# Longitudinal Survey Methods for the Household Finances and Consumption Survey

Peter Lynn, Institute for Social and Economic Research, University of Essex

Report prepared for the European Central Bank

V 2.1  February 2012

# Longitudinal Survey Methods for the Household Finances and Consumption Survey

Peter Lynn

## Contents

## 1. Introduction

This report discusses a number of survey design and implementation issues relevant to the introduction of a longitudinal component of the European Household Finances and Consumption Survey (HFCS). The objective of the report is to aid the development of plans for a longitudinal component by informing discussions amongst the members of the Eurosystem Household Finances and Consumption Network (HFCN). The report does not make specific recommendations of a design to adopt. Rather, it highlights issues that require consideration, outlines alternative design options and describes the advantages and disadvantages of a number of options. The content of the report has been shaped by a briefing by Carlos Muñoz-Sanchez of the European Central Bank (ECB), by examination of documents on the HFCN website, and by discussions at a meeting of the HFCN in Frankfurt on 30 November 2011.

The views expressed in this report are those of the author. The report does not necessarily reflect the views and opinions of the ECB or the HFCN.

## 2. Study Populations and Survey Samples

Unlike cross-sectional surveys, longitudinal surveys explicitly recognise that populations are dynamic. They aim to collect data over periods of time in order to be able to detect and measure the population dynamics. This requires the time dimension to be incorporated into the definition of the study population. However, this is not always a straightforward thing to do.

If a longitudinal survey is intended to provide inferences about the totality of dynamics taking place in a particular population over a particular period of time, then it must be based on a sample that represents all of the types of units that could potentially contribute to such dynamics. The total set of units that are members of a population at any point during a period of time can be classified to four types, as shown in Table 1.

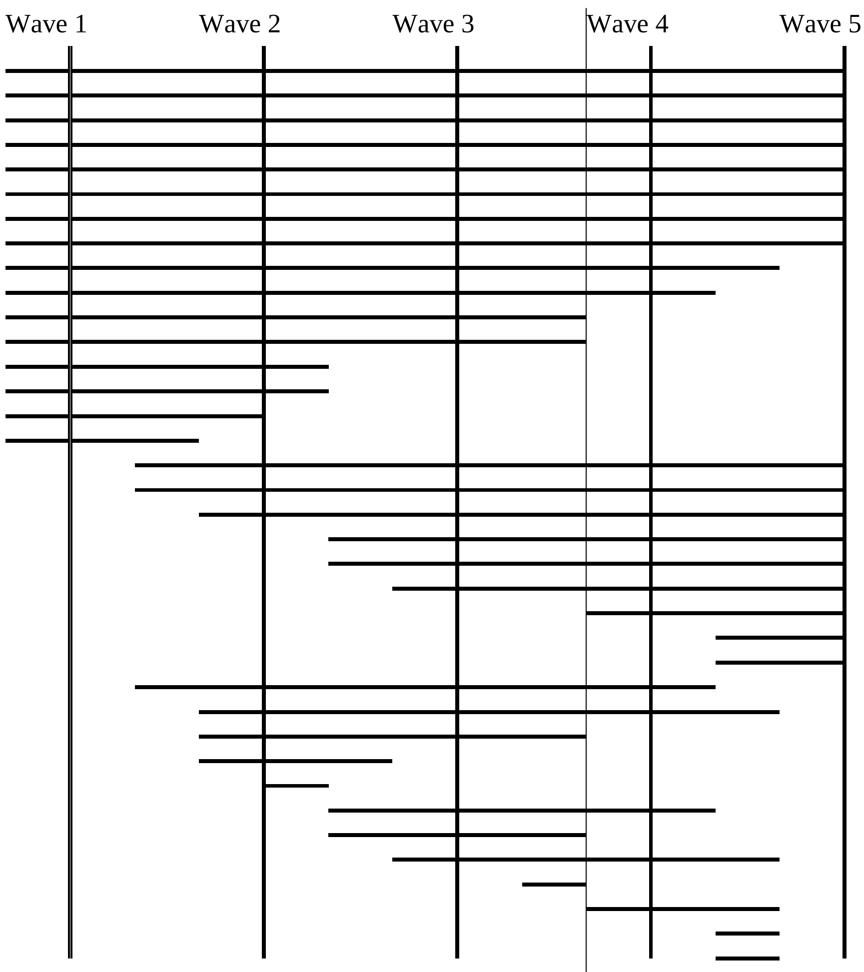
**Table 1: Types of units in a longitudinal population**

| Type | | |
|---|---|---|
| 1 | Stayers | Remain in population throughout study period |
| 2 | Leavers | In population at start of period, but no longer in population at end of period |
| 3 | Joiners | Not in population at start, but in population at end of period |
| 4 | Transients | Neither in population at start nor at end of period, but in population for a time during the period |

For a survey to represent all the units that are members of the study population at any point during the survey duration, it is necessary to include not only a sample of units that are continuously in the population during the period (“stayers”), but also samples of units that are in the population initially but leave it before the end of the period (“leavers”), of units that are not in the population initially but join it before the end of the period (“joiners”), and of units that are not population members either at the start or the end of the period, but are in the population for a limited time within the period of the survey ("transients").

Obtaining a representative sample of stayers and leavers is typically unproblematic, as this only involves sampling from the cross-sectional population at the time of the start of the survey period. Sampling joiners and transients is, however, often challenging. The issues involved can be illustrated by considering the relationship between the population dynamics and the realities of a longitudinal survey. Figure 1 shows population membership over time for a sample of population members. Each member is represented by a horizontal line. The timing of planned survey waves is shown by a series of vertical lines and the survey period is defined as the period from wave 1 until wave 5.

**Figure 1: Population Dynamics and Survey Waves**




Each horizontal line represents the time spent in the study population by one unit

The first 8 population members shown in Figure 1 are stayers; the next 8 are leavers; the following 9 are joiners; and the final 12 are transients. Selecting a sample from the cross-sectional population at the time of wave 1 and following them until the end of the survey period would result in representation of only the first 16 population member, out of 37. To achieve good coverage it is necessary to include samples of joiners and transients too, but these can of course not be identified at the start of the survey period, so the sampling must take place later, during the course of the survey.

Some approaches to sampling these groups, including rotating panel designs, involve selecting samples of those who are in the cross-sectional population at the time of subsequent waves, but Figure 1 illustrates that while these approaches may improve coverage, they still provide imperfect coverage. For example, suppose that a sample is selected at the time of wave 3 of population members who had not been in the population at the time of wave 1. This would provide coverage of 6 of the 9 joiners and 6 of the 12 transients. The extent of coverage can of course be improved by increasing the frequency of sampling. So, if at wave 4 another sample were selected, of population members who were not in the population at the time of wave 3, this would provide coverage of one additional joiner and one additional transient, giving a total coverage of 7 out of 9 and 7 out of 12 respectively. However, even adding an extra sample at every wave will not provide complete coverage, as some transients enter the population after one wave and leave before the next. To provide complete coverage a system of continuous real-time sampling is required. This may sometimes be possible, but often is not.

But does under-representation of transients and joiners matter? It may or may not, depending on the estimation objectives of the survey and the characteristics of the transients and joiners. For example, in a survey focused on labour market dynamics temporary migrant workers may have an important impact which would be missed entirely if transients were not included in the survey. Or in a survey of consumption patterns in a particular region, highly mobile people may similarly play an important role. In order to decide whether and how to sample transients and joiners, it is important to attempt to assess the approximate size of the sub-population and whether sub-population members are likely to have different characteristics from stayers and leavers in terms of the key survey estimates.

If transients and joiners are important, then the main choices available to ensure their inclusion are either to use a fixed panel with regular additional samples, or a rotating panel design. The latter has the disadvantage that only short-term change and outcomes can be studied, but the advantage that sample size can be more easily controlled. Either approach should be equally effective at providing cross-sectional estimates in addition to longitudinal estimates. In the following section, some issues in the design of rotating panels are discussed.

