to investment income a) attributable to insurance policy holders, b) payable on pension entitlements, and c) attributable to collective investment fund shareholders. The relevance of property income attributed to insurance policy holders varies considerably among the European countries, ranging from non-existent to almost 11 % of gross household disposable income.

In EU-SILC, there are two related income components. The first is income received from private pension insurance plans as annuities or other regular payments in the form of interest and dividends. These are recorded under PY080 regular pensions from private insurance plans. The EU-SILC concept mentions only pension funds, not other life insurance funds, although the latter are included in the international recommendations (e.g. OECD, 2013).

The investment income from investment funds is related also to EU-SILC variable interest and dividends, inasmuch investment funds pay out interest or dividends to the owners. These also may be reinvested by the investment fund instead of distributing the profits to the owners([22]). In national accounts, the imputed income flow is recorded when the asset is held in the insurance company, whilst in micro statistics the income flow is recorded after the policy ends and the savings and the accrued return are paid to the household. Thus, the timing of the recoding is quite different, and in a cross-section the distribution of the received income is likely to be different, with income recorded to older recipients in EU-SILC compared to the imputed flow according to the national accounts concept. For the recipients of private pension as annuity, the receipt includes both run down of the asset value (decrease in wealth) and interest received. The whole amount is typically perceived as income and is used to finance consumption.

If property income received attributed to insurance policy holders has to be allocated to household subgroups, this should be based on information on ownership of insurance savings (insurance technical reserves) rather than income received from such savings, given that the period of receiving income is shorter and occurs later in life cycle compared to saving and asset holding period. For this purpose, the data on ownership of whole life insurance and private pensions from Eurosystem Household Finance and Consumption survey could be used.

Financial intermediate services indirectly measured (FISIM) and interest paid

FISIM is the acronym for financial intermediate services indirectly measured. Banks indirectly charge service charges by charging a higher average rate of interest on the funds that they lend than the average rate of interest that they pay to depositors (OECD, 2014). The margin between these two rates of interest is known as financial intermediation services indirectly measured (FISIM). One would expect it to be positive, i.e. an addition to disposable income, but in Bulgaria and the Netherlands it is negative. As proportion of GHDI, FISIM received tends to be a relatively small transaction. For micro-macro comparisons of income, FISIM received is relevant, it adds to the national accounts disposable income. The FISIM interest is not actual interest received by households, and therefore national accounts interest received should be replaced with total interest before FISIM in the comparisons.

In National Accounts, interest payable is deducted from household disposable income on the uses side, including interest paid on consumer loans. In this paper, to reconcile the SILC concepts with NA concepts, we have chosen to add back the actual total interest paid (incl. FISIM) to NA disposable income. The motivation for this is that imputed rents and therefore mortgage interest payments are not included in SILC income, leaving only interest paid on business loans in the current SILC concept. Total interest paid is a

([22]) For instance, owners of investment funds (mutual funds) may realise their profits as capital gains and not as annuities in the form of interest or dividends. The mutual fund reinvests the dividends and interest the fund earns, and only when the household sells its shares, the realised gains become their cash flow. Since capital gains are not in the income concept, the income received from investment funds may never show up as household income in EU-SILC. In national accounts, the profits from investment funds are included annually as imputed investment income of policy holders.

significant component in certain countries, for instance the Netherlands and Denmark, and there is quite some variation in its share of gross household disposable income. Another item of significance in property income paid are rents, which mainly consist of land rents. These are also added back to NA disposable income (i.e. not deducted) in the adjusted concept.

Non-life insurance

Households receive claims from non-life insurance and pay premiums to such insurances. These are "accident insurances" to cover against risks of theft, road accidents etc. At micro-level, the number of households paying premiums is larger than the number of households receiving claims. The net effect of non-life insurance on household sector disposable income in national accounts is small, because the claims and premiums are almost the same. Elements of non-life insurance may be recorded in EU-SILC variable on income received from individual private pensions (PY080G).

Withdrawal from income from quasi-corporations

This component is part of distributed income of corporations, and hence property income received in national accounts. The treatment of quasi-corporations varies across countries, and is a controversial issue in the reconciliations. The standpoint of this paper is that income received by households from quasi-corporations is included in mixed income or withdrawals from income from quasi-corporations as part of distributed income of corporations (or even in compensation of employees). There is a reclassification issues between income components, but the amounts are recorded as household sector disposable income.

