HILDA PROJECT DISCUSSION PAPER SERIES
No. 2/06, December 2006

Modelling Longitudinal Survey Response: The Experience of the HILDA Survey

Nicole Watson and Mark Wooden

The HILDA Project was initiated, and is funded, by the Australian Government Department of Families, Community Services and Indigenous Affairs

Acknowledgements

This paper was presented at the ACSPRI Social Science Methodology Conference, University of Sydney, Sydney, Australia, 10-13 December 2006. An earlier (and substantively different) version was presented at the Methodology of Longitudinal Surveys Conference, University of Essex, Colchester, 12-14 July 2006.

1. Introduction

Like all sample surveys, longitudinal (or panel) surveys are affected by unit non-response. A distinctive feature of longitudinal surveys, however, is that non-response is not a one-off event and tends to accumulate over time as further waves of interviewing are conducted. Longitudinal surveys also face the problems of tracking sample members who relocate and of dealing with the respondent fatigue that is associated with repeated survey participation (Laurie, Smith and Scott, 1999).

As a consequence, many longitudinal surveys find, even after just a few waves of interviewing, that relatively large fractions of the responding sample from the initial wave are no longer participating. This has certainly been the case for the world's leading household panel surveys. The University of Michigan's Panel Study of Income Dynamics (PSID), for example, had lost just over one-quarter of its original 1968 sample by 1975 (i.e., wave 8) (see Fitzgerald, Gottschalk and Moffitt, 1998, Table 1).[1] More recent household panel studies, typically employing more complicated designs, report higher rates of sample attrition. After eight years of interviewing, the German Socio-Economic Panel (GSOEP), which commenced in 1984, and the British Household Panel Survey (BHPS), which commenced in 1991, both reported the loss of about 34 per cent of their original sample (Spieß and Kroh, 2004, Figure 9; Taylor *et al*., 2005, Table 20), and in the case of the Dutch Socioeconomic Panel the rate of sample loss at the same stage was almost certainly in excess of 50 per cent.[2] Relatively high rates of sample loss have also been reported in the European Community Household Panel (ECHP), a multi-country study conducted over the period 1994 to 2001. Watson (2003), for example, reported five-year retention rates that varied from a high of 82 per cent in Portugal to a low of 57 per cent in Ireland (after excluding deaths and other movements out-of-scope). Finally, the Survey of Income and Program Participation (SIPP), run by the US Census Bureau, has reported cumulative rates of sample loss of up to 35 per cent (of households) over a four-year period (Westat, 2001, Table 2.5, p. 2-19).

Rates of attrition might be different in other types of longitudinal surveys employing different designs and covering different populations. Birth cohort studies, for example, often report very high response rates many years after the original sample was drawn (e.g., Wadsworth *et al.*, 2003; Hawkes and Plewis, 2006). Such studies, however, are distinctive in that interviewing is relatively infrequent, and hence respondent burden tends to be far less than in other panel surveys where interview waves are more frequent. Nevertheless, it is also true that frequent (e.g., annual) survey waves are not always associated with high rates of sample attrition. The National Longitudinal Study of Youth (NLSY), which follows cohorts of young people in the US until well into adulthood, for example, obtained a rate of attrition after eight years of interviewing from its 1979 cohort of just 8 per cent, and even after 21 years the rate of sample loss was still under 20 per cent (Bureau of Labour Statistics, 2003). The NLSY experience, however, appears to be the exception and not the norm, with most other youth cohort panels (e.g., the Youth Cohort Study of England and Wales and the various cohort studies that comprise the Longitudinal Surveys of Australian Youth) recording much higher rates of attrition.[3]

---

[1] This figure includes deaths. If deaths are excluded, the accumulated attrition rate declines to about 22 per cent.

[2] Winkels and Withers (2000) reported that only 42 per cent of the original sample remained after 11 years of interviewing.

[3] The 1998 LSAY cohort, for example, had lost around 45 per cent of its wave 1 respondents after just four years (i.e., by the end of wave 5).

There is also mounting evidence indicating that the problem of sample attrition has been getting worse over time (Atrostic, Baytes and Silberstein, 2001; de Leeuw and de Heer, 2002; Tourangeau, 2003). The recent experience of the NLSY seems to be in line with this conclusion, with the rates of attrition for the 1997 cohort noticeably higher than the rates of attrition recorded for the 1979 cohort. After the first five waves, the overall sample retention rate, while still a healthy 88 per cent, was over eight percentage points lower than the rate reported at the comparable stage of the 1979 cohort.[4] A similar deterioration over time has also been reported for the SIPP. Involving relatively short overlapping panels (ranging from 24 months to 48 months long), rates of cumulative sample loss over eight waves (32 months) averaged around 21 per cent for the panels commencing between 1984 and 1991 (Westat, 1998, Table 5.1, p. 45). For the 1992 and 1993 panels the rate of sample loss rose to around 25 per cent over the same time span, and for the 1996 panel the rate of loss was over 31 per cent (Westat, 2001, Table 2.5, p. 2-19).

High rates of sample attrition pose a serious problem for longitudinal studies. At a minimum, attrition reduces the precision of survey estimates, and at sufficiently high levels can threaten the viability of continuing a panel, especially if the initial sample size was relatively small. Further, since attrition tends not to be random, it may impart bias to population estimates. Survey administrators thus face pressures to both ensure that they employ design features and fieldwork procedures that will maximize sample retention and, since some non-response is unavoidable, deliver as much information as possible about non-respondents to assist data analysts to make inferences in the presence of missing data. Achieving both of these objectives requires good knowledge of the response process and the factors that give rise to sample attrition.

Surprisingly, published research on longitudinal survey response remains relatively scarce. Further, the body of research evidence that does exist is not as useful as it might otherwise be. The emphasis in most past work has been on relationships with personal and household characteristics, representing most, if not all, of the explanatory variables included in many studies (e.g., Becketti *et al*., 1988; Behr, Bellgardt and Rendtel, 2005; Burkam and Lee, 1998; Fitzgerald *et al*, 1998; Gray *et al*., 1996; Hawkes and Plewis, 2006; Kalsbeek, Yang and Agans, 2002; Watson, 2003). This should not be surprising. First, and most obviously, detailed information about the characteristics of non-respondents is something that is readily available for all persons who exit longitudinal surveys. Second, differences in attrition propensities may be suggestive of possible attrition bias, and thus can be used to help correct for such bias, and it is this purpose which has driven most analysis of longitudinal survey response. What such studies, however, cannot tell us much about is how survey design features might be adjusted with a view to minimizing sample attrition.