## 3. Rotating Panel Designs

Rotating panel designs involve pre-determined proportions of sample units being retired at each fieldwork occasion and replaced by a new sample. Typically, each unit will remain in the sample for the same number of waves. A wide variety of rotation patterns are possible. Figure 2 presents a simple example in which each unit remains in the sample for three waves, one-third of the sample being replaced each time. The three samples interviewed in any one time period can be combined to make cross-sectional estimates, though this combination must take account of population joiners and leavers who may be represented in one or two of the samples but not all three. For example, when combining the three samples interviewed at time period 3, units that entered the population between periods 1 and 2 will be represented only in samples 2 and 3 and not in sample 1. Consequently, assuming equal sample sizes in each rotational group, those period 1-2 entrants in samples 2 and 3 should be given an additional weighting adjustment of 3/2 relative to the population stayers in all three samples. Similarly, units that entered the population between periods 2 and 3 will only be represented in sample 3 and should therefore be given a weighting adjustment of 3/1. Without making these adjustments, the sample will under-represent joiners and population estimates will be biased if joiners have different characteristics to stayers. Surveys that use rotating panel designs usually collect the same data from each unit at each wave in order to permit combination in this way to make cross-sectional estimates.

**Figure 2: Rotating Panel Design: 1-1-1 Rotation Pattern**

| | Time period 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Sample 1 | Wave 1 | Wave 2 | Wave 3 | | |
| Sample 2 | | Wave 1 | Wave 2 | Wave 3 | |
| Sample 3 | | | Wave 1 | Wave 2 | Wave 3 |

Rotating panel designs are often used when the main estimation objectives are cross-sectional estimates and short-term estimates of net and gross change. Labour Force Surveys have a rotating panel design in many countries. Fixed panels with additional samples of population entrants are used when longer-term estimates of net and gross change are required, though such panels can also provide cross-sectional estimates.

In some common rotating panel designs, units are not included in the survey at every period between the time they first join the panel and the time they leave. Figure 3 presents an example of such a design. In this example each sample unit is included in the survey for two consecutive periods, is then excluded for the next two periods, and is then included again for two more periods. The reason for adopting such designs is connected with the trade-off between respondent burden and representation of recent births, and on the other hand, variance of estimates. Suppose that the time periods in Figure 3 are in fact calendar quarters and that a main objective of the survey is to provide estimates of change from one quarter to the next ("quarterly change"). Further suppose that there is a seasonal component to change, as is often the case with business statistics. Then there will be interest in comparing the quarterly change in any particular quarter with the quarterly change in the *same* quarter of the previous year. For example, the change in period 7 (from period 6) will be compared with the change in period 3 (from period 2).

The difference in levels between periods 3 and 7 may also be of direct interest. We can see that with the rotation pattern of Figure 3, half of the sample units in period 3 are also sample units in period 7. And half of the sample units that can be used to estimate change from period 2 to 3 can also be used to estimate change from period 6 to period 7. This high degree of *sample overlap* reduces the variance of estimates of net change, relative to the use of independent samples. A higher degree of overlap would reduce variance further, but this would also increase respondent burden. At the extreme, 100% overlap could be achieved with a fixed panel design. The fixed panel design would minimize the variance of estimates of net change. It would also allow the estimation of gross change between any pairs of time periods. The rotating panel design of Figure 3 can also support gross change estimates for pairs of periods in which the same units are interviewed, e.g. between periods 3 and 4 (sample 3), 3 and 6 (sample 2), 3 and 7 (samples 2 and 3), and 3 and 8 (sample 3), but not between 3 and 5, 3 and 9, etc.

**Figure 3: Rotating Panel Design: 1-1-0-0-1-1 Rotation Pattern**

| | Time period | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Sample 1 | W 1 | W 2 | | | W 3 | W 4 | | | | |
| Sample 2 | | W 1 | W 2 | | | W 3 | W 4 | | | |
| Sample 3 | | | W 1 | W 2 | | | W 3 | W 4 | | |
| Sample 4 | | | | W 1 | W 2 | | | W 3 | W 4 | |
| Sample 5 | | | | | W 1 | W 2 | | | W 3 | W 4 |

If the HFCS were to opt for a rotating panel design, the key design decisions would concern the rotation pattern and the proportion of units to be replaced on each occasion. Analysis can be complicated if different samples remain in the survey for different numbers of observation period, so a 'balanced' design is generally recommended, in which each sample is treated in the same way. Each sample should remain in the survey for *n* periods, and 1/*n* of each sample should be retired at each period.

If different samples are of different sizes, this does not cause estimation problems in principle, but in practice this will result in different cross-sectional estimates being based on samples with different non-response characteristics. Consequently, comparability of cross-sectional estimates over time relies upon the adequacy of non-response adjustment weighting. While good non-response weighting is highly recommended, it is preferable to avoid having to rely on it. Consequently, it is recommended that each sample should be of the same size, as far as possible.

This leaves the question of the rotation pattern. If we assume annual waves of interviewing, there may be no *a priori* reason to assume that certain combinations of years are more important than others in terms of measuring change. There is no reason to suppose, for example, that it is more important to have good estimates of change over 4-year periods than over 3-year periods. This would imply that the most efficient design is for all sample units to be included in the survey for each of *n* consecutive annual data collection periods. The optimum value of *n* should depend on the true rate of change of measures of interest (e.g. wealth accumulation or reduction in wealth) in the population and the nature of related outcome variables of interest. In principle, these could differ greatly between countries, but for reasons of comparability it is likely to be preferred to use an identical – so far as possible – design in each country.

A final consideration is the design of a rotating panel in the first few data collection periods, until it reaches equilibrium. For example, suppose for illustration that the design is a 1-1-1-1-1-1 design, in which each sample unit is included for six consecutive years (Figure 4), and that each sample will consist of 300 units, so that the total sample included at each survey period is 1,800. Then, one option is to select a new sample of 300 units each year, with each sample then remaining in the survey for six years. That will result in the sample consisting of 300 units in year 1, 600 in year 2, 900 in year 3, and so on until the full sample size (equilibrium status) is achieved only in year 6. This design has the disadvantage of a small sample in the early years, which may make it useable only for a limited subset of desired estimates. The advantage is that fieldwork may be more manageable, giving time for new longitudinal procedures to be developed and improved before having to cope with the full sample size.

An alternative design (Figure 5) would be to select the full sample size of 1,800 units in year 1 and to subdivide them into six subsamples, of which one would be designated to remain in the survey for only one year, one would remain for two years, one for three years, and so on. Then, each year thereafter a new sample of 300 units would be selected. This would deliver the full annual sample size immediately from year 1, though the non-response characteristics of the annual sample would be different in each of the first five years to the equilibrium position achieved from year six onwards.

**Figure 4: A 1-1-1-1-1-1 Rotating Panel Design with Gradual Start**

| | | Time period | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Sample 1 | 300 | W 1 | W 2 | W 3 | W 4 | W5 | W6 | | | | |
| Sample 2 | 300 | | W 1 | W 2 | W 3 | W 4 | W5 | W6 | | | |
| Sample 3 | 300 | | | W 1 | W 2 | W 3 | W 4 | W5 | W6 | | |
| Sample 4 | 300 | | | | W 1 | W 2 | W 3 | W 4 | W5 | W6 | |
| Sample 5 | 300 | | | | | W 1 | W 2 | W 3 | W 4 | W5 | W6 |
| Sample 6 | 300 | | | | | | W 1 | W 2 | W 3 | W 4 | W5 |
| Sample 7 | 300 | | | | | | | W 1 | W 2 | W 3 | W 4 |
| Sample 8 | 300 | | | | | | | | W 1 | W 2 | W 3 |
| Sample 9 | 300 | | | | | | | | | W 1 | W 2 |
| Sample 10 | 300 | | | | | | | | | | W 1 |
| Total | | 300 | 600 | 900 | 1200 | 1500 | 1800 | 1800 | 1800 | 1800 | 1800 |

