the review of income and wealth

Review of Income and Wealth
Series 63, Number 1, March 2017
DOI: 10.1111/roiw.12202

# MEASURING THE CONSISTENCY OF CROSS-SECTIONAL AND LONGITUDINAL INCOME INFORMATION IN EU-SILC

by Kristina Krell*

*Gesis Leibniz Institute for the Social Sciences, University of Bielefeld*

Joachim R. Frick

*DIW Berlin, SOEP (deceased December 16, 2011)*

and

Markus M. Grabka

*DIW Berlin, SOEP*

The EU-wide survey "Statistics on Income and Living Conditions" (EU-SILC) is extremely important for international social science research and policy advice. It is therefore crucial to ensure that the data are of the highest quality and international comparability. This paper is aimed at identifying unexpected developments in income levels, income mobility, and inequality in the EU-SILC data between 2005 and 2009. We examine the consistency of EU-SILC by comparing cross-sectional results with findings based on two-year longitudinal samples. Although the data represent similar populations, for several countries the results of this comparison differ widely. One important outcome is the high degree of variability over time in countries that obtain their income information from register data. This suggests methodological challenges in the clear designation of new subsample members, in the reweighting of the data, in imputation of missing values, and in other areas.

**JEL Codes**: C89, D31, I39

**Keywords**: consistency, data quality, EU-SILC, inequality, mobility

## 1. Motivation

The European Union Statistics on Income and Living Conditions (EU-SILC) is a Europe-wide survey on income, social exclusion, and living conditions, but also covers aspects such as housing, health, and education. It was launched in 2003 in an effort to achieve "crucial progress" to eliminate poverty by 2010, an objective agreed upon at the March 2000 meeting of the European Council in Lisbon (for an evaluation of recent progress on this objective, see European Council, 2010). First surveyed in 2004, EU-SILC initially covered only 13 EU states and the EFTA states of Norway and Iceland. In 2005, 11 more states were added.[1] The initial plan

*Note*: We would like to thank Heike Wirth, Christof Wolf, and Rainer Siegers, as well as the anonymous reviewers for highly valuable comments and discussions that contributed to improving this paper. All remaining errors are our own.

*Correspondence to: Kristina Krell, University of Bielefeld, Department of Sociology, P.O. Box 10 01 31, D-33501 Bielefeld, Germany (kristina.krell@uni-bielefeld.de).

[1]2004: Belgium, Denmark, Estonia, Greece, Spain, France, Ireland, Italy, Luxembourg, Austria, Portugal, Finland, and Sweden. Additionally in 2005: Lithuania, Latvia, Poland, Slovak Republic, Hungary, Czech Republic, Slovenia, Cyprus, Germany, the Netherlands, and the United Kingdom.

was to produce comparable data for each participating country with the broader aim of identifying effective methods of fighting poverty ("best practices") that could be adapted to conditions in other countries in order to facilitate the process of social convergence among all Europeans ("open method of coordination"; see European Commission, 2001).

However, achieving this objective requires that the micro-data collected by EU-SILC fulfill high quality standards across all participating countries, as stipulated in the data requirements of the European Parliament and Council (European Parliament and the Council, 2003, 2004). A potential bias in the data could arise in both the pre-data collection and the post-data collection phase. Possible causes include deviations in sampling strategies (Groves *et al*., 2009), follow-up rules for contacting (former) household members (Schonlau *et al*., 2011), and different data preparation procedures for editing outliers (Little and Smith, 1987) or handling item non-response (Rubin, 1987) and weighting (Groves and Couper, 1998).

We discuss the variety of survey methods within this comparative study and examine the consistency of EU-SILC by comparing cross-sectional income levels and inequality results with findings based on two-year longitudinal samples. We complement these findings by mobility analyses. EU-SILC contains two datasets for each wave: one including the cross-sectional data for a single wave, and the second including the longitudinal data for each year up to this wave. The cross-sectional and the longitudinal datasets cannot be merged case by case in the user's database version for reasons of data protection. Therefore, the cross-sectional and longitudinal analyses in this paper are conducted on the basis of two different data files.

Because of EU-SILC's high importance in international poverty research, this paper concentrates on analyzing and evaluating the level and development of households' disposable income as reflected in the "Laeken indicators" and on the dynamics of household and personal income and poverty over a two-year period (income and poverty mobility). The cross-sectional and longitudinal data for each country should produce nearly the same income, poverty, and inequality results, as the samples represent almost identical populations.

Nevertheless, as noted above, the comparison of cross-sectional and longitudinal results brings serious inconsistencies to light in some countries. In view of this fact, the present paper contributes to the discussion of different survey methodologies and their impacts by providing a descriptive exploration of problematic features in the data of the EU-SILC countries (see Glaser *et al*., 2015, for another very recent contribution). Of course, descriptive analyses have only limited explanatory power. Since it is impossible for us to perform an in-depth quality check for each EU-SILC country as was done for the German data by Frick and Krell (2010), this paper is designed to offer an impetus for further investigation of the data quality of EU-SILC.

The paper is organized as follows. Section 2 describes the database EU-SILC, discusses important methodological issues, and presents population and income definitions. In Section 3, we show the results of the empirical analysis and identify problematic issues in particular countries that are associated with the mode of data collection used. The analyses are descriptive and cover the cross-sectional population only (considering poverty and inequality measures), a comparison of results

based on the cross-sectional and longitudinal population, as well as the analysis of the longitudinal population only (considering income and poverty mobility). Section 4 concludes with a discussion of the problems identified and some suggestions for future research.

All associated figures and tables can be found in the online appendix.

## 2. EU-SILC: Methodological Issues and Data Processing

### 2.1. *Methodological Issues*

EU-SILC contains a wide range of socio-demographic variables such as gender, age, income, health, and education of household members, as well as indicators of household composition, living situation, and non-monetary deprivation (European Commission, 2006). In order to increase the comparability of results across all countries and to maintain sufficient quality standards, a series of guidelines was developed by Eurostat on sample design and size, imputation, weighting, data processing, etc. (European Parliament and the Council, 2003, 2004). Based on experience with the precursor study—ECHP, an input-harmonized survey with a common survey instrument whose rules had proven to be too rigid—the rules for EU-SILC permitted survey methodologies and data processing to be adapted to national conditions (output-harmonization).

The survey population comprises persons living in private households. Once a household is selected, each household member above the age of 16 is contacted to be interviewed.

Different types of *interview modes* are used. Most countries carry out their interviews face-to-face by Paper And Pencil Interviewing (PAPI) or Computer Assisted Personal Interviewing (CAPI). An exemption is Germany, where interviews are usually self-administered by the respondent. Nordic countries as well as the Netherlands and Slovenia (and to a limited extend also Ireland and Latvia; see Lohmann, 2011) derive their income information from registers and usually conduct telephone interviews in order to obtain the remaining information (see also Rendtel *et al*., 2004; Epland, 2006). They use the concept of a "selected respondent" who answers the household as well as the personal questionnaires for all household members. The sample is therefore a sample of persons, not households. But it is not only in register countries that proxy interviews are widely used. Ireland and Spain use substitution households which are contacted in case the originally selected household refuses to respond or cannot be contacted, but without giving sufficient information on the substitution process (Verma, 2007, p. 293). Because this undermines the probability nature of the sample and the affected households cannot be identified by the user, Iacovou *et al*. (2012, p. 4) do not recommend the use of the data from these latter two countries in analyses that aim to draw broader conclusions on their populations.

EU-SILC is designed as a *rotating panel* with a panel duration of four years; thus, a quarter of the sample is generally replaced each year by new subsample members. Each household is therefore contacted up to a total of four times. Most countries follow this recommendation—the exceptions are France (nine years panel duration), Norway (eight years panel duration), and Luxembourg (pure panel).

Following from this, the sample sizes of the cross-sectional and the corresponding longitudinal datasets differ by up to a quarter each year. Given that this "methodological" drop out of the sample happens randomly like the selection into the sample, these differences should have no implication for the representativeness of the data.

Measurement error is—like in every survey—a common phenomenon in EU-SILC (Verma *et al*., 2010). An important source of measurement error is unit non-response. Verma (2007) argues that the probability nature of the sample is restricted given that, for example, unit non-response rates in EU-SILC vary between 12 percent in Greece and 52 percent in Belgium. This issue has to be handled by *weighting*. The generation of weights comprises various steps including the creation of design weights, the consideration of non-response of original sample members, the generation of base weights for non-original sample members, a weight sharing for all household members, and finally the integration of the four rotating subsamples (Rendtel, 1995). This also includes the calibration, trimming, and scaling of the final weighting factors (for more details, see Verma *et al*., 2007). However, Iacovou *et al*. (2012, p. 6) note that the accompanying documentation "is unclear on whether these are only design weights (i.e., whether they correct only for the probability of selection into the sample), or whether they also adjust for non-response—and if so, in what countries."