A further weakness of many studies of longitudinal survey response is the treatment of response as the result of a relatively simple, one-step decision process by sample members. In fact, the response process is much more complex than this. Lepkowski and Couper (2002), for example, distinguish three distinct phases in the response process: (i) the sample member is located; (ii) contact with the sample member is established; and (iii) the sample member provides an interview. At each of these stages there is an opportunity for the sample member to be lost from the sample.

In this paper we estimate a model predicting response over the course of the first four waves of the Household, Income and Labour Dynamics in Australia (HILDA) Survey that seeks to rectify some of these deficiencies. In particular, in addition to sample member

[4] The initial samples for the two cohorts, however, were not identical. Specifically, the 1997 cohort was younger, ranging from 12 to 16 years, whereas the 1979 cohort was aged between 14 and 21 years.

characteristics, we incorporate into our models variables describing both the interview situation and survey design features. In addition, we explicitly distinguish between the contact and cooperation stages of the response process.

2. The HILDA Survey Data

The HILDA Survey is a nation-wide household panel survey with a focus on employment, income and the family. Modelled on household panel surveys undertaken in other countries, and described in more detail in Watson and Wooden (2004), it began in 2001 with a large national probability sample of Australian households occupying private dwellings. All members of those responding households in wave 1 form the basis of the panel to be pursued in each subsequent wave (though interviews are only conducted with those household members aged 15 years or older), with each wave of interviewing being approximately one year apart. Like many other households panels, the sample is extended each year to include any new household members resulting from changes in the composition of the original households. The original sample members form the initial group of permanent sample members (PSMs). New children of PSMs and person who have a child with a PSM also become PSMs. All other new sample members only remain in the sample for as long as they live with a PSM.

During waves 1 to 4, households were paid either A$20 or A$50 each year they participated, with the higher amount only paid when interviews were completed with all in-scope household members. In wave 5 the interview incentive was changed to $25 per head, with a $25 household bonus paid if all household members complete the person interview.

After adjusting for out-of-scope dwellings and households, and multiple households within dwellings, the number of households identified as in-scope in wave 1 was 11,693. Interviews were completed with all eligible members at 6872 of these households and with at least one eligible member at a further 810 households. Within the 7682 households at which interviews were conducted, 13,969 persons were successfully interviewed (out of a total of 15,127 eligible household members).

Details about the evolution of the responding sample over the first five waves are provided in Table 1. This table shows that 10,392 (or 74 per cent) of those persons initially interviewed in wave 1 were re-interviewed in wave 5. If deaths and movements out of scope are excluded, the five-wave sample retention rate rises to 78 per cent. The rates of response among persons joining the sample at later waves appear to be much lower, but this is largely because many of these new sample members are only added to the sample on a temporary basis.

Table 2 provides the wave-on-wave response rates for waves 2 through 5. These response rates exclude people who have moved out-of-scope (i.e., have moved overseas, died between waves, or were temporary sample members who moved out of the sample households). The table distinguishes between previous wave respondents (who account for the large majority of in-scope sample members in any year), previous wave non-respondents, children turning 15 years of age, and new sample members. As can be seen, the wave-on-wave response rate for previous wave respondents (calculated as the proportion of in-scope previous wave respondents who provided an interview) has gradually improved over time, rising from 86.8 per cent in wave 2 to 94.4 per cent in wave 5. Furthermore, these response rates compare favourably with that reported in other leading household panel studies, such as the BHPS (see Figure 1).

TABLE 1
Individual Response (N) by Wave, HILDA Survey

Wave first interviewed	*Wave 1*	*Wave 2*	*Wave 3*	*Wave 4*	*Wave 5*
Wave 1	13969	11993	11190	10565	10392
Wave 2	-	1048	705	594	572
Wave 3	-	-	833	543	482
Wave 4	-	-	-	706	494
Wave 5	-	-	-	-	819
TOTAL	13969	13041	12728	12408	12759

TABLE 2
Response Rates by Wave, HILDA Survey
(figures in parentheses provide the % share of the total sample)

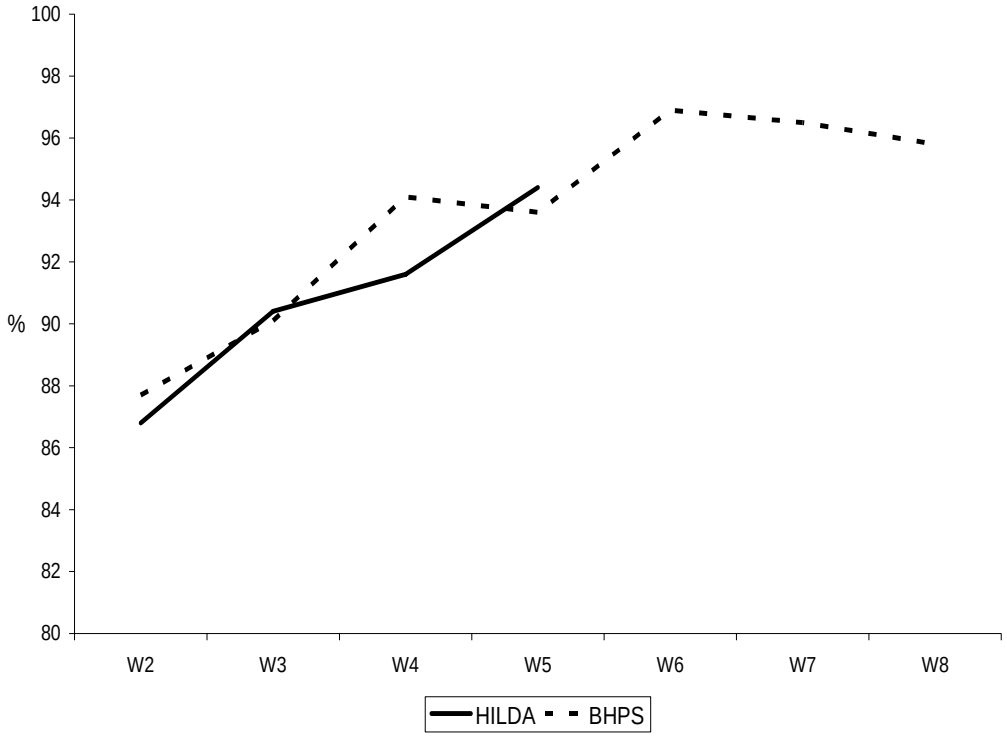
Sub-sample	*Wave 2*	*Wave 3*	*Wave 4*	*Wave 5*
Previous wave respondents	86.8 (86.1)	90.4 (76.7)	91.6 (72.9)	94.4 (69.5)
Previous wave non-respondents	19.7 (7.0)	17.6 (17.2)	12.7 (21.4)	14.7 (24.6)
Children turning 15 years	80.4 (1.9)	71.3 (2.0)	70.7 (2.0)	74.4 (1.9)
New sample entrants	73.3 (4.9)	76.0 (4.0)	70.4 (3.3)	81.7 (3.4)

