Cross-sectional and longitudinal weighting in a rotational household panel: applications to EU-SILC

Vijay Verma, Gianni Betti, Giulio Ghellini

Working Paper n. 67, December 2006

CROSS-SECTIONAL AND LONGITUDINAL WEIGHTING IN A ROTATIONAL HOUSEHOLD PANEL: APPLICATIONS TO EU-SILC

Vijay Verma[1], **Gianni Betti**[1], **Giulio Ghellini**[1]

ABSTRACT

This paper provides a comprehensive description of an integrated system of cross-sectional and longitudinal weighting for rotational household panel surveys. To be concrete and detailed, it develops the weighting procedures with reference to the EU-SILC *integrated design*. EU-SILC (Statistics on Income and Living Conditions) covers data and data sources of various types: cross-sectional and longitudinal; household and personal; economic and social; and from registers and interview surveys. The standard integrated design involves a rotational panel in which a new panel is introduced each year to replace one quarter of the existing sample; persons enumerated in each new panel are followed-up in the survey for four years (Verma and Betti, 2006). A common rotational sample of this type yields each year a cross-sectional sample as well as longitudinal samples of various durations. These sample data have to be weighted to make them more representative of the target populations they represent. The paper begins with a summary of the main features of EU-SILC and an overview of the integrated weighting system for the different types of data coming out of the rotational panel annually. It describes a step-by-step procedure for construction of *initial weights* to be applied to each new sample as it is introduced into the survey. An innovative feature of the weighting procedure is the concept of *base weights* (Verma and Clemenceau, 1996). Starting from the initial weight of each individual in the original sample, the person's base weight is constructed for each subsequent year to compensate for panel attrition. The final objective is of course to construct *cross-sectional weights* and *longitudinal weights* for use in data analysis. Procedures are described for constructing these two types of weights from the same base weights. This makes the whole system of weights internally consistent and integrated.

***Key words*:** weighting, household panel, rotational design, cross-sectional weights, longitudinal weights, EU-SILC.

[1] V. Verma, e-mail: verma@unisi.it (corresponding author); G. Betti e-mail: betti2@unisi.it, G. Ghellini, e-mail: ghellini@unisi.it. Department of Quantitative Methods, University of Siena, P.za S. Francesco, 7, 53100, Siena, Italy.

1. The EU-SILC framework

1.1. Introduction: context and content of the paper

The present paper provides a comprehensive description of an integrated system of cross-sectional and longitudinal weighting for rotational household panel surveys, specifically with reference to the EU-SILC integrated design.

As is well-known by this time, EU-SILC is the major new source of comparative statistics on income and living conditions in Member States of the European Union and some neighbouring countries. It has been developed as a flexible yet comparable instrument for the follow-up and monitoring of poverty and social exclusion at the EU and national levels. It covers data and data sources of various types: cross-sectional and longitudinal; household-level and person-level; economic and social; from registers and interview surveys; from new and existing national sources.

In previous papers and reports (Verma, 2001; Verma and Betti, 2006) we have elucidated the structure and main characteristics of EU-SILC surveys, and the various technical considerations involved in the design and implementation of samples for EU-SILC. Despite the diversity of arrangements permitted under EU-SILC, the standard *integrated design* recommended by Eurostat has been adopted by a big majority of the participating countries. This integrated design involves a rotational panel in which a new sample of households and persons is introduced each year to replace a part (normally one quarter) of the existing sample. Persons enumerated in each new sample are followed-up in the survey for four (or more) years. A common rotational sample of this type yields each year a cross-sectional sample as well as longitudinal samples of various durations. In most situations, these sample data have to be weighted to make them more representative of the target population of the survey. The complex structure of the sample means that the corresponding weighting procedures can also be quite complex. The weighting procedure described here is in fact based on detailed recommendations originally developed by one of the present authors (Verma, 2006). These recommendations have been adopted by Eurostat and are being implemented in EU-SILC national surveys. The present paper aims to provide a more systematic and clearer description of the weighting procedures, also introducing some refinements so as to enhance the consistency and completeness of original recommendations. The rest of Section 1 summarises the main features of EU-SILC, such as its scope and content, data structure, longitudinal follow-up (tracing) rules, and in particular the integrated rotational panel design adopted by most countries. The objective is to provide the necessary background for describing the recommended weighting procedures. Further details may be found in various EU-SILC regulations published in the Official Journal of the European Community, the many technical reports produced by Eurostat, and also in the references already cited above.

Section 2 provides an overview of the integrated weighting procedure. It clarifies the different types of weights required for the different types of data coming out of EU-SILC annually. With a 4-year rotational design, once established, one cross-sectional dataset and three longitudinal datasets, respectively of 2, 3 and 4 years duration are generated each year. We explain the concepts of 'initial' and 'base' weights, on the basis of which the required cross-sectional and longitudinal weights can be constructed as an integrated whole.

Section 3 deals with the construction of *initial weights*, by which we mean weights to be applied to each new sample as it is introduced into the survey. A step-by-step procedure is described, starting from design weights, followed by adjustments for non-response and calibration to external controls, and finally trimming and scaling as required to obtain the initial weights. Of course, these procedures are applicable to any survey and are not specific to a rotational panel. However, in a panel survey it is particularly important to construct good initial weights for each new sample as it is introduced, since these determine the quality of the longitudinal weights, as well as the cross-sectional weights, to be applied in subsequent years the sample remains in the survey.

An innovative feature of the weighting procedure is the concept of *base weights*, developed originally in the context of European Community Household Panel (Verma and Clemenceau, 1996). Starting from the initial weight of each individual in the original sample, the person's base weight is constructed - for each subsequent year the person remains in the survey - with the objective of compensating for panel attrition. The procedure for constructing base weights is described in Section 4.

Sections 5 and 6 describe the procedures for constructing, respectively, *longitudinal weights* and *cross-sectional weights*. These two types of weights can in fact both be obtained in a straightforward way from the same base weights. This makes the whole procedure internally consistent and integrated.

Finally, in Section 7 we comment on some important practical aspects in the implementation of the weighting procedures.

1.2. Scope and content of EU-SILC

As noted, EU-SILC aims to be a flexible yet comparable instrument covering data and data sources of various types.

In terms of the substantive content, four types of data are involved: (i) variables measured at the household level; (ii) information on household size and composition and basic characteristics of household members; (iii) income and other more complex variables measured at the personal level, but aggregated to construct household-level variables (which may then be ascribe to each member for analysis); and (iv) more complex non-income or 'social' variables collected and analysed at the personal level.

For set (i)-(iii) variables, a sample of households including all household members is required. Among these, sets (i) and (ii) are normally collected from a single, appropriately designated respondent in each sample household. Alternatively, some or all of these data may be compiled from registers or other administrative sources.

Set (iii) variables - concerning mainly, but not exclusively, the detailed collection of household and personal income - must be collected directly at the personal level, *covering all persons in each sample household*. In many countries, these income and related variables are collected through personal interviews with all adults aged 16+ in each sample household. This collection is normally combined with that for set (iv) variables, since the latter must also be collected directly at the personal level. These are the so-called *survey countries*. By contrast, in some countries, set (iii) variables are compiled from registers and other administrative sources, thus avoiding the need to interview all members (adults aged 16+) in each sample household. These are the so-called *register countries*.

Set (iv) variables are normally collected through direct personal interview in all countries. These are too complex or personal in nature to be collected by proxy, nor are they available from registers or other administrative sources. For the survey countries, this collection is normally combined with that for set (iii) variables as noted above, both being based on a sample of complete households, i.e. covering all persons in each sample household. However, from the substantive requirements of EU-SILC, it is not essential — in contrast to set (iii) variables — that set (iv) variables be collected for all persons in each sample household, since these variables need not to be aggregated to the household level. It is therefore sufficient to do this collection on a representative sample of persons. This option is normally followed in register countries, since for these countries interviewing all household members for set (iii) is not involved. In countries which choose to do so, the sampling process involves the selection of persons (usually one adult member aged 16+ per household) either directly or, optionally, through a sample of households.

1.3. Data structure

Hence different types of units of analysis are involved in EU-SILC for which sample weights have to be defined: (i) private households; (ii) all persons residing in sample households; (iii) all household members aged 16+; and optionally (iv) one selected adult per sample household. One may also be interested in special groups, such as children.[2]

Another dimension is that both cross-sectional and longitudinal data are required. The cross-sectional component covers information pertaining to the current and a recent period such as the preceding calendar year. It aims at providing estimates of cross-sectional levels and of net changes from one period (year) to another. The longitudinal component covers information compiled or collected through repeated enumeration of individual units, and then linked over time at the micro-level. It aims at measuring gross (micro-level) change and elucidating the dynamic processes of social exclusion and poverty. Both cross-sectional and longitudinal micro-data sets are updated on an annual basis. In EU-SILC a period of four years is taken as the minimum duration for longitudinal follow-up at micro level.

Combining the various types of units and the time dimension, the new data sets disseminated each year consist of the following: cross-sectional data pertaining to the most recent reference year for households and persons; data pertaining to three different longitudinal periods, covering 2, 3 and 4 years preceding the survey, only for persons.

1.4. The integrated design

Verma and Betti (2006) illustrate a typology of possible data source structures in EU-SILC. A single *integrated* source covering all components – cross-sectional and longitudinal, income and social - is by far the most common one adopted by countries up to now. Throughout, we will describe the weighting procedure with reference to this integrated design.

The basic idea is as follows. At any one time, the sample is made up of, say, 4 short-term subsamples or panels. Each year one new panel is added to stay in the survey for 4 years, and then dropped to be replaced by another new panel. Movers from the original sample are followed-up to their new location for up to the time their panel remains in the survey.[3] Each panel provides a longitudinal sample of the chosen duration. The units present at a given time from all the panels are appropriately put together to constitute the cross-sectional sample. Clearly, an important advantage of this scheme is that both cross-sectional and longitudinal data are obtained from the same common set of units. This overlap is highly economical, and also maximises internal consistency between longitudinal and cross-sectional statistics produced from the survey.

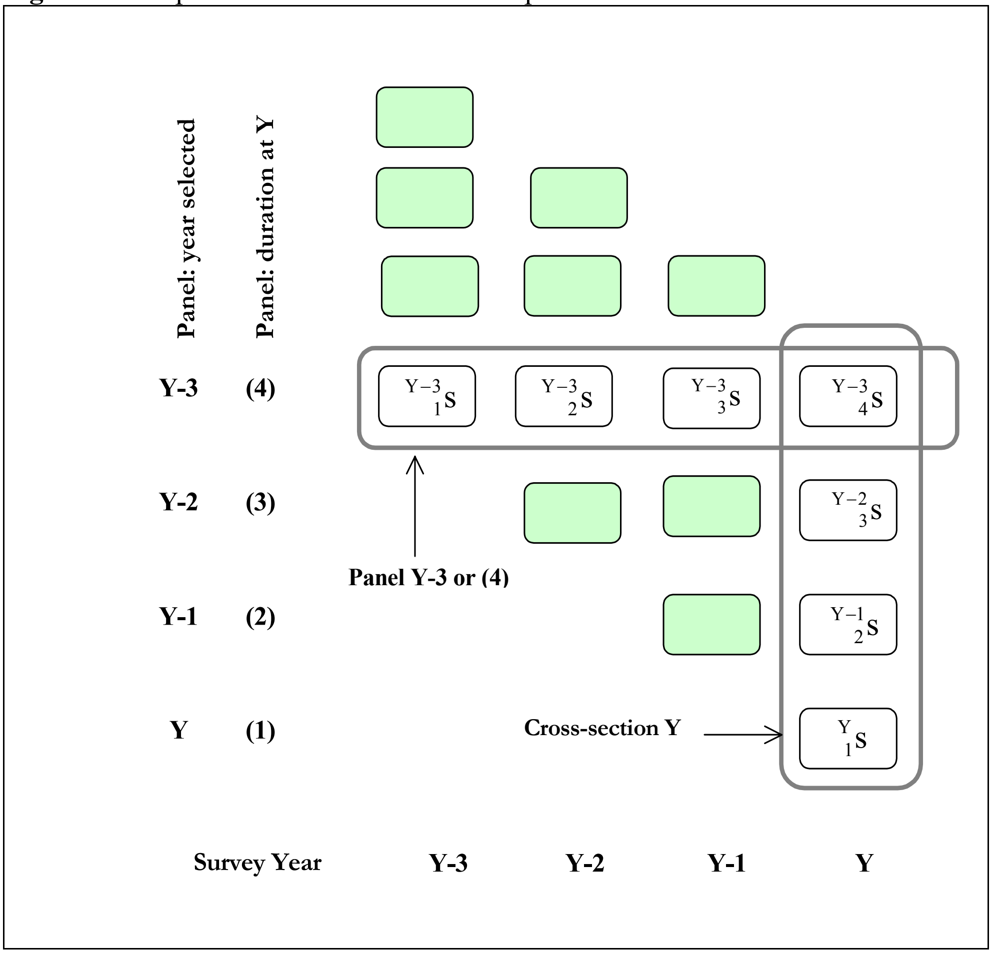
The rotational scheme is illustrated in Figure 1. It also specifies the notation we will use.

