Stephen P. Jenkins and Philippe van Kerm

How does attrition affect estimates of persistent poverty rates? The case of EU-SILC

Book section
(Accepted version)

Original citation: Originally published in Atkinson, Anthony B. and Guio, Anne-Catherine and Marlier, Eric, (eds.) Monitoring Social Inclusion in Europe. Luxembourg City, Luxembourg : Publications Office of the European Union, 2017

How does attrition affect estimates of persistent poverty rates? The case of EU-SILC

Stephen P. Jenkins **Philippe Van Kerm**

This version: 20 janvier 2015
Prepared for **BOOK CHAPTER**

summary for introductory chapter

Evidence about poverty persistence is an important complement to information about poverty prevalence at a point in time. The persistent at-risk-of-poverty rate is one of the primary indicators of social inclusion, and the only indicator that is derived using samples from the longitudinal component of EU-SILC. Sample drop-out from the longitudinal samples ('attrition') reduces sample size thereby decreasing the precision of estimates of persistent poverty indicators, and may be selective and lead to bias. We examine these issues. We show that rates of attrition from the four-year EU-SILC samples used to calculate persistent poverty rates vary substantially across Member States, and there is also substantial cross-national diversity in the characteristics of individuals lost to follow-up. We provide evidence that application of longitudinal weights does not fully account for the effects of attrition, and that different assumptions about the poverty status of attritors lead to wide bounds for estimates of persistent poverty rates for most Member States.

22. How does attrition affect estimates of persistent poverty rates? The case of EU-SILC[1]

Stephen P. Jenkins **Philippe Van Kerm**

22.1 Introduction

Over the last decade and through its Open Method of Coordination, the European Union (EU) has agreed a set of common objectives for monitoring and measurement of social protection and social inclusion, together with a set of indicators to assess national and EU progress towards these goals. Among the primary indicators of social inclusion is the persistent at-risk-of-poverty rate, defined as the proportion of persons in a country who are at risk of income poverty in the current year and who were at risk of income poverty in at least two of the preceding three years. Evidence about poverty persistence is an important complement to information about poverty prevalence at a point in time: it is widely agreed that poverty is worse for an individual, the longer he or she experiences it. Eurostat derives estimates of persistent at-risk-of-poverty rates using samples from the longitudinal component of the EU Statistics on Income and Living Conditions (EU-SILC) in which the fortunes of individuals are tracked over four consecutive years. Because not all of the individuals present in the first sample year provide four years of income data – there is attrition – estimates of persistent at-risk-of-poverty measure may not be reliable. In this chapter, we analyse the extent to which this is the case, and how the potential problems vary across EU member states.

Attrition is a potential problem for two reasons. First, it means that the sample size for the four-year sample used to calculate a persistent at-risk-of-poverty rate is smaller than the size of the sample of respondents in the first year of the four (wave 1), and a smaller sample size leads to *less precise inference* (larger standard errors and wider confidence intervals). Second, if the individuals who are lost to follow-up differ systematically from the initial respondent sample – the case of non-random or 'differential' attrition – the four-year sample may not be representative of the underlying population, thereby leading to *biased estimates* of persistent at-risk-of-poverty rates.

The longitudinal weights supplied with EU-SILC longitudinal data are intended to address the second problem. The idea is that, if differences in the chances of sample dropout can be fully characterised in terms of differences in individuals' observed characteristics, weighting will make the four-year sample representative of the initial sample. Individuals with characteristics associated with large dropout probabilities receive relatively large weights to compensate for the large fraction of similar individuals that have been lost. Individuals less likely to dropout receive relatively small weights. The weighting strategy works as long as observable characteristics predict dropout probabilities well and those who remain in the sample are not systematically different from those who attrit. However, problems arise if the chances of attrition also depend on unobserved characteristics that are systematically correlated with the

[1] Stephen P. Jenkins is at the LSE and affiliated with ISER (University of Essex) and IZA Bonn. Philippe Van Kerm is at the Luxembourg Institute of Socio-Economic Research. This work has been supported by the second Network for the analysis of EU-SILC (Net-SILC2), funded by Eurostat, and partially supported by core funding of the Research Centre on Micro-Social Change at the Institute for Social and Economic Research by the University of Essex and the UK Economic and Social Research Council (award ES/L0009153). The European Commission bears no responsibility for the analyses and conclusions, which are solely those of the authors. We thank Tony Atkinson, Carlos Farinha Rodrigues, Eric Marlier, and participants at the NetSILC-2 Lisbon conference for comments and suggestions.

chances of being persistently at-risk-of-poverty. Because such characteristics are unobserved, their impact is difficult to assess.

Our research builds on analysis of attrition in EU-SILC's forerunner, the European Community Household Panel (ECHP), undertaken by Behr, Bellgardt, and Rendtl (2005) and Watson (2003). EU-SILC differs substantially from the ECHP which ran between 1994 and 2001. Although both sources employ annual data collection, EU-SILC longitudinal data refer to information collected over a four-year period, rather than up to eight years. Instead of using a survey instrument with a cross-nationally harmonised design (household panel surveys in ECHP), EU-SILC uses output harmonisation. Countries are mandated to deliver a number of statistics conforming to particular specifications but have some discretion about the ways in which the information is collected. Most notably, some countries use household panel surveys to collect the longitudinal data; others use linked administrative registers.[2] In addition, there are many more countries contributing EU-SILC data than were in the ECHP: we use 23 countries in our analysis; there were only 15 countries covered by the ECHP.

Behr, Bellgardt, and Rendtl (2005) and Watson (2003) both drew attention to a substantial diversity in response rates in ECHP and, moreover, their conclusions were that, although the amount of attrition was relatively large, its effects on estimates of poverty rates and quintile transition probabilities were relatively benign. Indeed, Watson went so far as to state that 'fears that attrition has undermined the representativeness of the ECHP samples in later waves of the survey are largely unfounded' (2003: 361). Her results about representativeness are similar to those reported by Fitzgerald, Gottschalk, and Moffitt (1998) for the US Panel Study of Income Dynamics.