Note: Excluded from these figures are sample members who move overseas and then return. As a result the percentages in brackets do not necessarily sum to 100. The proportion of the eligible sample returning from overseas was 0.1% in wave 3, 0.3% in wave 4 and 0.5% in wave 5.

The response rates among previous wave non-respondents are very low. This is to be expected given these are people who have previously refused to participate. Indeed, many longitudinal surveys make no effort to recruit non-respondents back into the sample. In the HILDA Survey, however, a sample member is only permanently deleted from the panel if they make a call to the 1800 telephone number requesting that they be removed from the sample or if the sample member had behaved in a threatening manner towards an interviewer.[5] Attrition from the HILDA Survey sample is thus not monotonic, and as can be seen from Table 2, between 13 and 20 per cent of previous wave non-respondents issued into the field each year are successfully interviewed in the current wave. Note that we would expect this response rate to decline over time given the size of the denominator population is constantly growing. The fact that it rose in wave 5 is thus significant and most likely reflects the changed incentive payment structure and the increased resources devoted to the post-Christmas fieldwork phase in that wave.

[5] However, any sample member who has refused to participate in two successive waves is reviewed and a judgement formed about the cost effectiveness of attempting contact with them in the next wave.

FIGURE 1
Wave on Wave Response Rates: HILDA Survey and BHPS Compared



Note: To maximize comparability with the HILDA Survey response data, the BHPS data have been adjusted to exclude both proxy interviews and interviews obtained by telephone using a very short instrument.

Source: The BHPS data are derived from Tables 18a to 18g of Volume A of the BHPS User Guide.

For persons never previously interviewed (that is, children turning 15 years and new members of the household) the recruitment exercise is similar to recruiting new sample members at wave 1, but with the added advantage that these new sample members know much more about the survey due to their relationship with an existing sample member. As a result, we expect and obtain response rates that are much higher than the individual response rate obtained at wave 1. Individual response rates for new sample members have varied from a low of 70 per cent in wave 4 to a high of 80 per cent one wave later. The annual rates of response among children turning 15 years have been similar, varying between 71 and 80 per cent. Again it should be noted that for children turning 15 years (but not new sample members) a decline in response rates over time is expected. This is because the denominator population includes children from households where contact has been lost.

3. Estimation Approach

Pooled data from the first four waves of the HILDA Survey are used to estimate a model that employs information about the respondents and their interviews in wave *t-1* to predict response at wave *t*. As briefly mentioned earlier, the response process can be thought of as involving three stages: (i) locating the sample member; (ii) making contact with the sample member; and (iii) obtaining the cooperation of the sample member. Distinguishing between location and contact, however, is empirically difficult in the HILDA Survey data. We thus model survey response as the occurrence of two sequential events: (i) establishing contact;

and (ii) obtaining cooperation once contacted. This approach is essentially the same as that used by Nicoletti and Buck (2003), who analysed survey response over three waves of both the BHPS and GSOEP panels, and by Nicoletti and Peracchi (2005), who examined response behaviour over the first five waves of the ECHP.

We estimate three separate model specifications. Model I assumes the two events are independent and thus two separate probit equations are estimated. Following Nicoletti and Peracchi (2005), Model II relaxes this assumption and allows for conditional correlation in the error terms in the two equations. This is achieved by estimating a bivariate probit with sample selection. Model III again imposes the assumption of conditional independence but, following Nicoletti and Buck (2003), an unobserved random effect for the interviewer is introduced into both probit specifications.

The units of analysis are all individuals who were interviewed at wave *t-1* and deemed as in-scope at wave *t* (i.e., units are individual-wave combinations). Individuals were deemed as out-of-scope if they are known to have died, moved abroad permanently, or were new (i.e., temporary) sample members who no longer live with a PSM. All wave 1 non-respondents are obviously excluded (since no details of non-respondents at wave 1 are known). The pooled dataset contains 38,831 observations on 15,313 individuals (2423 are observed once only, 2262 are observed twice and 10,628 are observed three times). Of these 38,831 person-wave observations, 38,118 were contacted in the next wave and 34,751 were interviewed in the next wave. As many individuals from identical or like households are observed across multiple waves, the estimated standard errors in Models I and II assume the outcome variables are correlated across observations on the same individual from the same wave 1 households, but are independent across individuals from different households.[6]

4. Explanatory Variables

The probability of making contact with an individual is assumed here to be a function of: (i) whether the individual has changed address; (ii) the likelihood that the respondent is at home when the interviewer calls; (iii) the willingness of the individual to be found (as in Lepkowski and Couper, 2002); and (iv) interviewer effort and knowledge.

The first influence is relatively straightforward to measure and represented by a variable identifying whether the individual relocated between survey waves. Changing address has generally been found to exert a highly significant negative influence on the likelihood of response at the next wave (Becketti, *et al*., 1988; Lepkowski and Couper, 2002; Lillard and Panis, 1998; Watson, 2003; Zabel, 1998), which we hypothesise is mainly the result of the greater difficulty locating sample members who relocate.

The likelihood of finding the respondent at home is assumed to be a function of the number of calls made by the interviewer to the household in the previous wave, with evidence from other longitudinal surveys indicating that the sample members requiring the most effort in terms of the number of calls made at wave *t-1* are at greatest risk of attrition at wave *t* (Branden, Gritz and Pergamit, 1995; Lillard and Panis, 1998; Nicoletti and Buck, 2003; Nicoletti and Peracchi, 2005; Zabel, 1998). The usual explanation for this finding is that the number of calls needed is indicative of both how difficult it is to find the respondent at home and how evasive the respondent might be when setting up interview appointments

[6] This requires linking all individual respondents in waves 2 to 4 to a wave 1 household. For some households this assumption is unrealistic. The alternative assumption would be to only allow for correlated errors within individuals and not households. The assumption adopted here results in larger standard errors on most estimates.