Consumption of fixed capital

The difference between gross and net disposable income amounts to consumption of fixed capital, or depreciation. There is quite some variation in the importance of this item, ranging from 2-3 % to 10 % of gross household disposable income. In the household sector, consumption of fixed capital typically is related to operating surplus and mixed income, the counterparts of which in SILC are imputed rent and self-employment income. In SILC imputed rents are gross of depreciation, whilst self-employment incomes are defined to be net of depreciation.

It could be argued that one should use net rather than gross mixed income in the comparison with self-employment income. This would improve the coverage rates, because of deducting consumption of fixed capital. Conceptually, consumption of fixed capital is deducted from profits and losses from self-employment in SILC (Doc. 65, p. 316). In practice, the measurement may be quite different from national accounts, based e.g. on tax accounts, or not explicit when based on interview data on money drawn out of own business. Net mixed income nor consumption of fixed capital related to it are not publicly available from national accounts data. However, gross and net operating surplus and mixed income are available, as is total consumption of fixed capital (P51C) and this information can be used to approximate the consumption of fixed capital out of gross mixed income as (P51C) x (B3G/B2A3G) (Goda and Sanchez, 2017).

Annex 2: Adjustment factors of main income components, EU-SILC vs. National accounts, income reference year 2014 (EU-SILC survey year 2015)

	Wages and salaries	Self-employment income	Interest, dividends and profit sharing	Transfers received	Transfers paid
Belgium	0.99	1.48	6.27	1.24	1.33
Bulgaria	1.04	1.47	5.19	1.01	0.86
Czechia	1.13	1.76	19.50	1.30	1.67
Denmark	1.04	1.33	1.67	1.06	1.40
Estonia	0.89	6.62	14.50	1.18	1.00
Greece	1.19	1.80	24.15	1.29	0.92
Spain	1.13	1.88	2.96	1.17	1.30
Finland	1.02	1.40	1.38	1.07	1.13
France	1.07	1.02	0.74	1.16	1.32
Hungary	1.43	2.46	23.07	1.18	1.47
Italy	0.96	1.12	15.12	1.14	1.07
Lithuania	1.17	1.38	23.90	1.24	1.29
Latvia	1.09	5.82	12.24	1.10	1.02
Netherlands	1.00	1.51	1.31	1.13	0.98
Norway	0.93	0.35	1.06	1.04	1.05
Portugal	1.06	2.47	9.54	1.19	1.22
Romania	1.52	8.06	170.06	1.12	0.84
Sweden	0.96	2.15	1.55	1.05	1.23
Slovenia	1.08	2.85	1.20	1.09	0.96
Slovakia	1.14	4.54	7.95	1.52	1.12

Note: Only countries with sufficiently complete national accounts household sector data available are included.

Source: Author's elaboration from EU-SILC UDB 2015 and Eurostat, annual sector account tables (nasa_10_nf_tr).

Annex 3: Gini coefficients, at risk of poverty rates and thresholds after adjusting totals for micro/macro gap, income reference year 2014 (EU-SILC survey year 2015)