Patterns of attrition and their consequences may have changed over the last decade. Also, with many more countries with data, and output harmonisation rather than input harmonisation, there is much greater scope for differences across Member States. Our analysis of attrition and estimation of persistent at-risk-of poverty rates in EU-SILC data is therefore not only timely but also important given the place of this indicator in the EU's portfolio of social inclusion indicators.

The remainder of the chapter is organised as follows. In Section 2, we explain the data that we use, drawn from the 2011 longitudinal EU-SILC User DataBase. This discussion covers the definition of the persistent at-risk-of-poverty rate, how attrition arises, the weights that are available, and the samples that we use in the analysis. The extent of attrition across Member States, and how it varies with personal characteristics, is described in Section 3. In Section 4, we analyse the implications of sample dropout, again contrasting the situation across countries. We assess effects on representativeness by comparing estimates of at-risk-of-poverty rates from the full initial sample with estimates derived from the smaller four-wave sample. Section 5 contains a summary and conclusions.

For brevity, we report here only a subset of the analyses we have conducted. See Jenkins and Van Kerm (2015) for a complete set of results.

22.2. Data, definitions, sample selection, weighting

Our analysis is based on the 2011 EU-SILC longitudinal files. More specifically we use the scientific-use release of the longitudinal EU-SILC files made available to the NetSILC-2

[2] On this, see e.g. Lohmann (2011).

project, which are an update of UDB 2011-1, released August 2013. These files refer to data covering the four survey years 2008–2011. Because the reference period for EU-SILC income data is the calendar year preceding the year of data collection, the income years covered are 2007–2010.

22.2.1 At-risk of-poverty rates and persistent at-risk-of-poverty rates

Following EU official definitions, an individual's 'at-risk-of-poverty' status in a given income year is determined by the equivalised household disposable income of the household to which he or she belongs. (For further details of the sources included in household income and the equivalence scale, see Eurostat (2010).) A person is counted as being at-risk-of-poverty (henceforth poor) in a given year if his or her equivalised household disposable income is less than 60 per cent of the national median equivalised household income for that year.[3] The current at-risk-of-poverty rate (henceforth *current poverty rate*) for a particular country or group within a country is the proportion of persons in that country or group who are poor in the current income year.

The persistent at-risk-of-poverty rate (henceforth *persistent poverty rate* is the proportion of persons in the country or group who are currently poor and who were poor in at least two of the preceding three years. Thus in our longitudinal data, the persistent poverty rate refers to the proportion of individuals who were poor in 2010 as well as in at least two of the three previous years (2007–2009). This indicator is the principal official EU indicator on social inclusion for which estimation is based on the longitudinal component of EU-SILC, and hence the indicator that is most sensitive to attrition issues.

22.2.2 Samples

EU-SILC has a four-year rotating panel design. A fresh sample of households is drawn every year in every country, and the respondents in this sample are eligible for an interview in each of the following three years, contributing a total of up to four interviews. In each particular calendar year, data are available from four cohorts of respondents and contribute to the EU-SILC cross-section data. The 2011 EU-SILC longitudinal data (and similarly in preceding releases) consist of the three subsamples that provide data in 2011 and in at least one earlier survey year as well, i.e. the cohorts that entered the survey in 2008, 2009, or 2010.

To examine the magnitude and pattern of attrition, and to assess their implications for estimation of persistent poverty rates, we work with the 2008 rotation group sample which provide data over up to four years and is therefore the basis for calculation of the 2011 persistent at-risk-of-poverty indicator. We use the samples for 23 countries:. We exclude the samples for Luxembourg (because no rotation group was started in 2008), Norway (because some of the relevant sample weights were not available – see below), Denmark (because the 2011 database appears to exclude households that attritted before the fourth interview) and Sweden (because of unexplained differences in sizes between the 2010 and 2011 versions of the 2008 rotation group samples).

Our examination of the magnitude and effect of attrition relies on two overlapping subsamples. The first sample is composed of all individuals from all households in the rotation groups for survey years 2008–2011 that responded at wave 1 (wave 1 is the year in which households

[3] Throughout our analysis, poverty lines for each country and year are taken from Eurostat (2014). These thresholds are derived from the cross-sectional EU-SILC datasets.

entered the survey, i.e. 2008) irrespective of their subsequent participation. We refer to this full sample of the 2008 rotation group as the full *W1 Sample*. In principle, this sample should provide estimates close to those derived from the full 2008 cross-sectional sample. We return to this point later.

Our second subsample is composed of the subset of individuals from the W1 Sample that belong to a household succesfully interviewed in each one of the four survey years 2008–2011. This is the *four-wave Balanced Sample*, from which persistent poverty rates can be calculated.[4] We consider only individuals who were living in a household that was interviewed at wave 1: we discard children born after wave 1 as well as co-residents that joined a sample household after wave 1 since, by construction, these individuals do not have a full four-year set of responses. Individuals who change household over time remain in the balanced sample as long as their household of destination is successfully interviewed. Variations in practice and in the success of tracking such individuals and interviewing 'split-off' households has been shown to vary widely across countries by Iacovou and Lynn (2013): see Chapter 27. These are likely sources of the cross-country differences in attrition rates documented below.

22.2.3 Attrition

The differences in size and composition between the W1 Sample and the four-wave Balanced Sample reflect attrition. Not all individuals or households eligible for an interview after the first interview provide data in subsequent years. There are three reasons for this.

First, some individuals or households move out of scope after the first interview, for example because they die, or move abroad permanently, or move into an institution. Second, eligible individuals may not be followed by the data collection agency, or the agency may be unsuccessful in tracking them (with the chances greater for individuals that split off from a household, or where all members of a household move from the initially-sampled address). Third, individuals or households may refuse to participate in the survey in the second interview or later.[5]

The first kind of sample dropout reflects the dynamics of a population and is a natural feature that is built into the data collection design (based on representation of the population of individuals in private households in a particular country). By contrast, the second and third types of attrition are undesirable and, other things being equal, data collection agencies should aim to minimize them. Country-specific factors may also play a role, for example, whether up-to-date address registers exist, the prevalence of geographical mobility by households, general attitudes towards surveys, etc.