(though only Nicoletti and Buck (2003) and Nicoletti and Peracchi (2005) specifically identify attrition due to non-contact).

The likelihood of finding the respondent at home is also expected to be correlated with various respondent characteristics, including: age, sex, marital status, the number of people in the household and the age of the co-residents (the presence of young children is expected to be associated with a greater likelihood of finding the respondent at home)[7]; employment and labour force status; the type of dwelling; home ownership (i.e., living in rental accommodation compared with living in a home that the respondent owned and was purchasing); and the presence of a serious long-term health condition (defined as any health condition or disability that had lasted or was expected to last at least six months and prevented the respondent from undertaking any form of work). Area characteristics might also be relevant. We thus include a set of geographic dummies which identify whether the respondent lives in one of the major cities and if not how remote they are from a major urban centre.[8] Also included is a measure of the relative socio-economic disadvantage of the neighbourhood in which the respondent lives.[9]

Willingness to be found is represented by a range of variables describing the experience of the previous interview. We hypothesise that being from a partially responding household (at least one other member in the household refused to participate in the survey), not returning the self-completion questionnaire,[10] and being assessed by the interviewer as relatively uncooperative, suspicious and not having a good understanding of the questions, will all be negatively associated with the likelihood of making contact at the next wave. Willingness to be found might also be a function of respondent characteristics, and most notably country of birth and English language ability.

We are unable to include any direct measures of interviewer effort, but following Nicoletti and Buck (2003) include a measure of the interviewer's workload in wave *t* (the number of previous wave respondents allocated to the interviewer at the start of the fieldwork period). Nicoletti and Buck hypothesized that large workloads will reflect overworked interviewers and thus be negatively associated with both the probability of contact and response. In the HILDA Survey, however, better interviewers tend to be allocated larger workloads which should work to offset this effect. We thus experiment with both linear and quadratic specifications for this variable.

We also include variables identifying whether the interviewer conducting the interview at wave *t* is the same as at previous waves, given the widely held view that interviewer continuity can be beneficial for response in longitudinal studies, at least those administered on a face-to-face basis (Behr *et al*., 2005; Laurie *et al*., 1999; Hill and Willis, 2001; Nicoletti and Buck, 2003; Nicoletti and Peracchi 2005; Olsen, 2005; Waterton and Lievesley, 1987;

---

[7] The presence of older household members might also be expected to increase the likelihood of the respondent being at home. We, however, experimented with a measure of the presence of household members over the age of 64 years and then over 74 years and could find no evidence of any significant association with contact probabilities.

[8] The major cities are the mainland State capitals – Sydney, Melbourne, Brisbane, Perth, Adelaide – as well as Canberra, Newcastle, the Central Coast region of New South Wales, Wollongong, the Gold Coast in southern Queensland, and Geelong.

[9] Designed by the Australian Bureau of Statistics (ABS 2003), this variable has a mean of 1000 and a standard deviation of 100. The variable, however, is designed only to have ordinal meaning and so we have divided cases into quintiles, with the lowest quintile being the most disadvantaged.

[10] All persons completing a personal interview are also given an additional self-completion questionnaire. Interviewers attempt to collect the completed questionnaire at subsequent visits to the household but where this is not possible, respondents are asked to return it by mail. The proportion of interviewed respondents who return it (completed) averaged 92 per cent over the first four waves.

Zabel, 1998). That said, only Nicoletti and Buck (2003) explicitly test whether such effects matter more for contact probabilities than for response. The results they report appear to suggest interviewer continuity matters at both stages, but the relative magnitude is greater at the contact stage. We hypothesise that using the same interviewer as in the previous wave should enhance the probability of making contact given the interviewer's prior knowledge of the household. Note that we interact interviewer continuity with the survey wave in an attempt to identify whether the impact of interviewer continuity changes with the duration of that continuity.

The probability of an individual providing an interview once they have been contacted is a function of both their ability and willingness to respond. The list of variables used to proxy these influences, however, includes most of the variables included in the contact model.[11] In addition, we also include measures of: the length of the personal interview; the length of the household interview; the proportion of missing responses in each section of the questionnaire; whether the interview was conducted by telephone; whether an indigenous Australian; educational attainment; and equivalised household income.

Note that while it is widely assumed that the likelihood of respondent cooperation will be inversely related to expected interview length, we do not necessarily expect to find evidence of a negative association between survey response and actual interview length. This is because interview length is not exogenous to the interview process, and is instead a product of how willing respondents are to talk to the interviewer. Thus interview length may well be another measure of respondent interest in the study and so will be positively associated with cooperation propensities, and indeed this exactly what has been found in US studies (Branden *et al*., 1995; Hill and Willis, 1998; Zabel, 1998).

We also include variables measuring the extent of missing data on the grounds that item non-response is indicative of an unpleasant or negative interview experience. It thus follows that item non-response at one wave should also be predictive of unit non-response at the next, and the weight of evidence suggests this is so (Burkam and Lee, 1998; Zabel, 1998; Laurie *et al*., 1999; Loosveldt, Pickery and Billiet, 2002; Schräpler, 2002; Lee, Hu and Toh, 2004; Hawke and Plewis, 2006). Unlike previous research, however, we explicitly test whether the impact of item non-response differs depending on which part of the interview or questionnaire the non-response occurs.

Finally, the inclusion of a dummy variable identifying whether the interview was conducted by telephone deserves comment. It might be expected that providing respondents with multiple modes for responding should reduce respondent reluctance, a hypothesis that has received support from experiments conducted in conjunction with cross-section surveys (e.g., Voogt and Saris, 2005). However, in many longitudinal studies employing predominantly face-to-face survey methodologies (including the HILDA Survey), alternative response modes, such as telephones, are often used as a 'last resort' strategy to obtain responses. In such situations, a telephone interview in one year may be indicative of a relative lack of interest in the study and thus predictive of non-response in the next. Analyses of attrition in the NLSY (Branden *et al.*, 1995), the SIPP (Zabel, 1998) and the ECHP (Nicoletti and Peracchi, 2005) data all support this latter hypothesis.