In the above diagram, ${}^{Y-3}_{t}s$ is a typical panel sample introduced in year (Y-3); this sample (more strictly, the samples derived from it according to certain follow-up rules) is enumerated over four years (Y-3) to Y. Longitudinal samples of 2 to 4 years duration are constructed by putting together different panels of this type, as will be discussed later (Section 5). The full cross-sectional sample at year Y is composed of four panels ${}^{Y-t+1}_{t}s$, t = 1 to 4, as shown in Figure 1.

[2] Following the EU-SILC terminology, we will refer to these different types of units and to the associated data files as follows: H ('household'), R ('register' covering all members), P ('personal' covering all adults aged 16+), S ('selected respondent'), and Q (children or other special groups).

[3] According to *Commission Regulation (EC) No 1982/2003 as regards the Sampling and Tracing Rules* rotational design refers to "the sample selection based on a number of subsamples or replications, each of them similar in size and design and representative of the whole population. From one year to the next, some replications are retained, while others are dropped and replaced by new replications".

Figure 1. The panel and cross-sectional samples



1.5. Sample follow-up (tracing) rules

Follow-up rules are required in a household panel to preserve its representative nature over time. In the case of EU-SILC, the relevant Commission Regulation (European Community, 2003) specifies these formally as follows.

“ (a) Initial sample: refers to the sample of households or persons at the time it is selected for inclusion in EU-SILC.

(b) Sample persons: means all or a subset of the members of the households in the initial sample who are over a certain age.

(c) Age limit used to define sample persons: … In countries with a four-year panel using a sample of addresses or households, all household members aged 14 and over in the initial sample shall be sample persons. In countries with a four-year panel using a sample of persons, this shall mean the selection of at least one such person per household. ...

(e) Sample household: means a household containing at least one sample person. A sample household shall be included in EU-SILC for the collection or compilation of detailed information where it contains at least one sample person aged 16 or more. ...

(g) Co-residents (non-sample persons): all current residents of a sample household other than those defined above as sample persons. ...

(i) Age: refers to the age at the end of the income reference period.”

Ideally, all persons, irrespective of age, in initial households should be followed-up over the panel duration. The age limit of 14 has been introduced in EU-SILC regulations merely for practical reasons. The impact of this limitation is generally not important because we can expect very few children under 14 to move to a new address ‘alone’, without being accompanied by one or more adults in their original household at the time of selection. Hence for the present discussion, we will consider as sample persons *all individuals in the initial households, irrespective of age, and assume that all these are supposed to be followed-up to their new address if they move*. The few cases where the slightly restrictive rules adopted in EU-SILC do not require a follow-up can simply be treated as non-respondents.

The above applies to the normal situation in survey countries where all adults in the household are supposed to be followed-up. Additional problems can arise in ‘register countries’ where only one adult per household is selected for follow-up. We comment on this in the concluding Section 7.3.

2. Overview of the sample weighting procedure

2.1. The importance of weighting sample data

As noted by Kish (1992), “why, when and how” to weight are among the fundamental and most common questions in estimation and analysis using sample survey data. The answer to these questions may depend on the context, the data source and the type of data and analysis involved. Sometimes there are sharp arguments as to whether or not it is appropriate or necessary to weight at all. In the case of an intensive and complex survey of limited size, such as a household panel like EU-SILC, we believe that the answer to the question is quite clear: in most situations it is both necessary and useful to weight the sample data to compensate for imperfection of the achieved sample.

Firstly, weighting is introduced to compensate for differences in sampling probabilities. Such weighting is essential if those differences are large. Another important reason making weighting necessary is the high rates of non-response and attrition over time in panel surveys. Non-response and attrition is often not only large but also selective, such as being higher for households at the extreme ends of the income distribution.

Further improvement in the representativeness of the sample can be made through its calibration against more reliable external information such as age-sex composition and geographical distribution of the population, and the distribution of households by type and size. Such calibration can reduce biases in the sample due to non-response, non-coverage and other distortions, and also reduce variances. In the EU context at least, a great deal of external information is available in most countries for calibration of the sample.

In household panel surveys, two additional factors make the weighting procedures complex: one, the survey deals with units of different types (households, all members, adults); and two, both cross-sectional and longitudinal components are involved, each with its own weighting requirements.

In all situations, the legitimate objective of weighting the sample data is to reduce the resulting mean-square-error of the estimates from the survey. In special situations (such as optimal allocation and calibration) weights are introduced to reduce variance. However, in most situations the introduction of unequal weights tends to inflate variances, and the purpose of weighting is to reduce biases in the estimation. Compromise is often required to balance these effects and control the mean-square-error. Even a small number of extreme values, in particular at the upper end of the weight distribution, can inflate the variance significantly. In recommending the rather complex weighting procedures for households panels in this paper, we must emphasise the importance of checking for extreme values or large variations being introduced in the weights. Alternative procedures have been suggested to limit the range of weight values encountered, such as ‘shrinkage’ of weights (Spencer and Cohen, 1991). In our experience, the trimming of extreme values is often the simplest.

2.2. Initial and follow-up weights in a household panel

There are two broad aspects of the weighting procedure in a household panel: initial weighting at the time of introduction of the panel units into the survey; and modification of the sample and of the initial weights in the following period.

Initial weights

Irrespective of the details of the design, any panel introduced into the survey begins with a sample of households, and hence also of persons who are members of those households. We assume that by design, these samples are probability samples. We term weights assigned to the achieved sample as it first appears in the survey as the *initial weights*. These will be discussed in more detail in Section 3.

The procedures for constructing the initial weights are in fact the same as those used for all types of surveys, including purely cross-sectional surveys. The development of initial weights is generally performed in stages: three or four stages may be involved, possibly some of them involving multiple steps. Some useful references in the literature include Kish (1992), Kalton and Kasprzyk (1986), Little (1986), and, for detailed application to ECHP, Verma and Clemenceau (1996).

Follow-up sample and weighting

The initial sample provides the set of individuals, called *sample persons*, who are followed-up over duration of the panel. The sample needs to be modified to reflect changes in the target population over time. At least three types of adjustments are involved. The first concerns changes in the target population in private households: individuals leaving the population due to death or out-migration, and new persons joining the population through births or in-migration. Unless supplementary samples of new entrants into the population are added, the panel cannot reflect in-migration fully. The second arises from non-response and other losses in the sample. The weights of units remaining in the sample need adjustment in order to reduce the impact of non-response on sample representativeness. The third source of change in the sample is the entrance of non-sample persons into the households included in the sample. Many household panel surveys, including EU-SILC, collect a great deal of information on these non-sample members or *cohabitants*, while they are living with some sample person, in order to measure the circumstances (such as income) of the whole household. It is desirable (efficient) to exploit the information at the individual level collected for these cohabitants. Appropriate weighting schemes are required for their incorporation into the sample, which of course must continue to reflect the corresponding target population. The basis of these procedures is provided by the *weight share* method expounded by Ernst (1989). Some other useful references in the literature include Lavallée (1995), Kalton and Brick (1995), Lavallée and Caron (2001), Deville and Lavallée (2006), and, in the specific context of ECHP, Verma and Clemenceau (1996).

2.3. Structure of the weighting procedure; notation

In this section, we provide an overview of the proposed weighting procedure. Another important objective is to introduce the notation we use throughout.

Consider the rotational structure represented in Figure 2. At a given time, the total sample is made up of a number of panels. The following notation will be used for panel-specific weights (w):

${}_{t}^{Y}w_{j}^{(U,X)}$	where the identifiers of the weight are as follows
Y	refers to the year the panel was selected; alternatively, it may be a more convenient to identify the panel by its current age (T).
t	$1 \leq t \leq T$, refers to the panel's age in general.

U identifies different types of units: H (households); R (all household members); P (all members aged 16+); S (where relevant, the selected respondent); and Q (children or other population subsamples of interest).

X identifies the type of weights at wave 1 (D, N, F, I, see below). Thereafter we deal with base weights X = B.

j is a particular unit (household or person). It is understood that (j) belongs to sample s identified in the same way as w: $j \in {}^{Y}_{t}s^{(U)}$.

When certain identifiers are nor needed, they are dropped for simplicity. For instance, most commonly we deal with base weight (X = B) of individual persons (U = R), for any unspecified unit (j), so that the weight can be identified as ${}^{Y}_{t}w$ and the corresponding sample as ${}^{Y}_{t}s$.

When the only needed reference is to t, we will employ the simplified and commonly used notation w_t and s_t, respectively for the weight and the sample.

A longitudinal sample, ${}^{Y}_{t}s^{(L)}$, is defined as the set of individuals (j) who have been enumerated in the survey throughout the period 1 to t inclusive:

$$ {}^{Y}_{t}s^{(L)} = {}^{Y}_{1}s \cap {}^{Y}_{2}s \cap \ldots \cap {}^{Y}_{t}s \,; \tag{1} $$

when convenient, we use simplified notation corresponding to s_t, namely $s_t^{(L)}$.

<u>Initial and base weights</u>

As described in Section 3, the final wave 1 weights are constructed involving a step-by-step procedure defining:

- D, design weights, in inverse proportion to the unit's selection probability;
- N, the above weights modified to reflect non-response at wave 1;
- F, the resulting weights calibrated to 'fit' certain specified control distributions;
- I, trimming and scaling of the above, giving the initial wave 1 weights.

For individuals remaining in the sample during subsequent periods t >1, we define a sequence of base weights (B), which take into account attrition of the panel. To start with, the base weight is defined to be identical to the wave 1 initial weight:

$$ {}^{Y}_{1}w_j^{(U,B)} \equiv {}^{Y}_{1}w_j^{(U,I)} \,. \tag{2} $$

<u>Cross-sectional and longitudinal weights for the full sample</u>

The panel-specific weights (w) are 'intermediate' quantities. These are put together to construct weights (v) for the full sample comprised of different panels. Weights v are of different types (C, L2, L3, L4; see Figure 2) and represent the actual weight variables used in analysis. We indicate these as:

$$ {}^{Y}_{X}v_j^{(U)} \,. $$

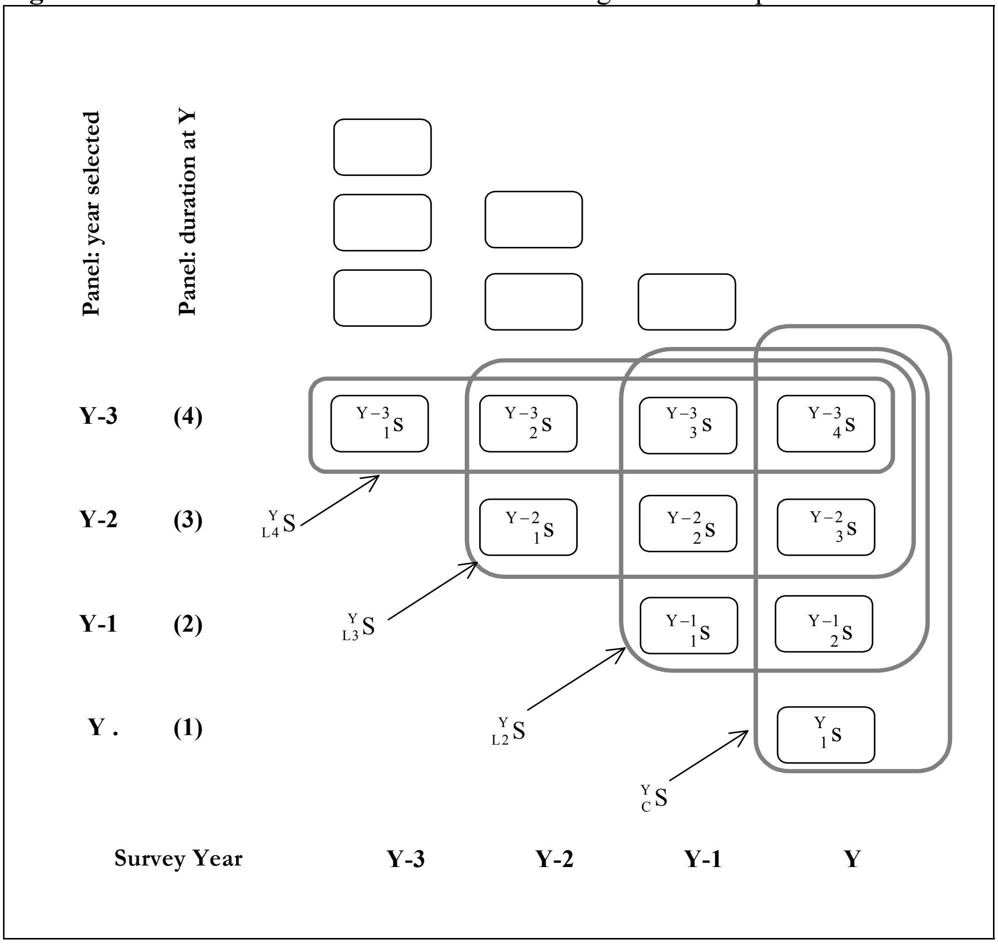
Firstly, for cross-sectional weights (X = C), for reference year Y, we have:

$$ {}^{Y}_{C}v_j^{(U)} \,, $$

defined for the full cross-sectional sample, i.e. all units in the survey at year Y:

$$ {}^{Y}_{C}S = {}^{Y}_{1}s \cup {}^{Y-1}_{2}s \cup {}^{Y-2}_{3}s \cup {}^{Y-3}_{4}s \,. $$

Figure 2. The annual cross-sectional and three longitudinal samples



It can be seen that each year, results for three new longitudinal samples of different durations – 2, 3 and 4 years – become available. We identify the corresponding weights, respectively, X = L2, L3 and L4. The constituent samples, S, to which these apply are:

$$
\begin{aligned}
{}^{Y}_{L2}S &= ({}^{Y-3}_{3}s \cap {}^{Y-3}_{4}s) \cup ({}^{Y-2}_{2}s \cap {}^{Y-2}_{3}s) \cup ({}^{Y-1}_{1}s \cap {}^{Y-1}_{2}s) \\
&= ({}^{Y-3}_{3}s \cap {}^{Y-3}_{4}s) \cup ({}^{Y-2}_{2}s \cap {}^{Y-2}_{3}s) \cup {}^{Y-1}_{2}s^{(L)} ,
\end{aligned} \quad (3)
$$

$$
\begin{aligned}
{}^{Y}_{L3}S &= ({}^{Y-3}_{2}s \cap {}^{Y-3}_{3}s \cap {}^{Y-3}_{4}s) \cup ({}^{Y-2}_{1}s \cap {}^{Y-2}_{2}s \cap {}^{Y-2}_{3}s) \\
&= ({}^{Y-3}_{2}s \cap {}^{Y-3}_{3}s \cap {}^{Y-3}_{4}s) \cup {}^{Y-2}_{3}s^{(L)} ,
\end{aligned} \quad (4)
$$

$$
{}^{Y}_{L4}S = ({}^{Y-3}_{1}s \cap {}^{Y-3}_{2}s \cap {}^{Y-3}_{3}s \cap {}^{Y-3}_{4}s) = {}^{Y-3}_{4}s^{(L)} . \quad (5)
$$

The quantity ${}^{Y}_{t}s^{(L)}$ has been defined in (1).

Note that generally there is a large overlap in terms of units and data among the longitudinal samples and, in an integrated rotational design, also with the cross-sectional sample.

Nevertheless, each sample $S = ({}^{Y}_{C}S, {}^{Y}_{L2}S, {}^{Y}_{L3}S, {}^{Y}_{L4}S)$, with corresponding X = (C, L2, L3, L4), has its own distinct set of weights. Furthermore, in general, the set of weights differ by the type of unit U = (H, R, P, S, Q).

3. Initial weights

3.1. A step-by-step procedure

We begin from the construction of initial weights given to households and persons as they first enter the survey. Throughout, we consider one panel at a time. This weighting procedure is of course the same as that in any comparable cross-sectional survey, but we aim at bringing out some important theoretical and practical points and make some recommendations.

In developing the weights, the best possible use has to be made of the information available, both internal to the sample and from external sources. This includes information on: (i) coverage and selection probabilities; (ii) characteristics and circumstances of non-responding units; (iii) the structure and characteristics of the target population.

Such information can be used in a systematic manner to apply weights in a step-by-step procedure. We may identify the following four steps:

- Design weights. Each household and person in the sample is weighted in inverse proportion to its probability of selection.
- Non-response weights. These are introduced to reduce the impact of non-response on the structure and characteristics of the achieved sample. These weights are derived on the basis of items of information which are available for both responding and non-responding units.
- Calibration weights. This refers to the adjustment of the distributions of the weighted sample according to various characteristics to agree with more reliable information on those distributions available from sources external to the survey. In a household survey, distributions both in terms of households and persons may be involved.
- Trimming and scaling. This refers to adjustments made to avoid extreme values in the distribution of weights. The objective is to avoid large increases in variance which result from the presence of extreme weights, even though such adjustment may introduce some bias. For many purposes, only the relative values of weights are relevant, and their overall scaling is immaterial. However, an appropriate scaling of these quantities is usually necessary for clarity and convenience.

Some basic features of the weighting procedure described below may be noted:

- The weights are applied in a sequence. That is, at any step after the first, the weights are computed from sample values already weighted according to the results of all preceding steps. Thus weighting for non-response is determined from the sample results weighted by the design weights. This simplifies the computations, and shows more clearly the contribution of each step in adjusting the weights.
- In a household panel survey, typically the data are used simultaneously for analysis at the household and the personal levels. It is desirable, therefore, to use a weighting procedure which ensures consistency between analyses involving the two types of units. The recommended procedure is 'integrative weighting' at wave 1, which ensures a uniform weight for all members of a household, the same as the weight of the household.
- The specific variables involved at each step and the sources of the data used may vary from one survey to another. Nevertheless, certain variables can be expected to be important in most circumstances, such as geographic location of the household, household size and composition, and distribution of the population by age, sex and other basic characteristics.
- It is possible in principle to combine more than one step into a single procedure. However, it is desirable that each step is separately implemented, where possible.

- The final weight of a unit is the product of the weighting factors determined at each step. The resulting weights at each step may be appropriately scaled, such as to average 1.0 per sample unit.

3.2. Design weights

Design weight of a unit is inversely proportional to its probability of selection into the sample, though allowances can sometimes be made for known exclusions or under-coverage of parts of the study population.

In a panel survey, the design weights are defined only at wave 1, when the unit is selected into the sample. The unit, or derivations from it, at subsequent waves may be considered subject to 'indirect sampling' (Deville and Lavallée, 2006), but in the present case the concept of base weights described in the next section provides a clearer and simpler alternative.

There are essentially two ways of selecting the sample for a household panel. The first is to select households - or addresses, dwelling or other structures which contain households - directly from appropriate lists. The procedure directly gives the probability of selection of the household (P_h). In the notation introduced in Section 2.3, the household design weight is:

$$_1w_h^{(H,D)} = k/P_h \,,$$

where k is an arbitrary scaling factor to be chosen.

In so far as all persons in the household are automatically taken into the sample once the household is selected, the members' selection probabilities and hence weights are identical to those of their household. The above equality applies also to any subpopulation in the household, such as persons aged 16+ (P), or children (Q):

$$_1w_{j\in h}^{(P,D)} = {}_1w_{j\in h}^{(Q,D)} = {}_1w_{j\in h}^{(R,D)} = {}_1w_h^{(H,D)} \,. \quad (6)$$

In situations where one adult in the household is selected as the respondent for a personal interview, we have the selection probability and weight of selected respondent (S) as:

$$P_{s\in h} = P_h/n_h \,;\; {}_1w_{s\in h}^{(S,D)} = {}_1w_h^{(H,D)}.n_h \,, \quad (7)$$

where n_h is the number of household members eligible to be a selected respondent.

The other way of obtaining a sample of household is to first select a sample of individuals meeting certain specified criteria, and then construct a household around each such selected individual. A household is selected through its association with one or more such individuals. Normally, the latter are selected from a list of adults. In so far as each eligible household contains at least one such person in the list, the household receives a non-zero probability of selection. The population of households not containing such a persons will not be covered in the survey.

The probability of selection of a household (P_h) relates to that of the individuals (P_s) through which it could have been selected as:

$$P_h = \sum_{s\in h} P_s \,.$$

Concerning weights we have:

$$_1w_s^{(S,D)} = k/P_s \,,$$

$$_1w_h^{(H,D)} = k\Big/\sum_{s\in h} P_s \,.$$

The person weights (R) of household members are identical to this household weight, as in equation (6).

Usually, persons eligible for selection within a household all have the same probability of selection, so that:

$$\sum_{s\in h} P_s = n_h P_s \,,$$

giving:

$$ {}_1w^{(H,D)} = \left(\frac{1}{n_h}\right) \cdot {}_1w_s^{(S,D)} . \tag{8} $$

3.3. Non-response adjustment

In a panel, the largest loss of the sample due to non-response generally occurs at the *first wave* when the household is introduced into the survey. Good, efficient procedures to re-weight the responding cases is, therefore, a critical requirement at wave 1. However, the possibilities are often constrained by lack of information: non-response adjustment has to be based on characteristics which are known for *both* responding and non-responding households.

Here the concern is with non-response at the stage of the *household interview*, which obtains information at the household level and enumerates the population in the household. The procedure involves estimating response rates or propensities to response as functions of characteristics available for both responding and non-responding households, including characteristics of the area where the household is located. This is also true when a sample of selected respondents has been used, except that personal characteristics of the selected individuals are also useful as determinants.

There are two commonly used procedures for non-response weighting.

<u>Weighting within classes</u>

The first is to modify the design weights by a factor inversely proportional to the response rate within each weighting class:

$$ {}_1w_{j\in K}^{(H,N)} = {}_1w_{j\in K}^{(H,D)} \cdot \frac{\overline{R}}{R_K} , \tag{9} $$

where R_K is the response rate in weighting class K, and $\overline{R}$ their average (this is just a constant introduced for convenience). The response rates should be computed with data weighted by the design weights:

$$ R_K = \frac{\sum_{j\in Kr} {}_1w_j^{(H,D)}}{\sum_{j\in Ks} {}_1w_j^{(H,D)}} , $$

where the set K_r represents the responding units in the class, and K_s the selected units in the same class.

By 'class' we mean an appropriately determined grouping of units. It is common to use sampling strata or other geographical partitions as weighting classes. Numerous, very small weighting classes can result in large variations in R_K values, and should be avoided. On the other and, if only a few broad classes are used, little variation in the response rates across the sample may be captured – making the whole re-weighting process ineffective. On practical ground, classes of average size 100-300 units may be recommended.

<u>Weighting according to response propensities</u>

An alternative is to use a regression-based approach. When many auxiliary variables are available, this approach is preferable to the previous one. Using an appropriate model such as logit regression, response propensities can be estimated as a function of auxiliary variables (X) which are available for both responding and non-responding cases:

$$ R_h = \Pr(r_h = 1 \mid X_h) , $$

where r_h is a (0,1) response indicator, equal to 1 if household h has been successfully enumerated; X_h is an appropriate set of auxiliary variables related to the households' response propensity, and R_h is that propensity predicted using regression.