22.2.4 Sampling weights

Sampling weights are designed to adjust for biases arising from cross-sectional non-response and subsequent longitudinal attrition. The EU-SILC longitudinal files include five types of sample weights (Museux 2006), of which two are relevant to our analysis.

[4] More precisely, it is based on the subsample with valid (non-missing) data on household income in the EU-SILC data files in all years. However, because missing information on income is imputed (and we use the imputed values), all households contain non-missing data on income in the EU-SILC data files.
[5] Chapter 27 by Iacovou and Lynn discusses the difficulty in consistently identifying the causes of attrition in EU-SILC across the different countries.

The first set of weights is the *individual-level base weights* (variable *rb060*). In wave 1, this is the design weight adjusted for non-response and calibrated. In later waves, it is the base weight of previous year adjusted for non-response. When individuals leave the sample, they are attributed with a weight of zero for each wave thereafter. Our analysis of the W1 Sample uses *rb060* to ensure the sample accounts for non-proportional sampling design (and initial non-response), and for differential attrition, and is calibrated to population totals in 2008.

The second set of weights that we use, *rb064*, is the individual-level longitudinal weights created for analysis of data for the four survey years 2008–2011 and, of course, the weights are only relevant for the single rotation group that provides data for these four years. For analysis of the four-wave Balanced Sample, we contrast results obtained with *rb064* (constructed to ensure that the balanced sample remains representative of the original 2008 population), with *rb060* at their 2008 values (so they correct for initial non-response and sampling rates, but not for differential attrition), and with *rb060* at their 2011 values (in which case they are similar to *rb064*).

We also create our own bespoke set of longitudinal weights (discussed later). The advantage of these weights is that we can use them to engage in a number of counterfactual exercises that we cannot undertake with the weights that are supplied. We show below that these weights generally closely reproduce estimates derived using the official longitudinal weights although our bespoke weights are derived using variables available in the longitudinal data files, and we do not have access to all the factors employed by statistical offices when producing longitudinal weights (*rb064*), nor do we attempt to calibrate our weights to known population totals, for example, as derived from other data sources or from the full EU-SILC cross-section files.

22.3. How much attrition is there? Who drops out?

In this section, we document how much attrition there was in the 2008–2011 EU-SILC longitudinal data, and which types of individual were most likely to be lost to follow-up. We discuss attrition – or its complement, sample retention – in terms of differences between the full Wave 1 Sample and the smaller four-wave Balanced Sample.

22.3.1 How much attrition is there overall?

The overall retention rate for each country is the fraction of the country's full W1 sample that belongs to the Balanced Sample. More precisely we calculate the retention rate as the proportion of individuals belonging to a respondent household at wave 1 which remains in a respondent household in each of the three subsequent waves. These rates are shown in Figure 1.

<Figure 1 near here>

There are very large differences in retention rates across countries, ranging from greater than 90 per cent to nearer 40 per cent. The UK stands out as having a particularly low retention rate, nearly 10 percentage points smaller than the next smallest rate, 50 per cent for Slovenia. There is a cluster of three countries with remarkably large retention rates: those for Romania and Bulgaria are all near 90 per cent. The method of data collection, whether by household panel survey or linked administrative registers, has no consistent association with the extent of attrition. There are countries using registers for data collection (identified by the circles in Figure 1) with both small and large retention rates.

22.3.2 Attrition's effect on the precision of estimates

The fall in sample sizes associated with attrition means that, questions of representativeness and hence bias aside (on which see below), estimates of persistent poverty rates are estimated less precisely. Standard errors are larger, and confidence intervals are wider. The effects of differences in sample size on the sampling variability of estimates can be gauged by noting that the persistent poverty rate is a proportion (p), and there is a standard formula for the standard error of a proportion. The standard error of p is given by $d\sqrt{p(1-p)/N}$, where N is the sample size and d is a design effect arising because of the complex survey design. We examined how standard errors vary with N, using values of N and p which cover the range of estimates observed in EU-SILC. Following Goedemé (2013), we set d equal to 1.8, i.e. survey design effects such as stratification and clustering (e.g. of individuals into households, and households into primary sampling units) increase the standard error by 80 per cent compared to the standard error for a simple random sample of the same size.

Our calculations may provide some cheering news for analysts. Even with substantial attrition and hence relatively small sample sizes, standard errors for persistent poverty rates at the national level may be sufficiently precise. For example, if the persistent poverty rate is around 20 per cent and the sample size is 2,500, the standard error for the rate is around 0.015, so the estimated rate is more than ten times larger than its standard error, and the 95% confidence interval is roughly [17%, 23%]. If the sample size were instead 1,000, then the standard error increases to around 0.025, so the ratio of estimate to standard error is around 8, and the confidence interval is approximately [15%, 25%]. If, instead, the persistent poverty rate is only 5 per cent, then a sample size of 1,000 implies a standard error of around 0.012, so the ratio of estimate to standard error falls to just over 4. Ratios of around 2 or more are often interpreted as indicating statistical significance of sufficient degree.

Estimates of persistent poverty rates for subgroups within a country may not be precisely estimated, however. For subgroups of particular policy interest, for example individuals living in households headed by a lone parent, sample sizes are likely to number a few hundred at most. With a sample size of 100 and a persistent poverty rate of 20 per cent, the standard error is around 0.06, implying a ratio of estimate to standard error of just over 3 and a 95% confidence interval of approximately [8%, 32%] which is rather wide. In this case, it would be hard to detect statistically significant changes over time in the subgroup persistent poverty rate. The same problem would arise if the persistent poverty rate were smaller than 20 per cent. To add to this cautionary note, we should say that we suggest later that, even for large sample sizes (such as for countries as a whole), confidence intervals may be sufficiently wide to encompass differences between estimates that are unbiased and those that are biased because of differential attrition.