[11] To help identify Model II, dwelling type (in addition to the variable identifying individuals that move) was excluded from the response equations in all models. It was not significant when included in the response equation in Model I.

5. Results

The results of our model estimation are presented in Table 3. Before focusing on specific coefficients we make four general observations. First, a Wald test suggests that the assumption of conditional independence should be rejected (ρ = .356; chi-squared = 7.47; p-value = 0.006). The estimated coefficients, however, are little different when the conditional independence assumption is relaxed. Second, unobserved interviewer effects, while statistically significant, are relatively small in magnitude. In both the contact and response models the estimated proportion of the total variance attributable to an unobserved random interviewer effect is less than 7 per cent.[12] The comparable percentages reported by Nicoletti and Buck (2003) varied from 13 to 51 per cent in the contact models and from 13 to 31 per cent in the response models. The implication again is that focusing on Model I results will not be misleading. Third, most of the covariates (but not all) have coefficients that are either in line with expectations or are, *ex post*, intuitively sensible (and given the large sample size, most are highly significant). They are also extremely robust to model specification. Fourth, despite the array of significant coefficients, the overall explanatory power of these models is relatively poor. This is a desired outcome and presumably reflects the large random component in survey non-response.

What then do our estimates reveal? Focusing first on the probability of making contact, we observe the following:

(i) The optimal interviewer workload for maximizing contact rates is 124 previous wave respondents, which is only slightly larger than the sample mean. With larger or smaller workloads, the likelihood of contact is reduced.

(ii) Interviewer continuity does not appear to be of any great benefit to contact probabilities. Only by wave 4 is there any evidence that interviewer continuity is beneficial for contact probabilities, and even then the positive differential is both small (statistically insignificant) and restricted to cases where the interviewer has been visiting the same respondent for all four survey waves.

(iii) Indicators of the interview experience tend to be strong predictors of response at the next wave. The number of calls made, belonging to a part responding household, and not returning the self-completion questionnaire in the previous wave, are all negatively associated with making contact with the respondent in the next wave. A cooperative respondent is also much more likely to be contacted at the next wave.

(iv) Consistent with other studies, moving house is a strong negative predictor of the propensity to make contact with a respondent in the next wave. The estimates in Model I suggest the mean predicted probability of making contact with a mover is 95.5 per cent, compared with 99.1 per cent for non-movers.

(v) Renters are typically harder to establish contact with than homeowners. This negative effect is even greater if living in unit or flat.[13]

[12] The estimated values of rho together with its standard error (in parentheses) were as follows – Contact: ρ = 0.067 (0.016); Response given contact: ρ = 0.062 (0.009).

[13] As might be expected, contact rates are lowest for people living in 'other dwellings' which consists mainly of highly mobile structures such as caravans, tents and houseboats. These, however, represent a very small fraction of the sample (less than one per cent).

TABLE 3
Coefficients for Contact and Response Probit Models, HILDA Survey

	Contact			*Response*		
Variable	*Model I*	*Model II*	*Model III*	*Model I*	*Model II*	*Model III*
Survey design features						
Interviewer workload (/10^2)	0.624 ***	0.629 ***	0.625 **	0.203	0.245	0.287
Interviewer workload squared (/10^4)	-0.251 **	-0.253 **	-0.281 **	-0.009	-0.026	-0.100
Interviewer continuity (base = T=2, no ivwr continuity)						
T=2, ivwr same as last wave	-0.069	-0.062	-0.054	0.013	0.015	-0.043
T=3, ivwr not same as last wave	0.418 ***	0.409 ***	0.406 ***	-0.002	0.016	0.001
T=3, ivwr same as last wave only	0.302 ***	0.303 ***	0.274 ***	0.071	0.087	0.046
T=3, ivwr same as last 2 waves	0.345 ***	0.351 ***	0.359 ***	0.203 ***	0.219 ***	0.147 ***
T=4, ivwr not same as last wave	0.502 ***	0.489 ***	0.502 ***	0.021	0.041	0.047
T=4, ivwr same as last wave only	0.412 ***	0.411 ***	0.368 ***	0.140 ***	0.159 ***	0.112 **
T=4, ivwr same as last 2 waves only	0.441 ***	0.449 ***	0.412 ***	0.266 ***	0.283 ***	0.246 ***
T=4, ivwr same as last 3 waves	0.647 ***	0.639 ***	0.669 ***	0.379 ***	0.400 ***	0.323 ***
Previous wave interview situation						
Number of calls made	-0.022 ***	-0.024 ***	-0.022 ***	-0.029 ***	-0.029 ***	-0.030 ***
Part responding HH	-0.215 ***	-0.218 ***	-0.214 ***	-0.530 ***	-0.535 ***	-0.538 ***
Cooperative	0.336 ***	0.353 ***	0.339 ***	0.307 ***	0.320 ***	0.328 ***
Understanding	0.085	0.091	0.081	0.099 **	0.101 **	0.093 **
Suspicious	-0.060	-0.067	-0.058	-0.272 ***	-0.271 ***	-0.298 ***
Didn't return SCQ	-0.372 ***	-0.373 ***	-0.385 ***	-0.382 ***	-0.397 ***	-0.399 ***
Prop missing in labour force section				-1.115	-1.124	-1.527
Prop missing in income section				-2.082 ***	-2.059 ***	-2.009 ***
Prop missing in family/relationship section				-0.202	-0.214	-0.166
Prop missing in ivwr obs section				-0.823	-0.768	-0.537
Prop missing in special section				0.108	0.120	0.204
Prop missing in satisfaction/moving section				-0.427	-0.390	-0.277
Prop missing in SCQ if returned				-1.029 ***	-1.017 ***	-1.021 ***
Prop missing in childcare/housing				-0.758 **	-0.754 **	-0.659 **
Prop missing in HH roster				1.132	1.119	1.440
PQ length (/10^2)				0.644 **	0.625 **	0.601 *
PQ length squared (/10^4)				-0.848 **	-0.829 **	-0.783 **
PQ length missing				0.170	0.162	0.133
HQ length (/10^2)				0.897 *	0.891 *	0.893 **
HQ length squared (/10^4)				-0.678	-0.680	-0.754
HQ length missing				0.228 **	0.226 **	0.239 ***
Telephone interview				0.156 *	0.156 *	0.159 **
Respondent characteristics						
Moved	-0.754 ***	-0.757 ***	-0.760 ***			
Dwelling type (base = Separate/semi-detached house)						
Unit/flat	-0.112 *	-0.108 *	-0.145 **			
Non-private dwelling	-0.059	-0.052	-0.133			
Other dwelling	-0.397 **	-0.411 **	-0.378 ***			
Missing dwelling type	0.309	0.314	0.322			
Renter	-0.230 ***	-0.223 ***	-0.225 ***	-0.003	-0.027	-0.011
Female	0.122 ***	0.120 ***	0.121 ***	0.006	0.011	0.008
Age (/10)	0.065 ***	0.066 ***	0.069 ***	0.183 ***	0.188 ***	0.186 ***
Age squared (/10^2)				-0.018 ***	-0.018 ***	-0.018 ***