The weighting of each responding unit is adjusted by the inverse of the estimated response propensity, in the same way as by the inverse of response rates in the previous method (Little, 1986). A very important point when using the regression approach is to ensure that weights assigned are confined to be within reasonable limits. The regression can predict zero or even negative values, which of course must be rejected. The problem is more general than that: extreme values should not be permitted. It is for this reason that it is very important to check the distribution of the resulting weight adjustments, and apply trimming or similar procedures to remove extreme values. This is discussed further in Section 3.5 below.

Non-response affects the enumeration of households and of all household members in exactly the same way. Hence the non-response adjustment to weights retains the basic equality in equation (6).

<u>Some practical aspects</u>

Generally it is useful to apply the adjustment in two steps:

(i) for non-contact (of households and/or of selected individuals); and

(ii) for non-response, once a contact with the households or the persons concerned has been made.

For both steps, especially for (i), area-level characteristics provide a main part of the auxiliary variables explaining non-response. This is because they are more easily available for both responding and non-responding units.

In dealing with the effect of non-response, it is of crucial importance to correctly distinguish between non-eligible and non-responding units. In fact, selected units which turn out to be non-eligible or non-existent must be excluded and not counted as non-respondents. Sometimes, imputation has to be made for units with unknown eligibility-response status, i.e. when it is not clear whether they are non-eligible or non-respondents.

For this purpose it is useful to distinguish the following categories:

1. Households that are out-of-scope: all individuals in these households are, of course, also out-of-scope.
2. Households that are successfully contacted and enumerated. These contain persons that are:
   - 2.a in scope;
   - 2.b out-of-scope;
   - 2.c persons for whom we do not know whether or not they are in-scope.
3. Households that are known to be in-scope, but are non-respondents. These households may contain persons who are actually out-of-scope, but this information is not available.
4. Households that are not contacted, and the eligibility status of the whole household is not known.

In order to impute a definite status for an individual in (2.c) or (3), we may assume that the propensity to be in-scope for these persons can be determined on the basis of the combined group (2.a+2.b), controlling for appropriate auxiliary variables. A logit regression model with appropriate control variables may be used. However, a simpler approach – for instance assigning individuals in (2.c) and (3) an eligibility status with probability equal to the proportion in that status in group (2.a+2.b), would suffice if the incidence of missing information is small.

Persons in group (4) are more likely to be out-of-scope than group (2) or (3) alone. It seems reasonable to assume that the status for group (4) can be imputed on the basis of *groups (1)-(3) combined*. A regression or a simple approach on the same lines as above may be used.[4]

[4] Some surveys allow substitution of non-responding units with new units (despite the general undesirability of this practice). In such cases, non-responding original units for which successful substitutions have been made are to be considered as 'responding units' in the computation of response rates for the purpose of determining non-response weights.

3.4. Calibration to external data sources

Calibration is carried out in an attempt to ensure that weighted sample sums of specified control variables or categories equal to the known population totals for those variables.

In so far as the first wave of a panel like EU-SILC is subject to high rates of non-response and possibly also to random and systematic distortions of the sample, we consider calibration to be particularly important in the first wave. Once the initial sample has been so adjusted, recalibration at subsequent waves may be done more selectively.

A robust and convenient method of adjusting the weighted sample distribution to a number of external controls simultaneously is the classical *iterative proportional fitting* or *raking* (Deming, 1943). Using special algorithms such as those implemented in the INSEE program CALMAR, upper and lower bounds can be imposed on the weight adjustment during calibration. (The limits, however, cannot be made too narrow as the iterative procedure involved becomes slow in converging, and may fail to converge altogether beyond a certain limit.) Deville and Särndal (1992) develop a family of calibration estimators of which the standard general regression estimator (GREG) is a first approximation. A problem with GREG and similar procedures is that they can yield negative weights, which of course does not make sense. In our experience the classical raking procedure, with simple trimming of any weight values outside desirable limits, continues to provide a practical and acceptable approach in many situations. (Concerning trimming, see Section 3.5 below).

Both household and person-level control variables are useful in calibration. Useful variables tend to be similar to those used for non-response adjustment, assuming availability: geographical location, tenure status, household type and size, age-sex composition of the population, etc. In some situations, additional variables may be available.

However, the crucial requirement in calibration is to ensure that the external control variables are strictly comparable to the corresponding survey variables, the distribution of which is being adjusted.

It is desirable that at wave 1, all persons in the same household receive the same weight, so that the weight given to each member is the same as the weight of the household, as in equation (6). Such uniform within-household weighting can be retained, even when the external controls in the calibration are at both the household and the personal level, by special technique known as 'integrative calibration'.[5]

3.5. Trimming and scaling

Trimming

It is important to ensure that no step in the weighting procedure results in extreme values of the weights. More precisely, what is required is to ensure that large variation in the weight values is not introduced as a result of the adjustments. This is because that variation inflates variance of the estimates from the survey (Kish, 1992). In fact, control of extreme values and large variations in weights is desirable at *each stage* of adjusting the weights - after non-response adjustment, and then again after calibration.

A common approach, as indicated earlier, is to trim extreme values.[6] However, there is no rigorous procedure for general use for determining the limits for trimming. While more sophisticated approaches are possible, it is desirable to have a simple and practical approach. Such an approach may be quite adequate for the purpose, at least in situations where the main problem is caused by a limited number of extreme values assigned during the adjustment process.

---

5 The procedure is described in, for instance, CALMAR documentation from INSEE (document no. F9310, November 1993). See also a brief description in Verma and Clemenceau (1996), Section 5.

6 For a more technical discussion of the issue, see Potter (1988, 1990).

After calculation of non-response weights, we have recommended and used the following simple procedure: any computed non-response weights outside the following limits are recoded to the boundary of these limits:

$$\frac{1}{L} \leq \left\{ \frac{{}_1w_h^{(H,N)}}{{}_1\overline{w}^{(H,N)}} \Big/ \frac{{}_1w_h^{(H,D)}}{{}_1\overline{w}^{(H,D)}} \right\} \leq L\,, \tag{10}$$

where ${}_1\overline{w}^{(H,D)}$, ${}_1\overline{w}^{(H,N)}$ are respectively the mean values of household design and non-response weights, and L is some appropriate upper bound for the adjustment in weights. L = 3 could be a reasonable value for this parameter.[7]

After calibration, we can follow the same form of check and correction for extreme values:

$$\frac{1}{L} \leq \left\{ \frac{{}_1w_h^{(H,F)}}{{}_1\overline{w}^{(H,F)}} \Big/ \frac{{}_1w_h^{(H,D)}}{{}_1\overline{w}^{(H,D)}} \right\} \leq L\,, \tag{11}$$

where each quantity is divided by its mean, so as to appropriately scale the weight values being compared.

Note that in both (10) and (11), the limits imposed are in terms of the ratio of the adjusted to the original design weights, i.e. the factor by which the design weights are being modified.

Any trimming or similar adjustment can be applied identically to household and person level weights, retaining the equality (6) for the trimmed weights, for instance:

$$ {}_1w_{j\in h}^{(R,F)} = {}_1w_h^{(H,F)}\,. $$

Scaling

As will be noted in Sections 5 and 6, the final longitudinal and cross-sectional weights actually used in data analysis may be scaled such that their sum is proportional to the size of the target population. Such scaling allows data from different countries and surveys to be put together for combined analysis without further adjustment to the weights (Verma, 1999).

However, at the present stage of the weighting procedure, we are considering one out of a number of panels which constitute the total sample at any given time. These panels are to be subsequently put together to constitute the full sample. Unlike the previous case, this is not an aggregation over different populations, but of samples representing the same population. In such aggregation, it is appropriate that each sample contributes inversely proportional to its variance, i.e., approximately in direct proportion to its sample size. Hence the weights should be scaled such that their sum is proportional to the *sample size* of the panel concerned – which is obtained most simply by scaling the weights to average 1.0 per unit. Such scaling allows data from different panels to be put together to construct the total sample without further adjustment to the weights.

Finally, we note that it is desirable to retain equality of the type in equation (6) also for the final rescaled weights for households and persons. To be precise, we need to choose whether to determine the required scaling factor with reference to the household or the person level sample. The latter is more appropriate in so far as the individual person is the more commonly used unit of analysis, as is the case for instance in the analysis of income distribution and poverty. Hence we define the required re-scaling factor as:

$$F_R = n_R \Big/ \sum_{j=1}^{n_R} {}_1w_j^{(R,F)}\,,$$

where n_R is the number of individuals enumerated in households of the panel sample at wave 1. The rescaled initial weights are:

7 Since trimming alters the mean value of the weights, the above adjustment may be applied iteratively, with the mean re-determined after each cycle. A very small number of cycles should suffice normally.

$$ {}_1w_j^{(R,I)} = {}_1w_j^{(R,F)}.F_R,\ {}_1w_h^{(H,I)} = {}_1w_h^{(H,F)}.F_R, $$

so that the following equality is maintained:

$$ {}_1w_{j\in h}^{(R,I)} = {}_1w_h^{(H,I)}. \tag{13} $$

It is not necessary to define these quantities for other subpopulations (P, Q, S) at this stage, but these can be taken simply as:

$$ {}_1w_j^{(P,I)} = {}_1w_j^{(Q,I)} = {}_1w_j^{(R,I)}, \tag{14} $$

$$ {}_1w_j^{(S,I)} = \left\{ \frac{{}_1w_j^{(R,I)}}{{}_1w_j^{(R,D)}} \right\} . {}_1w_j^{(S,D)}, \tag{15} $$

that is, for all the different type of units, the same modification (the first factor in the last equation) is applied to the original design weights.

4. Base weights

4.1. Longitudinal population and longitudinal sample

In the following we consider a panel selected fresh at time t = 1 from population P_1, and then enumerated for a total of T (say 4), waves, t = 1 to T. In this section we are primarily concerned with base weights (B) of the total population of persons (R) and for convenience will use the following simplified notation unless required otherwise (see Section 2.3):

$$ w_t = {}_tw_j^{(R,B)},\ j \in s_t^{(L)}, $$

where $s_t^{(L)} = s_{t-1}^{(L)} \cap s_t$, the longitudinal sample of persons from time 1 to t. We define the base weight at wave 1 as being identical to the initial weight defined in the previous section: $w_1 = {}_1w_j^{(R,B)} = {}_1w_j^{(R,I)}$.

From Section 3, these weights have been determined at wave 1 such that sample s_1 with weights w_1 *represents* population P_1, which we write as $(s_1, w_1) \longrightarrow P_1$.

Base weights are defined for persons but not for households. According to the following procedure, members of the same household may have different weights after wave 1.

By *longitudinal population* in the interval t = t_1 and t_2 is meant all persons who have remained in the target population throughout the period t_1 to t_2, inclusive. Let $P_t^{(L)}$ be the longitudinal population in the interval 1 to t. It comprises persons who were in the target population at wave 1, and have remained in the population up to and including time t. $P_2^{(L)}$ differs from P_1 by persons (say OUT_2) who have left the population between years 1 and 2: $P_2^{(L)} = P_1 - OUT_2$. This may for instance be due to death, migration out of the country or movement to an institution. Similarly, $P_3^{(L)}$ differs from P_1 by persons who have left between times 1 and 2 (OUT_2) or between times 2 and 3 (OUT_3):

$$ P_3^{(L)} = P_2^{(L)} - OUT_3 = P_1 - (OUT_2 + OUT_3). $$

In general, $P_T^{(L)} = P_1 - \sum_{t=2}^{T} OUT_t$. The longitudinal population $P_t^{(L)}$ differs from the actual (cross-sectional) population P_t at t, as the former does not include persons who are born or have migrated into the target population since time 1, and have remained in that population since that time.