**Figure 5: A 1-1-1-1-1-1 Rotating Panel Design with Immediate Start**

| | | Time period | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Sample 1 | 300 | W 1 | | | | | | | | | |
| Sample 2 | 300 | W 1 | W2 | | | | | | | | |
| Sample 3 | 300 | W1 | W2 | W3 | | | | | | | |
| Sample 4 | 300 | W1 | W2 | W3 | W4 | | | | | | |
| Sample 5 | 300 | W1 | W2 | W3 | W4 | W5 | | | | | |
| Sample 6 | 300 | W1 | W2 | W3 | W4 | W5 | W6 | | | | |
| Sample 7 | 300 | | W1 | W2 | W3 | W4 | W5 | W6 | | | |
| Sample 8 | 300 | | | W1 | W2 | W3 | W4 | W5 | W6 | | |
| Sample 9 | 300 | | | | W1 | W2 | W3 | W4 | W5 | W6 | |
| Sample 10 | 300 | | | | | W1 | W2 | W3 | W4 | W5 | W6 |
| Total | | 1800 | 1800 | 1800 | 1800 | 1800 | 1800 | 1800 | 1800 | 1800 | 1800 |

## 4. Sampling Births

In the case of the HFCS, births will consist of households consisting entirely of persons who were not resident in the country at the time the initial sample was selected, i.e. "immigrant households[1]". The challenge is therefore to devise a way of sampling immigrant households in each country. It is likely that some countries will be able to rely on official records and population registers to provide a sampling frame, but it should be noted that this will only cover legal immigrants. If the illegal immigrant population is

[1] Care should be taken in interpreting the shorthand label "immigrant households", as these will include households previously resident in the country but which were (temporarily) absent at the time of sample selection. Also, immigrant households could contain one or more new-borns, born in the country since the other household members entered the country.

likely to be anything other than negligible in size, it is recommended that consideration should be given to including them too. A generic method to do this is to reapproach all residential addresses that were sampled initially – regardless of whether the sampling frame was one of addresses or of households or persons – and to identify anyone resident at those addresses who meets the criterion of being an immigrant. This method is described in the context of a UK survey in Lynn (2011). It relies upon an assumption that the population of dwellings in the country is a closed population, which is obviously not strictly true but may be a reasonable approximation over relatively short periods of up to 4 or 5 years.

Often, a sampling frame that will provide acceptable coverage of the total (stock) population will be inadequate as a frame of new entrant (flow) population. This could be because the mechanism by which new entrants get added to the frame has a time delay, so it is not up to date. Or it could be because new entrants are simply more likely to have the kind of characteristics that are associated with absence from the frame. Thus, simply selecting a new sample from the same frame that was used to select the original wave 1 sample may not be adequate. Instead, this may need to be augmented by special procedures. If coverage of new entrants is poor, this could lead to coverage errors.

## 5. Fieldwork Aspects of Sample Coverage

As longitudinal surveys are concerned with representing micro-level change over time, it is this change which needs to be adequately represented by the survey data. This arises from a combination of sampling units over time, as discussed above, and collecting data over time for each of those units. In other words, the sample of changes (or more generally, events) depends firstly on the selection of units and secondly on the identification of all changes (events) experienced by each of those units over the survey period. For example, a longitudinal component of the HFCS may aim to measure change in financial assets over a period and the characteristics associated with such change. This obviously requires that each sample unit is followed and interviewed throughout the survey period (or until they leave the population). This aspect is dealt with in the section below on ‘avoiding sample attrition’. But it also requires that data is collected for relevant time periods. There are two potential pitfalls that should be avoided:

- A common approach is to consider the first interview to be a “baseline” interview, collecting data about the current situation. Each subsequent interview then collects a history of changes since the previous interview. Such an approach will produce data for the whole survey period for the initial sample, but any sample of population entrants must be treated a little differently. For them, the period from when they entered the country until the date of the first interview also falls within the survey period, so their first interview should not be the same as that for the initial sample. Instead, it should if possible also include collection of a history of changes since they entered the country. (See Figure 1: for each joiner or transient covered by the wave 3 sample, the part of the horizontal line that is to the left of the wave 3 vertical line would be excluded);

- Units who are found, at a particular survey wave, to have left the population, are often not then interviewed. This results in the loss of information relating to period between the previous interview and the time they left the population. This can be rectified by attempting to carry out a final interview on the first occasion that a sample unit is found to have left the population. For example, data could be collected by telephone or web from those who have emigrated, and “exit” interviews may be possible with a close relative of sample members who have died. (See Figure 1: for each leaver or transient, the part of the horizontal line that is to the right of the last intersecting vertical line would otherwise be excluded).

## 6. Households over Time

An important issue for the HFCS concerns the nature of the units to be sampled and followed. The survey is referred to as a survey of households and the data collected pertain to households (European Central Bank 2011). However, this does not rule out the possibility that the units of analysis could be individual persons, or indeed any other subset of household members, such as co-resident couples. The same data could be used to produce estimates of, for example, either the proportion of households with particular assets or the proportion of individuals who are in households with particular assets.

In the longitudinal context, it has long been argued and is widely agreed (Rose 2000a) that it is unhelpful to attempt to define continuing households over time. Household membership is inherently dynamic and attempts to make longitudinal measurements on a sample of household result in arbitrary decisions about when a household is or is not the "same" household as the one observed previously. Furthermore, the wide variety of possible changes over time in household composition makes it impossible to devise a set of rules that result in the sample continuing to be representative of the total population of households. In consequence, all household panel surveys take the approach of treating individual persons as the units of sampling and observation, while collecting data at each wave regarding the entire household of each sample individual (Rose 2000b). With this approach it is perfectly possible to use the data to study "household dynamics." The key point is that these dynamics affect, and are affected by, individual persons.

The impossibility of successfully defining all households over time can be illustrated by considering a simple example. Suppose our initial sample of households at wave 1 contains two households, each of which consists of two adults. Suppose that at wave 2, the composition of household A remains unchanged while household B has split into two 1-person households. If we randomly select one of those two new households to be the "same" household as the one we interviewed at wave 1 and interview only that one, then the sample as a whole will under-represent households of this type by a factor of 0.5 relative to households like household A. This could be overcome by weighting, but only because this is a simple situation in which neither of the members of household B are now co-residing with anyone else. If in fact one of these two people is now co-residing with another person, then that new household could have been included in the sample via the household of either of those people having been selected at wave 1. So, we now need to know the selection probability of a household that we neither sampled nor observed at wave 1. And if the person we interviewed at wave 1 is now co-resident with more than one other person, we need to know whether those people were in the same household as each other at the time of wave 1 and, if not, the selection probability of each of those unobserved households. The situation becomes considerably more complicated once we extend the argument to three or more waves.

Using a criterion other than random selection to determine which household is the "same" as the previously-observed one brings additional problems. Suppose we decide that if one household is still residing at the same address, then that household is the "same" household. If in our example one of the members of household B is, at wave 2, still at the same address at which we observed them at wave 1, we will interview that person rather than the other one. The consequence is that of all the wave 1 households in our sample that have experienced that particular dynamic (two people at wave 1; one living alone at the same address at wave 2 and one living alone elsewhere) we will always only observe the person who has remained at the same address. Our wave 2 sample will not include any households of the second type (one person, who at the time of the previous wave was living with one other person who is still living at the address where they previously lived together). Following rules of this kind therefore introduce systematic bias into our sample and into any estimate that is likely to be associated with the nature of the change in household composition.