They also criticize the fact that adjustment for non-response has not been carried out in a consistent manner across countries. "Different treatment of non-response, and especially correcting for it in some countries but not others, may lead to biased cross-country comparisons" (Iacovou *et al*., 2012, p. 6). Item and partial unit non-response are further sources of measurement error. Missing income components (item non-response, INR) and entirely missing incomes of certain household members (partial unit non-response, PUNR) are not treated in a harmonized way in each country (see Frick *et al*., 2012, for a general discussion about the handling of PUNR in population surveys). *Imputation* methods are briefly described in a series of the countries' quality reports as well as in a data accuracy report of Verma and Betti (2010). In case of PUNR, in some countries (Belgium, Czech Republic, Estonia, France, Cyprus, Italy, Lithuania, Austria, Poland, and United Kingdom) full-case imputation is used for missing incomes (starting with the 2009 wave). Other countries (Bulgaria, Germany, Greece, Latvia, Portugal, Romania, and Slovakia) adjust the household income by a simple flat correction factor that is derived from other households with similar characteristics. Ireland, Malta, Luxembourg, and Hungary, as well as all register countries, delete all households that are affected by PUNR, although PUNR cannot be considered to be random and this method leads to a severe loss of efficiency. Regarding the imputation methods themselves, there are brief descriptions in several of these countries' quality reports. The reports describe the use of deductive (net-gross-conversion, for instance), deterministic, or stochastic methods, and the use of cross-sectional information only (of other, completely interviewed households, i.e., hot-deck imputation procedure) versus a combination of cross-sectional and longitudinal information (of the same respondents from prior waves). A potential source for differences across countries could arise from the consideration of longitudinal information in the imputation process. When

panel information is used, the quality of the imputation is clearly enhanced (Spieß and Goebel, 2005). The only countries which consider this feature in their imputation process are Austria, Estonia (up to 2007), Cyprus, the Czech Republic (from 2007 onwards), and United Kingdom. Although it is well known that multiple imputation (Rubin, 1987) is superior to single imputation techniques, no EU-SILC country reported using this strategy.

### 2.2. *Data Processing*

The population analyzed consists of individuals in private households. On a cross-sectional basis, we use the income data and household characteristics from 2005 to 2009. On a longitudinal basis, four two-year periods are considered: 2005/2006, 2006/2007, 2007/2008, and 2008/2009. The panels are composed on the basis of the longitudinal datasets. Each of these two-year panels consists of all households that responded in both years that are considered. Each panel dataset—for instance 2008/2009—includes income and household information for the same households in 2008 and 2009, respectively.[2] The income reference period is usually the year before the observation, except in the case of Ireland, where the reference point is 12 months prior to the interview, and the U.K., where respondents are asked to report their current income.

For income analyses, we use the concept of equivalized disposable household income using the modified OECD scale. Incomes are displayed in 2005 prices. In contrast to Eurostat's published results, we only consider incomes that are positive or zero, which is in line with the recommendation of Verma (2007).

Non-cash incomes are not taken into account in our analyses. Although the Canberra group recommends the consideration of imputed rent (the estimated amount of money saved when living in one's own property or subsidized dwellings; Canberra, 2001), it has only been provided since 2007 in EU-SILC and comparability is limited given that the methods of deriving this fictitious income advantage are not harmonized across EU-SILC countries, despite the fact that the methods used affect the results (Frick and Grabka, 2003). We also refrain from considering other forms of in-kind pay such as company cars; for the incidence and relevance of such fictitious income advantages, see Frick *et al*. (2007).

The following empirical analyses separate the EU-SILC countries into three groups, based on similarities in their survey mode.

The "*register countries*" group includes countries that derive income information of the participating households partly or completely from registers. The interview is usually done by telephone. These are: Denmark, Sweden, Norway, Iceland, Finland, Slovenia, Netherlands, Latvia, and Ireland.

Countries that derive income information from interviews ("*survey countries*") and that conduct not more than 15 percent non-personal interviews on average are: Belgium, Greece, United Kingdom, Czech Republic, Estonia, Hungary, Bulgaria, Romania, and Slovak Republic.

[2]The mean income in one country in 2008 on cross-sectional basis refers to the whole population in this country in 2008, whereas the mean income in one country in 2008 on longitudinal basis 2008/2009 refers only to the population that lived in the country in 2008 as well as in 2009 (without the deceased/emigrated population in 2008, and the newborn/immigrated population in 2009).

Finally, "*proxy countries*", that is, Germany, Spain, France, Italy, Luxembourg, Austria, Portugal, Cyprus, Lithuania, and Poland conduct on average more than 15 percent of the interviews non-personally by proxy-interviews.

## 3. Empirical Analysis

In the following analysis, we make use of standard income inequality indicators such as the mean and the median, the well-known Gini coefficient, and the mean log deviation (MLD). The latter indicator belongs to the so-called "generalized entropy class" and is sensitive to changes at the bottom of the income distribution. In addition, we make use of the Foster–Greer–Thorbecke indices (FGT), especially the FGT(0), which represents the poverty head count ratio (Foster *et al.*, 1984). The threshold for relative income poverty is based on 60 percent of the median equivalized disposable household income. We compare these indicators by performing cross-sectional as well as longitudinal analyses to detect conspicuous or exceptional developments over time, which might suggest potential problems in the underlying micro-data.

### 3.1. *Cross-Sectional Analyses*

Figure 1 shows the development of equivalized disposable mean income of each country from 2005 to 2009. We only report changes that cannot be assumed to be of random nature (based on computed confidence intervals of two consecutive years that do not overlap). Naturally, in countries showing a rather small change in real GDP, one would also assume a fairly stable level of real mean income from one year to the next. This is, according to Figure 1, not the case in Germany, France, and Iceland. Germany shows a slump in mean income in 2006 of more than 4 percent from €18,512 to €17,371, but then a sharp increase of about 17 percent to €20,062 in 2007. For the same period, the German Socio-Economic Panel (SOEP), which is an alternative data source for income analyses in Germany, shows mean income variations between €19,200 and €19,500. Until 2007, the German EU-SILC sample was partly a quota sample, gradually replaced by a random sample—a fact that makes the results not reliable (for details on this case, see Frick and Krell, 2010). France shows an increase in income of about 23 percent from €18,000 in 2007 to €21,600 in 2008, which could be explained by the fact that from 2008 onwards, the French income data in EU-SILC are complemented by register data from tax sources. Therefore, Eurostat treats this difference as "break in series" and the results of 2007 and 2008 as not comparable. This would support the results of Lohmann (2011) and Rendtel *et al.* (2004) on the strength of the impact of the survey method on income measures. Information from national accounts for gross disposable incomes per capita does not indicate such a pronounced increase (for details on national accounts for gross disposable incomes per capita, see http://ec.europa.eu/eurostat/data/database). In Iceland, real mean income has risen by 29 percent from €26,700 in 2005 to €34,500 in 2008. In 2009 there is a plunge of 37 percent. Although Iceland experienced a strong weakening of the economy during this period, figures from the OECD income distribution database show only a decrease of about –7 percent in mean disposable