22.3.3 Who drops out?

We now consider which types of individuals are most likely to be included in the four-wave Balanced Samples. We classify individuals according to their characteristics when measured in Wave 1, and calculate retention rates separately for subgroups defined by those characteristics. The individual characteristics we use are poverty status, quintile group of equivalised disposable household income, age and sex, household type, labour market activity status and education level of the household head, and whether the interview questionnaire was completed by a proxy respondent (another household member filling out the questionnaire on behalf of the target respondent). Differences in attrition (retention) rates associated with individual characteristics exemplify the process of *differential* attrition (retention).

Figure 2 shows the univariate breakdowns for each country. Each panel of the figure has a common format. Each overall national retention rate is shown as a cross and subgroup retention rates are shown separately using a numerical code for each subgroup. If subgroup rate for a country is close to the national rate, then attrition is not associated with subgroup membership. Countries are ordered vertically in each chart by their overall retention rate.

For example, in Figure 2 panel (a), individuals classified as poor at Wave 1 are coded '11' (the code for non-poor is '10'). It can be seen that poor individuals are more likely to be lost to follow-up in around half the countries and, in a few countries, the differences from the national average are very large. For example, in Belgium and Iceland, poor individuals have a retention rate more than 10 percentage points less than the overall national retention. The difference is about 6 percentage points in the UK. Panel (b) tells a similar story. Retention rates do not vary greatly with income group, except that in a small number of countries, individuals in the poorest fifth are more likely to be lost. (The effects are more muted than in panel (a), probably because the poorest fifth includes more people than are counted as poor.)

<Figure 2 near here>

Figure 2, panel (c), shows that, in the vast majority of countries, young men (aged between 18 and 40 years) are more likely to attrit than the national average rate, as well as (to a lesser extent) young women. The differences in retention rates across age-sex groups is particularly marked in some countries. For example, in Malta and the UK, the range is around 20 percentage points between the smallest and largest rates.

Figure 2, panel (d), shows that, for many but not all countries, there are large differences in retention rates between household types in some countries, some of which are larger than shown in panel (c). The general picture is that single adult households (with and without) children are most likely to be lost to follow-up, whereas single or couple households with the head aged 60+ have substantially higher retention rates. These differentials are what one would expect given the positive correlation between geographical mobility and age. But dispersion in retention rates by household type is not inevitable: observe the relatively small differentials for the countries with large overall retention rates (at the top of the figure).

Figure 2, panel (e), shows that for most countries retention rates do not vary substantially with the labour market activity status of the household head, though there is a tendency for individuals with unemployed household heads to be more likely to be lost to follow-up and individuals with a retired household head to be less likely to be lost. (In both cases, the head may be the individual him- or herself.) This pattern is particularly marked in some countries. For example, in the UK, the retention rate is just below 20 per cent for individuals with an unemployed head but around 50 per cent for individuals with a retired head (a difference of some 30 percentage points). The corresponding differential is more than 20 percentage points in Malta.

There appears to be a more complex association between the education level of the household head and retention rates. For countries with relatively low overall retention rates, it is individuals whose household head has either of the two lowest educational levels who have the largest attrition rate. (Austria stands out as an example of this.) And for countries with relatively high overall retention rates, it is individuals whose household head has either of the two highest educational qualifications who have the largest attrition rate. (Look at the cases of Estonia or Slovakia, for example.)

Besides individual or household characteristics, fieldwork-related features are correlated with attrition. Figure 2, panel (g), shows that the retention rate for individuals for whom data were collected from a proxy respondent in wave 1 tend to have low retention rates. This is particularly strong in the Netherlands or Greece, for example. This is unsurprising because a proxy interview in the first wave is indicative of difficulties in securing a respondent's participation to start with. Under-representation of the 'proxy respondent' characteristic itself is unlikely to be a concern; rather, the concern is the extent to which being a 'proxy respondent' is associated with other relevant individual characteristics.

In sum, there is substantial diversity in the rates at which individuals from EU-SILC Wave 1 samples are also found in the four-wave Balanced Samples from which persistent poverty rates can be calculated. There is differential attrition in terms of observable characteristics.[6] The finding of diversity in retention rates was also reported by Behr, Bellgardt, and Rendtl (2005) and Watson (2003) in the ECHP, though specific results are difficult to compare because findings are summarised in different ways in the different studies (and there is no good 'one number' summary of the *amount* of differential attrition).

22.3.4 Generating bespoke sample weights from retention regressions

The purpose of longitudinal weights in general is to adjust the four-wave Balanced Sample so that the reweighted Wave 1 covariate distributions of the sample is the same as the Wave 1 distribution of covariates in the full Wave 1 Sample. Using simple multivariate probit regressions, we construct bespoke weights by multiplying the Wave 1 base weights of each observation *i* provided in EU-SILC, ω_i, by the inverse of the retention probability predicted by the combination of the fitted regression parameters and the values of the predictor variables. So, if $r_i = \Phi(X_i b)$ is the predicted retention probability of observation *i* belonging to the four-wave Balanced Sample for a given country given standard Normal distribution $\Phi(.)$, vectors of characteristics X_i and fitted regression parameters b, individual *i*'s bespoke longitudinal weight is $w_i = \omega_i \times (1/r_i)$.

EU-SILC longitudinal weight are constructed by national statistical institutes in a similar though not identical fashion. They use cumulative year-on-year retention probabilities (rather than a four-year probability we have). They may include Wave 1 characteristics or perhaps more detailed fieldwork information that is not available in the public release files; and more flexible specifications for the regression equations used to predict retention probabilities; and there may be adjustment and calibration to the marginal distributions observed in full cross-section samples. Our specification is basic, but it can be implemented straightforwardly using the data in the scientific-use EU-SILC files that are available to us.