Luckily, it is not just on practical and statistical grounds that household panel surveys treat individuals as the unit of observation and analysis. There are also good substantive reasons for doing so. When studying issues associated with resources, wealth, poverty and so on, it is individuals who are affected by these features and individuals who make influential decisions such as investment and consumption decisions. Thus, from a sociological perspective, individuals are the appropriate units of analysis. When studying income or wealth dynamics, both sociologists and economists treat individuals as the units of analysis, though key variables are measured at the household level (e.g. Hurst et al 1998, Jenkins 2011, Mazzocco et al 2006). Effectively, the household is treated as an attribute of an individual, and measures such as equivalised household income or wealth pertain to each individual in the household. With this approach, it is possible for analysis to reveal the full range of household dynamics and how these relate to, for example, investment decisions or dis-saving. The data can be used to identify households with unchanged composition, and those with particular types of changes (departure of an adult child, split of a couple, etc) and the wealth dynamics of each can be compared. With the alternative, of imposing an arbitrary rule that determines which household is the "same" as the one previously observed, some types of household changes will not be observed at all and it will not be clear what subset of the total the sample represents.

It can be argued that the simplification involved in the arbitrary rule approach may make little difference if changes are measured only over fairly short periods of time. However, even for short-period measures of change, it would be wise to test this assumption. Bricker et al (2011) report findings from the Survey of Consumer Finances 2007-2009 panel, in which an arbitrary rule method was employed. They note that in around 10% of the households interviewed in the second time period, the respondent had either gained or lost a spouse/partner since the earlier interview. Other compositional changes were of course also possible. The authors state that, "Eliminating families with such large compositional changes from the analysis does not affect the qualitative findings," but it is important to note that this is only a partial test of the assumption that the simplification has no impact. It does not address the impact of excluding certain types of households completely. For example, in cases where a couple split, the SCF 2007-2009 rule was to follow only the 2007 respondent and collect data regarding that person's 2009 household. The 2007 respondent was defined as "the financially most knowledgeable member of the economically dominant couple." Consequently, all the 2009 households of the financially less knowledgeable member of a recently-split couple are excluded from the survey, so there are no data with which to assess the impact of this exclusion. It seems that many studies of the dynamics of household finances simply ignore the possibly implications of household following rules. For example, Alessie et al (2004) analyse data from a Dutch panel with an arbitrary following rule (Nyhus, 1996), but they neither mention the rule in their description of the data nor mention any possible implications in their discussion of findings. Similarly, Chiappori and Paiella (2008) analyse data from the Italian Survey of Household Income and Wealth, but the neither describe the following rule nor mention that it could affect the findings. However, other researchers use data from surveys that follow individuals to directly study the effects of household composition changes on household finances (e.g. Jenkins 2009).

It is recommended that any panel element of the HFCS should consider individuals as the units to be sampled and followed and, for longitudinal analysis, as the units of analysis. This should not limit the type of analysis and estimation that is possible in any way, compared to the arbitrary rule approach, but it will require a careful understanding of the meaning of the estimates. For example, instead of making statements about the proportion of households who have experienced a particular change in their finances, statements could be made about the proportion of people who have experienced a particular change in their household finances. The measure is the same, but the unit of analysis is not.

It is unproblematic if the initial sample of individuals should be clustered within households, but equally unproblematic if it is not (e.g. a simple random sample of individuals sampled from a population register). At each wave, data should be collected relating to the entire current household of each sample member in order to be able to obtain, or derive, household-level measures such as equivalised wealth. Well-developed procedures and processes exist for carrying out surveys using this approach. A relevant example is the now-defunct European Community Household Panel Survey[2].

Of course, the arbitrary-rule approach could still be considered as an alternative, but an empirical assessment of the effects of the inherent assumptions on key estimates would be recommended. Before implementing such an approach, the HFCN should reassure themselves that the disadvantages of the inherent approximations are outweighed by whatever advantages are believed to stem from the approach.

[2] See http://epp.eurostat.ec.europa.eu/portal/page/portal/microdata/echp

## 7. Minimising Sample Attrition

All the reasons for non-response that apply to all surveys (Lynn 2008) apply also to longitudinal surveys. However, there are also additional considerations in the case of longitudinal surveys. Furthermore, the impact of nonresponse, and the techniques that can be used to tackle it, can be rather distinct in the case of longitudinal surveys.

### *7.1 Failure to locate*

Subsequent to the first wave of data collection, a major component of non-response and attrition on longitudinal surveys is caused by geographical mobility of sample members (Couper and Ofstedal 2009). Between waves, a proportion of sample members will move home, change employment, change telephone number or change email address. If the survey organisation is reliant upon any or all of these details to contact the sample member, then extra measures will need to be taken in order to be able to locate the sample member and make contact at the subsequent wave. Ideally, there should also be interviewers available to attempt contact with sample members in all the places to which sample members could move. (This is not an issue for a telephone, postal or web survey, but is important for a face-to-face interview survey.) For example, surveys of households in the UK with data collection at one-year intervals find that around 10% of persons on average change address between each consecutive pair of survey waves (Laurie et al, 1999). The importance of taking extra measures to keep in contact with sample members is greater the higher the level of mobility in the population and the less frequent the survey waves. Thus, more resource-intensive activities may be justified for a survey of a particularly mobile population or for a survey with large intervals between waves.

Survey organizations carrying out longitudinal surveys typically engage in three types of activities aimed at maintaining up-to-date contact details for sample members:

- In-wave activities to anticipate the need for tracing;
- Between-wave activities to keep in touch;
- In-wave activities to trace people found to have moved.

The combination of these types of activities is important. The objective is to obtain information prior to a move whenever possible, but also to maximise the chances of re-locating a sample member if a move takes place without prior notification.

***In-wave anticipatory activities***

A range of contact information should be collected at the first wave and at each subsequent wave. Any information that could be used subsequently to help trace a sample member who has moved will be helpful. Longitudinal surveys typically ask for any or all of the following:

- Phone numbers (home, office, mobile, other);
- Email address(es);
- Contact details (name, address, and telephone number) of one or more "stable contact" (parent, sibling, friend).

Information of this kind can then be used when needed to help establish contact at subsequent waves. To do this successfully requires maintenance of a secure user-friendly database that can be accessed by authorized field staff when needed. The survey organization must be able to update the database at all times – before, during and after survey waves – and to feed the latest information from the database to interviewers prior to each wave or upon request by field staff when it becomes apparent that a sample member has moved.

Additionally, it is useful to collect information that could help predict the propensity of the sample member to move in the near future, so that between-wave activities can be targeted. Respondents can be asked directly whether they expect to move in the near future and, if so, why and where they are likely to move to. Unless intensive efforts are made to track the respondents who say they are likely to move, the responses to such a question can be highly predictive of the ability to locate the respondent at the next wave, as Table 2 illustrates. Table 2 shows that the proportion of sample members who could not be located at wave 2 of the UK Household Longitudinal Survey (UKHLS) was 12% amongst those who had answered one year earlier, at wave 1, that they expected to move in the next 12 months. The equivalent proportion was only 2% amongst those who had not expected to move. This wave 1 survey question is therefore a very good predictor of the ease of locating a sample member at wave 2. As only 15% of wave 1 respondents had expected to move, it may have been possible to use relatively resource-intensive techniques to keep in touch with them in the months following the wave 1 interview, while less effort could have been made for the remaining 85% of wave 1 respondents. On the UKHLS cases were not treated differently between waves depending on whether they expected to move, hence the disparity in location rates at wave 2.

**Table 2: Association between wave 2 outcome and reported expectation regarding moving at wave 1, UK Household Longitudinal Survey**

| | Expect to move in next 12 months? | |
|---|---|---|
| | Yes | No |
| | % | % |
| Full interview | 64.9 | 77.3 |
| Proxy interview | 2.3 | 1.8 |
| Non-contact | 8.8 | 5.5 |
| Refusal | 12.2 | 13.5 |
| Unable to locate | 11.9 | 2.0 |
| *n* | *3,129* | *18,558* |

Notes: The UK Household Longitudinal Survey has a 12-month interval between waves. Wave 1 was carried out in 2009 and wave 2 in 2010.