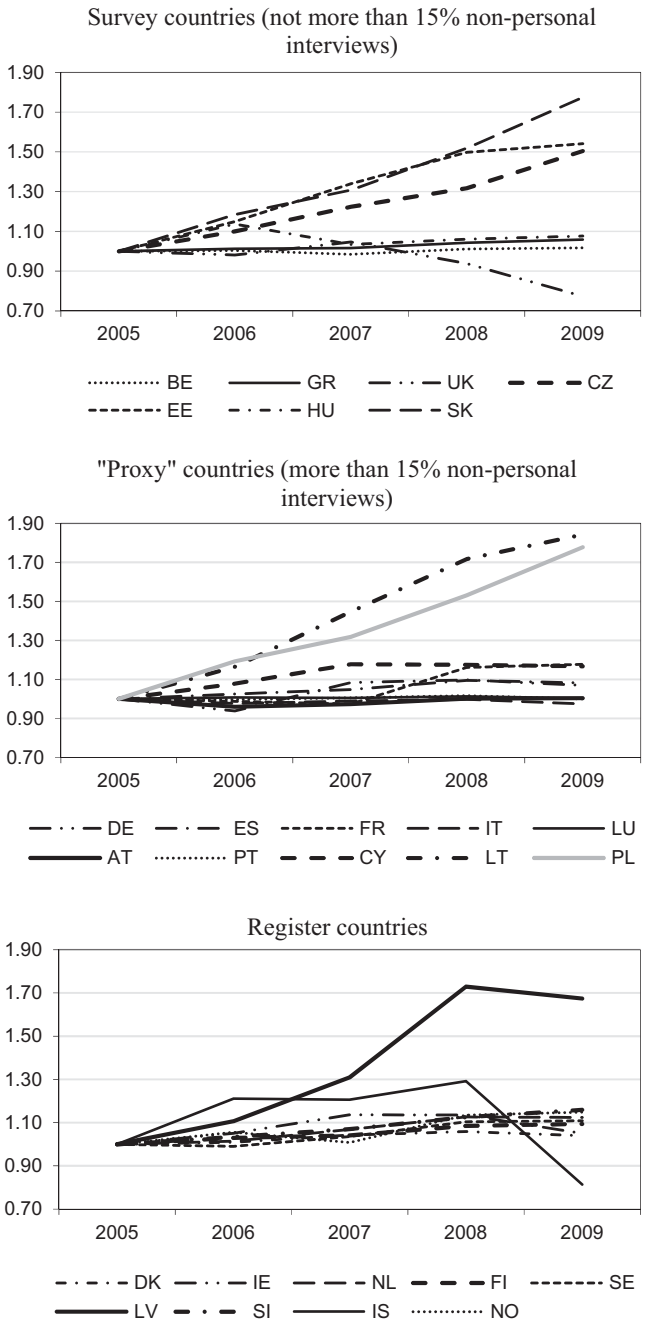


Figure 1. Development of real mean equivalized disposable household income (standardized to 2005 = 1).

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional data 2005 to 2009.

household equivalized income in the same period (see http://stats.oecd.org/index.aspx?queryid=46022, accessed by the authors on March 10, 2015). The results for the U.K. reflect the negative economic development from 2007 onwards. However, the 26 percent decrease in mean income seems implausibly large and is also not displayed in the national accounts provided by Eurostat.

The transition countries, which are the Eastern European as well as the Baltic states, are expected to show rather strong increases in real mean income, given that in Latvia, for example, real GDP rose by more than 10 percent each year between 2005 and 2008. This is confirmed, since all transition countries experienced a huge income increase of at least 23 percent from 2005 to 2008. However, Latvia far exceeds these levels with an extraordinary increase of nearly 120 percent in the same period. Hungary shows an income increase in 2006 followed by a decrease in 2007. Lelkes *et al*. (2009) suspect that Hungary might have had problems in the representativeness of its data in 2006 due to the low response rate of only 51 percent.

More than the development of the mean income, the developments of relative income poverty and inequality measures reveal the continuity and plausibility problems of some countries in EU-SILC. Figure 2 displays the development of the risk-of-poverty rate, FGT(0), over time. It becomes obvious that the transition countries, although showing a similar economic development, differ widely in the development of poverty. Whereas in most countries (e.g., Poland, Slovak Republic, Czech Republic) FGT(0) considerably has fallen, Latvia experienced a major increase in poverty by 35 percent between 2005 and 2009. Again, Hungary stands out by its erratic poverty measure, which rose from 13.4 percent in 2005 to 15.7 percent in 2006 before falling to 12.4 percent in 2007. The Southern European countries show by and large a homogeneous trend in poverty rates, with rises or falls of a maximum of one percentage point over time. Within the "proxy" country group, Germany shows a sharp increase of FGT(0) in 2007 from 12.4 to 14.9 percent, which is not in line with the economic situation in Germany in 2006, when unemployment fell substantially, which is usually interpreted as linked to a decrease in poverty (at least for Germany, see Grabka and Frick, 2011). Among the register countries, Sweden "breaks ranks." Here, poverty rose by 35 percent from 2005 (9.2 percent) to 2006 (12.1 percent), then fell in 2007 (10.7 percent) to rise to 13.2 percent in 2009. Another notable development can be observed in Ireland, where poverty rates fell from 19.7 percent in 2005 to 15.3 percent in 2009, although this is accompanied by a rather severe economic situation preceding the great recession in that country.

Examination of the income inequality—here displayed by the Gini coefficient in Figure 3—shows nearly the same picture as the poverty rate in the majority of countries.[3] Again, in 2006, Hungary shows an exceptional peak in income inequality, which rose by about 24 percent over the previous year, but Slovakia also experienced a sharp increase in inequality of nearly 10 percent in that year. Given that relative income poverty increased in Sweden, one could also assume a growth in inequality. However, among the register countries, it is not Sweden, but Norway

[3]Figures of the median income and the Mean Log Deviation (MLD) inequality measure present nearly the same picture as the mean income and the other inequality measures. They can be obtained from the authors.

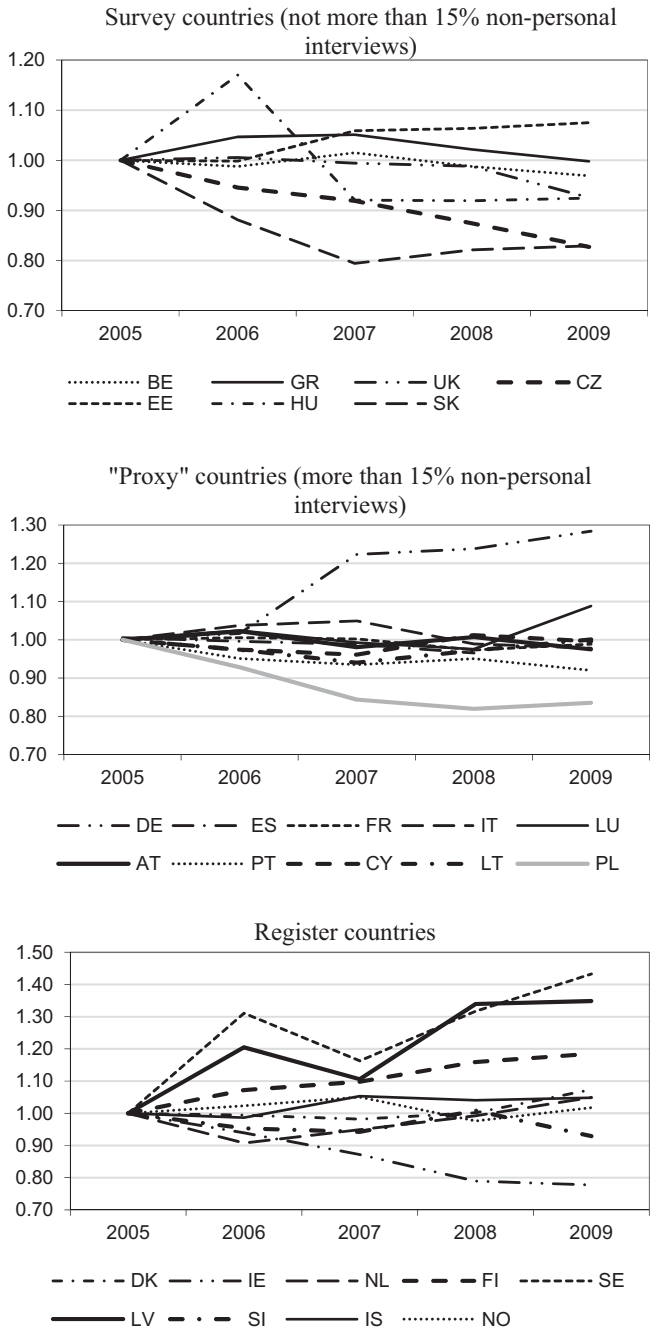


Figure 2. Development of relative income poverty (FGT(0)) (standardized to 2005 = 1).

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional data 2005 to 2009.