Longitudinal sample $s_t^{(L)}$ is defined as individuals who have been members of an enumerated household throughout the period 1 to t inclusive. (The person is not necessarily a member of the same household throughout this period). $s_2^{(L)}$ differs from s_1 by persons in the original sample who left the population between years 1 and 2 (out_2), and by persons still in the population (x_2) whose household was enumerated at t = 1 but not at t = 2: $s_2^{(L)} = s_1 - (out_2 + x_2)$. We assume that (out_2, w_1) is a representative sample of OUT_2; consequently:

$$\left[(s_2^{(L)} + x_2), w_1\right] \longrightarrow (P_1 - OUT_2) = P_2^{(L)} . \qquad (16)$$

4.2. The evolution of base weights starting with wave 1

We assume that all persons enumerated at wave 1 are eligible for follow-up to the next wave. Those not successfully followed-up are considered non-respondents (unless re-classified as out-of-scope at the later time).[8]

The objective is to determine new base weights, w_t, at $t \geq 2$, such that:

$$\left[\left(s_t^{(L)} + x_t\right), w_{t-1}\right] \longrightarrow \left(P_{t-1}^{(L)} - OUT_t\right) = P_t^{(L)} ,$$

is transformed with the new weight w_t to:

$$(s_t^{(L)}, w_t) \longrightarrow \left(P_{t-1}^{(L)} - OUT_t\right) = P_t^{(L)} . \qquad (17)$$

In order to determine base weight w_t from known w_{t-1}, we use the following procedure. Consider the set $\left(s_{t-1}^{(L)} - out_t\right)$ of persons enumerated at (t-1) who are still in-scope at t. For each person (j) in this set, we can define a binary variable:

$r_j=1$ if the person is in $s_t^{(L)}$, i.e. is successfully enumerated at t,

$r_j=0$ otherwise, i.e. the person is not successfully enumerated at t.

Using a logit model, for instance, we can determine the response propensity R_j of each person in the above set as a function of a vector of auxiliary variables X_j: $R_j = \Pr(r_j = 1 | X_j)$.

For any person (j) in $s_t^{(L)}$ (i.e., with $r_j = 1$) the required weight is $w_t = \dfrac{w_{t-1}}{R_j}$.

In distinction from non-response adjustment at wave 1 (Section 3.3), the set of auxiliary variables (X) can be rich in content because of the information available on the non-enumerated persons from preceding waves.

In so far as most non-response occurs at the household (rather than the personal) level, a majority of the relevant auxiliary variables (X) will be geographical and household level variables (region, household size and type, tenure), including constructed variables such as household income and household work status.

Many person-level variables are also likely to be useful (gender, age, employment status) – the sort of variables correlated with persons moving to a new address, setting up a new household, remaining traceable, etc.

$R_j \leq 1$ by definition. It is necessary to ensure that no negative, zero, or indeed very small values of R_j are allowed. As before, the following practical trimming limit is suggested:

---

[8] The following applies concerning EU-SILC. There are certain (small) categories of households and individuals, which, according to EU-SILC rules, are not followed-up. Examples are households not enumerated at wave 1 or at two consecutive waves thereafter, or even not enumerated at a single wave for some specified reasons. Also, persons below a certain age (under 14, or under 16 in some countries) are not followed up if they move 'alone', i.e. without being accompanied by an adult sample person. For the present purpose, all these categories are treated as non-respondents – even if these have not been recorded as such in the survey because of the particular tracing rules.

$$\frac{1}{L} \le \frac{\left\{ {}_t w_j^{(R,B)} \Big/ {}_t \overline{w}^{(R,B)} \right\}}{\left\{ {}_1 w_j^{(R,D)} \Big/ {}_1 \overline{w}^{(R,D)} \right\}} \le L$$

, with L = 3, for instance. Here each quantity is divided by its mean, and (R,D) refers to the individual's original design weight.

4.3. Determining eligibility and response status of units

Application of the above procedure requires that for each person enumerated at (t-1), the person's eligibility and response status at t is known. This means that it is possible to classify each person in $s_{t-1}^{(L)}$ into one of the following categories uniquely: the person

(1) is enumerated at t (i.e., is in set $s_t^{(L)}$);

(2) remains in the population, but is not enumerated at t (is in set x_t);

(3) has moved out of the population (is in set out_t).

In practice, for a proportion of non-enumerated persons, information is not available to determine whether they belong to category (2) or to category (3). Each such person has to be assigned to one or the other of these two groups on the basis of some appropriate model. The procedure is similar to that in Section 3.3 for wave 1, but generally simpler because fewer ambiguous categories are involved.

4.4. Base weights for other categories of persons

In the above, base weights have been defined only for the longitudinal sample starting from wave t = 1, i.e., for all individuals enumerated in the survey throughout the interval 1 to t:

$$j \in s_t^{(L)} = s_1 \cap s_2 \cap ... \cap s_t = s_{t-1}^{(L)} \cap s_t \text{, with } s_1^{(L)} \equiv s_1 . \qquad (18)$$

There are two small additional categories of persons who can be assigned non-zero base weights on the basis of the base weights of these longitudinal individuals in the same household, without affecting the weights of the latter. These are:

(i) Children born to sample women. They receive the weight of the mother.

(ii) Persons moving into sample households from outside the survey population. They receive the average of base weights of existing sample members in the household, including (i).[9]

At any time, the sample households include additional categories of members who have so far been assigned zero weights. The main group among these are the *co-residents*, defined as persons moving into a sample household from another non-sample household in the population. These are given zero base weight. The same applies to children born to non-sample women.

In order to construct longitudinal and cross-sectional weights described in the following sections (which are the actual target variables of interest), procedures are required to assign non-zero weights to some or all of the zero-weighted persons, so that they can be included in analysis of the survey data.

5. Longitudinal weights

5.1. Set of samples requiring longitudinal weights

Consider the sample data becoming available after survey reference year Y in Figure 2. The total sample is made up four panels, selected in years Y (most recent), (Y-1), (Y-2) and (Y-3). For convenience, we will also identify these by reference to their current duration in years, respectively as (1), (2), (3) and (4).

---

[9] In determining the response propensities from one wave to the next as described in Section 4.2, these additional categories can be included in the sample base being followed-up.

With the exception of the most recently introduced panel (1), the other three panels contribute to newly available longitudinal data sets. As described in Section 2.3, by putting together these panels we obtain three longitudinal data sets (${}^{Y}_{L2}S$, ${}^{Y}_{L3}S$, ${}^{Y}_{L4}S$), respectively of duration 2, 3 and 4 years.[10]

In terms of panels (2)-(4), the composition of these samples is as follows. Different types of panel segments are involved:

- Panels starting from their time of selection (t = 1):

A2: a 2-year longitudinal sample of panel (2), covering years (Y-1) to Y;

A3: a 3-year longitudinal sample of panel (3), covering years (Y-2) to Y;

A4: a 4-year longitudinal sample of panel (4), covering years (Y-3) to Y.

- Panels which are included from a later time (t >1):

B2: a 2-year longitudinal sample from panel (3), covering years (Y-1) to Y;

B3: a 3-year longitudinal sample from panel (4), covering years (Y-2) to Y;

C2: a 2-year longitudinal sample from panel (4), covering years (Y-1) to Y.

The concerned longitudinal data sets becoming available at Y are (see equations (3)-(5) in Section 2.3):

$${}^{Y}_{L2}S = A2 + B2 + C2;$$

$${}^{Y}_{L3}S = A3 + B3;$$

$${}^{Y}_{L4}S = A4;$$

with $A2={}^{Y-1}_{1}s \cap {}^{Y-1}_{2}s = {}^{Y-1}_{2}s^{(L)}$; similarly $A3={}^{Y-2}_{3}s^{(L)}$, $A4={}^{Y-3}_{4}s^{(L)}$;

$$B2={}^{Y-2}_{2}s \cap {}^{Y-2}_{3}s\,;\; B3={}^{Y-3}_{2}s \cap {}^{Y-3}_{3}s \cap {}^{Y-3}_{4}s\,;$$

$$C2={}^{Y-3}_{3}s \cap {}^{Y-3}_{4}s\,.$$

5.2. Longitudinal weights as distinct from base weights

The longitudinal weights discussed in this section are the actual weights used in longitudinal analysis based on the total available longitudinal sample. For a period t_0 to t, $1 \le t_0 < t$, these are defined for all units (from all the available panels) who have been enumerated in the survey throughout this period.

As noted earlier, base weights are for the longitudinal sample of each panel, starting from time t = 1 when the panel is first introduced into the survey. In the special case of $t_0 = 1$, there is essentially no difference between these base and the required longitudinal weights, except that the former are defined for each panel while different panels need to be put together for the latter.

With $t_0 > 1$, the longitudinal sample over the period t_0 to t includes certain additional categories of individuals not present in the longitudinal sample of original sample persons over the full period 1 to t, to which non-zero weights have to be assigned. The differences from cases (i)-(ii) of Section 4.4 is that assigning non-zero weights to these additional categories will require *adjusting the weights of other members in the sample*. The additional categories include:

(iii) *longitudinal co-residents*, entering the household at or before t_0 and remaining as household members over the period (t_0 to t);

[10] There are also other sequences of longitudinal data embedded in the data set shown in the diagram: a 3-year longitudinal sample from (Y-3) to (Y-1) in panel (4); and three 2-year samples, consisting of (Y-3)+(Y-2) and (Y-2)+(Y-1) in panel (4), and (Y-2)+(Y-1) in panels (3). These panels correspond to data which have been produced in previous years.

(iv) *returnees*, defined as sample persons who left the household temporarily but later returned to the household at or before t_0 and remained a household member over the period (t_0 to t). [11]

Adjustment has also to be made for persons entering the target population before time t_0 but subsequent to the time of selection of the panel (t = 1). Of course, these persons cannot be represented in the panel concerned, but they are a part of the longitudinal population over the period (t_0 to t_1) and therefore should be represented in the sample.

As will be described below, this can be done on the basis of other (later introduced) panels in the rotational design being put together to construct the total longitudinal sample (Ardilly and Lavallée, 2003).

Among those entering the target population, we will need to distinguish between those joining existing households (containing persons from the existing population), and those forming new households. Finally, we may mention that the longitudinal weights also need to take into account the 'non-standard' structure of the total sample in the first years of the survey (the first 3 years in a 4-year panel, as shown in Figure 2).

5.3. Longitudinal weights by age of panel

Our main objective is to determine longitudinal weights for units (of the total population R) in the six sets S = (A2, A3, A4, B2, B3, C2) for a certain reference year Y, as defined above.

<u>Sets A2, A3, A4</u>

In all cases of type A, the longitudinal population and the longitudinal sample are, for the panels concerned, exactly the same as those considered in the construction of the base weights in Section 4. For example, for A2, using the terminology of Section 4:

Longitudinal population $P_1 - OUT_2$

Longitudinal sample $s_1 - (out_2 + x_2)$.

Hence for any (j) in sets A2-A4, the required longitudinal weights involved are identical to the base weights defined earlier. Using simplified notation on the left for convenience we may write this, respectively for the three sets, as:

$$^{Y}v_j = {}^{Y-1}_{2}w_j^{(R,B)},\ j \in A2 = {}^{Y-1}_{2}s^{(L)},$$

$$^{Y}v_j = {}^{Y-2}_{3}w_j^{(R,B)},\ j \in A3 = {}^{Y-2}_{3}s^{(L)},$$

$$^{Y}v_j = {}^{Y-3}_{4}w_j^{(R,B)},\ j \in A4 = {}^{Y-3}_{4}s^{(L)}.$$

<u>Set B2</u>

Now consider set B2, i.e. a longitudinal sample covering the second (t = 2) and third (t = 3) year of panel (3). The population at t = 2 may be written as:

$P_2 = P_1 - OUT_2 + BORN_2 + IN_2^{(old)} + IN_2^{(new)}$, where, between time t = 1 and t = 2, OUT_2 are persons who have left the target population, $BORN_2$ are children born to women in the population, $IN_2^{(old)}$ are persons who have moved into the population (from abroad or the non-household sector), but into existing households; $IN_2^{(new)}$ are persons who have moved into the population to set-up new households consisting only of such in-migrants. The sample at t = 2 may be written as $s_2 = s_1 - (out_2 + x_2) + born_2 + in_2^{(old)} + co_2$.

Quantities (out_2, $born_2$, $in_2^{(old)}$) are defined in the same way as the corresponding population quantities. There is no term in the sample corresponding to the population term

[11] Strictly speaking, to be considered a returnee, the person leaving the household for a period must still have remained within the target population. Otherwise, on departure the person is treated in the weighting process as an out-migrant and on return as an in-migrant into the population.