### *Between-wave keeping in touch activities*

Between survey waves, sample members can be contacted and requested to provide up-dates to their contact information, in case any details have changed. This can be done by mail, email or phone, or by a combination of these methods. Email contact is the cheapest and simplest to administer, so this method of contact could be used for all sample members for whom a valid email address is available, in addition to contacting all sample members by mail or phone. Emails can include a hotlink to a survey website with a simple reply form as well as allowing reply by email or phone. Mail contact is often preferred as this gives the survey organisation an opportunity to combine the request for contact details with a “thank you” letter, survey findings, or other information, and possibly a gift or incentive (Laurie et al 1999). Some surveys send birthday cards or other greetings cards to sample members. To inform the survey organisation of changed details, sample members can be provided with a reply-paid card, an email address, or a freephone number – or, preferably, all three.

Recently, researchers have investigated the most effective ways to obtain up-dated contact details from sample members between survey waves. Two main findings have emerged. The first is that more frequent contact results in a higher proportion of sample members replying. McGonagle et al (2011) carried out an experiment on the American Panel Study of Income Dynamics (PSID), which has two-year intervals between waves. They randomly allocated sample members to two groups, one of which received only one between-wave mailing and the other of which received two between-wave mailings with a four-month interval between them. They found that the second group were more likely to supply address updates and less likely to remain unlocated at the following wave.

The second finding regarding between-wave mailings is that it is more effective to ask sample members to supply new details only if their details have changed than to ask all sample members to confirm whether or not their details have changed (Fumagalli et al 2010). This may be because movers feel that the request is relevant to their specific situation and is therefore salient.

***In-wave tracing activities***

When a sample member cannot be located at their last known address during the course of a wave of a face-to-face interview survey, the field interviewer is often best-placed to trace them. There are several local potential sources of information that the interviewer can exploit. She can ask the new residents of the previous address whether they know the new address of the sample member. She can also make enquiries with neighbours and nearby relatives, some of whom may previous have been mentioned by the respondent as "stable contacts" (see above) or may indeed be or have been sample members themselves if they had previously lived in the same household as the sample member who requires tracing. There may also be other local resources upon which the interviewer can draw, such as registers and records that may be held at the local library or town hall. In most countries the majority of residential moves are relatively local, so there is a good chance that the sample member will still be in the same area.

If the interviewer is unable to trace the sample member in the field, the next step is to refer the case back to the survey office. Office staff may be able to obtain information from any one of a number of sources, including:

- Commercial search agencies;
- Online resources such as electoral registers;
- Internet searches;
- Social networking sites (see );
- Phoning or otherwise contacting stable contacts or other related persons on the sample data base.

This stage of "office tracing" is an important part of sample maintenance for any longitudinal survey and most longitudinal surveys will either have one member of staff dedicated to these tasks or a number of people who are able to carry out these tasks as and when necessary in addition to their other responsibilities. The approach to organising office responsibility for these tasks will tend to depend on whether or not the survey has continuous fieldwork and on the sample size, as these features determine the flow of office tracing work.

### *7.2 Refusals*

In addition to the universal causes of refusals that affect all surveys, there are two special features of importance with longitudinal surveys. The first is that participation requires a considerable commitment on the part of sample members – not just a single interview, but several, over a period of time. This is sometimes referred to as high *respondent burden*. In consequence, special incentives or motivation may be needed to compensate. Typically (but not always), longitudinal surveys offer sample members a small payment for each interview, or some other form of gift, as well as putting particular effort into making the sample member feel like an important, irreplaceable, component of the study and persuading them that the study itself is valuable. Survey "branding" can play an important role in fostering loyalty to the survey on the part of respondents.

The second special feature of longitudinal surveys relevant to refusals is that, after the first wave, sample members have already experienced the survey interview and therefore have a very good idea of exactly what it consists of, what kinds of questions will be asked, and how difficult, embarrassing, sensitive, uninteresting and time-consuming they find it. This is very different from a typical survey situation, where a sample member will have only a rather vague and general impression of what they are being asked to do at the time when they are being asked to co-operate. Consequently, on a longitudinal survey it is very important to try to make the interview experience as pleasant as possible for the respondent. If a respondent finds that the negative aspects of the interview outweigh the positive aspects, they will be less likely to be willing to take part again at the next wave.

One important design question for longitudinal interview surveys is whether it is advantageous, where possible, to assign the same interviewer to a respondent at each wave. In terms of the effect on response rate, evidence is rather equivocal. Most studies that purport to demonstrate an effect of this sort are non-experimental and, in consequence, confound interviewer stability with area effects. Two experimental studies provide an exception to this. The first was carried out on the British Household Panel Survey. Campanelli and O'Muircheartaigh (2002) concluded that the apparent differences in co-operation between cases approached by the same interviewer and those approached by a different interviewer could be accounted for by non-random interviewer attrition. The second experimental study was carried out on the Omnibus survey of the UK National Centre for Social Research and included some extensions to the previous study. Lynn, Kaminska and Goldstein (2011) found no evidence of any effect of changing the interviewer, except where an experienced interviewer was replaced by an inexperienced one. In conclusion, then, there is no direct evidence that maintaining interviewer continuity improves response rates, *ceteris paribus*. It may be helpful for some sample members and not for others.

A general tactic that can be used on longitudinal surveys to reduce the risk of refusals is to “tailor” the approach to each sample member to suit their particular circumstances and preferences. From wave 2 onwards, a lot of information is known about each sample member prior to each approach for interviewer. This includes both the survey responses and the process data from earlier waves (for example, at what times of day contact attempts were successful or unsuccessful; whether a refusal conversion attempt was needed and what were the reasons for the initial refusal). This information can be used to identify a promising approach at the next wave. Aspects of the approach that could depend upon this prior information include the timing of contact attempts, the wording of any advance letter or other written communication, the nature of any incentive offered, the messages that are emphasised by the interviewer when introducing the survey, and even the mode of data collection. Fumagalli et al (2010) found that even a seasoned panel reacted positively to receiving a tailored mailing between waves. Respondents were sent a short report of survey findings, but for one experimental group the report was, to some extent, tailored to the respondent’s circumstances (there were three versions of the report, for young people, busy employed people, and others). Response rates at the subsequent wave were higher for the group that received the tailored report.

## 8. Weighting

For longitudinal surveys, as for any survey, weighting is used to adjust for differential selection probabilities (design weights) and for differential non-response (non-response adjustments). But some of the factors to be considered are unique to the longitudinal survey situation.

### *8.1 Design Weights*

On a longitudinal survey, selection probabilities can vary on three different dimensions:

- Between units within the initial sample;
- Between units within subsequent samples;
- Between different samples

The issues with probabilities varying between units within a single sample are no different from those on any cross-sectional survey. The additional considerations for a longitudinal survey arise from the combination of samples selected at different times. Four situations can arise, depending on the purpose and nature of the additional sample(s). We will illustrate the ideas here with examples involving two samples, an initial one and one added at a later date, but the ideas extend easily to larger numbers of additional samples.

***Refresher sample of the "same" population***

Many surveys select an additional sample using the same population definition as the one used for selection of the initial sample. Rotating panel surveys are an example of surveys that do this. Though the population definition is the same, the population membership is of course not the same due to the dynamic nature of populations. If a design of this kind were used for the HFCS, it would be likely that many persons would have two chances of selection, as they would be in a resident household both at the time the initial sample was selected and at the time the refresher sample was selected – "stayers" in the terminology of section 2 above. But other people would have only one chance of selection, being either joiners or leavers. These differences need to be taken into account in developing the survey weighting.

The overall sampling fraction need not be the same on the two occasions. Without loss of generality, assume the sampling fraction for the initial sample is ${n_1}/{N_1}$, while that for the subsequent sample is ${n_2}/{N_2}$. In practice it is often the case that $N_1 \cong N_2$ but it is common that $n_1 \neq n_2$. If, for simplicity of illustration, we assume that each of the two samples is an equal-probability sample, this means that for analysis that combines the two samples, selection probabilities are as follows:

Stayers: $P_1 = n_1/N_1 + n_2/N_2$

Leavers: $P_2 = n_1/N_1$

Joiners: $P_3 = n_2/N_2$

Design weights for such analysis should be in inverse proportion to these probabilities, viz:

$$w_i \;\alpha\; 1/P_i$$

It should be noted that $P_2$ and $P_3$ only require knowledge of the selection probability of each person in the sample in which they were actually sampled. However, $P_1$ additionally requires knowledge for each sampled person of the probability that they would have been selected in the other sample, where they were not selected. This means that all members of the refresher sample must be asked survey questions to establish if they were members of the population at the time the initial sample was selected – in other words, to establish whether they are stayers or joiners. If the initial sample also had varying sampling fractions within the sample – for example, over-sampling a particular region – then the members of the refreshment sample must also be asked questions to establish their probability of selection into the initial sample – for example, whether they lived in the over-sampled region at that time.