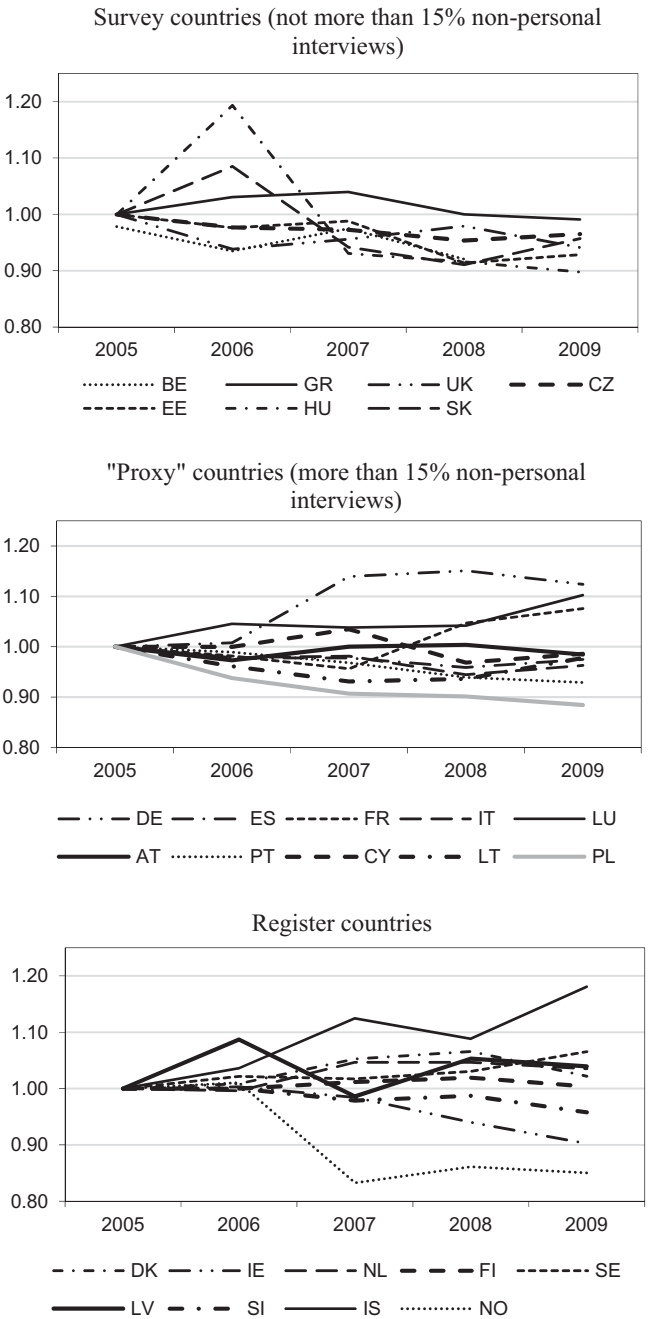


Figure 3. Development of Gini coefficient (standardized to 2005 = 1).

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional data 2005 to 2009.

that shows an extreme decrease in inequality of nearly 20 percent from 2006 (0.284) to 2007 (0.234). It is noteworthy that register countries such as Norway, but also Iceland, whose income information should be counted among the most detailed of all the countries in our sample, show some of the most extreme developments. In principle, major changes over time in income levels, poverty, and inequality cannot be ruled out, but such dramatic developments in a two-year period as shown in this data—especially in Latvia, Germany, Sweden, Iceland, and Norway—must be studied carefully.

Analyzing the German EU-SILC data in Frick and Krell (2010) revealed that poverty thresholds and in turn poverty rates are substantially affected by misrepresentation of certain population groups, such as people from the former East Germany or people with low education, as well as by inadequate weighting factors. Given that a quarter of the sample is replaced in most countries every year, such a break may also occur within two years. Thus, the results presented could only be interpreted as an initial indication of potential inconsistencies or problems in the EU-SILC micro-data.

### 3.2. *Cross-Sectional vs. Longitudinal Results*

In this section, income levels, poverty, and mobility results for the cross-sectional and longitudinal populations are compared (see also Glaser *et al.*, 2015). The rationale for this comparison is that one could assume rather similar results given that only ordinary changes occur within one year in the underlying population due to birth, death, and migration, which should affect the results only to a minor degree. However, migration could in principle have a relevant effect on these analyses, but no prominent migration happened in Europe in the observed period of 2005 to 2009. The rotational design of EU-SILC could also theoretically have an effect given that each year, a quarter of the total population is replaced, but if the new subsample is a proper representation of the underlying population, no relevant distortion should be expected. However, any new subsample is usually affected by non-response; thus, proper weighting is a basic prerequisite when non-response is suspected to be non-random (Groves and Couper, 1998).

Figures 4 to 7 show indicators that are constructed as follows. Four two-year panels are analyzed. For each year, we divided the results for the longitudinal population (e.g., 2005/2006) by the results for the cross-sectional population (e.g., 2005). Therefore, for 2006, 2007, and 2008, there are two values because two longitudinal populations are affected (e.g., 2005/2006 and 2006/2007 for the year 2006). Each symbol therefore represents one fraction. In cases where the two considered 90 percent confidence intervals do not overlap, symbols are filled in the figures. Because the samples that are compared are not independent of each other, we avoid the term "significantly different" in these cases. The symbols are expected to lie above but close to the line for mean income and below the line for poverty and inequality, because the longitudinal population does not contain population groups that are affected to an above-average degree by poverty: for instance, the newborn, the deceased, but also immigrants (e.g., Zaidi, 2006; Lelkes, 2007; BMAS, 2012).

As one can see in these figures, this is true for most countries and years. To emphasize some positive examples, the Polish, Romanian, Cyprian, and Italian

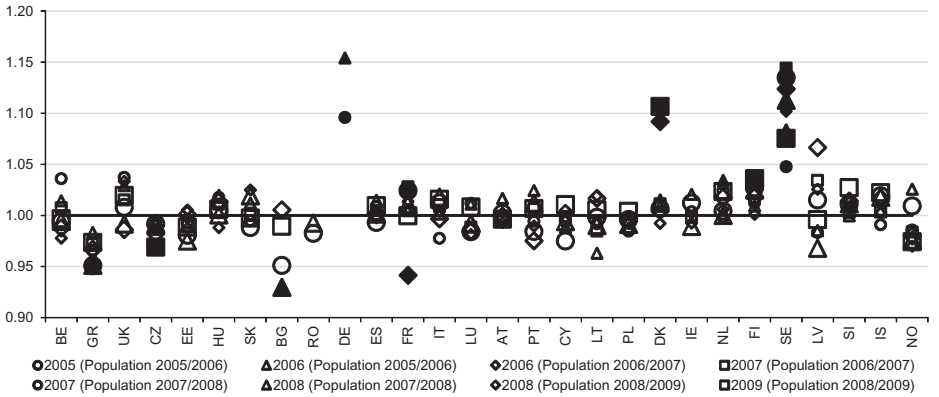


Figure 4. Mean equivalized disposable household income—ratio of longitudinal and cross-sectional results 2005–2009.

Symbols are filled when confidence intervals of cross-sectional and longitudinal estimations do not overlap. For better legibility of this figure, we do not show the value of CY (2008), which is a ratio of 1.7 and SK (2008) with a ratio of 0.03.

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional and longitudinal data 2005 to 2009.

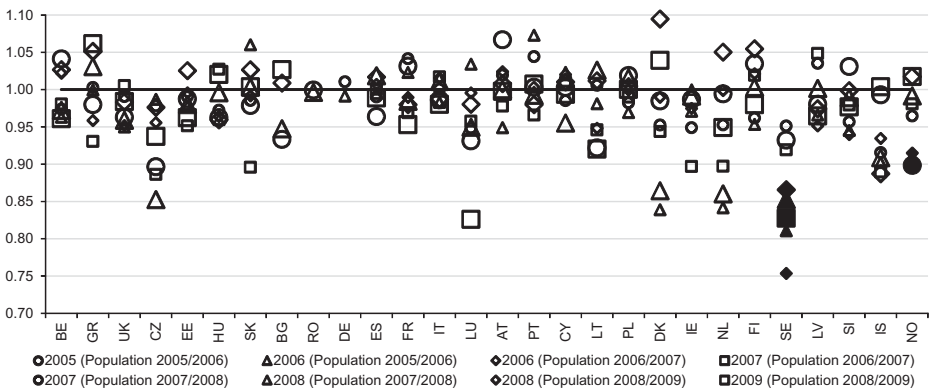


Figure 5. Poverty rate FGT(0)—ratio of longitudinal and cross-sectional results 2005–2009.

Symbols are filled when confidence intervals of cross-sectional and longitudinal estimations do not overlap.

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional and longitudinal data 2005 to 2009.

data show fractions of just ±2–4 percent for all four presented indicators, which are most likely the result of the aforementioned and expected demographic shifts. However, in a number of countries, the difference is more than 10 percent, but in many cases the confidence intervals are quite large and thus overlap. It is interesting that again countries are affected by extreme deviations that derive their income information from registers—Denmark, Norway, and Sweden in particular. Deviations in the poverty rate and in the MLD in Sweden and Norway range from 5 to

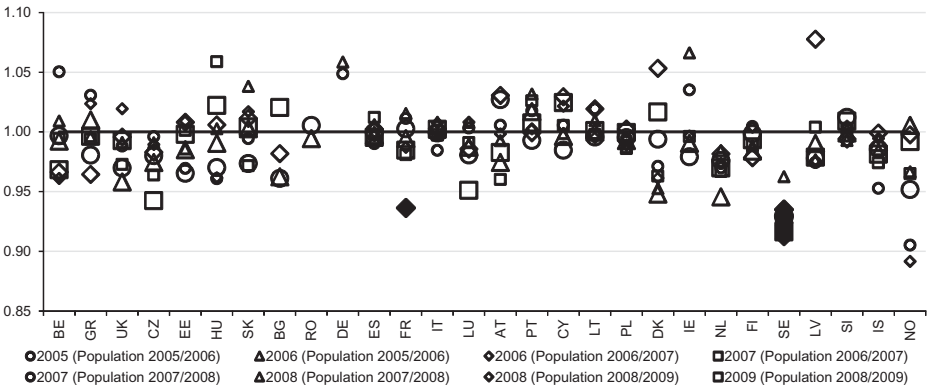


Figure 6. Gini coefficient—ratio of longitudinal and cross-sectional results 2005–2009.

