This is the accepted version of the manuscript. The definitive version is published as follows:

Watson, N., and Wooden, M. (2014), 'Re-engaging with survey non-respondents: Evidence from three household panels', *Journal of Royal Statistical Society, Series A (Statistics in Society)*, 177(2), 499-522. [doi:10.1111/rssa.12024]

Re-engaging with survey non-respondents: Evidence from three household panels

Nicole Watson and Mark Wooden

University of Melbourne, Australia

Address for correspondence: Nicole Watson, Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Vic. 3010, Australia.

Email: n.watson@unimelb.edu.au

Summary. Previous research into the correlates and determinants of non-response in longitudinal surveys has focused exclusively on why it is that respondents at one survey wave choose not to participate at future waves. This is very understandable if non-response is always an absorbing state, but in many longitudinal surveys, and certainly most household panels, this is not the case. Indeed, in these surveys it is normal practice to attempt to make contact with many non-respondents at the next wave. This study differs from previous research by examining re-engagement with previous wave non-respondents. Drawing on data from three national household panels it is found that the re-engagement decision is indeed distinctly different from the decision about continued participation. Further, these differences have clear implications for the way panel surveys should be administered given the desire to enhance overall response rates.

Acknowledgements

The research reported on in this paper was supported by an Australian Research Council Discovery Grant (#DP1095497). Central to the analysis are unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey, the British Household Panel Survey (BHPS) and the German Socio-Economic Panel (SOEP). The HILDA Survey is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs and managed by the Melbourne Institute of Applied Economic and Social Research (University of Melbourne); the BHPS was funded by the Economic and Social Research Council and managed by the Institute for Social and

Economic Research (University of Essex); and the SOEP, until 2002 at least, was funded by the German National Science Foundation and Federal Ministry of Education and Research, and is managed by the German Institute for Economic Research. We also thank Peter Lynn, Markus Grabka, Noah Uhrig, the journal's editors and two anonymous referees for their helpful comments. The findings and views reported in this paper, however, are those of the authors and should not be attributed to any of the aforementioned individuals or organisations.

1. Introduction

In the area of survey methods, no issue has arguably attracted more attention from researchers than non-response, and this is even more so for research on longitudinal survey methods. As a result there is now a relatively large body of literature concerned with identifying the correlates and determinants of non-response in longitudinal surveys (for reviews, see Uhrig, 2008; Watson and Wooden, 2009). Virtually all of this research, however, is concerned with understanding why it is that respondents at one survey wave choose not to participate at future waves. This is understandable if non-response at one wave is synonymous with attrition (i.e., permanent exits from the sample), as would occur if no attempts are made to contact non-respondents at subsequent survey waves, but in many longitudinal surveys, and certainly most household panels, this is not the case. Indeed, in these surveys, with the exception of those non-respondents who make explicit requests not to be contacted again, it is normal practice to attempt to make contact with non-respondents at the next wave, often with considerable success. In the data sources examined in this paper, for example, as much as 34% of all wave 1 respondents who drop out of the survey can be re-engaged in a subsequent wave. Success, however, is also highly variable across different studies and samples, leading to the obvious question of what might explain such differences. But independent of the rate of success, the question still remains as to whether the potential exists for longitudinal surveys to substantially raise re-engagement rates. Even more importantly, what are the levers for bringing this about and do they differ from those that have been found to influence attrition?

Such considerations suggest that analyses of the decision to continue participation need to be complemented with analyses of the decision taken by non-respondents to recommence participation. Yet to the best of our knowledge there have been no published studies that have specifically examined re-engagement with longitudinal survey sample members from one wave to the next (though Burton *et al*. (2006) consider re-engagement within a wave). It is this which is the main contribution of this paper. Specifically, we draw on data from three major national household panel surveys to estimate logit models predicting the likelihood that non-respondents at one survey wave can be persuaded to respond at the next.

The use of data from three different surveys is another feature of this paper that sets it apart from most preceding work. Research findings based on a single survey are always subject to the criticism that it can be difficult to know to what extent they can be generalised to other surveys covering different populations and employing different design features. We respond to this by estimating similarly specified models using data from three separate household panel studies conducted at different times in three different countries (Australia, Britain and Germany).

2. Previous Research, Theory and Hypotheses

Driven by the desire to both better target scarce survey resources and to produce weights (or imputed values) that help reduce the extent of bias arising from non-random response, longitudinal survey researchers have long been interested in identifying, and quantifying the impact of, variables that predict why sample members elect not to participate in the next survey wave (or, in many cases, discontinue all further participation in the study). This typically involves estimating a discrete choice probability model of survey response of the form originally suggested by Hausman and Wise (1979), but often extended to allow for many time periods, rather than the two-period form of the original Hausman-Wise specification. In some instances, proportional hazard models, where the outcome variable is the number of years until attrition, have been estimated (e.g., Becketti *et al*., 1988), but for the most part the models estimated have been of the discrete choice variety (e.g., Fitzgerald *et al*., 1998; Nicoletti and Peracchi, 2005; Hawkes and Plewis, 2006; Vandecasteele and Debels, 2007; Watson and Wooden, 2009).

The target sample in this literature is, without exception, respondents to a previous survey wave. This is understandable if non-response is always an absorbing state, and hence survey participation follows one of two patterns – continued participation or monotone attrition. However, as is widely recognised (see Kalton and Brick, 2000; Lynn, 2009), most longitudinal studies are designed to allow a much greater variety of response patterns, both because they permit new entrants into the sample after the initial wave and because attempts are made to contact many non-respondents at the next (and later) waves.

Despite this, only rarely has any explicit consideration been given to the possibility that the magnitude of relationships between response probabilities and hypothesised predictors and correlates might vary with the type of response pattern. Indeed, we are only able to identify a handful of studies that use data spanning multiple waves from longitudinal surveys to test for such a possibility (Burkam and Lee, 1998; Hawkes and Plewis, 2006; Plewis *et al*., 2010). Each of these studies estimate multinomial logit models that distinguish between two types of non-response: (i) attrition; and (ii) wave non-response, or what has also been described as non-monotonic attrition (Burkam and Lee, 1998), arbitrary response patterns (Chambers and Welsh, 1993) or irregular response patterns (Voorpostel, 2009). Nevertheless, and like all other research on non-response to longitudinal surveys, these studies still define their target sample to be respondents at a point in time.

Another feature of most previous research is that it is largely data driven and seemingly not guided by any overarching theoretical framework designed to help explain why response behaviour varies across sample units. Indeed, most empirical studies have tended to focus on the role of respondent characteristics, and less on the features of the data collection process, even though it is the latter that is usually thought to be most amenable to intervention. In reaction to this weakness in the literature, Lepkowski and Couper (2002) proposed a theoretical framework for thinking about survey non-response in a longitudinal setting, which, in turn, drew on the framework proposed by Groves and Couper (1998) for cross-sectional survey participation. In their model, survey participation is a function of both factors that are under the control of the survey manager (survey design and interviewer characteristics) and those that are not (sample member characteristics and the broader socio-economic environment), as well as the nature of the interaction between the sample member and the interviewer. We use this framework to guide our thinking, though unlike them we do not attempt here to distinguish between the different conditional processes of location, contact and cooperation. Further, we focus most attention on those influences that are either under the control of the survey manager or are central in shaping the interview experience.

In line with the empirical literature, the theoretical framework proposed by Lepkowski and Couper (2002) also makes no distinction between previous wave respondents and previous wave non-respondents. The implicit assumption, therefore, is that the processes that influence the decision to participate in a survey are independent of the sample member's participation history. Yet the notion of consistency, a heuristic which Groves and Couper (1998) identify as an important influence on survey response behavior, challenges this assumption. According to this heuristic, persons who agree to an initial request to participate in a survey based on their specific cost-benefit calculation (especially if there is an explicit understanding of the survey's longitudinal nature) will be much more likely to accede to subsequent requests to participate in later survey waves. But this same reasoning also suggests that it will be relatively difficult to convince persons who decide not to participate at one point in time to change their mind at later waves, which, in turn, may require approaching these sample members in a very different manner.

Consider, for example, the role of interviewer continuity. The notion of consistency would suggest that assigning the same interviewers to individual respondents in successive survey waves will be beneficial for response. And indeed this view is both the conventional wisdom, and receives empirical support from a number of studies (see Watson and Wooden, 2009, pp. 162-163), though experimental evidence suggests that the relationship may be more complex than usually assumed (Lynn *et al*., 2011). According to the consistency heuristic, respondents will find it more difficult to refuse a request to participate from the same person with whom they made the initial commitment. In addition, by ensuring interviewers are matched to individual respondents through time enables interviewers to build rapport with these respondents and to develop a relationship of trust (Laurie *et al*., 1999), or, to use the terminology of Groves and Couper (1998), a 'liking' for the interviewer, which in turn enhances the likelihood of ongoing participation. Now consider the situation of sample members who have withdrawn their cooperation at the previous wave. In these situations, sending back the same interviewer at the next wave may be precisely the wrong thing to do. Indeed, in some instances the respondent may have declined to participate because they were uncomfortable with the interviewer. And even in situations where the respondent-interviewer relationship is not a causal factor, persuading a respondent to reverse the decision they took at the previous wave will often require a different approach from that used last wave, which in many cases will be more likely where the interviewer is different.