The greater the dispersion of subgroup attrition rates around a national average, the greater the variance in a country's sample weights. This can also have an adverse impact on the precision of estimates of persistent poverty rates – a factor that we did not take into account earlier when showing the connection between standard errors and sample size. We assumed in our illustration that the design effect was constant across countries. However, the variance of the sample

[6] We have also undertaken probit regression analysis to examine the association between the probability of retention and each characteristic. Multivariate analysis helps tease out the associations between retention rates and a particular characteristic, holding other characteristics constant. Our multivariate analysis of the correlates of retention propensities is reported in Jenkins and Van Kerm (2015: Figure 4). It turns out that many of the associations uncovered by the univariate analysis are also found in the multivariate analysis, and so we do not discuss these results further here.

weights is one of the factors that influence the design effect, mediating the relationship between sample size, standard errors and the poverty rate. So, although application of sample weights may adjust for bias associated with differential attrition, it may come at a cost in terms of sampling variability when attrition is heavily differential and therefore sample weights have substantial variability.

We use our bespoke weights for some counterfactual exercises that can not be undertaken with the EU-SILC weights.

22.4. What effects does differential attrition have?

In this section, we provide indirect evidence about potential bias in estimates resulting from differential attrition. See Jenkins and Van Kerm (2015) for additional analysis using different approaches.

22.4.1 Indirect evidence of attrition bias: comparisons of estimates of Wave 1 poverty rates

We follow Behr, Bellgardt, and Rendtl (2005) and assess the magnitude and potential impact of attrition by comparing our original base sample with the sample that remains after attrition. In our application, this means comparisons of statistics derived from the W1 Samples with the same statistics derived from the four-wave Balanced Samples. The benchmark statistic is the Wave 1 (2008) poverty rate. If there are differences in estimates, this suggests that the differences in the samples will also lead to bias in estimates of persistent poverty statistics (which can not be calculated for both samples, of course). We refer to this as indirect evidence because it is not directly about the persistent poverty rate.

Figure 3, panel (a), contains four series of estimates of Wave 1 poverty rates for each country (countries are ranked as before). The ‘All W1’ sample consists of all individuals in the Wave 1 sample; there are two series calculated using our Balanced Samples but using different weights (the EU-SILC longitudinal weights and our bespoke ones). As an additional reference point, we show the estimates of the 2008 poverty calculated using the full cross-section sample (i.e. based on multiple cohorts rather than simply one) that are calculated and reported by Eurostat (2014).

In an ideal world, without random sampling variability, attrition or other data problems, all series would provide the same estimate for the poverty rate in 2008. Comparison of the Eurostat series (black filled circles) with the ‘All W1’ series (hollow diamonds) indicate that there are discrepancies associated with differences between full cross-sectional and longitudinal data files that we are unable to resolve using the data available to us. Although the differences in estimates are small for some countries (e.g. Slovakia, Bulgaria, France, Poland, Estonia, Hungary, Italy, Slovenia), there are differences that are relatively large for some other countries, nearly five percentage points in some cases (sometimes a negative difference, sometimes a positive one). See e.g. Romania, Iceland, Greece, the Netherlands, and the UK. There appears to be no systematic relationship with type of data collection or persistent poverty rate. Because these differences relate to aspects that we cannot control (correspondences between full cross-sectional and longitudinal files are primarily connected to sampling variability since the former are composed of about four times as many observations), we note the inconsistencies and pass on to the other comparisons.

<Figure 3 near here>

Specifically, we are interested in the extent to which the longitudinally-weighted estimates from the four-wave Balanced Samples match the estimates from the 'All W1' samples (compare estimates denoted by a cross and by a hollow circle, respectively), and then the extent to which estimates using our bespoke weights match the estimates based on the Eurostat weights (compare estimates denoted by a hollow circle and by a cross). These comparisons are easiest to make if one looks at Figure 3, panel (b), in which each longitudinal sample estimate is expressed as a ratio of the corresponding 'All W1' sample estimate. Longitudinal estimates that lie outside the boundaries demarcated by the vertical dashed lines differ by more than 10 per cent from the 'W1 sample' estimates. These cases are a signal that differential attrition is likely to lead to bias that is not fully accounted for by weighting. (The boundaries are sufficiently wide to allow for differences arising from sampling variation.)

The figure shows that, for 18 of the 23 countries, the longitudinal estimates based on the Eurostat weights are within 10 per cent of their full Wave 1 Sample counterparts. However, for 5 countries, the estimates are outside the boundaries, and hence there is indirect evidence that unaccounted-for differential attrition is leading to bias. For three countries (the Czech Republic, Slovenia, and the UK), longitudinal sample estimates are between 80 per cent and 90 per cent of their corresponding Wave 1 sample estimate and, for three countries (Finland, Iceland), the longitudinal estimates are even smaller, less than 80 per cent of their corresponding Wave 1 sample estimate. However, again, the evidence about potential bias from differential attrition is less strong for these countries if our bespoke longitudinal weights are used rather than the Eurostat ones. With the former, the longitudinal estimates for only two countries differ by more than 10 per cent from their Wave 1 Sample counterparts: these are the estimates for Iceland and Sweden.

One other country stands out in this exercise: the Netherlands. Its poverty rate estimate based on the balanced sample with EU-SILC longitudinal weights is more than twice the estimate obtained on the All Wave 1 sample and the estimate obtained from the balanced sample with our bespoke weights. (The relative difference is so large that it does not fit the horizontal scale of Figure 7 (b).) Eurostat's estimate is in between the two estimates although it is almost twice as large as our calculations. This suggests, first, that the other three cohorts forming the full cross-section dataset differ widely from the 2008–2011 cohort we focus on. Second, the EU-SILC longitudinal weights have been calibrated to fit the characteristics of the full cross-section data. This casts serious doubt about the representativeness of analysis based on the Dutch longitudinal data, since, as we shall illustrate, the large adjustments to the longitudinal sample weights lead to wide sampling variability.

In sum, there is suggestive evidence of bias from differential attrition for a number of countries, but strong(er) conclusions are difficult to draw because of the inconsistencies across the different sets of estimates. We observe that our bespoke longitudinal weights generally do a good job of reproducing estimates derived using Eurostat longitudinal weights but, again, there a few marked differences.