***Boost sample of a population subgroup***

Sometimes a sample may be selected of a sub-population of particular analytical interest, but subsequently to the initial sample, for example because funding becomes available for this extension only later. With a design of this kind, the issues are analogous to those with the refreshment sample design outlined above. Suppose that the target population for the second sample is the resident population of a particular region, which we shall refer to as region A. And suppose again that the sampling fractions for the two samples are respectively $n_1/N_1$ and $n_2/N_2$ and that each sample is an equal-probability sample. Then, selection probabilities will vary between three types of persons as follows:

Persons in the population at $t_1$ and in region A at $t_2$ : $P_1 = n_1/N_1 + n_2/N_2$

Persons in the population at $t_1$ but not in region A at $t_2$ (they may or may not be in the overall population at $t_2$): $P_2 = n_1/N_1$

Persons not in the population at $t_1$ but in region A at $t_2$ : $P_3 = n_2/N_2$

In this case, in order to know the overall selection probability for each sample member, it is necessary to establish for every member of the initial sample whether or not they are resident in region A at the time the boost sample was selected and for every member of the boost sample whether they were a member of the overall population at the time the first sample was selected, regardless of whether or not they resided in region A at that time. It is easy to forget to include these important extra questions in the survey questionnaire.

***Sampling dependent on the existing sample***

Some longitudinal surveys add new sample members at each wave based on their relationship to existing sample members. An example is the practice on the German Household Panel Survey of adding to the sample new household members of each continuing sample member (Spiess et al 2008). In this case, the selection probability of each person is dependent on the selection probability of each other current household member, which is in turn dependent on the selection probability of each other previous household member of each of those people. It is simply not possible to collect adequate co-residency histories for each household member and each other person who had lived in the same household as that person at any time since the start of the survey. Consequently, selection probabilities cannot be known for any sample design with this kind of dependency in the selection mechanism. Instead, model-based approaches with untestable assumptions must be used to estimate the probabilities. Moving away from design-based inference to model-based inference is a big step for a survey to take.

***Sample of population entrants since the time of the initial sample***

If an independent sample can be selected, with known probabilities, from the set of people who have entered the population since the time the initial sample was selected (e.g. immigrants and births), then combining this sample with the existing sample is unproblematic. There is no overlap between the two populations represented by the two samples and therefore no analogue to the first of the three situations set out above regarding refresher samples of the "same" population and boost samples of a population subgroup, in which persons have two chances of selection. Instead there are just two simple situations:

Persons in the population at $t_1$ (stayer or leaver): $P_2 = n_1 / N_1$

Persons not in the population at $t_1$ but in the population at $t_2$ (joiner): $P_3 = n_2 / N_2$

However, it is often difficult to identify a way to select an independent sample of joiners. Instead, surveys often resort to mechanisms that are dependent on the existing sample. If the mechanism involves a simple one-to-one relationship between persons, then it may be possible to know the selection probabilities. This is the case with the common practice on household panel surveys of including a newborn child as a sample member if their mother is an existing sample member. In this case, the selection probability of the child equals that of the mother, which must already be known as the mother is a sample member. A variant of this rule, used on some surveys, is to include a newborn child as a sample member if either of his or her biological parents are themselves a sample member.

The selection probability is already a little more complicated in this case. If the initial sample design involved selecting households and then including as sample members all the persons in those households, newborns whose parents were living in two different households at the time of the initial sample selection will have twice the probability of selection of newborns whose parents were already living together in the same household at the time of the initial sample selection. The design weight should therefore take that into account. Similar rules to those described here for newborns could be used for age-restricted samples. For example, if a survey aims to cover all residents aged 16 or over, 16 year-olds could be added to the sample at the first survey wave after their sixteenth birthday if their mother is a sample member.

Another approach to adding population entrants to the sample is to include as sample members any joiners (e.g. immigrants) who are located in the same household as one or more existing sample members. This approach has appeal as it involves an inexpensive method of sampling. However, it suffers from a similar problem to that outlined in the previous section regarding sampling dependent on the existing sample. The selection probability of each joiner depends on the probability of at least one member of that household being a sample member. Therefore it is necessary to know the joint selection probabilities of all household members, including any who may not themselves be sample members. This approach to sampling population entrants would in any case provide only partial coverage of joiners: only those who co-reside with non-joiners soon after joining would be covered.

### *8.2 Non-response adjustments*

In general, non-response may differ, both numerically and in terms of the correlates of non-response, between each different analysis base. Consequently, a separate non-response weight adjustment may be advisable for each analysis base. For cross-sectional surveys, analysis bases are typically:

- the whole sample (often this is the *only* analysis base);
- respondents to specific instruments (e.g. interview and self-completion);

- units related 1:n (e.g. persons and households, or employees and companies).

But for longitudinal surveys, each of the above can apply to any possible combination of waves. There are $2^t - 1$ combinations of waves at which a sample member may have responded if there are $t$ waves in total, so if $t > 3$, it is not really feasible to create $2^t - 1$ sets of weights (of each type). For example, if $t$=10, then $2^t - 1 = 1{,}023$.

Instead of producing weights for every combination of waves represented in the data set, a practical approach is to produce weights for a limited subset of the possible combinations of waves. The choice of wave combinations should be guided by the main uses of the data. For example, if the main objective of the survey is to estimate change relative to baseline data that were collected at wave 1, then there is very little point in producing weights for combinations of waves that do not include wave 1. If a module of questions on a particular topic is included only at waves 1, 4, 7 and 10, then that particular combination should be a strong candidate for weighting. For almost all longitudinal surveys, the complete set of waves should be one of the combinations for which weights are produced. The only exception would be if, by design, there are no units that were eligible for data collection at every wave. Lynn and Kaminska (2010) suggested some criteria for choosing which wave combinations warrant the production of a weight:

- Survey Design: Any wave combinations not present by design can be ignored (e.g. if survey policy is to not re-approach w1 non-respondents);
- Analytic Use: Some combinations may be more important than others (e.g. if a particular module of questions is included every 4 waves);
- Levels of Non-response: If responding samples differ only very slightly between combinations, the suboptimality in weighting only for the more restrictive combination will be small;
- Correlates of Non-response: A similar argument could be made regarding the predictors of non-response;

- Impact on Estimates: The impact of the above could be assessed by comparing estimates produced with alternative weights. This would be resource-intensive and would depend on the estimates selected.

However, it is important to be aware that carrying out analysis based on respondents to a particular set of waves using weights designed for a different set of waves is sub-optimal. Consider a 3-wave survey and suppose that only one set of longitudinal weights is provided, designed to make the set of persons who responded to all three waves representative of the 3-wave longitudinal population. Suppose we want to estimate some parameter of change between wave 1 and wave 3, for which we only need to use data collected at waves 1 and 3. We could use all units with response patterns XXX or X0X (where X indicates a response and 0 a non-response). But the longitudinal weights will be set to 0 for units with a response pattern of X0X. No longitudinal weight is defined for these units. In consequence, a proportion of the available cases (900 out of 9,296) will be dropped from the analysis because of the unavailability of an appropriate weight. For this estimation, it would have been better to produce a set of weights to represent the X?X population (units that are in the population at the times of both waves 1 and 3 regardless of whether or not they are also in the population at the time of wave 2). These would be non-zero for all units in the XXX and X0X samples.