By a similar logic, a change in survey mode, while it might be harmful for cooperation rates among previous wave responders, may be beneficial for previous wave non-responders. Furthermore, opportunities to develop rapport between interviewer and sample member will vary with survey mode, which in turn will impact on response. We thus expect that face-to-face contact will be associated with higher response rates than where contact is attempted by mail or telephone.

The consistency heuristic also suggests that the likelihood of persuading a previous wave non-respondent to reverse their previous decision will be more likely the sooner the attempt to convert them is made.

The reason for the previous wave non-response is also likely to have a bearing on re-engagement rates. Again the consistency heuristic leads us to expect less success when dealing with sample members who have explicitly refused to participate, as compared with sample members who did not participate for other reasons (such as ill-health or a failure to make contact within the interview period) given those other reasons are more likely to be of a temporary nature.

A further consideration is the role that other household members may play. The presence of other household members who continue to cooperate with the survey may have a tendency to favourably influence the non-respondent's attitude towards the study, in the same way that non-responders have been found to adversely affect the likelihood that responding sample members will continue to respond (Watson and Wooden, 2009).

And while the focus of this paper is not on individual (or household) characteristics, it should not necessarily be assumed that such characteristics will always be associated with response in the same way. Perhaps the most obvious example here is geographic mobility. It has been well established that movement is a major risk factor for locating survey respondents at the next survey wave (e.g., Gray *et al*., 1996; Lepkowski and Couper, 2002; Plewis *et al*., 2008; Watson and Wooden, 2009; Fumagalli *et al*., 2010). Less clear is whether mobility will also be detrimental to re-engagement with survey non-respondents. While every move arguably makes the task of locating sample members more difficult, it needs to be borne in mind that previous wave non-respondents have already disengaged with the study, and in some cases this may have been because of previous moves that took them beyond the easy reach of interviewers. Subsequent moves may thus have the opposite effect. In addition, moves are often associated with changes in household composition, which might mean some sample members shift into environments that are less hostile to survey participation.

In summary, we hypothesise that:

(i) interviewer-respondent continuity, on average, will not be beneficial for response when re-engaging with previous wave non-responders;
(ii) changes in survey mode will be associated with lower response rates among previous wave respondents but not among previous wave non-respondents;
(iii) face-to-face interviewing will be associated with higher response rates among both previous wave respondents and non-respondents (when compared with other survey modes);
(iv) the likelihood of success at converting a previous wave non-respondent will decline with the length of time the sample member has been a non-participant;
(v) success in converting previous wave non-respondents will be correlated with the reason for the previous wave non-response, with success rates being much lower for refusers than for those with other reasons for not responding;
(vi) partial household response in the prior wave will be associated with higher response rates among previous wave non-respondents and lower rates among previous wave respondents; and
(vii) the magnitude and direction of relationships with respondent characteristics, and more specifically changes in address, may vary with whether the sample member responded at the previous wave or not.

3. The Household Panels

The research reported on here makes use of data from three national household panel surveys: (i) the Household, Income and Labour Dynamics in Australia (HILDA) Survey; (ii) the British Household Panel Survey (BHPS); and (iii) the German Socio-Economic Panel (SOEP). Given the focus here on the importance of survey design in influencing response outcomes, it is important to have an understanding of how the design and management of these studies differ. In what follows we, therefore, briefly describe the key common design features and the main areas of contrast of these surveys. This is then followed by a presentation of summary data describing the patterns of response over time reported by each. Our analysis is mostly restricted to the experience over the first nine waves, a decision dictated by the age of the youngest of the three – the HILDA Survey – which only commenced fieldwork in 2001.

3.1 Key Features

While the three studies are conducted independently using different survey instruments and different fieldwork procedures, they have many common design features that aid a comparison. The key similarities (and differences) are summarised (in the context of the Cross-National Equivalent File) by Frick *et al.* (2007). Of most significance, each of these studies:

i. collects information on a broad range of socio-economic variables, but with a focus on household composition, income, employment and housing;

ii. commenced with a population sample that was clustered by location and was intended to be broadly representative of the national population resident in private households;
iii. conducts survey waves on an annual basis;
iv. follows members of the original sample households and their offspring;
v. interviews all adult members of the household (15 years or older in the HILDA Survey, and 16 years or older in both the BHPS and SOEP);
vi. augments the sample in later waves with persons who join a sample household (though in the case of both the HILDA Survey and the BHPS these persons typically only remain in the sample for as long as they co-reside with a 'permanent' sample member); and
vii. attempts to contact persons who have been non-respondents at previous waves.

There are, however, important differences in the studies, including in survey design and fieldwork procedures. First, the studies differ in age, with the SOEP commencing in 1984, the BHPS in 1991 and the HILDA Survey, as already noted, in 2001.

Second, survey modes vary. The HILDA Survey and BHPS are mostly conducted in-person by an interviewer. Both the HILDA Survey and the BHPS also provide for telephone interviews, though the BHPS telephone interview is much shorter than the full instrument used in face-to-face interviews and is only delivered as a method of last resort with the main purpose of retaining the sample member (see Laurie *et al*., 1999). In contrast, the SOEP permits much greater flexibility in the choice of mode across respondents. Thus an oral interview is used in some cases and a self-administered instrument (distributed by the interviewer or by mail) in others. Importantly, the SOEP tends to rely heavily on mailed-out questionnaires when dealing with previous wave non-respondents.

Third, while all three studies provide respondents with small gifts to encourage a sense of reciprocity and social interaction between the respondent and interviewer, they vary substantially in the use of cash and cash-like incentives (Laurie and Lynn, 2009). The SOEP provides respondents with a lottery ticket (which in 2005 was valued at around €1.50, equivalent to about £1), the BHPS provides a store gift voucher (worth £10 in 2005), and the HILDA Survey provides cash (in 2005 this involved the payment of A$25, equivalent to about £11 at the time, to each respondent who completed an interview, and a further A$25 household bonus if all household members respond).

Fourth, the interviewer workforces are managed quite differently. The HILDA Survey is distinct in breaking its fieldwork into three phases, with the workloads increasingly concentrated on the most successful and experienced interviewers in the later phases. Non-responding cases are thus handled by multiple interviewers during the course of the fieldwork. In contrast, in the BHPS and SOEP all cases are usually handled by the one interviewer. The size of the workloads are also quite different, with the average face-to-face interviewer on the HILDA Survey handling anywhere from 51 to 66 households each survey wave. This compares with 25 to 30 households for face-to-face interviewers working on the BHPS and just 12 to 18 households for the SOEP.

Finally, each study adopts different practices in determining when a refusal to cooperate in one wave constitutes a request to be permanently removed from the sample, and when to cease pursuing serial non-responders. This is reflected in the different proportions of the non-responding sample that are issued to field at the next wave. This can range from as much as 78% in the early waves to less than 10% in later waves, with the SOEP being much less likely to attempt to re-contact non-respondents than either the BHPS or the HILDA Survey. We will return to this issue in the next section.

3.2 Response Rates and Patterns

Given the differences discussed above, together with the obvious differences in populations and timing, it would only be coincidental if response rates were similar across the three studies. And indeed the initial response rates at wave 1 are different, varying from just 66% of all households in the HILDA Survey (measured as the percentage of all in-scope households providing at least one adult respondent) to 74% in the BHPS. In the SOEP, wave 1 response rates of 61% and 68% are reported for their two original samples (the Western German or "A" sample and the foreigner or "B" sample), but these rates are not directly comparable with the rates reported by the other two studies given that the SOEP only included households in their wave 1 responding sample if interviews were obtained with all eligible adults in the household. Consistent with evidence from other surveys (e.g., De Leeuw and De Heer, 2002), the German experience also suggests that obtaining cooperation from new sample members is becoming increasingly difficult over time, with the initial household response rates (now defined as any household providing at least one person interview) to the nationally representative refreshment samples introduced in 1998 (sample "E"), 2000 (sample "F") and 2006 (sample "H") falling to 54%, 52% and 41%, respectively. Partly for this reason, we include the largest of these refreshment samples – sample F – in all of the analyses that follow.

Turning to non-response at later waves, in Figure 1 we report the wave-on-wave response rates (proportion of responding persons from one wave who respond at the next, after excluding deaths and movements out of scope) for the life of the initial samples of all three studies, together with the F sample of the SOEP. As can be seen, with the exception of SOEP F, the annual response rates follow a similar pattern, rising from around 87% in wave 2 to around 95% by wave 8. Thereafter response rates tend to plateau and, after enough years, may even start to decline. The experience of the more recent SOEP F sample appears to be somewhat different, with the response rates at all waves being much lower than in the other surveys and samples.