22.4.2 Is attrition bias within the range of sampling variability?

Our ability to draw strong conclusions is also complicated by the fact that all estimates are subject to sampling variation, and this may overwhelm any differences in bias due to differential attrition. We illustrate this point in Figure 4, which shows estimates of persistent poverty rates from the four-wave Balanced Sample calculated using the Eurostat weights (crosses) and our bespoke weights (hollow diamonds), and their associated 95% confidence intervals. For reference, also shown (using hollow circles) are the estimates published by Eurostat for the same period (Eurostat 2014) and estimates obtained without any weights (hollow squares).[7]

The main messages of Figure 4 are, first, that confidence intervals for persistence poverty rates calculated using Eurostat and our bespoke longitudinal weights overlap substantially in the vast majority of countries. There are some clear differences, to be sure, most notably for the Netherlands, but also for several other countries (such as Slovakia) – the countries for which we identified differences between the series earlier.

The second lesson is that confidence intervals for persistent poverty rates can be relatively wide. The ranges shown in the figure are of course similar to those suggested by our calculations earlier, but the lesson here is that the effects of differential attrition would have to be relatively large for differences to be significant in the statistical sense and, for example, to change the ranking of countries by persistent poverty rates.

The wide confidence interval for the Netherlands with the EU-SILC longitudinal weights connects to the discussion from Figure 3 and illustrates how large adjustments to sampling weights influences sampling variability. In light of these results, it is unclear whether the benefits in terms of bias reduction from calibrating sample weights to more reliable external information (here the cross-section data) are not offset by the increased sampling variability. A more detailed analysis in terms of mean squared error (which summarises both bias and variance in a single statistic) would be relevant here.

The cross-country ranking of persistent poverty rates shown in Figure 4 is broadly the same as the ranking that we reported in earlier work (Jenkins and Van Kerm 2014), though we should point out that the estimates are not directly comparable because the sets of countries differ (the 21 used in the earlier paper are not a subset of the 23 used here), and the EU-SILC data have been revised since the earlier study.

<Figure 4 near here>

22.5. Summary and conclusions

Rates of attrition from the four-year EU-SILC samples used to calculate persistent poverty rates vary substantially across Member States. The loss of sample size associated with attrition may lead to increases in standard errors for persistent poverty rates, and wider confidence intervals, that are sufficiently large – especially for population subgroups – that it is not possible to derive statistically robust conclusions about changes in persistent poverty rates over time or differences between subgroups.

There is substantial cross-national diversity in the characteristics of individuals lost to follow up. Differential attrition abounds in the EU-SILC. Moreover, we provide indirect evidence that application of longitudinal weights is essential yet it may not fully account for the effects of attrition, and that different assumptions about the poverty status of attritors lead to wide bounds in estimates of persistent poverty rates for most Member States. Thus, overall, we are less sanguine about the impact of attrition of EU-SILC-based estimates of persistent poverty than Watson (2003) was about the estimation of cross-sectional statistics using the ECHP.

[7] Puzzlingly, there are some distinct differences between the published estimates and estimates based on the Eurostat longitudinal weights. The former are larger than the latter for Spain and Cyprus (around three percentage points in the latter case).

We have been unable to pin down with confidence the effects of sample attrition on bias and precision in estimates of persistent poverty rates, but we have produced sufficient evidence to support a conclusion that researchers and data producers need to be mindful of these issues. Researchers analysing persistent poverty should at the least provide standard errors and confidence intervals for their estimates of rates, and their changes over time or differences between countries. (For EU-SILC applications, see e.g. Goedemé 2013 and Chapter 26 by Berger, Goedemé and Osier.) Sampling variability is not identical to the uncertainty introduced by attrition, but accounting for the former should help remind readers of the effects of the latter.

National data collectors and Eurostat should continue their efforts to reduce loss to follow-up. This would be all the more important if a decision is made to extend the panel dimension of EU-SILC to more than four years. Extending the panel duration is attractive on substantive grounds, but minimizing attrition and minimizing cross-country differentials in its patterns is essential to reap the benefits of panel lengthening. Our analysis has provided additional information about the groups at greatest risk of not providing income data for four years, and whom should therefore receive special attention. Reducing inconsistency across countries in the application of following rules would also have payoffs for sample retention overall and reduce differential attrition (see Chapter 27 by Iacovou and Lynn). If more information about the details of the data collection process were available in the EU-SILC User Database, this might be used to derive better weights to account for attrition or to build more successful parametric models. We have also drawn attention to a number of apparent inconsistencies in estimates between cross-sectional and longitudinal components of the EU-SILC, and it would be good to have these resolved.

References

Behr, A., Bellgardt, E., and Rendtl, U. (2005). ‘Extent and determinants of the extent of panel attrition in the European Community Household Panel’, *European Sociological Review*, 21 (5): 489–512.

Eurostat (2010). ‘Description of target variables: cross-sectional and longitudinal. 2008 operation (Version January 2010)’, Document EU SILC 065. Luxembourg: Eurostat.

Eurostat (2014). Persistent at-risk-of-poverty rates. Online database. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_li21&lang=en

Fitzgerald, J., Gottschalk, P., and Moffitt, R. (1998). ‘An analysis of sample attrition in panel data The Michigan Panel Study of Income Dynamics’, *Journal of Human Resources*, 33 (2): 300–344.

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*, 110 (1): 89–110.

Iacovou, M. and Lynn, P. (2013). ‘Implications of the EU-SILC following rules, and their implementation, for longitudinal analysis, ISER Working Paper 2013-17. Colchester: Institute for Social and Economic Research, University of Essex.