$IN_2^{(new)}$, since entrants to the population after the sample selection who set-up entirely new households are not represented in the sample. On the other hand, there is no representation in the population for the sample term co_2. It represents co-residents, defined as persons who were already members of the population, and have simply moved from a non-sample private household to a sample household (i.e. to a household containing at least one sample person). Re-arranging, we can write:

$$\left(P_2 - IN_2^{(new)}\right) = P_2^{(L)} + \left(BORN_2 + IN_2^{(old)}\right),$$
$$s_2 = s_2^{(L)} + \left(born_2 + in_2^{(old)}\right) + co_2 .$$

Here, the first population term on the right, $P_2^{(L)} = \left(P_1 - OUT_2\right)$, is the longitudinal population considered earlier in the determination of base weights. The first sample term on the right, $s_2^{(L)} = \left(s_1 - out_2 - x_2\right)$, is the longitudinal sample considered earlier; it estimates the longitudinal population with the base weights.

Extending the concept of base weights to cover additional categories of persons as noted earlier:

(i) All children, $born_2$, born to sample mothers during the preceding year, are assigned the base weight of the mother; with these weights, they represent $BORN_2$.

(ii) The mean of the base weights of persons $\left(s_2^{(L)} + born_2\right)$ in the household is assigned to each person in the group $in_2^{(old)}$ in that household; with these weights they represent $IN_2^{(old)}$.

(iii) Co-residents, co_2, continue to be assigned a zero base weight.

This means that with these weights s_2, the total sample at t = 2, represents the population P_2 at t = 2, excepts for entrants $IN_2^{(new)}$ who have formed entirely new households without including any member from the existing population. We represent this as:

$$\left(s_2, {}_2^{(3)}w\right) \longrightarrow \left(P_2 - IN_2^{(new)}\right)$$

where ${}_2^{(3)}w = {}^{Y-2}_{\;\;2}w_j^{(R,B)}$, $j \in {}^{Y-2}_{\;\;2}s$, meaning the base weight at age t = 2 for panel (3), which was introduced into the survey in year (Y-2). Now, using exactly the same logic as used earlier in the construction of the base weights, it follows that:

$$\left((s_2 - out_3 - x_3), {}_3^{(3)}w\right) \longrightarrow \left((P_2 - OUT_3) - IN_2^{new}\right).$$

Since all co-residents are zero weighted, we can re-write the above as:

$$\left[\left\{s_2 - (out_3 + x_3 + co_2) + co_{23}\right\}, {}_3^{(3)}w\right] \longrightarrow \left((P_2 - OUT_3) - IN_2^{new}\right), \quad (19)$$

where co_2 are the co-residents at t = 2, and among those co_{23} are also present in the household at t = 3. The left side is the longitudinal sample between times t = 2 and t = 3, i.e. individuals (irrespective of whether or not they are from the original, t = 1, sample) who are present at both waves. The first term on the right, $(P_2 - OUT_3)$, is the corresponding longitudinal population, i.e. individuals who are members of the population at both times. With the already computed base weight ${}_3^{(3)}w$, the sample quantities on the left estimate the population values, except for persons $IN_2^{(new)}$ who entered and formed entirely new households between the sample selection and t = 2, and are not represented in the sample.

<u>Set B3</u>

For B3 – a longitudinal sample covering three years (t = 2 to t = 4) of a panel, the procedure is essentially the same, and we obtain the expression:

$$\begin{aligned} &\left[\left\{s_2 - (out_3 + out_4 + x_3 + x_4 + co_2) + co_{234}\right\}, {}_4^{(4)}w\right] \\ &\longrightarrow \left[P_2 - (OUT_3 + OUT_4)\right] - IN_2^{new} \end{aligned} \quad (20)$$

where co_{234} are co-residents present in the household at t = 2 to 4. Further adjustment is still required to B2 and B3 weights in order to incorporate longitudinal co-residents (co_{23} and co_{234} respectively) with non-zero weights.[12]

5.4. Adjustment for returnees and longitudinal co-residents

Set C2; returnees

Now we consider C2, i.e. a longitudinal sample covering the third and fourth years (t = 3, 4) of panel (4). The additional complication in this case is the following.

Sample s_3 at t = 3 may contain returnees to the sample: sample persons at t = 1 who were non-respondents at t = 2, but were again enumerated at t = 3. If they also continue to be present at t = 4, they need to be included in the longitudinal sample being considered (since, as defined, they belong to the corresponding longitudinal population).

On return at t = 3 after the absence, these persons receive a zero base weight. These units need to be given a positive weight if they are to be included in the analysis. With the procedure described below, we can assign a non-zero weight to returnees not present in year (2) but again present in year (3). This requires, in compensation, a (slight) deflation of the base weights of the units in the longitudinal sample, from ${}^{(4)}_{3}w$ to say ${}^{(4)}_{3}\tilde{w}$, using the procedure indicated below. The longitudinal sample of interest is not affected by any returnees between t = 3 and 4, so that the above can be used to estimate the revised weights:

$$ {}^{(4)}_{4}\tilde{w} = {}^{(4)}_{4}w.\left\{ \frac{{}^{(4)}_{3}\tilde{w}}{{}^{(4)}_{3}w} \right\}. \tag{21} $$

With this modification we can write the required expression similar to (19). The procedure is a little more involved algebraically, but the result is in the same form:

$$ \left[\left\{ s_3 - (out_4 + x_4 + co_3) + co_{34} \right\}, {}^{(4)}_{4}\tilde{w} \right] \longrightarrow (P_3 - OUT_4) - \left[IN_2^{new} + IN_3^{new} \right]. \tag{22} $$

The left side is the longitudinal sample between times t = 3 and t = 4, i.e. individuals who are present at both waves, irrespective of whether they are from the original sample (t = 1) or were present at t = 2.

The first term on the right is the longitudinal population between t = 3 and t = 4. It is represented by the corresponding longitudinal sample with weights ${}^{(4)}_{4}\tilde{w}$, except for entrants IN_2^{new} and IN_3^{new} forming entirely new households during the two years preceding t = 3.

Procedure for weight adjustment to incorporate returnees

Consider a longitudinal population N and the corresponding longitudinal sample n over a three-year period (1)-(3). For simplicity, let us only consider persons who remain members of N throughout the period. Let n_1, n_2 and $n_3 = n + r$ be the achieved samples in years 1-3, where r are the returnees, i.e. persons in n_1 who were not enumerated in n_2 at year (2), but were enumerated in n_3 at year 3. Starting with weights w_1 at t = 1, base weights for the longitudinal sample can be constructed as described in Section 4.2. We may write the base weights w_2 at t = 2 as:

$$ w_2 = w_1 / P(n_2 \mid n_1), $$

where the denominator stands for the propensity of a sample unit at t = 1 to remain in the sample at t = 2. Similarly:

$$ w_3 = w_2 / P(n \mid n_2). $$

Using a similar procedure, we can estimate the propensities of units in n_3 to be present in n. The adjustment (deflation) of the base weights of units in n allows for the incorporation of the returnees $r = (n_3 - n)$ into the sample with non-zero weights. The adjusted weights are:

[12] Note that in the common simplified notation used in equations (19) and (20), the various sample quantities on the left refer to different panels – to panel (3) in the former, and to panel (4) in the latter.

$$\tilde{w}_3 = w_3 P(n \mid n_3) = \frac{w_1 P(n \mid n_3)}{P(n_2 \mid n_1) P(n \mid n_2)}, \text{ for units in longitudinal sample n,} \quad (23)$$

$$\tilde{w}_3 = \frac{w_1}{P(n_3 \mid n_1)}, \text{ for returnees } r = (n_3 - n). \quad (24)$$

Incorporating longitudinal co-residents

As noted, it is desirable to incorporate longitudinal co-residents into longitudinal analysis. By these we mean co-residents present throughout the longitudinal interval of interest. Since, henceforth, such persons are zero weighted, they need to be assigned a non-zero longitudinal weight. The weight-share procedure described in Section 6.3 can be used for this purpose.

No modification need to be made to the weights of members in a household which does not contain any longitudinal co-residents. Otherwise, in a household containing one or more longitudinal co-residents, we compute the mean of longitudinal weights as the sum of base weights of longitudinal sample persons divided by the number of all longitudinal persons in the household, including longitudinal co-residents. This uniform weight is then assigned to all longitudinal household members, including longitudinal co-residents.

5.5. Construction of the target variables

We return to the three longitudinal data sets ${}^{Y}_{L2}S$, ${}^{Y}_{L3}S$ and ${}^{Y}_{L4}S$ for which the sample weights are required. The construction of those weights simply involves putting together the constituent panels. The components and the weight variables for the data sets are shown in Table 1. Generally, new entrants are represented in some but not all the constituent panels. To compensate for the missing ones, the weight of the included new elements can be appropriately inflated as follows. Let $W_{(A2)}$ be the sum of weights $\sum_{j \in A2} {}^{(2)}_{2}w_j$ excluding entrants $IN^{(new)}_{Y-2}$ and $IN^{(new)}_{Y-1}$. We define similar quantities for other components, e.g. $W_{(C2)}$ as the total of ${}^{(4)}_{4}\tilde{w}_j$, $j \in C2$.

Table 1. Coverage of new entrants in the panels comprising the longitudinal samples

Panel: year introduced (1)	Panel duration at Y (2)	Units (3)	Base weights* (4)	Whether entrants represented $IN^{(new)}_{Y-2}$ (5)	$IN^{(new)}_{Y-1}$ (6)	$IN^{(new)}_{Y}$ (7)
Longitudinal data-set of 2 year duration becoming available after year Y: ${}^{Y}_{L2}S$ = A2 + B2 + C2						
Y-1	(2)	$j \in A2$	${}^{(2)}_{2}w$	√	√	-
Y-2	(3)	$j \in B2$	${}^{(3)}_{3}w$	√	x	-
Y-3	(4)	$j \in C2$	${}^{(4)}_{4}\tilde{w}$	x	x	-
Longitudinal data-set of 3 year duration becoming available after year Y: ${}^{Y}_{L3}S$ = A3 + B3						
Y-2	(3)	$j \in A3$	${}^{(3)}_{3}w$	√	-	-
Y-3	(4)	$j \in B3$	${}^{(4)}_{4}w$	x	-	-
Longitudinal data-set of 4 year duration becoming available after year Y: ${}^{Y}_{L4}S$ = A4						
Y-3	(4)	$j \in A4$	${}^{(4)}_{4}w$	-	-	-
√ = YES, x = NO, - not relevant						

* for panel (4), modified base weights are defined in the text.

Since population $IN_{Y-2}^{(new)}$ is represented only in panels (2) and (3), we inflate the weight of each element in category $in_{Y-2}^{(new)}$ by the factor:

$$\frac{W_{(A2)} + W_{(B2)} + W_{(C2)}}{W_{(A2)} + W_{(B2)}} \tag{25}$$

to compensate for its absence in component C2. With equal weighted panel sizes, this factor becomes 3/2. In accordance with Section 3.5, the weights in each panel are assumed scaled to average 1.0, so that the sum of weights is equal to the (unweighted) sample size.

Similarly, the weight of each element in category $in_{Y-1}^{(new)}$ is multiplied by the factor:

$$\frac{W_{(A2)} + W_{(B2)} + W_{(C2)}}{W_{(A2)}}.$$

With equal weighted panel sizes, this factor equals 3.

With these modifications for entrants, weights in col. (4) are the required longitudinal weight variables, expect for the following final adjustments.

(1) Scaling the weights to average 1.0 for each panel separately automatically ensures that the contribution of each panel is proportional to its sample size. Also, the average weight remains 1.0 or close to it when the panels are merged to form the full sample. However, sometimes weights are scaled to sum to the populations size (or to equal inverse of the selection probabilities). In this case the weights in the individual panels have to be proportionally reduced to obtain the full sample weights. Ideally, before putting the panels together, these weights should be re-scaled such that for each panel the sum of weights is proportional to the panel sample size.

(2) For multi-country analysis using combined data, it is desirable to scale the weights such that their sum is proportional to the longitudinal population size of the country concerned. When (as is often the case), size of the longitudinal population is not available, the sum of weights may be scaled to be proportional to the population at a certain time such as Y.

(3) In accordance with Section 3.4, each panel is assumed calibrated against external controls at the time of its entry into the sample. This is the most important application of calibration. Additional or repeated calibration at subsequent times is often not worthwhile. In any case, in many situations information on the *longitudinal* population is not available for the purpose.