Figure 2 provides some indication of what these rates of response mean for sample retention, by reporting the percentage of wave 1 respondents interviewed at each subsequent wave until wave 9. As can be seen, the experiences of the HILDA Survey and the BHPS have been close to identical. Non-response in the SOEP A and B samples also initially followed a similar path to the other two studies, before diverging from wave 5 onwards. The high rates of non-response to the more recent SOEP F sample, however, lead to retention rates for this sample that are lower at all points in time, with less than 50% of the original sample of respondents still participating at wave 9.

Figure 2, however, still tells us little about the pattern of response, and, more specifically, the extent to which non-response is associated with sample attrition or is only a temporary phenomenon. Table 1 thus presents summary data which take respondents from wave 1 and divides them into four broad groups based on the observed patterns of response over the first eight waves. These four groups are: (i) continuous participation (i.e., persons who responded in all eight waves); (ii) monotone attrition (i.e., persons who withdrew cooperation at one wave and never participated at a subsequent wave); (iii) other response patterns, which can be thought of as comprising two main groups, intermittent responders and returners (persons who withdrew cooperation for a wave or two but were then subsequently persuaded to resume their participation); and (iv) persons who either died or moved abroad during the first eight waves. Note that the observation of response patterns is potentially truncated at wave 8.

That is, it is not always obvious when a non-responding case is a case of monotone attrition or not. We have reserved one wave of data (wave 9) to help determine the status.

Table 1 shows that continuous participation over the 8-wave window varies from a low of 45% for the SOEP F sample up to 59% for the BHPS. Of the non-response patterns, monotonic attrition is clearly the most important. Nevertheless, more variable response rates are also reasonably common, especially in the HILDA Survey and the BHPS, where they account for 12.0% and 9.7%, respectively, of all wave 1 respondents. While these fractions may not seem large, they represent around one in every three sample members who fail to respond to at least one survey wave (after excluding sample members who either died or moved out of scope). In the SOEP such variable response patterns are less common, reflecting the greater tendency in this survey for non-respondents at one wave to be treated as panel attritors.

More direct evidence on the extent to which the different panels have been successful in re-engaging with previous wave non-respondents is provided in Table 2. In this table we again use the sample of persons who responded at wave 1 (after excluding respondents who either died or moved out of scope), but now report the fraction of persons who did not respond at the previous wave but did respond in the current wave. Thus we can see that the HILDA Survey was successful in converting almost 22% of (in-scope) non-respondents in wave 2 to responding status in wave 3. This proportion then tends to gradually decline with each successive wave.

The decline in this rate over time is driven, at least in part, by the gradual rise in the number of cases that the survey managers determine are no longer worth pursuing. As shown in Figure 3, the SOEP is much more averse to chasing non-respondents than the other two studies, with the proportion of non-respondents issued to field the next wave falling to below 20% by wave 4, and to less than 10% by wave 7. The BHPS and HILDA Survey, on the other hand, commenced with very high non-respondent issue rates; in the order of 75 to 78%. These rates then fell quite rapidly over the following waves, before plateauing at around 33% in the BHPS and 25% in the HILDA Survey. As a result of the differences in the proportion of non-respondents re-issued to field, we find, in the bottom panel of Table 2, that when we restrict the denominator to non-responding cases from the previous survey wave who were issued to field in the current wave, the pattern of a decline over time is less obvious.

In summary, despite some noticeable differences in the way the surveys are administered, the longitudinal response rates follow similar patterns in each of the four samples considered. The similarity in the experiences of the HILDA Survey and BHPS is especially remarkable. Re-interview rates in the SOEP samples, however, have clearly been below those subsequently achieved in both the HILDA and BHPS samples.

4. Modelling survey re-engagement

We now move to an examination of the factors associated with within-sample differences in the success in re-engaging survey respondents, and how these influences vary across the four samples under consideration.

Let p_{hit} be the probability that $R_{hit} = 1$ given $R_{hit-1} = 0$, where R_{hit} is a dichotomous variable indicating whether sample member i in household h responds to the current survey wave t or not. The model that we wish to estimate then takes the form:

$$p_{hit} = F(\alpha + Z_{hi}\delta + X_{hit}\beta + \upsilon_h + \varepsilon_{hit});\ \ h = 1, \ldots, N;\ \ i=1,\ldots n_h;\ \ t = 3, \ldots, 9 \qquad (1)$$

where $F(w) = e^w / (1 + e^w)$ is the cumulative logistic distribution, Z_{hi} is a vector of fixed observable individual characteristics and X_{hit} is a vector of time varying characteristics. Since sample members are clustered within wave 1 households (*h*), we allow for a random effect (υ_h) to be associated with all individuals belonging to the same wave 1 household, and assume these random effects are independent across households and are normally distributed with a mean of zero. The combined effect of any omitted time invariant household-specific covariates that cause some individuals in that household to respond is reflected in this random intercept term. This model is estimated in STATA using the adaptive Gauss-Hermite quadrature method (as discussed in Rabe-Hesketh and Skrondal, 2008, pp. 258-261), with the number of quadrature points set to ensure stability in the parameter estimates. (The number of quadrature points for each model are 24 for the HILDA sample, 48 for the BHPS sample, 36 for the SOEP AB sample and 48 for the SOEP F sample. These were chosen to ensure the parameter estimates changed by less than 0.01% compared to models with a third more or less quadrature points.) Note that since our sample is restricted to respondents at wave 1 who did not respond at a subsequent wave, our observations span a maximum of seven periods beginning with wave 3.

A potentially major weakness in our approach is that the four samples used for this analysis are all restricted to non-respondents at *t-1* who are *issued to field* for the current wave (time *t*), and the decision about which cases to issue to field is non-random. Survey managers make judgments about the relative benefits and costs of attempting to interview non-responding cases, with those cases deemed to be high cost (which will be inversely correlated with the expected likelihood of response) most likely to not be issued to field. This suggests that we should also be modelling the decision about which cases are issued to field. Unfortunately, it is not obvious how we could identify this two-equation system, with the variables that might influence the decision to issue also expected to directly affect response probabilities. As a result, some of our estimates may be affected by sample selection bias. We also expect the extent of bias to be more pronounced in the SOEP samples and least pronounced in the BHPS sample given the marked differences across studies in re-issue rates.

Another weakness with our approach is that while each of the samples analysed here provide considerable detail about its respondents, we do not have data on time varying characteristics for survey waves in which sample members do not respond. We thus simplify the analysis by measuring most individual characteristics at wave 1. This, of course, is not a problem for time invariant characteristics, such as sex, but is potentially quite restrictive for variables that vary markedly over time, such as employment status. Nevertheless, our approach ensures the variables are measured on a consistent basis irrespective of the response history of sample members.

A list of the explanatory variables used in our analysis is provided in Table 3 (together with summary descriptive statistics). Most important are the variables used to test our hypotheses.

First, we include simple dummy variables that control for both the effect of interviewer continuity and interviewer experience. While interviewer experience does not feature in our hypotheses, it nevertheless seems important to identify not only whether the assigned interviewer changes between waves, but among cases where the interviewer does change whether that interviewer has prior survey experience or not. Most obviously it is widely accepted that experienced interviewers are generally more successful in obtaining cooperation from sample members than inexperienced interviewers (for supporting evidence see: Couper, 1991; Groves and Couper, 1998; Durrant *et al*. 2010). For this analysis, we are only able to identify whether the interviewer has worked on the survey in previous waves; we do not have measures of general experience (such as years employed as an interviewer). This survey-specific measure of interviewer experience, however, may well be the more appropriate measure given the type of survey work undertaken by many commercial market research companies (e.g., cross-section opinion polls and market research surveys) is a relatively poor complement to the type of survey work involved in an academic research-led household panel.

We also expect interviewer continuity to be constrained by sample member mobility. Our final specification thus includes a series of dummy variables that separately identify whether the assigned interviewer is the same as in the preceding wave, is a different but experienced interviewer or is a different but inexperienced (or 'new') interviewer, and which are interacted with a variable identifying whether the sample member changed addresses between waves.

Second, we test for mode effects by including variables indicating whether the assigned interviewer was a face-to-face interviewer or a telephone interviewer, and whether survey mode had changed since the prior wave. These variables, however, are only relevant for the HILDA and SOEP samples. As noted earlier, while the BHPS does use the telephone to conduct interviews, this mode is only used as a last resort, and even then a much shortened version of the survey is conducted (averaging just under 7 minutes in wave 3 – the first year it was introduced – compared with over 40 minutes for the full interview when delivered in person). We have thus treated all telephone responses in the BHPS as non-responses. Further, it needs to be borne in mind that 'telephone interviewers' in the SOEP generally only mail-out a questionnaire for the sample member to complete conditional on obtaining an indication of willingness to co-operate; they typically do not conduct interviews.