TABLE 3 (cont'd)

	Contact			*Response*		
Variable	*Model I*	*Model II*	*Model III*	*Model I*	*Model II*	*Model III*
Country of birth (base=Australia)						
Main English speaking country	-0.134 **	-0.133 **	-0.136 **	-0.068 *	-0.072 *	-0.082 **
Other o/s; Speaks English well	-0.262 ***	-0.269 ***	-0.268 ***	-0.160 ***	-0.169 ***	-0.157 ***
Other o/s; Not speak English well	-0.499 ***	-0.493 ***	-0.499 ***	-0.221 ***	-0.237 ***	-0.212 ***
Indigeneous Australian				-0.033	-0.033	-0.038
Marital status (base = Married)						
Defacto	-0.252 ***	-0.245 ***	-0.243 ***	0.029	0.018	0.036
Separated	-0.320 ***	-0.316 ***	-0.339 ***	0.139 **	0.125 *	0.143 **
Divorced	-0.366 ***	-0.359 ***	-0.390 ***	0.063	0.050	0.056
Widowed	-0.035	-0.031	-0.043	0.113 *	0.111 *	0.103 *
Single	-0.346 ***	-0.344 ***	-0.352 ***	0.052	0.038	0.044
Number of children aged 0-14	-0.050 **	-0.048 **	-0.052 ***	-0.001	-0.003	-0.001
Number of adults	-0.003	-0.005	-0.004	-0.108 ***	-0.107 ***	-0.109 ***
Education (base = Year 11 and below)						
Year 12				0.061 **	0.059 *	0.057 *
Certificate				0.060 **	0.059 **	0.052 *
Diploma				0.171 ***	0.168 ***	0.167 ***
Graduate				0.265 ***	0.262 ***	0.258 ***
Equivalised HH income ($/10^5$)				0.045	0.043	0.037
Equivalised HH income squared ($/10^{10}$)				-0.009	-0.009	-0.009 *
Employment / labour force status (base = Not in labour force)						
Unemployed	-0.084	-0.087	-0.075	-0.033	-0.044	-0.034
Employed part time (1-34 hrs)	0.200 ***	0.201 ***	0.213 ***	-0.002	0.006	0.002
Employed full time (35-54 hrs)	0.187 ***	0.191 ***	0.197 ***	-0.135 ***	-0.128 ***	-0.131 ***
Employed full time (55+ hrs)	0.068	0.064	0.083	-0.142 ***	-0.138 ***	-0.140 ***
Area of residence (base = Major city: Sydney)						
Major city: Melbourne	0.022	0.024	-0.037	-0.037	-0.034	0.060
Major city: Brisbane	0.042	0.031	-0.038	0.088	0.088	0.059
Major city: Adelaide	0.189 *	0.175 *	0.232 *	0.024	0.032	-0.038
Major city: Perth	-0.169 *	-0.162 *	-0.303 ***	0.042	0.036	0.031
Major city: other	0.202 *	0.190 *	0.179	-0.016	-0.010	0.056
Inner regional	-0.061	-0.063	-0.094	0.048	0.044	0.067
Outer regional	-0.114	-0.117	-0.138	0.032	0.026	0.026
Remote and very remote	-0.352 ***	-0.360 ***	-0.423 ***	-0.036	-0.054	-0.035
Index of disadvantage (base = Lowest quintile)						
2nd lowest quintile	-0.030	-0.021	-0.009	-0.025	-0.026	-0.009
Middle quintile	0.062	0.069	0.069	-0.056	-0.053	-0.066 *
2nd highest quintile	0.080	0.086	0.084	-0.039	-0.036	-0.025
Highest quintile	0.147 **	0.159 **	0.169 **	0.036	0.041	0.029
Serious long-term health condition	0.347	0.353	0.356	-0.216 **	-0.204 **	-0.227 **
Constant term	1.567 ***	1.544 ***	1.682 ***	0.656 ***	0.582 ***	0.698 ***
Pseudo log-likelihood	-2745	-13054	-2722	-10314	-13054	-10133
Pseudo R^2	0.228			0.090		
N	38778	38756	38761	38043	38756	38026

Notes: * $0.10 \geq p > 0.05$, ** $0.05 \geq p > 0.01$, *** $p \leq 0.01$.

(vi) Contact probabilities are significantly higher for women, married persons, and English speakers, and rise with age. The presence of a severe long-term health problem attracts a large positive coefficient but the estimate is very imprecise and so not statistically significant.

(vii) Counter to expectations, the number of children is a negative predictor of contact, whereas the number of adults in the household has no significant association with the likelihood of making contact. The negative effect of children is a perplexing result but nevertheless is rigorous to alternative specifications. For example, similar results are obtained when using a simple dummy variable representing the presence of children, or when disaggregating children based on their age.

(viii) With the exception of persons working very long hours (55 hours or more per week), it is easier establishing contact with employed persons than with the unemployed and other non-workers. This finding stands in contrast to results often reported in cross-section surveys and possibly reflects more extensive social networks among the employed which, in turn, makes it easier to trace sample members.

(ix) Clear straightforward associations with location are hard to find, though it is very clear that contact rates are, other things equal, relatively low in the remotest parts of Australia. This almost certainly reflects the high cost of getting interviewers to these areas. We also find that people living in the areas of least socio-economic disadvantage have significantly higher contact rates.

One other influence that is expected to influence contact probabilities, but not included in Table 3, is the amount of contact information collected at the preceding wave. In particular, apart from the usual contact details (e.g., mobile telephone numbers, email addresses, business contact details), all respondents were asked to provide contact details for friends or relatives who might know their whereabouts in the event that we cannot easily find them in the future. We thus included the number of names provided in a separate model (not reported here) using data from waves 2 and 3 (this information is not available for wave 1). Respondents who provided two contacts instead of one were found to be more likely to be contacted in the next wave, while respondents who refused to provide any contacts were less likely to be contacted in the next wave. However, contrary to the BHPS experience, neither of these differences is statistically significant.