(4) For units in any subpopulation such as children (Q), we can take the longitudinal weight to be identical to the units' sample weight in the total population (R): $v_j^{(Q)} = v_j^{(R)}$.

(5) After successful enumeration of the population (R) within households, additional non-response may be involved in the subsequent personal interview survey with adults (P). Adjustment to the weights is required to compensate for this *within-household non-response*. A simple approach is to calibrate the achieved P-sample to match the corresponding R-sample over a set of important characteristics. The raking procedure described in Section 3.3 can, for instance, be used for the purpose. We begin with the above determined R-weights applied to the achieved P-sample. Then the weights are adjusted (to give the required P-weights) such that P-sample, with these adjusted weights, agrees with R-sample (with R-weights) on the specified control characteristics. The above calibration corrects for overall distortions in the P-sample due to within-household non-response. Further correction needs to be applied to the data obtained for the particular households affected by within-household non-response, as described in Section 7.4.

(6) In countries where P-data are obtained from registers there is generally no within-household non-response of the above type. In such cases, we can take the longitudinal P-weight of a unit (in the P-sample) to be identical to the units' weight in the R-sample: $v_j^{(P,L)} = v_j^{(R,L)}$. In these countries personal interview is conducted over a sample of

selected respondents (S), normally one adult per sample household. Here the S-weights firstly need to take into account the differences in the design weights between the two samples:

$$v_j^{(S)} = \left[\frac{{}_1 w_j^{(S,D)}}{{}_1 w_j^{(R,D)}} \right] . v_j^{(R)} . \tag{26}$$

Next, to control for the effect of sampling within households, we may calibrate the S-sample to match the corresponding R-sample over a set of characteristics, using for instance the raking procedure of Section 3.3. For this purpose, we can begin with the above defined S-weights applied to the S-sample, and adjust these weights to achieve agreement with the full R-sample (weighted by R-weights) on specified control characteristics.

(7) As to whether such calibration is also required to compensate for S-sample attrition depends on the follow-up rules and the field procedures followed (see Section 7.3 concerning the former). Two types of procedures can be envisaged:

(a) One option is to determine the follow-up of the R-sample entirely on the basis of the achieved S-sample – i.e. follow-up household and the associated household members only for successfully interviewed selected respondents (S-sample). In this case, there is no relative non-response between the two samples, and weights derived in (6) require no further adjustment.

(b) In our view, the above procedure is unnecessarily damaging to the quality of the income and other data collected on the R-sample from registers and other sources not themselves subject to non-response. This arises from imposing on the R-sample the interview non-response to which the S-sample is subject. The alternative is to follow-up the full R-sample independently of the outcome of the S-sample. Under this scenario, one should construct the whole set of base, longitudinal and cross-sectional weights described in Sections 4-6 for the S-sample in its own right.

5.6. Start-up of the integrated design

Figure 2 also shows how the integrated design may be started up. For a 4 years panel design, we may begin with 4 new panels in year 1, drop one of these and introduce a new one in year 2. The whole longitudinal sample, consisting of three panels, is of type A2, and the required longitudinal weights are given directly by the base weights as described in Section 5.3.

In year 3, we drop another of the original panels and introduce a new one. The longitudinal sample of 2-year duration consists of one panel of type A2 and two panels of type B2, say B2(1) and B2(2). Part $IN_{Y-2}^{(new)}$ is represented in all these panels and no reweighting is required. The reweighting for part $IN_{Y-1}^{(new)}$ is similar to that in the 'normal' case discussed in Section 5.5. The weight of any unit in this part is multiplied by:

$$\frac{W_{(A2)} + W_{(B2(1))} + W_{(B2(2))}}{W_{(A2)}} . \tag{27}$$

From year 4 onwards the sample structure becomes normal for the integrated design.

6. Cross-sectional weights

6.1. Special aspects of cross-sectional weighting

Some special features of cross-sectional weights include the following.

- In the case of EU-SILC at least, the first objective is to obtain good quality cross-sectional estimates. This enhances the importance of constructing good cross-sectional weights: by 'good' we mean weights chosen to reduce mean-square-error.
- For the sake of efficiency, cross-sectional analysis should use all available sample cases at the time concerned. This requires that all units enumerated at that time, irrespective of whether or not they are original sample persons, be given non-zero positive weights.
- It is more convenient and consistent to assign a uniform cross-sectional weight to all current members of a household, the same as the weight of the household, rather than to allow the weights to vary across persons in the same household.
- It is both more desirable and more feasible to calibrate the sample weights of the cross-sectional sample each wave on external control distributions on relevant population characteristics, compared to the longitudinal samples. Much more information is usually available for cross-sectional calibration than for longitudinal calibration.

In the rotational integrated design, cross-sectional weights can be constructed from the base weights (Section 4) in a straightforward way, on lines similar to the construction of longitudinal weights in Section 5. Many of the details of the procedure are similar for the two types of weighting, and this section can therefore be brief. The full cross-sectional sample is constructed by putting together results from different panels for the same year (Y), as shown in Table 2 and illustrated in Figure 2.

Table 2. Coverage of new entrants in the panels comprising the cross-sectional sample

Panel: year introduced	Panel duration at Y	Units	Base weights*	Whether entrants represented $IN_{Y-2}^{(new)}$	$IN_{Y-1}^{(new)}$	$IN_{Y}^{(new)}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Y	(1)	$j \in {}_1^Y s$	${}_1^{(1)}w$	√	√	√
Y-1	(2)	$j \in {}_2^{Y-1} s$	${}_2^{(2)}w$	√	√	x
Y-2	(3)	$j \in {}_3^{Y-2} s$	${}_3^{(3)}\tilde{w}$	√	x	x
Y-3	(4)	$j \in {}_4^{Y-3} s$	${}_4^{(4)}\tilde{\tilde{w}}$	x	x	x
√ = YES, x = NO						

* for panels (3) and (4), modified base weights are defined in the text.

6.2. Adjustment of the base weights

Base weights form the basis of the required cross-sectional weights. With extension of the base weights to additional categories of persons (Sections 4.4 and 5.2), further adjustment is still required to the incorporation of returnees into the sample with non-zero weights. With the procedure described in Section 5.4, the weights can be further adjusted for returnees as follows.

We have already computed the adjusted weights ${}_3^{(4)}\tilde{w}$, referring to panel (4) in its 3rd year. Exactly in the same way, we can estimate ${}_3^{(3)}\tilde{w}$, which refers to the 3rd year of panel (3). For panel (4), it is necessary to adjust the weights for (i) returnees between its 2nd and 3rd years, and then further for (ii) returnees between its 3rd and 4th years. An estimate of (i) is as before (equation (21)):

$$ {}^{(4)}_{4}\tilde{\tilde{w}} = {}^{(4)}_{4}w.\left\{ \frac{{}^{(4)}_{3}\tilde{w}}{{}^{(4)}_{3}w} \right\}. $$

For the further adjustment (ii), the procedure is the same as that in Section 5.4. We estimate the propensity P of individuals in ${}^{(4)}_{4}s$ at wave 4 to also have been enumerated at wave 3 as $P = \left({}^{(4)}_{3}s \cap {}^{(4)}_{4}s\right) / \left({}^{(4)}_{4}s\right)$, giving ${}^{(4)}_{4}\tilde{\tilde{w}} = P.\left({}^{(4)}_{4}\tilde{w}\right)$.

The above are the modified base weights for persons other than returnees. For the returnees in the 4th year who were present in year 2 but not in year 3, we may use the slightly approximate expression[13]:

$$ P = \left({}^{(4)}_{2}s \cap {}^{(4)}_{4}s\right) / \left({}^{(4)}_{2}s\right), \text{ giving } {}^{(4)}_{4}\tilde{\tilde{w}} = {}^{(4)}_{2}w / P. $$

6.3. Construction of cross-sectional weights

Coverage of in-migrants

New entrants into the population during the three years preceding Y are represented only in some of the constituent panels. To compensate for the missing ones, base weights of new entrants in the panel(s) where they have been included are inflated as follows.

Let $W_{(S1)}$ be the sum of base weights $\sum\left\{{}^{(1)}_{1}w_j\right\}$, $j \in {}^{1}_{1}s$, similarly $W_{(S2)} = \sum\left\{{}^{(2)}_{2}w_j\right\}$, $W_{(S3)} = \sum\left\{{}^{(3)}_{3}\tilde{w}_j\right\}$, $W_{(S4)} = \sum\left\{{}^{(4)}_{4}\tilde{\tilde{w}}_j\right\}$, in all cases excluding new entrants $in^{(new)}$.

Since population $IN^{(new)}_{Y-2}$ is not represented in panel (4), we inflate the weight of each element in $in^{(new)}_{Y-2}$ category in every of the panel (1)-(3) by the factor:

$$ \frac{W_{(S1)} + W_{(S2)} + W_{(S3)} + W_{(S4)}}{W_{(S1)} + W_{(S2)} + W_{(S3)}}. \tag{28} $$

When the weighted panel sizes are equal, this factor equals 4/3.[14]

Similarly, weights of present units in category $in^{(new)}_{Y-1}$ are multiplied by the factor:

$$ \frac{W_{(S1)} + W_{(S2)} + W_{(S3)} + W_{(S4)}}{W_{(S1)} + W_{(S2)}}, \tag{29} $$

which equals 2 when the weighted panel sizes are equal.

For units in category $in^{(new)}_{Y}$, the unit weights are inflated by the factor:

$$ \frac{W_{(S1)} + W_{(S2)} + W_{(S3)} + W_{(S4)}}{W_{(S1)}}, \tag{30} $$

which equals 4 when the weighted panel sizes are equal.

Weight sharing

Within a household, $j \in h$, each member has been assigned a weight w_j, except for co-residents for whom w_j = 0. The average of these weights over all household members (including co-residents) is assigned to each member (including co-residents). We recommend to *apply this averaging process to all households, including households not containing a co-resident*. This procedure ensures that non-zero weights are assigned to all persons enumerated at the cross-section concerned, and that within each household, all individuals enumerated have

[13] We have assumed that individuals missed in only one wave may return to the sample; those missed in two or more waves consecutively are not allowed to return to the sample. This in fact is in line with the follow-up rules adopted in EU-SILC. The procedure, of course, can be extended in a straightforward way to remove the above limitation.

[14] As was noted earlier, in putting different panels together, it is best to scale the weights in each panel separately to average 1.0. In this case the sum of weights is simply the panel sample size.

the same weight. The result is to assign the same uniform weight to all current members of a household. The sum of base weights (after any adjustment as above) remains unchanged for the household: this sum is simply shared among all current members.

In brief outline the weight-sharing procedure is as follows (Ernst, 1989). If u_k is a random variable associated with each unit k in the population of size N, then $\hat{X} = \sum_{k=1}^{N} x_k . u_k$ provides an unbiased estimator of the population total $X = \sum_{k=1}^{N} x_k$, provided $E(u_k) = 1$. In the conventional Horvitz-Thompson estimator, we take u_k as inverse of the selection probability of the sample unit, and take $u_k = 0$ for all units not in the sample. The weight-share approach is based on the observation that an unbiased estimate is also produced by defining a new set of weights w_j for all units (j) in the population in the following form:

$$w_j = \sum_{k=1}^{N} a_{jk} . u_k \text{, with } \sum_{k=1}^{N} a_{jk} = 1,$$

where the constants a_{jk} are independent of u_k, but are otherwise determined according to freely chosen rules. This is because with the above definitions we have $E(w_j) = \sum_{k=1}^{N} a_{jk} . E(u_k) = \sum_{k=1}^{N} a_{jk} = 1$, thus satisfying the above requirement for the estimator to be unbiased. This form involves redefining the weight of each unit (j) in the population as a (weighted) average of the weights u_k for units in a certain set determined according to some specified rules. For instance, for each individual (j) in household h in the population, the set may refer to all members in the individual's household. The factors a_{jk} corresponding to each of the individual's household member may be taken as equal to $1/n_h$ (where n_h is the household size), and as equal to 0 corresponding to all persons in the population who are not members of (j)'s household. With these definitions the redefined weights w_j simply become the average of the original weights of individuals in the set (household) of (j).