Third, we include dummy variables indicating the number of years of successive non-participation. In interpreting the effects of these variables, however, bear in mind that our samples are restricted to prior wave non-respondents who are issued to field. The decision to issue cases is, in turn, a function of how long a sample member has ceased to cooperate. As a result, and as shown in Table 3, we can see that relatively few sample members are identified as having more than three successive waves of non-participation.

Fourth, we include a dummy variable indicating whether or not another household member had responded at the previous survey wave.

Fifth, we include indicators of the type of non-response from the previous wave. It should be noted, however, that the categories of non-response are not directly comparable across the three surveys. The SOEP data, for example, describes many non-responses as the result of 'processing problems'. No such analog exists for the other two studies, and nor is it entirely obvious what this category covers. Very differently, the BHPS data includes responses obtained by proxy interview and by telephone interview, which we have treated as non-responses given so little information is collected when these modes are used. The proxy interviews are classified as 'other' non-response as we do not know the reason a proxy was used. The telephone interviews are separately classified as these occur at the end of the fieldwork period as part of a refusal conversion process that may eventually lead to a full interview in later waves.

We also include a set of control variables comprising respondent (and observable household) characteristics and a limited number of other fieldwork characteristics. The choice of respondent characteristics is guided by previous research (especially as summarised in Watson and Wooden, 2009) but restricted by the desire to have broadly comparable measures for all four samples. We thus include controls for sex, age (specified as a quadratic), whether native or foreign born, education (a simple dummy variable indicating whether the respondent had completed a university level qualification), employment status, the presence of work limiting health conditions, household composition (number of adults and number of children in the household), home ownership, equivalised gross annual household income (measured in the relevant local currency, and specified as a quadratic, with the equivalence scale used being the simple 'modified' scale recommended by the OECD), and marital / relationship status.

All of the respondent characteristics mentioned to this point, except for age, are measured at wave 1. Apart from age, the only other time-varying individual characteristic included in our model is whether the respondent had changed address (either to a traced or untraced location) since the previous survey wave, which as noted above is interacted with interviewer continuity and experience.

We also include a measure of interviewer workload (the number of households allocated to the interviewer at the start of fieldwork). As discussed in Watson and Wooden (2009), a priori the effect of this variable is uncertain. Nicoletti and Buck (2004) argue that large workloads will be synonymous with 'overwork' and thus might be expected to be associated with relatively poor response outcomes. On the other hand, if larger workloads are allocated to better interviewers then the opposite correlation may be observed. Readers will also observe from Table 3 that average workloads seem very different across the three studies; indeed in the SOEP it appears to average over 400 interviews per wave per interviewer. These very large numbers, however, are a reflection of the high reliance SOEP places on in-office telephone interviewers for dealing with previous wave non-respondents, which in turn is conducive to high volume (but typically with poor returns).

Finally, we include wave dummies which will help control for wave-specific variations in fieldwork procedures as well as any general change in respondent attitudes that evolve over time. These dummies should also control for the effects on the probability of engagement that might arise from any change in the extent of selectivity within the sample.

5. Results

5.1 Models of Survey Re-Engagement

Logit estimates and mean marginal effects for the key variables of interest in our models of survey re-engagement are presented in Table 4. The mean marginal effect is the difference in the predicted probability of response given a one unit change in the covariate while fixing all other covariates equal to their observed values. Since our models are non-linear and include both fixed and random components, the estimated predicted probability is calculated by summing the linear predictors of the fixed and random components (being $\widehat{w}_1 = \hat{\alpha} + Z_{hi}\hat{\delta} + X_{hit}\hat{\beta}$ and $\widehat{w}_2 = \hat{\upsilon}_h$ respectively) and then converting this total ($\widehat{w} = \widehat{w}_1 + \widehat{w}_2$) to a probability metric via $\hat{p}_{hit} = e^{\widehat{w}}/(1 + e^{\widehat{w}})$.

The first point to note is the relatively high proportion of the residual variance attributed to household effects, as reflected in the rho (ρ) parameters. The proportion of the residual variance attributed to household effects is 45% in both the HILDA Survey and the BHPS, 70% in the SOEP AB sample and 76% in the SOEP F sample. It could be argued that since we are using individual-level panel data we should also allow for individual-specific effects. Estimation of three-level models that incorporate both household and individual random effects, however, found no evidence of a significant individual random effect. In models fitted to all four samples, only the household random effect is ever significant, and hence we only report results from models incorporating a household effect.

We use the results presented in Table 4 to assess the seven hypotheses developed earlier. As interviewer continuity is constrained by sample mobility we assess the first and seventh hypotheses together. We hypothesised that interviewer continuity would not be beneficial in re-engaging previous wave non-respondents but that sample member mobility may be. We find evidence to support both of these hypotheses in the HILDA Survey, BHPS and SOEP F samples. Where the household has not moved, sending a different experienced interviewer rather than the same interviewer increases the mean predicted probability of response among previous wave non-respondents by values ranging between 4 percentage points, in the HILDA sample, and 10 percentage points, in the BHPS sample. Sending a new interviewer can also achieve a higher conversion rate among non-movers than sending back the same interviewer as in the previous wave, a result that was statistically significant in two of the four samples analysed. Individuals that move are also more likely to be persuaded to recommence participation, but only in the HILDA, BHPS and SOEP F samples; there is no evidence of this in the SOEP AB sample. In terms of the type of interviewer who attempts a conversion following a move, the experience of both the HILDA Survey and BHPS (but not the SOEP) is that a different interviewer is more likely to achieve an interview regardless of whether they are new to the study or not. That is, for this type of sample member there appears to be no return to survey-specific experience.

One potential concern in interpreting these results is that when deciding which interviewers to allocate to respondents each survey wave, survey managers may attempt to match the least cooperative cases (i.e., previous wave non-respondents) with their best performing interviewers. As a result, what we have interpreted as being a return to a change of interviewer may just reflect differences in interviewer abilities. In part, we have attempted to control for this by distinguishing between experienced and inexperienced interviewers. Further, and more importantly, in all of these studies interviewer allocation is constrained by costs – interviewers are usually matched to areas close to where they reside.

Our second hypothesis was that response rates among previous wave non-respondents would be improved by changes in interview mode. We find that this is not the case. Changes in interview mode (determined on the basis of initial assignment) appear to have no beneficial effects, and indeed in the case of the SOEP F sample a change in interview mode is harmful for re-engagement probabilities.

Consistent with our third hypothesis, telephone interviewers are far less successful than face-to-face interviewers in persuading non-respondents to recommence participation. Indeed, the mean predicted probability of a telephone interviewer converting a previous wave non-respondent is between 9 and 34 percentage points lower than face-to-face interviewers. The magnitude of this effect is largest in the SOEP (and especially SOEP F), which we believe reflects the fact that telephone mode in this survey usually means self-completion of a paper form that has to be returned by mail.

One potential problem with interpreting these findings with respect to survey mode and change in survey mode is that the choice of mode is usually not exogenous. Of most concern is the possibility that the hardest-to-get case are deliberately assigned to telephone. This is perhaps most obvious with respect to SOEP, where a relatively large proportion of previous wave non-respondents are allocated to "phone" each year. This criticism, however, will be less valid in the case of the HILDA Survey where the major factor driving changes in survey mode is the relocation of sample members to locations not covered by the network of face-to-face interviewers.

In support of our fourth hypothesis, we find that the likelihood of re-engagement tends to decline with the number of successive waves of non-participation. This is most obvious with respect to the BHPS sample, where the likelihood of re-engagement declines more or less in a linear fashion with the number of waves a sample member has been a non-respondent. The estimated patterns are less clear for the other samples, though in all cases except the SOEP AB sample, the results indicate that the greatest impact on re-engaging non-respondents is felt early on, implying that attempts at re-engagement will be most effective after just one wave of non-response.

Also as hypothesised, persons who refused to participate at a previous wave are typically less likely to rejoin the survey than other types of non-respondents. This finding is consistent with research on the effectiveness of within-wave conversion (Burton *et al*., 2006). The magnitude of this effect, however, varies substantially across studies. Success in converting non-contacts ranges from just 5 percentage points higher than refusal conversions, in the case of the SOEP F sample, up to 31 percentage points higher, for the HILDA Survey. For persons giving ‘other’ reasons for non-response the advantage over refusal conversions also varies widely – from 4 percentage points to 24 percentage points – though in both the SOEP samples the estimates are very imprecise (and hence statistically insignificant). As mentioned earlier, the short telephone interviews in the BHPS are used as a tool to retain the sample and the results show that nearly a quarter of those completing a short telephone interview in one wave provide a full interview in the next wave.

Relatedly, we also tested whether particular types of non-response are more likely to be followed by the same type of non-response in the subsequent wave. This involved the estimation of a multinomial logit model (assuming response is correlated across individuals from the same wave 1 household) where the outcome variable identified the different types of response outcomes. For the most part, the findings strongly supported the hypothesis of persistence across survey waves in type of response outcomes. In the interest of brevity, however, we have not reported the results of these separately estimated models here.