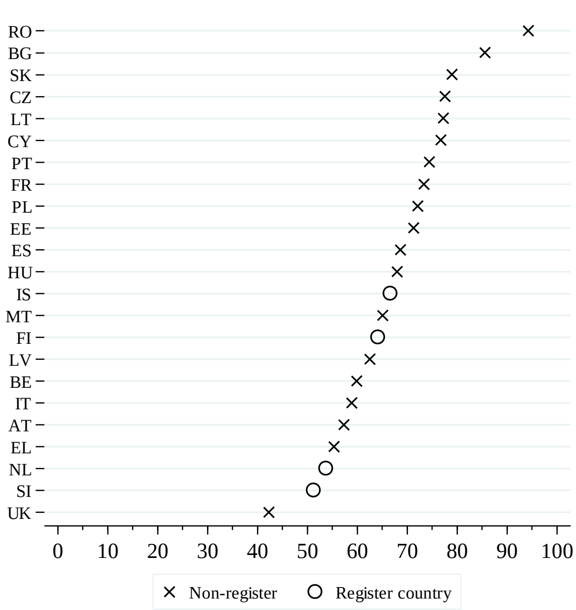
Jenkins, S. P. and Van Kerm, P. (2014). ‘The relationship between EU indicators of persistent and current poverty’, *Social Indicators Research*, 116 (2): 611–638.

Jenkins, S. P. and Van Kerm, P. (2015). ‘How does attrition affect estimates of persistent poverty rates? The case of EU-SILC’, Eurostat Net-SILC2 Working Paper, forthcoming.

Lohmann, H. (2011).‘Comparability of EU-SILC survey and register data: the relationship among employment, earnings and poverty’, *Journal of European Social Policy*, 21 (1): 37–54.

Museux, J. M. (2006). 'Weighting for EU-SILC'. Presentation at Methodological Workshop on 'Comparative EU-Statistics on Income and Living Conditions: Issues and Challenges', Helsinki, 6–8 November 2006.
Watson, D. (2003). 'Sample attrition between waves 1 and 5 in the European Community Household Panel', *European Sociological Review*, 19 (4): 361–378.

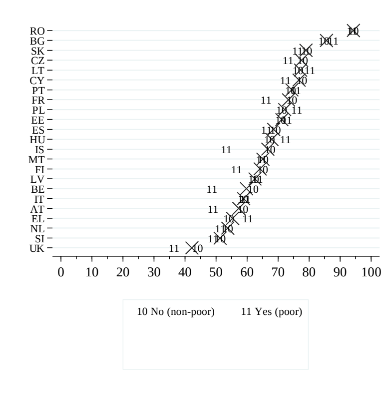
Figure 1. Retention rates by country, 2008-2011 (per cent)



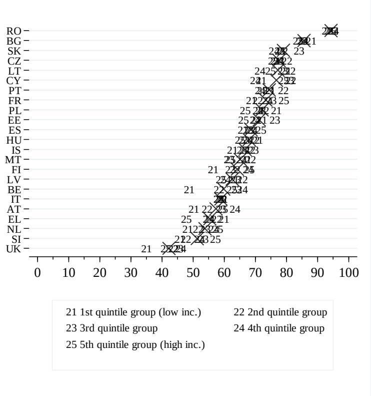
Note: The retention rate is the proportion of individuals belonging to a respondent household at Wave 1 (2008) which remains in a participating household in each of the three subsequent waves. Only these individuals are used for the calculation of the 2011 persistent poverty rates. Unweighted proportions of wave 1 sample.
Source: Calculations from 2011 EU-SILC Longitudinal data; 2008 rotation group only.

Figure 2. Retention rates by characteristic by country, 2008-2011
(per cent)