Turning now to the likelihood of response conditional on making contact, we make the following observations:

(i) Unlike the GSOEP and BHPS experience, large interviewer workloads do not appear to have a detrimental impact on the likelihood of gaining an interview with a respondent. Indeed, interviewers with the highest workloads have tended to achieve the best response rates, but these differences are statistically insignificant. As hypothesized earlier, we believe the absence of a negative association reflects the way work is allocated to interviewers on the HILDA Survey.

(ii) Interviewer continuity is positively associated with response and the strength of this relationship increases with each wave the interviewer remains the same. Further, this result is only slightly weakened by controlling for a random interviewer effect in Model III. Based on the parameter estimates in Model III, the mean predicted probability of response in wave 4 when the interviewer is not the same as in wave 3 is 91.3 per cent. This compares with 92.2 per cent if the interviewer is the same for the last two waves, 93.8 per cent if the same for the last three years, and 94.6 per cent if the same for all four waves.

(iii) The number of calls made at wave *t-1* has the expected negative association with survey response at wave *t*. Thus a respondent from a household that required just one call at *t-1* is calculated to have a mean predicted response of 92.8 per cent. If five calls are required (the mean number in the sample) the predicted response rate falls to 91.2 per cent. At households requiring as many as 20 calls (the maximum in the data is 24 calls) the response probability at wave *t* is just 83 per cent.

(iv) Possibly the most important predictor of response in the next wave is whether the household was partly responding last wave. The estimates in Model I indicate the mean predicted probability of response for individuals in partly responding households is 81.9 per cent, compared with 91.1 per cent for individuals in fully responding households.

(v) Another important predictor is whether the individual returned the self-completion questionnaire after their interview last wave or not. The mean predicted probability of response when this questionnaire was returned is 92.1 per cent, compared to 85 per cent when it was not returned.

(vi) As has been found in other studies of attrition, the interviewer's assessment of the respondent in one wave is predictive of response in the next. Respondents who were cooperative, not suspicious and appeared to understand the questions were more likely to provide an interview in the next wave.

(vii) Also consistent with most previous research, the proportion of missing data items tends to be negatively associated with the likelihood of responding at the next wave. However, this effect is not uniform across all components of the survey. Specifically, these negative associations are only significant when the item non-response occurs in the income or childcare and housing sections of the interview, or if it occurs in the self-completion questionnaire.

(viii) In contrast to the analysis of attrition in the NLSY, SIPP and ECHP, conducting the previous wave interview by telephone was not predictive of attrition in the next wave. Indeed, the coefficient is positive (though only of marginal significance).

(ix) The highest level of participation in the next wave is observed with a personal interview of around 38 minutes duration, which is slightly in excess of the duration targeted (35 minutes). The implication thus is that both very short and very long interviews will result in less cooperation next wave. We believe that short interviews signal either the lesser relevance of the survey content to the respondent or respondent disinterest in the survey. Long interviews, on the other hand, obviously impose a greater time cost on the respondent. Household interview length, on the hand, is positively associated with the likelihood of participation in the next wave, a result we believe reflects the importance of topic salience.

(x) Response probabilities are lowest among both the young and elderly (people aged around 50 are the most cooperative). Response probabilities are also relatively low for persons born overseas (and especially those who speak English poorly), the least educated, and those with a serious long-term health condition, and decline with the number of adults living in the household. Sex, marital status, the number of children, Aboriginality, (equivalised) household income and location are all insignificant predictors of the probability of response.

(xi) While employed people are somewhat easier to establish contact with, they are less likely to respond if they work full-time hours (35 or more per week).

6. Conclusion

There are perhaps three main conclusions that can be distilled from the empirical analysis reported in this paper. First, the factors that influence contact are quite distinct from those that influence cooperation, so it follows that the empirical modeling process should reflect this. That said, it should still be borne in mind that contact rates in most well-resourced surveys are likely to be very high; in the HILDA Survey they average around 98 per cent each wave. Inevitably it is the cooperation stage where the risk of sample loss is greatest.

Second, while this paper was partly motivated by interest in identifying survey design features that influence response probabilities, it is actually very difficult to test the influence of design features without controlled experiments. Our analysis, however, is not without any implications for survey design. Most obviously, we found evidence that quite large interviewer workloads are beneficial for response. This is almost certainly a function of the practice in the HILDA Survey of assigning more work to better performing interviewers. Our results also support the contention that interviewer continuity improves response, though the magnitude of the effect is small and restricted to response conditional on contact. More surprising, we found no evidence that obtaining more contact names at one wave significantly improves either contact or response probabilities at the next, though the direction of effect is in the expected direction.

Third, while there is undoubtedly (and thankfully) a large random component to survey non-response, it is nevertheless very clear that there are strong associations between many observable characteristics of both respondents and the interview process and experience that are predictive of non-response. Indeed, arguably the most striking feature of the analysis reported on here is just how many different variables are included in the analysis. Such information can be used to assist in tailoring approaches and targeting special measures at the next wave. It also can provide variables for inclusion in attrition models used in the construction of population weights or as instruments at the analysis stage.

References

Atrostic, B.K., Baytes, N. and Silberstein, A. (2001), 'Nonresponse in U.S. Government Household Surveys: Consistent Measures, Recent Trends, and New Insights', *Journal of Official Statistics*, vol. 17, no. 2 (June), pp. 209-226.

Australian Bureau of Statistics (2003), *Information Paper: Census of Population and Housing – Socio-Economic Indexes for Areas, Australia* (ABS cat. no. 2039.0), ABS, Canberra.

Becketti, S., Gould, W., Lillard, L. and Welch, F. (1988), 'The Panel Study of Income Dynamics after Fourteen Years: An Evaluation', *Journal of Labor Economics*, vol. 6, no. 4 (October), pp. 472-492.

Behr, A., Bellgardt, E. and Rendtel, U. (2005), 'Extent and Determinants of Panel Attrition in the European Community Household Panel', *European Sociological Review*, vol. 21, no. 5 (December), pp. 489-512.