In testing the sixth hypothesis, we find no evidence that non-respondents from partially responding households are more likely to respond at future waves. The most likely explanation for this result lies in decision-making by survey managers and fieldwork companies about which households to issue to field each wave. It is almost inevitable that survey managers would issue households in the current wave if at least one person responded in the last wave, irrespective of the strength of refusals of other household members. Very differently, in households where nobody responded last wave, this decision to issue to field is based on the likelihood of re-interviewing at least one household member. We also find that in the case of the SOEP, non-respondents from partially responding households are actually less likely to respond at future waves. This suggests the fieldwork procedures in SOEP are somewhat more protective of the ongoing co-operation of the other household members than the other studies.

It is also worth noting that while wave dummies (the estimates of which are not reported in Table 4) were jointly significant in all samples except SOEP F, the estimated coefficients did not suggest any obvious pattern or trend over the duration of the panel. There was thus no evidence to suggest that the probability of re-engagement was changing over time in a systematic way (at least once other influences were controlled for), as might have been expected if, for example, the samples were become increasingly selective on characteristics associated with the likelihood of response.

Finally, we also estimated additional specifications which included interaction terms between fieldwork variables and respondent characteristics, as well as between selected fieldwork variables. Most interactions were either insignificant or had at most small effects on the estimated coefficients on the key variables of interest. A few interactions moderated the effect of the of the fieldwork variables but did not alter the direction of the effect. Among the interactions between fieldwork variables that are of most interest is the interaction between interviewer continuity and the type of non-response. Specifically, we re-estimated the models reported in Table 4 after including a variable that interacted different (or new) interviewer with whether the outcome at the previous wave was a refusal (as distinct from any other form of non-response). *A priori*, it might be expected that the benefit from assigning a different interviewer to a non-respondent from the previous wave would be greatest where that sample member had explicitly refused at the previous wave. And this is exactly what is found in the HILDA Survey sample – a very large positive and highly significant coefficient is estimated. But again our key finding is one of diversity across studies. For the BHPS the comparable estimate was both modest in size and insignificant, while for the SOEP the estimates were negative, and in the case of SOEP AB large and highly significant. Again we can point to differences in fieldwork procedures and protocols as explanation for the differences. In particular, if previous wave non-respondents tend to be re-issued where there is some commitment from the non-respondent to participate at a later wave, as is most likely in the SOEP samples, then it follows that the better strategy would be to send back the interviewer with whom this agreement was made.

5.2 Comparisons with Models of Survey Continuation

The question remains as to whether these results from modeling the decision by survey non-respondents to recommence survey participation are distinct from those obtained from more conventional models of the decision by respondents to continue or discontinue participation. We thus estimated logit models of whether or not survey respondents at time *t-1* are still participating at time *t* using a specification analogous to that reported in Table 4. The model set up is thus the same as in equation (1), with the exception that response ($R_{hit} = 1$) is conditional on response at the previous wave ($R_{hit-1} = 1$) rather than on non-response. The list of explanatory variables has also been restricted to be as close as possible to that used in our model of re-engagement. There are, however, three obvious differences. First, there is no need for measures of the previous wave outcome in the models of continuation (since all sample members were previous respondents). Second, the closest analog to the variable denoting whether an interview was obtained with another household member at the previous wave is one indicating whether at the previous wave any other eligible household members did not participate. Third, the number of consecutive waves as a non-respondent has been replaced by the number of consecutive waves as a respondent.

As with the models of survey re-engagement, the proportion of the residual variance attributed to household effects in these models of survey continuation is very high – 66% in the HILDA sample, 63% in the BHPS sample, 83% in the SOEP AB sample and 72% in the SOEP F sample. Indeed, for all samples except SOEP F, these proportions are considerably higher than those observed for re-engaging with the sample member. Again we tested a three-level model which included both household and individual random effects and again found that only household random effects were significant.

The detailed results are reported in Table 5, and it is a straightforward matter to compare the estimated coefficients reported in this table with the relevant coefficients in Table 4. One major point of difference is immediately apparent – the very different roles played by interviewer continuity, experience and sample member mobility. As expected, both interviewer continuity and experience are beneficial for continuation probabilities. When the household has not moved, changing to a different experienced interviewer can reduce the response rates by between 1 percentage point, as observed in the HILDA Survey, and 5 percentage points, based on the SOEP F sample. An interviewer who is completely new to the study will have a similar effect. When the household moves, the probability of re-interviewing the sample member can fall by a further 2 to 12 percentage points due to the difficulties in tracing the sample members to their new locations. In contrast, we observed very different associations with re-engagement probabilities. Maximising the chances of recruiting a non-respondent back into the panel appears to typically require assigning a different interviewer.

We also find, as hypothesised, that a change of survey mode reduces the likelihood of response, as does being approached by a telephone interviewer. Further, previous wave respondents who are part of a household that did not fully cooperate in the previous wave are less likely to participate at the current wave.

6. Discussion

While all longitudinal surveys justifiably focus their resources on encouraging survey respondents to continue to respond at later waves, non-response is both unavoidable and accumulates over time. While high response do not necessarily guarantee a reduction in nonresponse bias (see Groves and Peytcheva, 2008), maintaining the viability of a panel typically requires persuading some non-responding sample members to re-engage with the survey. Despite this, we are unaware of any previous research that has modelled re-engagement with survey non-respondents at later survey waves. Instead, research into longitudinal survey non-response and attrition has focused exclusively on survey respondents and the differences between those who respond at future survey waves and those who do not. This perhaps would not be a significant shortcoming if the factors that influence the decision to continue participation were the same factors that influence the decision to re-engage, with effects of broadly similar magnitude. The key insight from this study is that such an assumption is not warranted, and that the re-engagement decision is distinctly different from the decision about continued participation.

Our findings also have clear implications for survey management practices. First, when approaching previous wave non-respondents, survey managers should consider the feasibility of ensuring that interviewers are different to those assigned to these sample members in previous waves. Or at a minimum, they should ensure that the approach adopted in the previous wave that proved unsuccessful is fully documented and that the interviewer when returning to the same household has a tool bag of several fresh approaches that could be used for this wave. Second, efforts at re-engaging with non-respondents need to be initiated as soon as possible (that is, at or prior to the very next wave). Third, the cases likely to be most productive in re-engaging attempts are the non-contacts and persons who gave other non-refusal reasons for non-response (such as poor health). Nevertheless, refusals constitute a sizeable portion of previous wave non-respondents who are issued to field – ranging from 38% in the BHPS to 77% in the HILDA Survey – and effort spent converting these cases is still worthwhile, with conversion rates of between 13% (in the BHPS) and 29% (in the SOEP AB sample). Fourth, though this is hardly news for experienced survey managers, the benefits from continuing to track movers, even if over quite a long period, may be worth the costs. Many of these persons can be easily recruited back into the panel if they can be found. That said, it needs to be recognised that persons who drop out from the survey in one wave are likely to do so again. The proportion of re-engagers who drop out at the very next survey wave, for example, varies from a low of 12%, in the case of the BHPS, to 37% in the SOEP AB sample.

Finally there is an important caveat to this research. As mentioned earlier, the samples used for this analysis are restricted to cases that are issued to field for the current wave, and the decision to issue to field will not be independent of response behavior in past waves. Some of our estimates may therefore be affected by selection bias.

References

Becketti, S., Gould, W., Lillard, L. and Welch, F. (1988) The Panel Study of Income Dynamics after fourteen years: an evaluation. *J. Labor Econ.*, **6**, 472-492.

Burkam, D. T. and Lee, V. E. (1998) Effects of monotone and nonmonotone attrition on parameter estimates in regression models with educational data. *J. Hum. Resour.*, **33**, 555-574.

Burton, J., Laurie, H. and Lynn, P. (2006) The long-term effectiveness of refusal conversion procedures on longitudinal surveys. *J.R. Statist. Soc. A*, **169**, 459-478.

Chambers, R. and Welsh, A. (1993) Log-linear models for survey data with non-ignorable non-response. *J.R. Statist. Soc. B*, **55**, 157-170.

Couper, M. P. (1991) Modelling survey participation at the interviewer level. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 98-107.

De Leeuw, E. and De Heer, W. (2002) Trends in household survey response: a longitudinal and international comparison. In *Survey Nonresponse* (eds. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little), pp. 41-54. New York: John Wiley and Sons.

Durrant, G. B., Groves, R. M., Staetsky, L. and Steele, F. (2010) Effects of interviewer attitudes and behaviors on refusal in household surveys. *Public Opin. Q.*, **74**, 1-36.

Fitzgerald, J., Gottschalk, P. and Moffitt, R. (1998) An analysis of sample attrition in panel data: the Michigan Panel Study of Income Dynamics. *J. Hum. Resour.*, **33**, 251-299.