Symbols are filled when confidence intervals of cross-sectional and longitudinal estimations do not overlap.

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional and longitudinal data 2005 to 2009.

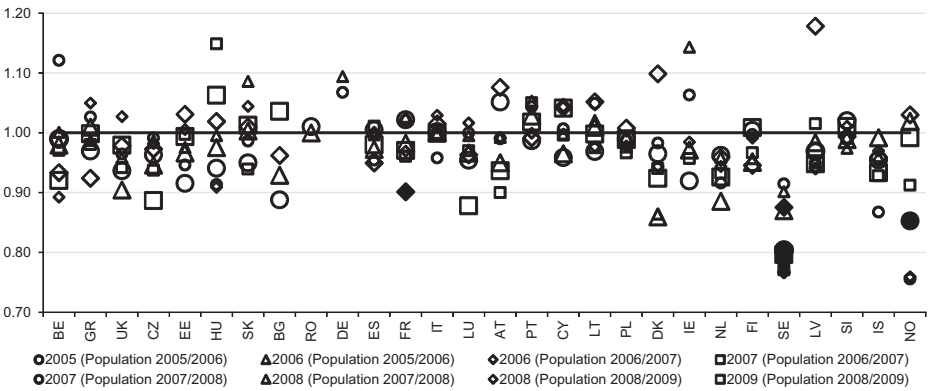


Figure 7. Mean Log Deviation—ratio of longitudinal and cross-sectional results 2005–2009.

Symbols are filled when confidence intervals of cross-sectional and longitudinal estimations do not overlap.

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional and longitudinal data 2005 to 2009.

25 percent (see also Jenkins and van Kerm, 2014). Cyprus seems to have serious errors in the weighting variable for the 2008/2009 panel, which leads to a deviation in mean income of 70 percent in 2008 (€30,700 in the 2008/2009 longitudinal population versus €18,000 in the 2008 cross-sectional population). The results for Slovakia are even more extreme, with an annual mean income of €145 for 2008 in the 2008/2009 population (both deviations are not shown in the graph for reasons of legibility). In this case, it is obvious that the income data for this panel population are incorrect.

There are several means of explaining the differences between cross-sectional and longitudinal results. One possible reason, aside from the underrepresentation

of poorer population subgroups in the balanced panel, is that income extremes are more likely to drop out of the panel, which may cause a different distribution of income in the observed sample if not reweighted properly. Analyses of Glaser *et al.* (2015) show a rather strong relation between non-response rates and incoherence (pp. 34 ff.). They see a potential problem also in weighting procedures in the longitudinal data, as the longitudinal weights "rely mostly on a year by year base weight adjustment by estimated response propensities. Further adjustments like calibration are usually not carried out for longitudinal weights" (p. 38).

The large discrepancies primarily found in the register countries suggest that there is a problem in the representativeness of the sample or in the weighting factors rather than in data collection of income itself or in imputation (given that register countries do not impute at all). Iacovou and Lynn (2013, p. 2) point to the fact that "[a]ccording to the regulations, co-residents [of sampled individuals in register countries] only need to be followed in the next wave of the survey if they remain living in a household containing at least one sample person. Thus, in the register countries, if a household splits into two parts between the first and second waves, only one of the two parts would be followed (as only one part contains the sample person), whereas in non-register countries both parts would be followed (as both contain at least one sample person). In consequence, many members of initially-observed households will not be followed in register countries," and further that "in the register countries, no adults are followed who do not continue to live in the parent household" (Iacovou and Lynn, 2013, p. 11). This fact could undermine the representativeness of the samples in register countries, if no proper reweighting is applied. Glaser *et al.* (2015) conclude from their results that the selected respondent design applied in the register countries needs to be reviewed in more detail in order to figure out the design's effect on data coherence. Further, at least in Sweden, calibration of the base weights seems to be based only on the marginal distribution of age and gender (Glaser *et al.*, 2015). Second, as mentioned in Section 2, households that are affected by PUNR are completely removed in register countries and a number of other countries. Some population groups—e.g., people with very high or very low incomes—are more likely to refuse to reply (Frick and Grabka, 2010). Again, the data may be biased without appropriate weighting. Sharp differences in the weighted population sizes of cross-sectional and longitudinal samples (e.g., 32 to 48 percent in Iceland, Greece, the Czech Republic; and 10 to 20 percent in Norway, the U.K., Lithuania, and Portugal) suggest that there do indeed seem to be problems with the weighting procedure in several countries (detailed tables are available in the online supporting information). Unfortunately, there is no way of linking cross-sectional data across waves or cross-sectional and longitudinal data from EU-SILC to find out which individuals or households exactly left the sample in a given two-year period, but this information is needed to evaluate a potential bias of the panel attrition.

In order to shed some light on the relevance of inconsistencies between cross-sectional and longitudinal results, we now focus on child poverty given that this topic is of outstanding relevance for public policy. Besides Jenkins and van Kerm (2011), who also carried out consistency checks for the poverty rate on a subgroup level preceding their analyses on current poverty and persistent poverty in Europe, others including the European Commission's Social Protection

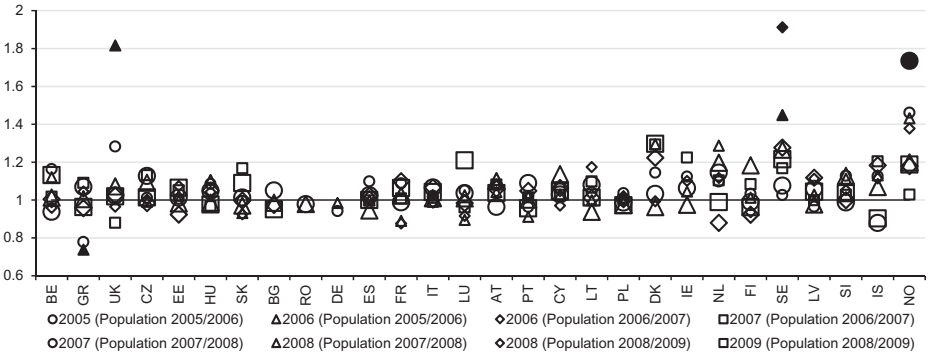


Figure 8. Child poverty rates—ratio of longitudinal and cross-sectional results 2005–2009.

Symbols are filled when confidence intervals of cross-sectional and longitudinal estimations do not overlap.

*Source*: EU-SILC (version of September 2011), persons living in private households, cross-sectional and longitudinal data 2005 to 2009.

Committee have emphasized the need for adequate income data to be able to analyze child poverty rates as well as their persistent poverty risk (Social Protection Committee, 2012), thereby also highlighting the importance of consistent cross-sectional and longitudinal data on current social policy questions.

Figure 8 again represents the ratios between cross-sectional and longitudinal results for the child poverty rate (the share of children in families with a disposable income below the poverty threshold). As expected, the deviations are much more pronounced than for the total population given that a smaller subset of the total sample is used for the calculations. Also, the respective confidence bands get much broader. Insofar, significant deviations are scarcer. However, we still find some outstanding outliers. These encompass Greece, Sweden, the U.K., and Norway. All of them show at least one substantial deviation in all four periods under investigation. Most notable is again the finding for Sweden, where in 2006 the cross-sectional result points to a child poverty rate of 14.6 percent, while the result for the panel 2006/2007 arrives at 7.6 percent only. Thus, with almost similar populations child poverty can be nearly twice as high when using cross-sectional data instead of longitudinal information. It is obvious that policy recommendations derived from such findings can be misleading.

### 3.3. *Income and Poverty Mobility*

Compared to the FGT measures that describe the risk of being poor in a country, income and poverty mobility measures give an additional indication on whether chances of climbing up or falling down the income ladder are high or low in comparison to other countries. As this feature makes the use of panel data especially desirable for research, the following section seeks to explore mobility analyses with EU-SILC in more detail.