Country	Gini-coefficients				At risk of poverty, %				Change in AROP threshold			
	Original	Simple scaling	Calibrated	Pareto	Original	Simple scaling	Calibrated	Pareto	Original	Simple scaling	Calibrated	Pareto
Belgium	26.2	27.9	31.0	34.8	14.9	10.8	14.4	13.3	12 993	13 437	13 149	12 431
Bulgaria	37.0	39.5	39.9	42.2	22.0	22.8	22.4	22.5	1 999	2 187	2 167	2 099
Czechia	25.0	29.3	34.7	35.0	9.7	8.5	7.6	7.1	4 454	5 224	4 697	4 460
Denmark	27.4	28.9	29.6	30.5	12.2	10.5	10.5	10.5	17 019	15 080	14 977	14 781
Estonia	34.8	44.6	45.7	44.0	21.6	16.4	20.6	19.0	4 733	4 751	4 748	4 452
Greece	34.2	41.7	43.4	50.2	21.4	21.1	20.3	20.1	4 512	6 618	6 167	5 357
Spain	34.6	36.1	37.1	43.2	22.1	21.7	22.0	21.4	8 011	9 599	9 392	8 634
Finland	25.1	25.7	26.9	28.5	12.4	10.9	12.3	11.9	14 258	14 629	14 494	14 179
France	29.2	27.0	29.2	28.7	13.6	12.0	12.8	12.9	12 849	13 246	12 828	12 946
Hungary	28.2	36.1	36.7	34.5	14.9	16.4	13.9	12.9	2 734	3 709	3 540	3 403
Italy	32.4	37.7	38.1	45.1	19.9	21.3	21.4	19.8	9 508	10 603	10 523	9 554
Lithuania	37.9	50.9	53.0	53.5	22.2	21.9	23.1	21.5	3 108	3 740	3 555	3 385
Latvia	35.4	45.8	47.0	47.9	22.5	23.5	24.3	23.0	3 497	4 143	4 056	3 743
Netherlands	26.7	29.2	31.0	32.7	11.6	11.0	13.0	12.0	12 775	13 867	13 712	13 246
Norway	23.9	23.4	23.6	22.3	11.9	12.2	13.1	12.0	24 890	22 782	23 037	23 294
Portugal	34.0	42.6	44.3	48.7	19.5	20.8	22.5	20.0	5 061	6 379	6 199	5 564
Romania	37.4	45.3	49.4	49.8	25.7	26.1	24.5	23.2	1 389	3 270	2 899	1 799
Sweden	25.1	27.2	27.6	29.0	14.3	13.8	15.2	14.3	15 943	16 303	16 359	15 914
Slovenia	24.5	27.2	28.4	33.5	14.3	13.7	13.5	14.0	7 400	8 449	8 263	7 730
Slovakia	23.7	34.2	40.2	33.6	12.3	12.0	8.0	7.8	4 158	5 779	4 970	4 635

Original = Unadjusted EU-SILC estimate. Simple scaling = simple proportional scaling of main income components. Calibrated=wages and salaries and transfers received calibrated to modified NA totals, self-employment income and interest and dividends proportionally scaled. Pareto = self-employment income and interest and dividends pooled and top 5 % of the conditional distribution replaced with values that follow Pareto distribution, taxes proportionally scaled. Countries marked in red: Pareto adjustment not validated.

Note: Only countries with sufficiently complete national accounts household sector data available are included. Countries marked in colour: Pareto-imputation not validated, for information only.

Source: Author's elaboration from EU-SILC UDB 2015.

Getting in touch with the EU

In person
All over the European Union there are hundreds of Europe Direct Information Centres. You can find the address of the centre nearest you at: https://europa.eu/contact

On the phone or by e-mail
Europe Direct is a service that answers your questions about the European Union. You can contact this service
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by e-mail via: http://europa.eu/contact

Finding information about the EU

Online
Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU Publications
You can download or order free and priced EU publications at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents
For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU
The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

PDF: KS-TC-19-004-EN-N

Reconciliation of EU statistics on income and living conditions (EU-SILC) data and national accounts

The coherence of household survey data with national accounts has been studied extensively in recent years, following the "Beyond GDP" initiatives. This paper compares income aggregates in EU-SILC and national accounts, adjusts for the main conceptual differences, and discusses factors that could influence the observed discrepancies. Following a proposal by Atkinson, Guio and Marlier (2017), sensitivity of key social indicators to the micro/macro-discrepancies is then examined by adjusting the micro data totals to match the reconciled macro aggregates. Three adjustment methods are tested (simple proportional scaling, calibration to margins, Pareto imputation), and their impact on the measures of income inequality and at risk of poverty compared. In line with other studies, the micro/macro gaps are found to vary significantly across countries, and are more substantial in property and self-employment income compared to wages and salaries and transfers received. The observed gaps are likely to be mostly due to measurement errors and conceptual differences. Adjusting the micro data with the gaps results in significant increases in inequality and median income levels, but more subdued changes in at risk of poverty rates. The results are sensitive to the adjustment methods as well as proper assessment of the micro/macro gaps. Caution is warranted if distributional indicators are computed from macro-adjusted micro data.

For more information
https://ec.europa.eu/eurostat/

Publications Office

PDF: ISBN 978-92-76-08674-1