Frick, J., Jenkins, S. P., Lillard, D. R., Lipps, O. and Wooden, M. (2007) The Cross-National Equivalent File (CNEF) and its member country household panel studies. *Schmollers Jarhrbuch*, **127**, pp. 627-654.

Fumagalli, L., Laurie, H. and Lynn, P. (2010) Experiments with methods to reduce attrition in longitudinal surveys. ISER Working Paper no. 2010-04. Colchester (UK): Institute for Social and Economic Research, University of Essex.

Gray, R., Campanelli, P., Deepchand, K. and Prescott-Clarke, P. (1996) Exploring survey non-response: the effect of attrition on a follow-up of the 1984-85 Health and Lifestyle Survey. *The Statistician*, **45**, 163-183.

Groves, R. M. and Couper, M. P. (1998) *Nonresponse in Household Interview Surveys*. New York: John Wiley and Sons.

Groves, R. M. and Peytcheva, E. (2008) The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opin. Q.*, **72**, 167-89.

Hausman, J. A. and Wise, D. A. (1979) Attrition bias in experimental and panel data: the Gary Income Maintenance Experiment. *Econometrica*, **47**, 455-73.

Hawkes, D. and Plewis, I. (2006) Modelling non-response in the National Child Development Study. *J. R. Stat. Soc. A*, **169**, 479-492.

Kalton, G. and Brick, M. (2000) Weighting in household panel surveys. In *Researching Social and Economic Change: The Uses of Household Panel Studies* (ed. D. Rose), pp. 96-112. London: Routledge.

Laurie, H. and Lynn, P. (2009) The use of respondent incentives on longitudinal surveys. In *Methodology of Longitudinal Surveys* (ed. P. Lynn), pp. 205-233. Chichester (UK): Wiley.

Laurie, H., Smith, R. and Scott, L. (1999) Strategies for reducing nonresponse in a longitudinal panel survey. *J. Offic. Stat.*, **15**, 269-282.

Lepkowski, J. M. and Couper, M. P. (2002) Nonresponse in the second wave of longitudinal household surveys. In *Survey Nonresponse* (eds. R. M. Groves, D. A. Dillman, J. L. Eltinge and R. J. A. Little), pp. 259-272. New York: John Wiley and Sons.

Lynn, P. (2009) Methods for longitudinal surveys. In *Methodology of Longitudinal Surveys* (ed. P. Lynn), pp. 1-19. Chichester (UK): Wiley.

Lynn, P., Kaminska, O. and Goldstein, H. (2011) Panel attrition: how important is it to keep the same interviewer? ISER Working Paper no. 2011-02. Colchester (UK): Institute for Social and Economic Research, University of Essex.

Nicoletti, C. and Buck, N. (2004) Explaining interviewee contact and co-operation in the British and German household panels. In *Harmonisation of Panel Surveys and Data Quality* (eds. M. Ehling and U. Rendtel), pp. 143-166. Wiesbaden: Statistisches Bundesamt.

Nicoletti, C. and Peracchi, F. (2005) Survey response and survey characteristics; microlevel evidence from the European Community Household Panel. *J. R. Statist. Soc. A*, **168**, 763-781.

Plewis, I., Ketende, S. and Calderwood, L. (2010) Assessing the accuracy of response propensities in longitudinal studies. CCSR Working Paper no. 2010-08. Cathie Marsh Centre for Census and Survey Research, University of Manchester.

Plewis, I., Ketende, S. C., Joshi, H. and Hughes, G. (2008) The contribution of residential mobility to sample loss in a British cohort study: Evidence from the first two waves of the UK Millennium Cohort Study. *J. Offic. Stat*., **24**, 365-385.

Rabe-Hesketh, S. and Skrondal, A. (2008) *Multilevel and Longitudinal Modeling using Stata.* College Station: Stata Press.

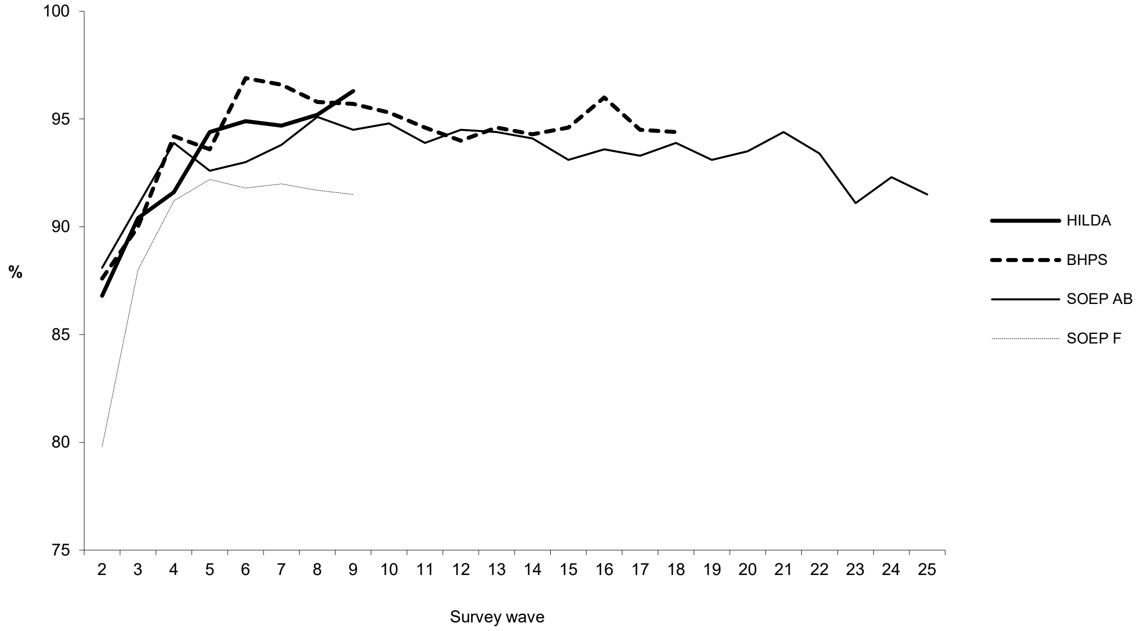
Uhrig, S. C. N. (2008) The nature and causes of attrition in the British Household Panel Survey. ISER Working Paper no. 2008-05. Colchester (UK): Institute for Social and Economic Research, University of Essex.

Vandecasteele, L. and Debels, A. (2007) Attrition in panel data: the effectiveness of weighting. *Eur. Sociol. Rev.*, **23**, 81-97.

Voorpostel, M. (2009) Attrition in the Swiss Household Panel by demographic characteristics and levels of social involvement. FORS Working Paper 1/09. Lausanne: Swiss Foundation for Research in Social Sciences.

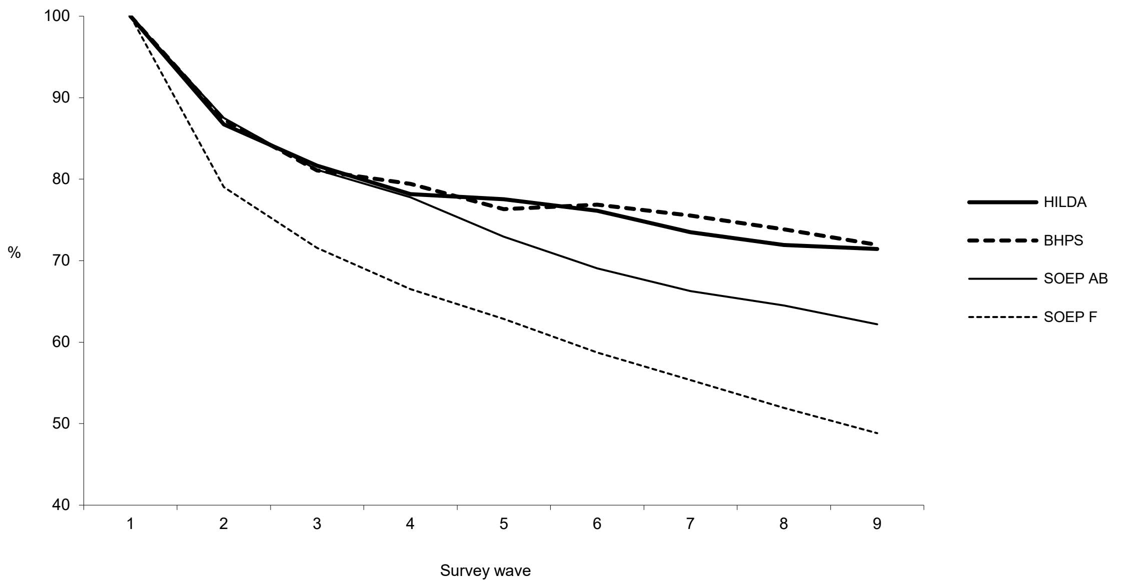
Watson, N. and Wooden, M. (2009) Identifying factors affecting longitudinal survey response. In *Methodology of Longitudinal Surveys* (ed. P. Lynn), pp. 157-182. Chichester (UK): Wiley.