Figure 9 shows the normalized average jump from one income quintile to another in all countries, comparing two consecutive two-year panels (Fields and

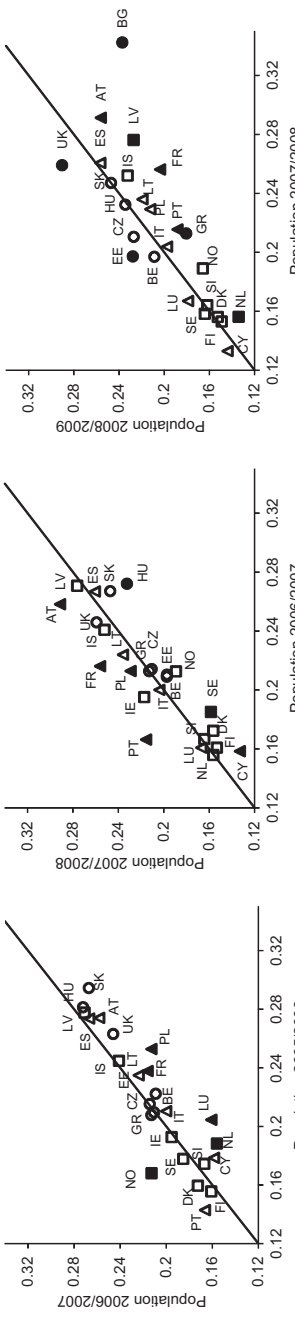


Figure 9. Comparison of mobility measures in three consecutive two-year-panels: Normalised Average Jump.
Symbols are filled when confidence intervals of the two mobility estimations do not overlap.
*Source*: EU-SILC (version of September 2011), persons living in private households, longitudinal data 2005 to 2009.

Ok, 1999). Ideally, the symbols should lie close to the diagonal, given that, again, there are only rather small changes in the underlying population due to the rotational design of EU-SILC. A substantial deviation could therefore be interpreted as a proxy for data problems. We assigned each country group a different kind of symbol, so it becomes visible that there is—by and large—less mobility in register countries, which are mostly Scandinavian states, than in other countries.[4]

In terms of the Normalized Average Jump, mobility varies strongly and substantially in Poland, France, Austria, and Portugal with alternately falling and rising measures between the periods, and the Netherlands with constantly decreasing measures of mobility over time. Bulgaria provides longitudinal income information from 2007/2008 onwards. There, the Normalized Average Jump decreases by almost 30 percent between the populations of 2007/2008 and 2008/2009.

The picture is quite similar when considering another income mobility measure as suggested by Shorrocks (1978); see Figure 10. But here, Norway shows extreme fluctuations between 2005/2006 and 2007/2008. In the first period (2005/2006), this indicator was 0.09 and increased substantially to more than 0.2 in 2006/2007. The second graph in Figure 10 shows that this value fell substantially to only about 0.13 in 2007/2008. Such dramatic jumps are probably not a result of true changes in the underlying population, but an indication of weak data quality. Slovakia's outlier between 2007/2008 and 2008/2009 may be caused by the inexplicable income data error in the 2008/2009 panel.

The results are largely confirmed when analyzing poverty mobility, that is, the share of the population that went from poverty to non-poverty or vice versa (Figure 11). The fact that Hungary's income measures in 2006 are already questionable may explain that poverty mobility fell considerably in the 2007/2008 panel. It is again noteworthy that many register countries—in particular Norway, but also the Netherlands and Sweden—showed evidence of wide variation in poverty mobility, also given that Rendtel *et al*. (2004) found higher stability of mobility measures when using income information from registers compared to survey income information.

A factor that influences mobility and may cause a bias is the type of interview. Face-to-face-interviews usually produce the most reliable income information, aside from register data. If the households are interviewed by the same person each year, this can be beneficial for the quality of the answers due to a steady increase in trust between the interviewer and the respondent (Tourangeau and Smith, 1996). Having no interviewer assistance, however, can cause some respondents to round their income or give inaccurate information if they do not understand the question. Furthermore, respondents are then more likely to leave the survey (Schräpler, 2004). Proxy interviews can also be problematic. Assuming that a person provides his exact income in one year but does not take part in the survey in the following year, another household member may respond by proxy. If the respondent is only roughly informed about another household member's income situation and only gives a rough estimate, this can cause artificial mobility. With countries such as Austria, France, Poland, and Portugal, especially the "proxy

[4]There is no symbol for Germany in EU-SILC because no longitudinal dataset has been available since 2006/2007. The same is true for the Irish data of 2008/2009.

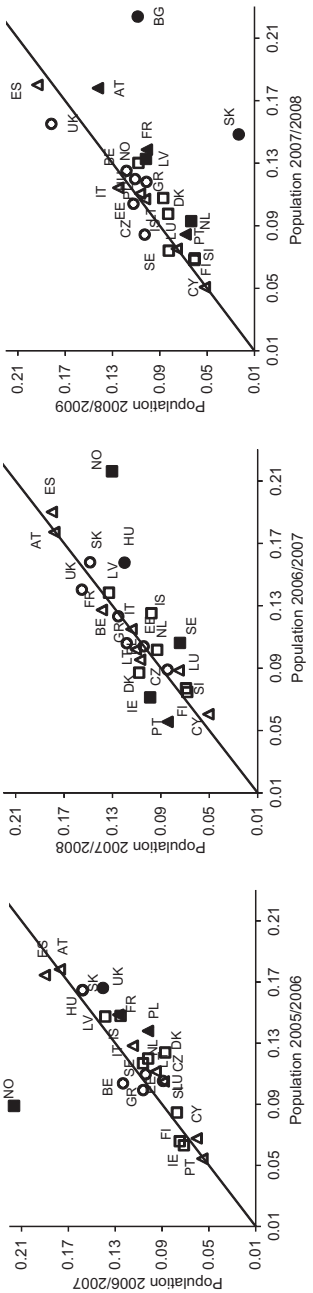


Figure 10. Comparison of Shorrocks (1978) mobility measures in three consecutive two-year-panels.
Symbols are filled when confidence intervals of the two mobility estimations do not overlap.
*Source:* EU-SILC (version of September 2011), persons living in private households, longitudinal data 2005 to 2009.

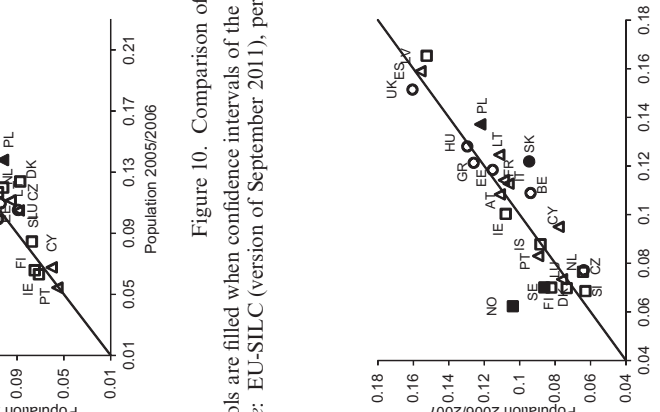


Figure 11. Comparison of mobility measures in three consecutive two-year-panels: Poverty mobility—share of mobile population (in percent).
Symbols are filled when confidence intervals of the two mobility estimations do not overlap.
*Source:* EU-SILC (version of September 2011), persons living in private households, longitudinal data 2005 to 2009.

countries" conducting more than 15 percent of the interviews non-personally (on average) show very erratic mobility measures. This can be seen as an indicator for non-personal interviews being a source of data inaccuracy. Glaser *et al*. (2015) suspect that using proxy interviews is a way of solving language problems within a selected household, leading to lower data quality (p. 46).

If income information is retrieved from registers, the problem of rounding or false answers should not exist, but the elimination of households that are affected by partial unit non-response can cause problems (see Frick *et al*., 2012). Since it can be assumed that PUNR is not random, and some of the register countries (particularly Norway) are among the countries with the highest inconsistencies in the mobility measures, there is reason to suspect that this procedure could cause biased mobility measures, if no appropriate weighting takes place.

Imputation can influence mobility results as well. Imputation methods using only cross-sectional information on other similar respondents with complete interviews often overestimate mobility (e.g., Frick and Grabka, 2010). There is evidence that imputation is of higher quality if longitudinal information on individuals with missing income components can be used (Spieß and Goebel, 2005; Rässler and Riphahn, 2006; Watson and Starick, 2011). Because of this, the imputation method and the consideration of cross-sectional versus longitudinal data in the imputation process may have an influence on aggregate income mobility. In fact, there are detailed guidelines for documentation on the imputation method as well as on the type of data used in EU-SILC. Unfortunately, some of the income flags in EU-SILC are nevertheless confusing and inconsistent over time and between countries, so we have not been able to check this hypothesis so far in a satisfactory way. A first problem when making use of imputation flags in EU-SILC is that this information is not consistently provided over time. Since 2008, a five-digit flag variable is claimed by Eurostat that contains information on the imputation method used. However, not all EU-SILC member states follow this rule. In case of the German part of EU-SILC, one finds a flag variable with six digits.