Calibration

As noted, there is a greater need and possibility of calibration on external control distributions on relevant population characteristics in the cross-sectional compared to a longitudinal context. As a rule, it is recommended to calibrate the cross-sectional sample weights each wave. Again, integrative calibration will ensure that uniformity of the person weights within each household is retained. Calibration should normally be applied to the whole, pooled sample; separate calibration of individual panels is generally not necessary and often not even desirable.

Trimming

Following any operation like calibration, the resulting weights should always be checked for extreme values, followed by trimming or other appropriate treatment as necessary.

Scaling

The basic principles noted in Section 5.5 apply. Scaling the weights to average 1.0 for each panel separately, automatically ensures that the contribution of each panel is proportional to its sample size. For pooling across countries or other domains, it is desirable to scale the weights such that their sum is proportional to the population size.

Weights for other types of units

With weight sharing (and integrative calibration, uniform scaling etc., as applicable), the equality of individual (R) and household (H) weights can be retained. For units in any subpopulation such as children (Q), we can also take the cross-sectional weight to be identical to the units' cross-sectional weight in the total population (R): $v^{(Q)}_{j\in h} = v^{(R)}_{j\in h} = v^{(H)}_{h}$.[15]

For data from the personal interview survey with adults (P), adjustment is required to compensate for any additional non-response at that stage. The simple approach is to calibrate the achieved P-sample to match the corresponding R-sample over a set of important characteristics, described in Section 5.5. Further correction needs to be applied to the data obtained for the particular households affected by within-household non-response, as noted in Section 7.4.

In countries where the personal interview is conducted over a sample of selected respondents (S), the S-weights firstly need to take into account the differences in the design weights. Next, to control for the effect of sampling within households, we may calibrate the S-sample to match the corresponding R-sample over a set of characteristics (Section 3.3; see also Section 7.3).

6.4. Start-up of the integrated design

We refer again to the start-up of the integrated design in Figure 2. All panels in year 1 are freshly introduced. The required cross-sectional weights are given directly by the initial or base weights for year 1. Panels can be put together to construct the full cross-sectional sample, without further rescaling if the weights of each panel have been scaled to average 1.0.

In year 2, category $IN^{(new)}_{2}$ is represented in only one of the four panels (say with sum $W_{(S1)}$), and base weights of the cases belonging to it have to be inflated as in equation (30). No other adjustment is required.

In year 3, category $IN^{(new)}_{3}$ is represented in only one of the four panels, and $IN^{(new)}_{2}$ in two of the four. The weights of sample units in these categories have to be inflated, respectively, by expressions similar to equations (30) and (29). As in the longitudinal case, from year 4 onwards the cross-sectional sample structure becomes normal for the integrated design.

7. Some variations and practical aspects

In this concluding section, we note some further details and practical issues.

7.1. Information on migration prior to entering the survey

A precise application of longitudinal and cross-sectional weighting procedures (Sections 5 and 6) requires the identification of in-migrants into the target population of private households in the country (from abroad or from the institutional sector).

Information is also required as to whether an in-migrant formed a 'new' household, or moved into an already existing ('old') household containing a person from the existing population. Information on migration status is also required for co-residents who move into sample households at subsequent waves. The above requires the following items to be specified.

(1) For *all persons* enumerated in the household at the first wave: whether they entered the target population within the past 3 years, and if so when.

[15] It has been suggested that, for special purposes, the weights for certain subpopulations (Q) may be subject to further, more precise calibration. See, for instance, Eurostat (2004) in relation to child-care data.

(2) At each subsequent wave, for each new co-resident (a non-sample person moving into the household that wave): whether they have moved from another household in the target population, or from outside that population (another country or an institution).[16]

Let $in^{(new)}_{Y-t+1} = 1$ indicate an in-immigrant during the interval (Y-t) and (Y-t+1) into a 'new' household as defined above (this indicator is equal to 0 for all other persons). From Section 6, it can be seen that for the panel introduced at Y, this indicator is required for each person for the preceding three years (Y-t+1), t = 1, 2, 3. This can constructed from question (1) above as follows.

- If any current member of the household entering the survey has been a member of the target population for $t' > 3$ years, then the indicator $in^{(new)}_{Y-t+1} = 0$ for all members, all t = 1, 2, 3.
- Otherwise, let $t' \leq 3$ identify the interval for the set of current household members who are the earliest to have moved into the target population. Then $in^{(new)}_{Y-t'+1} = 1$ for each person in the above-mentioned set. For other $t \neq t'$ for this set, and for all t for any remaining member, this indicator equals 0.

Question (2) identifies co-residents entering the sample household from outside the target population at any time from wave 2 onwards. For persons who have moved into the household from outside the target population, index $in^{(old)} = 1$ for the waves they remain in the sample. As noted earlier, these persons are assigned a non-zero base weight (Section 5.3). By contrast, co-residents moving in from within the target population are treated differently. They originally have zero base weight. The longitudinal co-residents among them receive non-zero longitudinal weights, and all of them receive non-zero cross sectional weights, on the basis of weight sharing (Sections 5.4, 6.3).

7.2. Variations from the integrated design

A vast proportion of the countries conducting EU-SILC have used the integrated design described in Section 1.4. The main departures arise from the need to combine EU-SILC with an existing survey or sample (Finland, Germany, Norway). France uses the same structure as the integrated design, but a panel duration of 9 (in place of 4) years. (The panel duration is also longer – 8 years – in Norway). Luxembourg uses a panel of indefinite duration similar to the pure panel ECHP design, but complements this with annual samples to compensate for panel attrition (Clemenceau *et al.*, 2006).

When the design departs from the standard, the weighted procedures described in this paper of course require adaptation. However, the basic approach, and many details as well, remain valid and useful for developing alternative procedures. When panels of long or indefinite duration are used, special care is needed to retain cross-sectional representativity of the sample. More careful weighting for panel attrition, and possibly also periodic supplementation of the sample, become important.

7.3. On follow-up rules in register countries

The analysis of income distribution, inequality and poverty – which is the primary aim of a survey like EU-SILC – is normally carried out with individual person as unit of analysis.

Longitudinal analysis, such as of persistent poverty, therefore requires that all persons in the original sample, irrespective of age or other characteristics, be followed-up in subsequent waves of the panel. As noted earlier (Section 1.5), restricting the follow-up to persons over a certain age, such as 14, normally does not have a major effect since only children under the specified age who move alone (that is, without being accompanied by an adult in the original household) are not followed-up. Usually such cases are rare and can be treated as non-response.

[16] In principle, the same set of information is required for returnees for their proper statistical treatment – see footnote (11) in section 5.2.

In so-called register countries, however, a more serious issue needs to be considered. Here we are dealing with two populations:

(i) all members of each sample household to be covered in the measurement of income and related characteristics, which can be obtained from registers or other sources not requiring detailed personal interviews; and

(ii) selected respondents, typically one per household, who are interviewed in detail for non-income or 'social' variables.

Which of these two populations – all persons or selected respondents – constitute the 'sample persons' to be followed-up in subsequent waves?

It is important to note that while for longitudinal analysis of variables covered under (ii) it is sufficient to follow-up selected respondents, that is not adequate for the longitudinal analysis of income and poverty.

Following-up only the selected respondents would mean that a representative sample of the whole population, especially of children, is not maintained. The follow-up rules for (i) have to be more inclusive.

7.4. Within-household non-response

By this is meant the failure to obtain the detailed interview with eligible individuals even when their household has been successfully enumerated. As a rule, this problem arises only in survey countries.

In Sections 5.5 and 6.4 we noted that in practice, the incidence of such non-response tends to be quite low – once the household interview has been successfully completed, the personal interview also tend to be obtained. The modest effect of within-household non-response on the composition of the sample can be controlled by calibrating the interviewed over the eligible sample for the personal interview in terms of age-sex and other characteristics.

However, the impact of such non-response on the particular households affected is generally much more important, and is not ameliorated by a calibration such as the above at the aggregate level.

Correction is required at the micro-level to the data of the households affected. This is because income and other characteristics at the household level have to be constructed by taking into account the contribution of all members of the household.

When the incidence of within-household non-response is low, the missing personal interviews may be imputed altogether. However, usually it is more appropriate and convenient to introduce a special weighting factor for the household concerned as a compensation for the information missing as a result of within-household non-response. We have not considered the construction of such special weighting factors in this paper, but for an example see Eurostat (2001).

Acknowledgements

The weighting procedures described here are based on the authors' detailed recommendations developed under a project with Eurostat, and we wish to thank Eurostat for providing the opportunity for their development. In fact, those recommendations are being implemented in EU-SILC national surveys. The present paper aims at providing a more systematic and clearer description of the weighting procedures, also introducing refinements to our original recommendations so as to enhance their consistency and completeness.

REFERENCES

ARDILLY, P., LAVALLÉE, P. (2003), The weight share method and the European Survey on Income and Living Conditions. *Proceedings Symposium 2003, Challenges in survey taking for the next decade*. Ottawa: Statistics Canada.

CLEMENCEAU, A., MUSEAUX, J-M., BAUER, M. (2006), EU-SILC: issues and challenges. Presented at conference *Comparative EU Statistics on Income and Living Conditions: Issues and Challenges*, Helsinki.

DEMING, W. E. (1943), *Statistical Adjustment of Data*, Chapter VII, Wiley (Paperback Dover1964).

DEVILLE, J-C., LAVALLÉE, P. (2006), Indirect sampling: The foundations of the generalized weight share method. *Survey Methodology*, **32**(2), pp. 165-176.

DEVILLE, J-C., SÄRNDAL, C-E. (1992), Calibration estimators in Survey Sampling. *Journal of the American Statistical Association*, **87**, pp. 376-382.

ERNST, L. (1989), Weighting issues for longitudinal household and family estimates. In *Panel Surveys* (eds. Kasprzyk D. *et al.*), pp. 135-159. Wiley.

EUROPEAN COMMUNITY (2003), Regulation (EC) No 1177/2003 of the European Parliament and of the Council. *Official Journal of the European Union*, pp. L165/1-9.

EUROSTAT (2001), ECHP UDB Construction of Variables, doc.PAN 167/01.

EUROSTAT (2004), First ideas on weighting for child care data, EU-SILC doc 135/04.

KALTON, G., BRICK, J. M. (1995), Weighting schemes for household panel surveys. *Survey Methodology*, **21**(1), pp. 33-44.

KALTON, G., KASPRZYK, D. (1986), The treatment of missing survey data. *Survey Methodology*, **12**, pp. 1-16.

KISH, L. (1992), Weighting for unequal P_i. *Journal of Official Statistics*, **8**, pp. 183-200.

LAVALLÉE, P. (1995), Cross sectional weighting for longitudinal survey of individuals and household using the weight share method. *Survey Methodology*, **21**(1), pp. 25-32.

LAVALLÉE, P., CARON, P. (2001), Estimation using the generalized weight share method: the case of record linkage. *Survey Methodology*, **27**(2), pp. 155-168.

LITTLE, R. J. A. (1986), Survey non-response adjustment for estimates of means. *International Statistical Review*, **54**, pp. 139-157.

POTTER, F. J. (1988), Survey of procedures to contrast extreme sampling weights. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 453-458.

POTTER, F. J. (1990), A study of procedures to identify and trim extreme sampling weights. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 225-230.

SPENCER, B., COHEN, T. (1991), Shrinkage weights for unequal probability samples. *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 625-630.

VERMA, V. (1999), Combining national surveys for the European Union. *Bulletin of the International Statistical Institute*, **58**.

VERMA, V. (2001), *EU-SILC Sampling Guidelines*. EU-SILC doc 51/01. Luxembourg: Eurostat.

VERMA, V. (2006), *EU-SILC Weighting Procedures: an outline*. Report to Eurostat on Advanced Estimation Methods project.

VERMA, V., BETTI, G. (2006), EU Statistics on Income and Living Conditions (EU-SILC): Choosing the survey structure and sample design. *Statistics in Transition*, **7**(5), pp. 935-970.

VERMA, V., CLEMENCEAU, A. (1996), Methodology of the European Community Household Panel. *Statistics in Transition*, **2**(7), pp. 1023-1062.