Figure 1. Percentage of respondents re-interviewed at the next survey wave: HILDA Survey, BHPS and SOEP compared



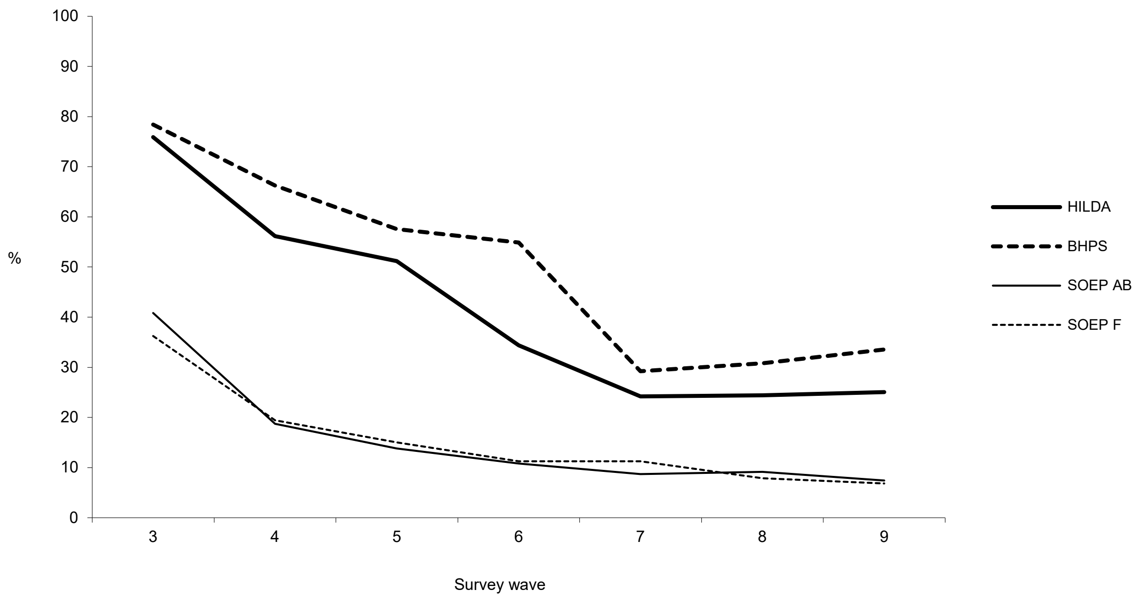
Notes: (1) The denominator excludes deaths and persons known to have moved out of scope. The latter includes persons who move abroad and temporary sample members who cease co-residing with a permanent sample member. (2) For reasons of data comparability, responses obtained in the BHPS by proxy and via the short telephone interview have been treated as non-responses.

Figure 2. Percentage of wave 1 respondents re-interviewed by survey wave: HILDA Survey, BHPS and SOEP compared



Notes: (1) All samples exclude any respondent who died or moved abroad during the first 9 waves. (2) For reasons of data comparability, responses obtained in the BHPS by proxy and via the short telephone interview have been treated as non-responses.

Figure 3. Percentage of previous wave non-respondents issued to field by survey wave: HILDA Survey, BHPS and SOEP compared (wave 1 participants)



Notes: (1) All samples exclude any respondent who died or moved abroad during the first 9 waves. (2) For reasons of data comparability, responses obtained in the BHPS by proxy and via the short telephone interview have been treated as non-responses.

Table 1. Patterns of response by survey, waves 1 to 8 (wave 1 participants)

Survey	*Number*	*Continual participation (%)*	*Monotone attrition (%)*	*Other response patterns (%)*	*Deaths and movements out of scope (%)*
HILDA	13,969	56.9	21.8	12.0	9.2
BHPS	9,912	58.9	20.8	9.7	10.6
SOEP AB	11,610	54.1	29.5	6.2	10.2
SOEP F	10,324	44.9	42.6	6.8	5.7

Notes: (1) For wave 1 participants, movements out of scope comprise respondents who moved abroad during the next seven waves. (2) For reasons of data comparability, responses obtained in the BHPS by proxy and via the short telephone interview have been treated as non-responses.

Table 2. Number of previous wave non-respondents and percentage converted: HILDA Survey, BHPS and SOEP compared (wave 1 participants)

Survey	*W3*	*W4*	*W5*	*W6*	*W7*	*W8*	*W9*
Number of previous wave non-respondents							
HILDA	1,685	2,327	2,769	2,848	3,029	3,362	3,560
BHPS	1,125	1,676	1,824	2,097	2,050	2,166	2,319
SOEP AB	1,303	1,965	2,315	2,818	3,223	3,517	3,699
SOEP F	2,042	2,767	3,259	3,616	4,019	4,348	4,680
% of all non-respondents converted							
HILDA	21.8	15.3	15.6	8.7	5.1	5.6	7.1
BHPS	17.2	14.7	8.2	11.0	4.3	3.5	3.6
SOEP AB	13.3	7.5	3.0	3.6	3.5	3.2	2.0
SOEP F	10.4	4.2	3.9	2.2	2.2	1.6	1.5
% of all non-responding cases issued to field converted							
HILDA	28.7	27.2	30.4	25.2	21.0	22.9	28.1
BHPS	21.9	22.2	14.2	20.0	14.7	11.4	10.8
SOEP AB	32.5	39.9	21.9	33.5	40.4	35.3	26.5
SOEP F	28.8	21.8	26.0	19.3	19.5	20.4	22.2

Notes: (1) All samples exclude any respondent who died or moved abroad during the first 9 waves. (2) For reasons of data comparability, responses obtained in the BHPS by proxy and via the short telephone interview have been treated as non-responses.

Table 3. Explanatory variables: Means and standard deviations (in parentheses)

Variable	*HILDA*		*BHPS*		*SOEP AB*		*SOEP F*	
Fieldwork / Interviewer characteristics								
Interviewer continuity / experience by move								
Same interviewer + No move (base category)	0.379	(0.485)	0.663	(0.473)	0.318	(0.466)	0.424	(0.494)
Diff experienced + No move	0.259	(0.438)	0.177	(0.382)	0.490	(0.500)	0.483	(0.500)
New interviewer + No move	0.171	(0.377)	0.064	(0.245)	0.013	(0.113)	0.010	(0.098)
Same interviewer + Move	0.053	(0.224)	0.055	(0.228)	0.060	(0.237)	0.030	(0.171)
Diff experienced + Move	0.095	(0.293)	0.032	(0.177)	0.116	(0.320)	0.052	(0.223)
New interviewer + Move	0.043	(0.202)	0.009	(0.095)	0.003	(0.055)	0.001	(0.035)
Unchanged interview mode	0.839	(0.367)			0.540	(0.499)	0.565	(0.496)
Telephone interviewer	0.121	(0.326)			0.685	(0.464)	0.710	(0.454)
No. of non-responding waves								
One (base category)	0.609	(0.488)	0.455	(0.498)	0.888	(0.316)	0.824	(0.381)
Two	0.218	(0.413)	0.218	(0.413)	0.075	(0.263)	0.077	(0.266)
Three	0.093	(0.290)	0.137	(0.344)	0.018	(0.132)	0.042	(0.200)
Four	0.041	(0.199)	0.092	(0.289)	0.009	(0.095)	0.026	(0.160)
Five or more	0.039	(0.194)	0.097	(0.296)	0.010	(0.101)	0.031	(0.173)
Previous wave interview with other hh member	0.250	(0.433)	0.298	(0.458)	0.213	(0.410)	0.273	(0.446)
Previous wave outcome								
Refusal (base category)	0.769	(0.421)	0.379	(0.485)	0.581	(0.493)	0.763	(0.425)
Short telephone interview			0.087	(0.282)				
Non-contact	0.079	(0.269)	0.052	(0.223)	0.078	(0.268)	0.029	(0.169)
Other non-response	0.123	(0.329)	0.222	(0.416)	0.005	(0.069)	0.007	(0.082)
Lost in tracking	0.029	(0.167)	0.260	(0.438)	0.050	(0.217)	0.029	(0.166)
Processing problems					0.286	(0.452)	0.172	(0.378)
Interviewer workload	71.0	(25.9)	45.0	(42.7)	420.2	(280.2)	431.4	(318.1)
Survey wave								
W3 (base category)	0.174	(0.379)	0.141	(0.348)	0.222	(0.415)	0.226	(0.418)
W4	0.177	(0.382)	0.178	(0.383)	0.151	(0.358)	0.163	(0.370)
W5	0.192	(0.394)	0.168	(0.374)	0.134	(0.340)	0.149	(0.356)
W6	0.131	(0.338)	0.185	(0.388)	0.126	(0.332)	0.123	(0.329)
W7	0.097	(0.296)	0.096	(0.295)	0.120	(0.325)	0.138	(0.345)
W8	0.108	(0.310)	0.107	(0.309)	0.132	(0.339)	0.104	(0.306)
W9	0.121	(0.326)	0.124	(0.330)	0.116	(0.320)	0.096	(0.295)
Respondent characteristics (all except age are measured at w1)								
Female	0.480	(0.500)	0.459	(0.498)	0.476	(0.500)	0.494	(0.500)
Age	44.0	(16.3)	45.5	(19.0)	42.3	(16.4)	46.0	(16.8)
Foreign born	0.298	(0.457)	0.092	(0.290)	0.299	(0.458)	0.123	(0.329)
University graduate	0.126	(0.332)	0.050	(0.218)	0.073	(0.260)	0.159	(0.366)
Employed	0.651	(0.477)	0.611	(0.487)	0.633	(0.482)	0.624	(0.485)
Health limitations	0.144	(0.351)	0.151	(0.358)	0.293	(0.455)	0.276	(0.447)
No. of adults in household	2.7	(1.2)	2.4	(1.0)	2.7	(1.2)	2.4	(0.9)
No. of children in household	0.8	(1.1)	0.7	(1.0)	0.8	(1.0)	0.6	(0.9)
Home owner	0.724	(0.447)	0.654	(0.476)	0.359	(0.480)	0.583	(0.493)
Equivalised annual household income (/10^5)	0.356	(0.485)	0.097	(0.066)	0.154	(0.123)	0.259	(0.208)
Marital status								
Married (base category)	0.539	(0.499)	0.520	(0.500)	0.546	(0.498)	0.586	(0.493)
De facto	0.100	(0.300)	0.075	(0.264)	0.032	(0.176)	0.168	(0.374)
Separated / divorced	0.080	(0.272)	0.132	(0.339)	0.087	(0.281)	0.058	(0.234)
Never married	0.281	(0.449)	0.272	(0.445)	0.336	(0.472)	0.188	(0.391)
N (observations)	7,330		6,191		2,311		3,225	
N (individuals)	3,943		2,564		1,875		2,488	
N (wave 1 households)	2,497		1,767		1,185		1,566	