Finally, Figure 12 shows the comparison of two consecutive two-year panels for a given year. The panels 2005/2006 and 2006/2007 both cover the year 2006. The same is true for the year 2007, which was covered by the panels 2006/2007 and 2007/2008, and 2008 (2007/2008 and 2008/2009). The income and inequality measures are now computed for those years that were covered twice in the longitudinal populations, and the respective results were divided. So again, we will get a fraction with the expected value of 1 if analyses based on both populations yield identical results.

The populations are nearly equal and therefore should not show big differences. Indeed, the results are more concentrated around 1 than in the previous analyses. Considering the whole EU-SILC sample, the number of substantial deviations in 2006 does not even exceed the number of random deviations that is expected to occur in a random sample (although, of course, the measures within a country are not independent of each other). But still, in certain countries and years (Hungary and Austria in 2007, Denmark and France in 2008), deviations seem to pile up in more than one inequality and income measure. The differences reach values of 18 percent, and in Denmark even 28 percent (although here with overlapping confidence intervals).

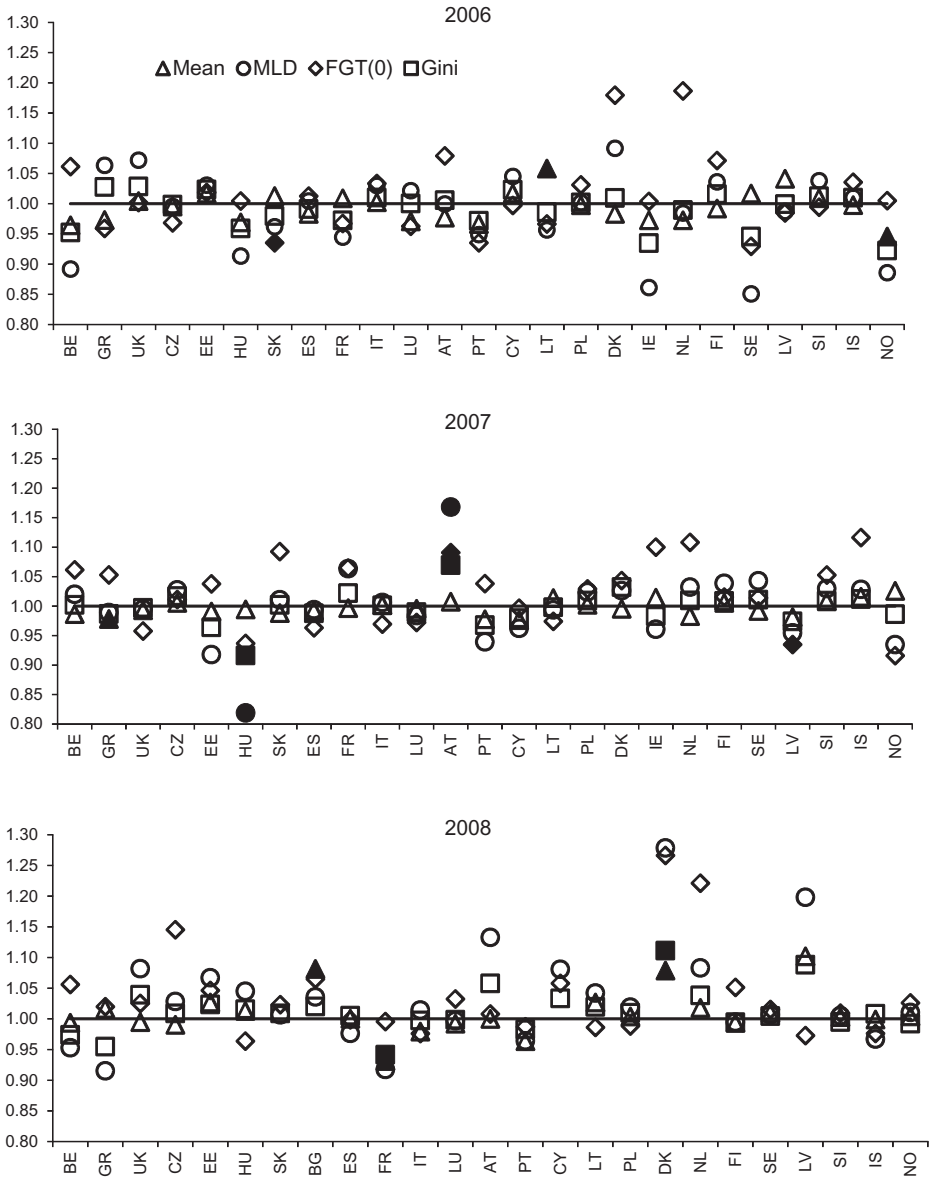


Figure 12. Ratio of income, inequality, and poverty measures in two consecutive two-year-panels.

"Mean" indicators are not displayed for Cyprus and Slovak Republic in 2008. Symbols are filled when confidence intervals do not overlap.

*Source*: EU-SILC (version of September 2011), persons living in private households, longitudinal data 2005 to 2009.

## 4. Conclusion

EU-SILC is a data source which has high international relevance for social science research and policy advice. Especially when it comes to comparative analyses based on income measurement at the EU level, EU-SILC is an

unparalleled data source of major importance for research. It is therefore imperative that the EU-SILC data fulfill the highest quality standards with respect to comparability and consistency over time. However, our results show that there are pronounced variances in income, inequality, and mobility development over the period 2005–09 in both the cross-sectional and longitudinal data. Our research approach is based on the straightforward assumption that results from cross-sectional and longitudinal data should show a rather high degree of overlap, given that the underlying populations differ only very slightly due to migration, birth, and death. Large deviations could therefore be interpreted as first indications of poor data quality, which should be analyzed with caution.

However, in some countries—strikingly often in register countries like Norway, Sweden, and Denmark, but also in Germany, France, Bulgaria, and the Czech Republic—results deviate substantially between cross-sectional and longitudinal analyses. Especially in Sweden and Norway the results seem to be dubious and we recommend analyzing these countries' data with particular caution. Mobility analyses based on two-year panels—in which there is an overlap of at least one wave of data collection—also point to rather large deviations from one year to the next and seem to be problematic especially for proxy countries. Failure to consider these potential implausibilities in income research may have a severe impact on the reliability of the results. Given the political significance of EU-SILC, this fact should not be underestimated.

Problems in countries with income registers suggest that it is not the method of data collection that is potentially flawed but rather the sampling procedure or the weighting of the sample. Given the fact that in most countries, a quarter of the sample is replaced every year, special attention must be paid to the representativeness of the data. One of the Net-SILC2 project's claims is a common definition of tracing rules in EU-SILC (Glaser *et al.*, 2015, p. 44), a point that especially affects the register countries applying the selected respondent design. Also, the project members call for a stronger weighting scheme in the countries with low data coherence and propose to "depart from design based longitudinal weights and ensure coherence by calibration of longitudinal weights" (p. 45).

For reasons of data protection, there is currently no means of linking cross-sectional and longitudinal data, although this is essential for the analysis of the effect of panel attrition and a possible attrition bias on data quality or any systematic selectivity analysis. Last but not least, the possibility of linking the databases is needed for a deeper analysis of whether the problems identified in this paper originated in the cross-sectional data rather than the longitudinal data or vice versa. A highly welcome solution would be to provide special (restricted) data access to enable researchers to address such methodological issues. Additionally, future research should make use of alternative (administrative) data sources to evaluate potential reasons for the observed inconsistencies.

Given the potentially large influence of imputation on data quality, there should be further investigations of this aspect. Documentation materials covering imputation methods and the extent of imputation within the EU-SILC data that are comparable across countries and over time, would be extremely helpful. Furthermore, all countries should take advantage of longitudinal data for income

imputation and ideally use the same imputation techniques as a means of reducing potential bias (Frick and Grabka, 2010).

In sum, it is beyond question that EU-SILC is a highly valuable data source for social and economic research in Europe. Of course, shortcomings or simplifications in the sampling process or in the course of data processing are always rooted in cost considerations. Nevertheless, users of EU-SILC should be aware of the data limitations when conducting income analyses. Also, every EU-SILC country should continuously check its data quality and the appropriateness of its survey methods, and, in case of crude data, weighting or imputation errors, carry out revisions for all existing waves to ensure consistency across time.