Another important consideration is that sets of weights are usually produced at several different points in time during the life of a longitudinal survey. Often, this is done after each new wave of data is available as analysts will want to analyse the latest data without waiting until the next wave is completed. Thus, as a minimum, at each wave a set of weights will be produced representing the longitudinal population at all waves to date. This means that ultimately weights will be available for every “attrition sample”. For example, after five waves there will be weights for the X0000, XX000, XXX00, XXXX0 and XXXXX samples. If the survey has a policy of attempting to collect data only from previous wave respondents this will be all the weights that are needed. Otherwise, the task is to identify which other combinations of waves are sufficiently important to warrant the calculation of weights.

Once the relevant longitudinal populations have been identified for which weights are to be produced, it remains to identify a *method* of calculating nonresponse weight adjustments and a set of *auxiliary variables* that will define the weighting classes and weights. The criteria for both the method and the variables are no different in the longitudinal survey case from those for any other kind of survey. Essentially (Lynn, 2004), the objective is to choose a method and a set of variables such that when the method is used to create a set of classes defined by the variables, the resulting classes have the following properties:

- Response propensities vary over the classes;
- Values of key sample statistics (e.g. means, proportions, regression coefficients, etc) vary over the classes;
- Values of key sample statistics are similar for included and excluded units (sampled and not sampled, respondents and non-respondents) within each class.

On a longitudinal survey, we should bear in mind that the key sample statistics will tend to be measures of change and measures of association of other variables with measures of change. This is likely to have important implications for the creation of weighting classes. The auxiliary variables that correlate most strongly with these measures of change may well be survey variables from previous waves – and particularly (often) measures of change in previous periods. For this reason, non-response weighting for longitudinal surveys is often done sequentially. For non-response at wave 1, it is necessary to use data external to the survey as auxiliary data. But from that point on (unless the survey includes responses at later waves from units that did not respond at wave 1) response propensity at subsequent waves can be estimated conditional upon response at wave 1 (or other previous waves). In its simplest form, non-response (NR) weights for the attrition samples could be calculated as follows:

- Weights for wave 1 NR ($w_1$) use auxiliary data external to survey;
- Weights for wave 2 NR conditional upon wave 1 response ($w_{2|1}$) use wave 1 data as auxiliary data. The weight for wave 2 NR is $w_2 = w_1 \times w_{2|1}$;

- Weights for wave 3 NR conditional upon wave 2 response ( $w_{3|2}$ ) use wave 1 and 2 data as auxiliary data (perhaps including measures of change between waves 1 and 2). The weight for wave 3 NR is $w_3 = w_2 \times w_{3|2}$;
- Etc…

This simple form – approach (1) – would obviously have to be amended if units with wave non-response patterns were also to be included in the analyses. An alternative to this simple approach – approach (2) – is to re-weight each wave directly back to wave 1:

- Weights for wave 1 NR ( $w_1$ ) use auxiliary data external to survey;
- Weights for wave 2 NR conditional upon wave 1 response ( $w_{2|1}$ ) use wave 1 data as auxiliary data. The weight for wave 2 NR is $w_2 = w_1 \times w_{2|1}$;
- Weights for wave 3 NR conditional upon wave 1 response ( $w_{3|1}$ ) use wave 1 data as auxiliary data. The weight for wave 3 NR is $w_3 = w_1 \times w_{3|1}$.

There are advantages and disadvantages to these two approaches to attrition weighting:

- Approach (1) uses all available data. Auxiliary data will tend to be more recent to the event of non-response, so may provide stronger prediction. Measures of change can be used as auxiliary variables;
- The nature of non-response may change between waves. Approach (1) has potential to correct for this, whereas approach (2) assumes a fixed model across (some) waves.
- Approach (2) has fewer steps and therefore may tend to introduce less random variation
- The relative advantages of approach (1) will be greater the larger the sample size, as bias, rather than variance, becomes a more important component of overall accuracy of estimates.

Once non-response adjustments have been calculated, the final step in the weighting process is to produce a weight for use in analysis. Consider the example of producing a weight for analysis of a wave 3 attrition sample (XXX). The steps to produce an analysis weight would proceed as follows:

- Assume selection probabilities recorded at wave 1 or wave of entry. Appropriate design weight: *sw*
- Non-response weight derived as above: *nw3*
- Apply *sw* * *nw3* and compare with population distributions on known auxiliary variables. If necessary, calculate post-stratification weight: *pw*
- Then, analysis weight is: *w* = *sw* * *nw3* * *pw*

## 9. Imputation

In any context, the key issues with imputation are the choice of imputation method and the choice of auxiliary data and how they should be used. In a longitudinal survey context, a notable feature of imputation is often that the auxiliary data can include – or even be restricted to – previous period values of the variable for which an imputed value is being sought. In a cross-sectional survey, wealth may have to be imputed on the basis of variables such as income, occupation, sex, age and household composition. But in a longitudinal survey, wealth can be imputed on the basis of wealth at each previous wave. Often, wealth at the previous wave will be a much better predictor of wealth at the current wave than any set of current wave measures. The use of data from previous waves as auxiliary variables in an imputation process is often, therefore, advantageous.

Another issue in imputation that arises only in the longitudinal context is whether imputations should be revised in the light of additional data. For example, after the completion of wave 2 of a longitudinal survey, wealth reported at wave 1 may well be the best available predictor of wealth at wave 2 and therefore the best choice of auxiliary variable for income imputation. But later, when wave 3 data are available, it may turn out that wealth is reported at both waves 1 and 3 for some respondents for whom income was missing (and therefore imputed) at wave 2. For these respondents, use of the wave 1 and 3 data in combination may lead to a better (and different) imputation for wave 2 than use of wave 1 data alone. But by this time analysts will have already been using the wave 2 data with the imputations provided at the time of wave 2. The data provider is faced with a dilemma. The options are:

- Provide revised (hopefully better) imputations to replace the ones provided previously, on the basis that the best possible data should always be provided even if this means that there will be inconsistencies between analyses carried out at different points in time;
- Do not revise any imputations already provided in order to avoid inconsistencies between analyses. If this strategy is adopted, the advantages of using subsequent wave data as auxiliary variables for imputation may lead to the conclusion that imputations should not be made until the subsequent wave: e.g. imputations for wave 2 missing values will not be made available until the release of the wave 3 data. However, this can still lead to inconsistent estimates.
- Provide revised imputations and also continue to provide the original imputations, in order that analysts can check the sensitivity of their results to the imputation procedures.

The issue of whether or not to revise imputations is an important one for longitudinal surveys. Different surveys have adopted different policies and it is difficult to be prescriptive as the best policy will depend on the characteristics of the particular survey in question. However, it is generally felt to be important that the policy should be decided in consultation with users and at the earliest stage possible.

A third issue with imputation for longitudinal surveys is whether to impute values for the measures of levels at each wave, or for measures of change. There are arguments for the latter on the grounds that these are more likely to be the variables in which analysts are primarily interested. However, given the typically wide range of measures of change that may be used by different analysts, it may be preferable to impute the measures of levels (cross-sectional measures) and subsequently to derive measures of change using these imputed values. This will also ensure that different measures of change are mutually consistent. This is the approach taken by all the national household panel surveys currently in operation.

## 10. Microdata Anonymisation

As with any survey data collection exercise, data should be anonymised prior to release to analysts, in order to preserve respondent confidentiality. The principles behind this are set out in the International Statistical Institute's declaration on professional ethics (ISI, 2010). This is easy to achieve: it requires only the use of anonymous identification numbers on the data set, with the key (look-up table) relating these to disclosive identifiers being held in a secure setting, accessible only by those who need it for purposes of managing the data collection and respondent communication activities.