Table 4. Estimated coefficients, standard errors and mean marginal effects for selected variables†: Models of survey re-engagement

Variable	*HILDA*			*BHPS*			*SOEP AB*			*SOEP F*		
	Coeff.	*SE*	*MME*	*Coeff.*	*SE*	*MME*	*Coeff.*	*SE*	*MME*	*Coeff.*	*SE*	*MME*
Interviewer continuity / experience by move (base = Same interviewer + No move)												
Diff experienced + No move	0.281***	0.106	0.038	0.965***	0.138	0.095	0.379	0.325	0.045	1.577***	0.327	0.064
New interviewer + No move	0.119	0.122	0.016	0.455**	0.195	0.040	0.564	0.798	0.067	2.145**	0.972	0.101
Same interviewer + Move	0.380**	0.171	0.052	0.583***	0.197	0.053	-0.411	0.403	-0.046	1.198**	0.489	0.044
Diff experienced + Move	1.133***	0.145	0.172	1.179***	0.229	0.121	-0.317	0.401	-0.036	1.719***	0.499	0.073
New interviewer + Move	1.171***	0.198	0.179	1.534***	0.403	0.168	-2.014	1.623	-0.192	-3.197	2.897	-0.043
Unchanged interview mode	0.168	0.139	0.023				0.514	0.330	0.060	1.682***	0.325	0.072
Telephone interviewer	-0.696***	0.170	-0.088				-1.760	1.163	-0.226	-4.031***	1.032	-0.343
No. of non-responding waves (base = One)												
Two	-0.483***	0.101	-0.066	-0.407***	0.133	-0.039	-0.348	0.316	-0.040	-1.052***	0.371	-0.034
Three	-0.314**	0.149	-0.044	-0.588***	0.194	-0.054	-0.136	0.696	-0.016	-1.663***	0.613	-0.047
Four	-0.932***	0.252	-0.118	-0.881***	0.248	-0.076	-0.480	0.934	-0.054	-0.879	0.746	-0.030
Five or more	-0.455*	0.253	-0.063	-1.312***	0.337	-0.103	0.898	0.878	0.112	-0.724	0.770	-0.025
Previous wave interview with other hh member	-0.035	0.101	-0.005	-0.135	0.154	-0.012	-1.120***	0.290	-0.127	-1.982***	0.315	-0.067
Previous wave outcome (base = Refusal)												
Short telephone interview				1.870***	0.175	0.226						
Non-contact	1.849***	0.140	0.311	1.051***	0.208	0.105	0.725*	0.383	0.088	0.999**	0.469	0.052
Other non-response	0.871***	0.121	0.132	1.115***	0.156	0.113	1.819	1.298	0.238	0.738	0.916	0.035
Lost in tracking	0.038	0.235	0.005	-0.296	0.183	-0.022	0.324	0.433	0.038	-0.526	0.602	-0.018
Processing problems							0.329	0.234	0.039	0.276	0.364	0.012
Log likelihood	-3,678			-2,210			-1,238			-1,379		
Pseudo R-squared	0.066			0.141			0.064			0.116		
ρ	0.448	0.031		0.446	0.042		0.697	0.043		0.762	0.031	
Observations (N)	7,330			6,191			2,311			3,225		
Wave 1 households (N)	2,497			1,767			1,185			1,566		

† All models also include controls for respondent characteristics, interviewer workload, and survey wave (see Table 3).
* 0.10≥p>0.05, ** 0.05≥p>0.01, *** 0.01≥p.

Table 5. Estimated coefficients, standard errors and mean marginal effects for selected variables†: Models of survey continuation

Variable	*HILDA*			*BHPS*			*SOEP AB*			*SOEP F*		
	Coeff.	*SE*	*MME*	*Coeff.*	*SE*	*MME*	*Coeff.*	*SE*	*MME*	*Coeff.*	*SE*	*MME*
Interviewer continuity / experience by move (base = Same interviewer + No move)												
Diff experienced + No move	-0.247***	0.054	-0.010	-1.183***	0.071	-0.053	-0.653***	0.069	-0.027	-1.011***	0.087	-0.055
New interviewer + No move	-0.249***	0.058	-0.010	-0.281***	0.089	-0.009	-0.757***	0.126	-0.032	-0.529***	0.152	-0.026
Same interviewer + Move	-0.523***	0.084	-0.023	-1.505***	0.074	-0.075	-1.307***	0.094	-0.064	-0.656***	0.102	-0.033
Diff experienced + Move	-0.699***	0.073	-0.033	-1.673***	0.099	-0.087	-2.258***	0.132	-0.142	-1.040***	0.167	-0.057
New interviewer + Move	-0.592***	0.093	-0.027	-2.112***	0.160	-0.126	-2.217***	0.354	-0.138	-1.230**	0.489	-0.070
Unchanged interview mode	0.777***	0.104	0.041				0.535***	0.126	0.024	1.666***	0.149	0.112
Telephone interviewer	-0.712***	0.115	-0.037				-2.473***	0.248	-0.169	-6.519***	0.325	-0.789
No. of responding waves (base = One)												
Two	-0.108	0.126	-0.004	0.084	0.174	0.003	0.418*	0.215	0.010	0.0326	0.211	0.001
Three	-0.402***	0.110	-0.014	-0.109	0.153	-0.004	-0.524***	0.173	-0.016	-0.997***	0.172	-0.037
Four	-0.641***	0.123	-0.024	-0.366**	0.152	-0.014	-0.644***	0.186	-0.020	-1.201***	0.210	-0.046
Five	-1.113***	0.132	-0.048	-0.452**	0.184	-0.018	-1.869***	0.202	-0.076	-2.046***	0.217	-0.090
Six	-1.319***	0.147	-0.060	-0.498***	0.189	-0.020	-1.586***	0.212	-0.061	-2.163***	0.246	-0.097
Seven	-2.041***	0.174	-0.109	-0.704***	0.216	-0.030	-2.487***	0.251	-0.118	-4.018***	0.288	-0.256
Eight	-2.408***	0.191	-0.140	-0.859***	0.218	-0.039	-3.153***	0.259	-0.172	-4.148***	0.306	-0.270
Previous wave partial hh response	-0.949***	0.060	-0.051	-0.484***	0.074	-0.021	-0.157	0.105	-0.006	-0.166*	0.092	-0.008
Log likelihood	-16,959			-10,951			-12,479			-11,917		
Pseudo R-squared	0.048			0.064			0.107			0.230		
ρ	0.663	0.012		0.628	0.015		0.828	0.007		0.722	0.011	
Observations (N)	81,722			57,191			63,739			52,912		
Wave 1 households (N)	7,187			5,000			5,181			5,492		

† All models also include controls for respondent characteristics, interviewer workload, and survey wave (see Table 3).
*0.10≥p>0.05, **0.05≥p>0.01, ***0.01≥p