Branden, L., Gritz, R.M. and Pergamit, M. (1995), 'The Effect of Interview Length on Attrition in the National Longitudinal Study of Youth', National Longitudinal Surveys Discussion Paper no. 28, Bureau of Labor Statistics, US Department of Labor, Washington DC.

Bureau of Labour Statistics (2003), *NLS Handbook, 2003*, US Department of Labour, Bureau of Labour Statistics, Washington DC.

Burkam, D.T. and Lee, V.E. (1998), 'Effects of Monotone and Nonmonotone Attrition on Parameter Estimates in Regression Models with Educational Data', *The Journal of Human Resources*, vol. 33, no. 2 (Spring), pp. 555-574.

De Leeuw, E., and de Heer, W. (2002), 'Trends in Household Survey Nonresponse: A Longitudinal and International Comparison', in Groves, R.M., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (eds), *Survey Nonresponse*, John Wiley and Sons, New York.

Fitzgerald, J., Gottschalk, P. and Moffitt, R. (1998), 'An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics', *The Journal of Human Resources*, vol. 33, no. 2 (Spring), pp. 251-299.

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 Life Style Survey', *The Statistician*, vol. 45, no. 2, pp. 163-183.

Hawkes, D. and Plewis, I. (2006), 'Modelling Non-response in the National Child Development Study', *Journal of the Royal Statistical Society (Series A)*, vol. 169, part 3 (July), pp. 479-492.

Hill, D.H. and Willis, R.J. (2001), 'Reducing Panel Attrition: A Search for Effective Policy Instruments', *The Journal of Human Resources*, vol. 36, no. 3 (Summer), pp. 416-438.

Kalsbeek, W.D., Yang, J. and Agans, R.P. (2002), 'Predictors of Nonresponse in a Longitudinal Survey of Adolescents', *Proceedings of the American Statistical Association (Survey Research Methods Section)*, pp. 1740-1745.

Laurie, H., Smith, R. and Scott, L. (1999), 'Strategies for Reducing Nonresponse in a Longitudinal Panel Survey', *Journal of Official Statistics*, vol. 15, no. 2 (June), pp. 269-282.

Lee, E., Hu, M.Y. and Toh, R.S. (2004), 'Respondent Non-cooperation in Surveys and Diaries: An Analysis of Item Non-response and Panel Attrition', *International Journal of Market Research*, vol. 46, no. 3 (Autumn), pp. 311-326.

Lepkowski, J.M. and Couper, M.P. (2002), 'Nonresponse in the Second Wave of Longitudinal Household Surveys', in R.M. Groves, D.A. Dillman, J.L. Eltinge and R.J.A. Little (eds), *Survey Nonresponse*, John Wiley and Sons, New York.

Lillard, L.A., and Panis, C.W.A. (1998), 'Panel Attrition from the Panel Study of Income Dynamics: Household Income, Marital Status, and Mortality', *The Journal of Human Resources*, vol. 33, no. 2 (Spring), pp. 437-457.

Loosveldt, G., Pickery, J. and Billiet, J. (2002), 'Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey', *Journal of Official Statistics*, vol. 18, no. 4 (December), pp. 545-557.

Nicoletti, C. and Buck, N. (2003), 'Explaining Contact and Refusals in the British and German Household Panels', in N. Buck, C. Nicoletti, A. McCulloch and J. Burton, *Report on Attrition Analysis and Item Non-response*, Chintex Working Paper no. 16, February.

Nicoletti, C. and Peracchi, F. (2005), 'Survey Response and Survey Characteristics: Microlevel Evidence from the European Community Household Panel', *Journal of the Royal Statistical Society, Series A*, vol. 168, part 4 (November), pp. 763-781.

Olsen, R.J. (2005), 'The Problem of Respondent Attrition: Survey Methodology is Key', *Monthly Labour Review*, vol. 128, no. 2 (February), pp. 63-70.

Schräpler, J.P. (2002), 'Respondent Behavior in Panel Studies – A Case Study for Income-Nonresponse by Means of the German Socio-Economic Panel (GSOEP)', Paper presented at the 5th International Conference of German Socio Economic Panel Users, Berlin, 3-4 July.

Spieß, M. and Kroh, M (2004) 'Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (GSOEP)', Research Note No. 28a, German Institute for Economic Research (DIW), Berlin.

Taylor, M.F. (ed.) with Brice, J., Buck, N. and Prentice-Lane, E. (2005), *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

Tourangeau, R. (2003), 'The Challenge of Rising Nonresponse Rates', in R. Tourangeau, *Recurring Surveys: Issues and Opportunities*, Report to the National Science Foundation based on a Workshop held on March 28-29, 2003, National Science Foundation.

Wadsworth, M.E.J., Butterworth, S.L., Hardy, R.J., Kuh, D.J., Richards, M., Langenberg, C., Hilder, W.S. and Connor, M. (2006), 'The Life Course Prospective Design: An Example of Benefits and Problems Associated with Study Longevity', *Social Science and Medicine*, vol. 57, no. 11 (December), pp. 2193-2205.

Voogt, R.J.J. and Saris, W.E. (2005), 'Mixed Mode Designs: Finding the Balance Between Nonresponse Bias and Mode Effects', *Journal of Official Statistics*, vol. 21, no. 3 (September), pp. 367-387.

Waterton, J. and Lievesley, D. (1987), 'Attrition in a Panel Study of Attitudes', *Journal of Official Statistics*, vol. 3, no. 3 (September), pp. 267-282.

Watson, D. (2003), 'Sample Attrition Between Waves 1 and 5 in the European Community Household Panel', *European Sociological Review*, vol. 19, no. 4 (September), pp. 361-378.

Watson, N. and Wooden, M. (2004), 'The HILDA Survey Four Years On', *The Australian Economic Review,* vol. 37, no. 3 (September), pp. 343-349.

Westat (1998), *SIPP Quality Profile 1998* (SIPP Working Paper no. 30), US Department of Commerce, Bureau of the Census, Washington DC.

Westat, in association with Mathematica Policy Research (2001), *Survey of Income and Program Participation Users' Guide* (3rd ed.), US Department of Commerce, US Census Bureau, Washington DC.

Winkels, J.W. and Withers, S.D. (2000), 'Panel Attrition', in D. Rose (ed.), *Researching Social and Economic Change: The Uses of Household Panel Studies*, Routledge, London.

Zabel, J.E. (1998), 'An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labour Market Behavior', *The Journal of Human Resources*, vol. 33, no. 2 (Spring), pp. 479-506.