However, a related but more challenging issue is that of disclosure control. It is desirable to take steps to minimize the risk that the identity of any sample member could become known by a person who should not know it, for example a journalist or an estranged spouse.  This therefore applies to public-use data files deposited with data archives. It is not usually deemed necessary for data that is analysed by the primary data collection or management agency, provided the data are held securely. Recent advances in the field of statistical disclosure control for social surveys are reviewed by Shlomo (2010). Disclosure control methods usually involve two steps, a) assessing the risk of identification, based on a set of assumptions and constraints; b) identifying and implementing procedures to reduce any unacceptable risks to an acceptable level. The procedures can include any of all of the following:

- Suppressing certain variables entirely from the data set;
- Perturbing certain variables in a specified way, e.g. adding a random error;
- Rounding or banding certain variables;
- Deriving variables believed to be of interest to researchers and releasing only the derived variables and not the original ones.

Most methods for statistical disclosure control have been developed for cross-sectional data and applied in that context. In the longitudinal survey situation, an additional consideration is that indicators of change can become disclosive variables, in addition to the measures of levels at each wave. There are no explicit methods for dealing with this. It is up to individual data providers to assess the risk posed by observed change in the same way that they would for any measure of level or status at a single wave.

In practice, the majority of survey microdata can safely be released in full provided that the data are anonymised (no identifying information left on the data set) and that geographical identifiers are restricted to relatively high levels (NUTS level 3 is nearly always unproblematic; NUTS level 4 can often be released too, though this depends on the content of the data). Many major surveys now provide data under different arrangements, depending on the degree of disclosure risk. Typically, the majority of the data may be freely available on standard public release files, but more disclosive data, such as small area identifiers or detailed occupational classification for rare occupations, will only be available in a secure setting ("data enclave" or "virtual microdata laboratory") or through some other secure means. For example, in the UK the recently-established Secure Data Service (SDS, http://securedata.data-archive.ac.uk/) supply data through a secure server, to which users can log on from their office desktop PC. The user is unable to download data or disclosive output, but can run analysis directly on the data without having either to submit their syntax to a third party or to travel to a virtual microdata laboratory. The new UK Household Longitudinal Study makes all data freely available through the UK Data Archive, with the exception of national grid references which are available, to 1 metre accuracy, only through the SDS, allowing users to link the survey data to all kinds of geographically-referenced data.

## 11. Identification Variables

With a longitudinal survey, and especially one that may collect data from more than one person per household per wave, it is important to be able to identify on the data set the provider of each data record and the relationships between providers. This enables the data for a person to be linked between waves and the data for members of the same household to be linked at each wave. Household panel surveys typically employ a unique person identification number for each sample person. This number does not change over the time. It will be the same in the data files for every wave. At each wave, every household will be assigned a unique household identification number. These numbers relate only to the specific wave in question and do not have continuity across waves.

The usual way of identifying relationships between household members is to provide two data files for each wave, one at household level and one at individual level. The household level file will contain one record for each household and each record will contain all of the data items that are collected from one person on behalf of the whole household, along with the household identification number. The individual level file will contain one record for each individual, consisting of all data items that are collected for each individual, along with the person and household identification numbers. Even on surveys that primarily collect data that refer to the household, a minimum number of items are usually collected for each household member, e.g. sex, age, economic activity status and relationship to each other household member (or just relationship to a single reference person). The current HFCS collects a few items of this kind in the household listing and demographic sections of the survey instruments. Other data items required for each person on each wave are indicators of field outcome (response, refusal, not contacted, etc) and sample status (original sample member, temporary sample member, new entrant (temporary), new entrant (permanent), deceased, moved abroad, etc). These allow the analyst to determine, and specify, which sample units are eligible for any particular analysis.

Organisation of the data in this way makes it easy to carry out analysis at either individual or household level, or at the level of other units in-between such as couples. An example of how units are identified and how data are organized on a major longitudinal survey can be found in sections III.1 and III.2 of the BHPS User Manual at http://www.iser.essex.ac.uk/bhps/documentation/vola/vola.html.

## References


Bricker J, Bucks B, Kennickell A, Mach T, and Moore K (2011). Surveying the Aftermath of the Storm: Changes in Family Finances from 2007 to 2009. *Financial and Economic Discussion Series*, paper 2011-17. Washington DC: Federal Reserve Board. At: http://www.federalreserve.gov/pubs/feds/2011/201117/201117pap.pdf

Chiappori P A and Paiella M (2008) Relative risk aversion is constant: evidence from panel data. *Department for Economic Studies Discussion Paper* 5/2008, University of Naples, "Parthenope". At: http://economia.uniparthenope.it/ise/sito/DP/DP_5_2008.pdf

Couper M P and Ofstedal M B (2009) Keeping in contact with mobile sample members. Chapter 11 in Lynn P (ed), *Methodology of Longitudinal Surveys*, New York: Wiley

European Central Bank (2011) The Euro-area Household Finance and Consumption Survey: survey mode, oversampling wealthy households and other methods to reduce nonresponse bias. Paper presented at the 59th plenary session of the *United Nations Economic Commission for Europe Conference of European Statisticians*, Geneva, 14-16 June. At: http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2011/48.e.pdf

Fumagalli L, Laurie H, and Lynn P (2010) Experiments with methods to reduce attrition in longitudinal surveys. *ISER Working Paper* 2010-04. At: http://www.iser.essex.ac.uk/workpaps/iser/2010-04

Hurst E, Luoh M C, Stafford F P, and Gale W G (1998) The wealth dynamics of American families, 1984-94, *Brookings Papers on Economic Activity* 1998:1, 267-337.

International Statistical Institute (2010) Declaration on Professional Ethics. At: http://isi-web.org/images/about/Declaration-EN2010.pdf

Jenkins S P (2009) Marital splits and income changes over the longer term. Chapter 13 in M. Brynin and J. F. Ermisch (ed.s), *Changing Relationships*. London: Routledge.

Jenkins S P (2011) *Changing Fortunes: Income Mobility and Poverty Dynamics in Britain*. Oxford: Oxford University Press.

Spiess M, Kroh M, Pischner R, and Wagner GG (2008) On the treatment of non-original sample members in the German Socio-Economic Panel Study (SOEP) – tracing, weighting and frequencies, *SOEP Paper* 98. At: http://www.diw.de/documents/publikationen/73/diw_01.c.82421.de/diw_sp0098.pdf

Laurie H, Smith R, and Scott L (1999) Strategies for reducing nonresponse in a longitudinal panel survey, *Journal of Official Statistics*, 15 (2), 269–282. At: http://www.jos.nu/Articles/abstract.asp?article=152269

Lynn, P. (2004) Weighting. In: Kempf-Leonard, K. (ed.), *Encyclopedia of Social Measurement*. London: Academic Press.

Lynn P (2008) The problem of nonresponse. Chapter 3 in de Leeuw E D, Hox J J and Dillman D A (ed.s) *International Handbook of Survey Methodology*, New York: Lawrence Erlbaum.

Lynn P (2011) Maintaining cross-sectional representativeness in a longitudinal general population survey, *Understanding Society Working Paper* 2011-04. Colchester: University of Essex. At: `http://research.understandingsociety.org.uk/publications/working-paper/2011-04`

Lynn P and Kaminska O (2010) Criteria for developing non-response weight adjustments for secondary users of complex longitudinal surveys, paper presented at *2010 International Workshop on Household Survey Non-Response*, Nürnberg

Mazzocco M, Ruiz C, and Yamaguchi S (2006) Labor supply, wealth dynamics, and marriage decisions. *California Center for Population Research Working Paper* 065-06. At: `http://papers.ccpr.ucla.edu/papers/PWP-CCPR-2006-065/PWP-CCPR-2006-065.pdf`

Rose D (ed) (2000a) *Researching Social and Economic Change: The Uses of Household Panel Studies*, London: Routledge.

Rose D (2000b) Household panel studies: an overview. Chapter 1 in Rose D (ed.) *Researching Social and Economic Change: The Uses of Household Panel Studies*, London: Routledge.

Shlomo N (2010) Releasing microdata: disclosure risk estimation, data masking and assessing utility, *Journal of Privacy and Confidentiality*, 2(1), 73-91. At: `http://repository.cmu.edu/jpc/vol2/iss1/7`