## References


BMAS (Bundesministerium für Arbeit und Soziales), *Lebenslagen in Deutschland. Entwurf des 4. Armuts- und Reichtumsbericht der Bundesregierung*, Bonn, 2012.

Canberra Group, Expert Group on Household Income Statistics, *Final Report and Recommendations*, Ottawa, 2001.

Epland, J., "Challenges in Income Comparability. Experiences from the Use of Register Data in the Norwegian EU-SILC," Paper presented for the VII International Meeting on Quantitative Methods of Applied Sciences, University of Siena, September 11–13, 2006.

European Commission, "Involving Experts in the Process of National Policy Convergence," 2001 (http://media.europa-nu.nl/9353000/d/commissierapportomc.pdf; accessed February 17, 2015).

———, "EU-SILC Users Database Description," 2006 (epunet.essex.ac.uk/EU-SILC_UDB.pdf; accessed February 17, 2015).

European Council, "17 June 2010 Conclusions," EUCO 13/10, Brussels, 2010 (http://ec.europa.eu/eu2020/pdf/council_conclusion_17_june_en.pdf; accessed February 17, 2015).

European Parliament and the Council, "Regulation (EC) No 1177/2003 of the European Parliament and of the Council of 16 June 2003 concerning Community Statistics on Income and Living Conditions (EU-SILC)," 2003 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:165:0001:0009:EN:PDF; accessed February 17, 2015).

———, "Commission Regulation No 28/2004 of 5 January 2004," 2004 (http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:005:0042:0056:EN:PDF; accessed February 17, 2015).

Fields, G. S. and E. Ok, "Measuring Movements of Income," *Economica*, 66, 455–71, 1999.

Foster, J., J. Greer, and E. Thorbecke, "A Class of Decomposable Poverty Measures," *Econometrica*, 2, 761–66, 1984.

Frick, J. R. and M. M. Grabka, "Imputed Rent and Income Inequality: A Decomposition Analysis for the UK, West Germany and the USA," *Review of Income and Wealth*, 49, 513–37, 2003.

———, "Item Non-Response and Imputation of Annual Labor Income in Panel Surveys from a Cross-National Perspective," in J. A. Harkness, B. Edwards, M. Braun, T. P. Johnson, L. E. Lyberg, P. Ph. Mohler, *et al.* (eds), *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, Wiley & Sons, New York, 355–72, 2010.

Frick, J. R. and K. Krell, "Measuring Income in Household Panel Surveys for Germany: A Comparison of EU-SILC and SOEP," *SOEPpapers*, 265, DIW Berlin, 2010.

Frick, J. R., J. Goebel, and M. M. Grabka, "Assessing the Distributional Impact of 'Imputed Rent' and 'Non-Cash Employee Income' in Micro-Data," in European Communities (ed.), *Comparative EU Statistics on Income and Living Conditions: Issues and Challenges. Proceedings of the EU-SILC conference (Helsinki, 6–8 November 2006)*, Office for Official Publications of the European Communities, Luxembourg, 2007.

Frick, J. R., M. M. Grabka, and O. Groh-Samberg, "Imputation of Annual Income in Panel Surveys with Partially Non-Responding Households," *Sociological Methods and Research*, 41, 89–123, 2012.

Glaser, T., E. Kafka, N. Lamei, L. Lyberg, and M. Till, "Review of National Data Collection and Coherence of the Longitudinal Component," *Net-SILC2 Work Package (WP) CO&ME-LONGIT.1 Final Report*, 2015.

Grabka, M. M. and J. R. Frick, "Country Case Study—Germany," in St. P. Jenkins, A. Brandolini, J. Micklewright, and B. Nolan (eds), *The Great Recession and the Distribution of Household Income*, Oxford University Press, Oxford, 3-1–3-36, 2011.

Groves, R. M. and M. P. Couper, *Non-Response in Household Interview Surveys*, John Wiley & Sons, New York, 1998.
Groves, R. M., F. J. Fowler, Jr., M. J. Couper, J. M. Lepkowski, E. Singer, and R. Tourangeau, *Survey Methodology*, Wiley & Sons, New York, 2009.
Iacovou, M. and P. Lynn, "Implications of the EU-SILC Following Rules (and their Implementation) for Longitudinal Analysis," ISER Working Papers 2013–17, 2013.
Iacovou, M., O. Kaminska, and H. Levy, "Using EU-SILC Data for Cross-National Analysis: Strengths, Problems and Recommendations," ISER Working Papers 2012–03, 2012.
Jenkins, S. and P. van Kerm, "Patterns of Persistent Poverty: Evidence from EU-SILC," ISER Working Papers 2011–30, 2011.
———, "The Relationship Between EU Indicators of Persistent and Current Poverty," *Social Indicators Research*, 116, 611–38, 2014.
Lelkes, O., "Poverty among Migrants in Europe," *Policy Brief April 2007*, European Centre, 2007.
Lelkes, O., M. Medgyesi, I. Tóth, and T. Ward, "Income Distribution and the Risk of Poverty," in T. Ward, O. Lelkes, H. Sutherland, and I. G. Tóth (eds), *European Inequalities: Social Inclusion and Income Distribution in the European Union*, TÁRKI Social Research Institute, Budapest, 17–44, 2009.
Little, R. J. A. and P. J. Smith, "Editing and Imputation for Quantitative Survey Data," *Journal of the American Statistical Association*, 82, 58–68, 1987.
Lohmann, H., "Comparability of EU-SILC Survey and Register Data: The Relationship among Employment, Earnings and Poverty," *Journal of European Social Policy*, 21, 37–54, 2011.
Rässler, S. and R. T. Riphahn, "Survey Item Non-Response and its Treatment," *Allgemeines Statistisches Archiv*, 90, 213–28, 2006.
Rendtel, U., *Panelmortalität und Panelrepräsentativität*, Campus, Frankfurt, 1995.
Rendtel, U., L. Nordberg, M. Jäntti, J. Hanisch, and E. Basic, "Report on Quality of Income Data," CHINTEX Working Paper, 21, 2004.
Rubin, D. B., *Multiple Imputation for Nonresponse in Surveys*, Wiley & Sons, New York, 1987.
Schonlau, M., N. Watson, and M. Kroh, "Household Survey Panels: How Much Do Following Rules Affect Sample Size?" *Survey Research Methods*, 5, 53–61, 2011.
Schräpler, J.-P., "Respondent Behavior in Panel Studies—A Case Study for Income-Nonresponse by means of the German Socio-Economic Panel (SOEP)," *Sociological Methods & Research*, 33, 118–56, 2004.
Shorrocks, A., "The Measurement of Mobility," *Econometrica*, 46, 1013–24, 1978.
Social Protection Committee, "SPC Advisory Report to the European Commission on Tackling and Preventing Child Poverty, Promoting Child Well-Being," 2012 (http://www.epha.org/a/5260; accessed February 17, 2015).
Spieß, M. and J. Goebel, "On the Effect of Item Nonresponse on the Estimation of a Two-Panel-Waves Wage Equation," *Allgemeines Statistisches Archiv*, 1, 63–74, 2005.
Tourangeau, R. and T. W. Smith, "Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context," *Public Opinion Quarterly*, 60, 75–304, 1996.
Verma, V., "Issues in Data Quality and Comparability in EU-SILC," in Eurostat (ed.), *Comparative EU Statistics on Income and Living Conditions: Issues and Challenges. Proceedings of the EU-SILC conference (Helsinki, 6–8 November 2006)*, Office for Official Publications of the European Communities, Luxembourg, 2007.
Verma, V. and G. Betti, "Data Accuracy in EU-SILC," in A. B. Atkinson and E. Marlier (eds), *Income and Living Conditions in Europe*, Eurostat Statistical Books, Luxembourg, 57–77, 2010.
Verma, V., G. Betti, and G. Ghellini, "Cross-Sectional and Longitudinal Weighting in a Rotational Household Panel: Applications to EU-SILC," *Statistics in Transition—New Series*, 8, 5–50, 2007.
Verma, V., G. Betti, and F. Gagliardi, "An Assessment of Survey Errors in EU-SILC," *Eurostat Methodologies and Working Papers*, Office for Official Publications of the European Communities, Luxembourg, 2010.
Watson, N. and R. Starick, "Evaluation of Alternative Income Imputation Methods for a Longitudinal Survey," *Journal of Official Statistics*, 27, 693–715, 2011.
Zaidi, A., "Poverty of Elderly People in EU25," *Policy Brief August 2006*, European Centre, 2006.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Table A1:** Cross-sectional and longitudinal results (Supporting Information, PDF)