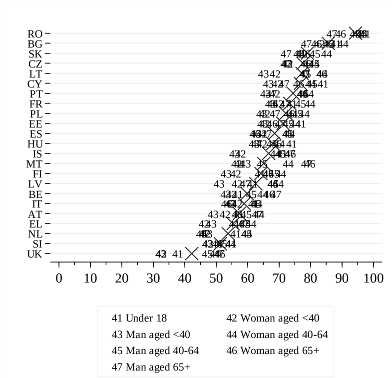
(a) Poverty status



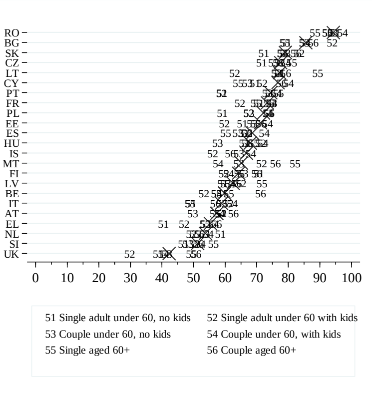
(b) Income quintile group



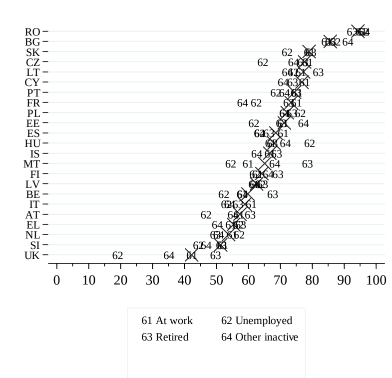
(c) Age-sex



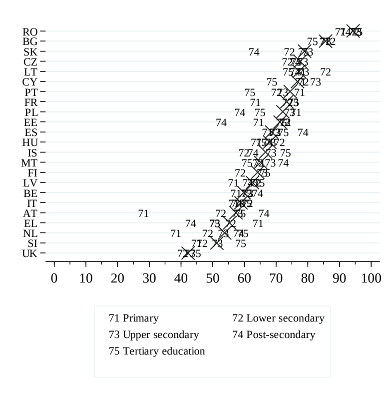
(d) Household type



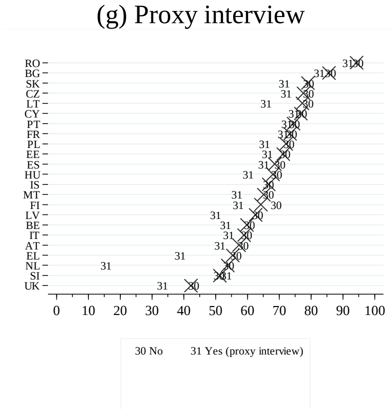
(e) Activity status of household head



(f) Education of household head



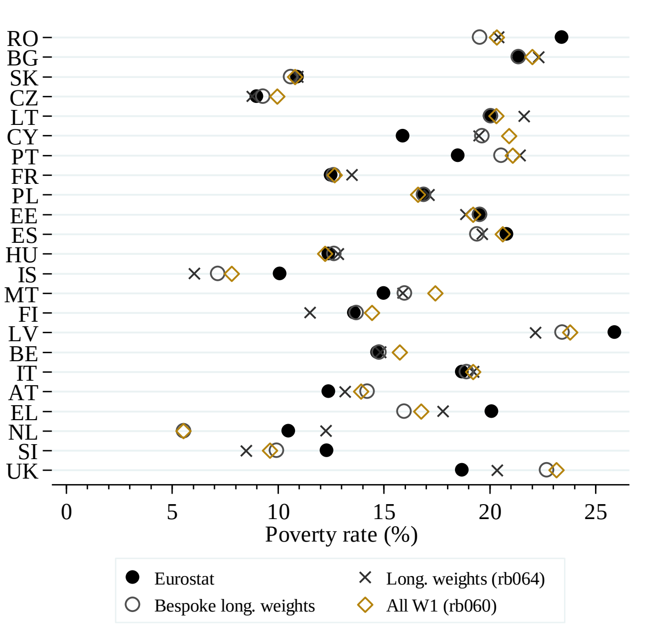
(g) Proxy interview



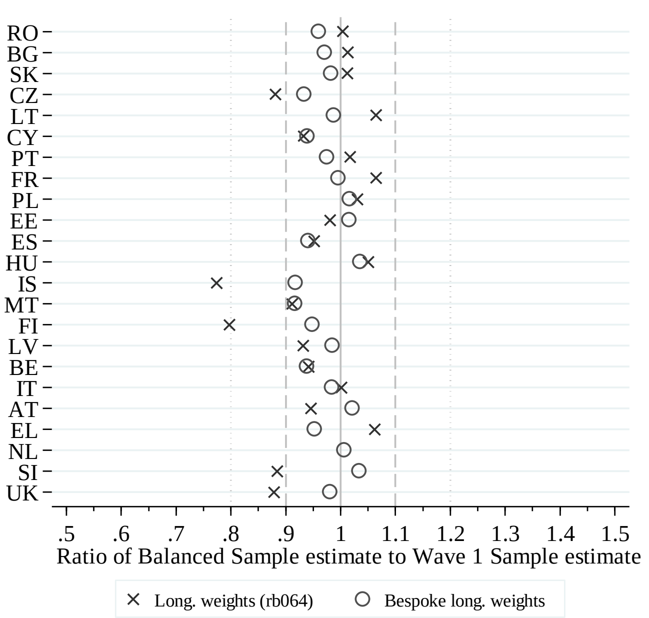
Note: Breakdowns are based on data observed in wave 1. Unweighted proportions of wave 1 sample. Crosses indicate the overall retention rate while numbers identify subgroup retention rates. Retention rates are defined as in Figure 1.
Source: Calculations from 2011 EU-SILC Longitudinal data, 2008 rotation group only.

Figure 3. Estimates of Wave 1 (2008) poverty rate, by subsample and sample weight

(a) Rates



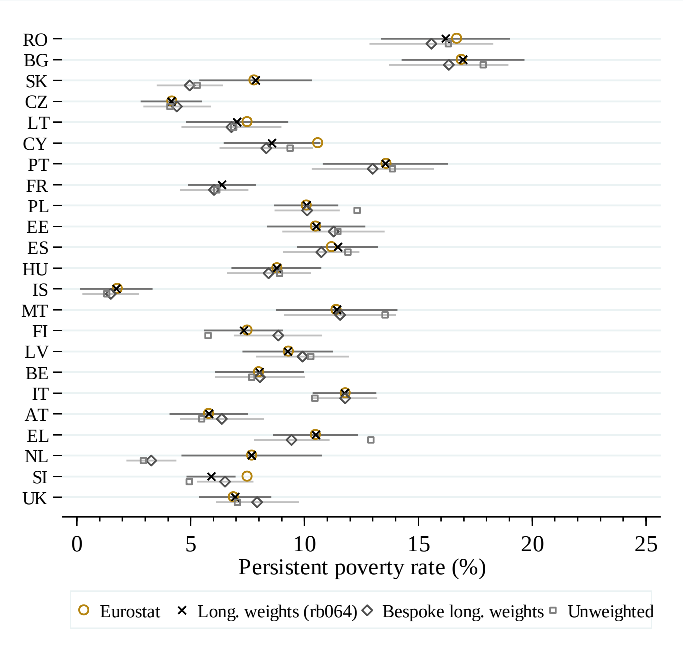
(b) Rates expressed as a fraction of the rate for the full Wave 1 Sample



Notes. All estimates are based on the balanced sample of full 4-year respondents, except 'All W1' which are based on all respondents at wave 1 (including subsequent attritors) and Eurostat's estimates based on cross-section data for 2008. The weighting of the samples are: 2008 base weights (*rb060*) for 'All W1', 2011 SILC longitudinal weights (*rb064*) and 2011 bespoke longitudinal weights for the Balanced Sample. Panel (b) shows estimates expressed as a fraction of the corresponding 'All W1' sample estimates. In the bottom panel, the estimate for NL is not shown as its estimate from the SILC longitudinal sample is an outlier (see text).

Source: Calculations from 2011 EU-SILC Longitudinal data; 2008 rotation group only. Eurostat estimates are from Eurostat (2014) and are computed using 2008 cross-section EU-SILC data.

Figure 4. Estimates of 2008-2011 persistent poverty rates with different sampling weights (per cent)



Source: Calculations from 2011 EU-SILC Longitudinal data (version 1, released August 2013). Eurostat estimates are from Eurostat (2014).