Identifying Factors Affecting Longitudinal Survey Response

Nicole Watson and Mark Wooden
Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne

Abstract

This chapter examines the factors that influence continued participation by sample members in longitudinal surveys. It is structured into two distinct parts. First, evidence from previous research that has modeled the response process within a multivariate framework is reviewed. Second, estimates of predictors of response from a new national household panel survey – the Household, Income and Labour Dynamics in Australia (HILDA) Survey – are presented. Following other recent treatments in the literature, the estimation model treats survey participation as involving two sequential events, contact and response.

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 *et al*., 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 *et al*., 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).

Of course rates of attrition might be different in other types of longitudinal surveys employing different designs and covering different populations. Birth cohort studies, for

[1] This figure includes deaths. If deaths are excluded, the accumulated attrition rate declines to about 22%.
[2] Winkels and Withers (2000) reported that after 11 years of interviewing only 42 per cent of the original sample remained.

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 do not need to be 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). Nevertheless, the NLSY experience 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.

There is also mounting evidence indicating that the problem of sample attrition, has been getting worse over time (Atrostic *et al.*, 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.[3] 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).

In general, a high rate of sample attrition poses 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. Of greater concern, 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, and it is this which is the subject of this chapter. More specifically, this chapter reviews the growing body of empirical evidence on the determinants of response to, and attrition from, longitudinal surveys. We are especially interested in those factors that are most amenable to manipulation by survey designers and administrators. As a result, the emphasis here is more on the role of *survey design features* and the *interview situation*, and less on the demographic characteristics of non-respondents, information which can be readily distilled from all longitudinal survey data sets and is regularly used in the construction of population weights that adjust for attrition. We then estimate a model predicting response over the course of the first four waves of the Household, Income and Labour Dynamics in Australia (HILDA) Survey, a relatively new household panel survey that has experienced significant attrition.

[3] 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.

2. Factors Affecting Response and Attrition

Following Lepkowski and Couper (2002), the response process can be divided into three conditional stages: (i) the sample member is located; (ii) contact with the sample member is established; and (iii) the sample member provides an interview.

2.1 Locating the Sample Member

Locating sample members is a straightforward exercise where the sample member has not moved. Determining the whereabouts of movers, however, is often far from straightforward. Moreover, changing address has consistently been found to exert a highly significant negative influence on the likelihood of response at the next wave (Becketti, *et al*., 1988; Behr *et al*., 2005; Lepkowski and Couper, 2002; Lillard and Panis, 1998; Watson, 2003; Zabel, 1998). Sample unit characteristics, of course, have a major bearing on mobility (Buck, 2000) and hence the likelihood of establishing location. Studies of attrition, however, do not explicitly model the decision to move and thus cannot separate the effect of a variable on attrition via mobility from its effect via other stages in the response process. Nevertheless, studies that model the location / contact outcome separately from the cooperation outcome generally find that movement per se does not have any negative effects on cooperation once the sample member has been found and contacted (Gray *et al.*, 1996; Lepkowski and Couper, 2002). Such findings are consistent with findings from cross-section surveys which have found slightly lower refusal rates among recent movers compared with members of households that have not moved (Groves and Couper, 1998, p. 139).

Where movement has taken place, the likelihood of successful location can be influenced by the tracking procedures that are implemented, the length of time between survey waves, and the extent and nature of contact between waves. Tracking is especially important (see Couper and Osftedal, Chapter X in this volume) and all well designed longitudinal surveys will ensure that information is collected during each interview which can be subsequently used to locate the respondent if they move (such as work, home and mobile numbers, email addresses, and contact details of relatives or friends). How successful interviewers are at extracting this information is potentially very important for subsequent attrition. For example, Laurie *et al*. (1999) analysed attrition over the first four rounds of the BHPS and found a strong positive association between response and the provision of a contact name at wave 1. Of course, even when armed with details of contact persons, it does not automatically follow that the details provided at the previous interview will be current or that the contact person knows the current location of the sample member. Other tracking mechanisms will thus need to be employed, such as searching telephone directories and electoral rolls, or asking neighbours or other local people who might know where the sample member may have moved to.[4]

2.2 Contacting the Sample Member

In many respects, the contact task in longitudinal surveys is little different from that in cross-section surveys and, as discussed in Groves and Couper (1998, Chapter 4), will be affected by such factors as the accessibility of the dwelling (in the case of face-to-face interviews), the use of answering machines (in the case of telephone surveys) and the availability of the sample member. Survey design features that can influence the rate of successful contact include the number of calls attempted and the timing of those calls (over the day, week or year), and the duration of the fieldwork period. Cross-section surveys, however, are often not well placed to assess how important some of these influences are because of both lack of information about non-respondents, though this problem can be overcome if survey samples

[4] For a review of the variety of tracking mechanisms employed in panel studies, see Ribisl *et al*. (1996).

can be matched to some external data source, such as Census records (e.g., Foster, 1998; Groves and Couper, 1998), and the potential endogeneity of some design features (e.g., number of calls made). In contrast, longitudinal surveys typically contain a lot of information about non-respondents after the initial wave, and the panel nature of the data permits stronger inferences about causality.

Nevertheless, empirical studies that both employ longitudinal data and distinguish the contact stage from the cooperation stage remain relatively scarce. The major studies that we have identified are: Gray *et al.* (1996), who analysed response to the two waves of a health and lifestyle survey conducted in Britain in 1984-85 and 1991-92; Thomas *et al.* (1991), who reported on attrition in the four-year follow-up to the 1993 Indonesian Family Life Survey; Lepkowski and Couper (2002), who examined response in two separate two-wave longitudinal surveys conducted in the US; Nicoletti and Buck (2003), who analysed survey response over three waves of both the BHPS and GSOEP panels; and Nicoletti and Peracchi (2005), who examined response behaviour over the first five waves of the ECHP. Variables that have featured prominently in the models of contact estimated in one or more of these studies include: residential mobility (discussed earlier); socio-demographic characteristics hypothesized to be associated with the likelihood of finding someone at home (such as age, sex, and household size and composition); characteristics of the region in which the sample member resides; measures of community attachment (such as home ownership and caring responsibilities); interview workloads; interviewer continuity; the number of calls made; and the length of the fieldwork period. Space precludes a detailed discussion of these findings and so we focus here on data collection characteristics.

It is axiomatic that making additional attempts to contact a sample member who has yet to be found should raise the probability of contacting that sample member. The evidence from longitudinal surveys, however, indicates 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 *et al*., 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 (though only Nicoletti and Buck (2003) and Nicoletti and Peracchi (2005) specifically identify attrition due to non-contact).

Of course, it is not just the number of calls made that matters, but when those calls are made (see Groves and Couper, 1998). In longitudinal surveys this is less of an issue given interviewers will typically have available to them a lot of information about sample members, including the call pattern and interview times from the previous wave, to determine appropriate call times. What interviewers often have less control over is how long they spend in the field. Longer fieldwork periods should enhance the likelihood of making contact; for example, by increasing the probability of reaching sample members who are away temporarily from their home address. Nicoletti and Peracchi (2005), in their analysis of response behaviour over the first five waves of the ECHP, were able to include a measure of the duration of the fieldwork period, since this varied both over time and across the different countries participating in the ECHP. As expected, this variable was found to be associated with higher contact probabilities.

It is also widely recognized that interviewers can have a major influence on survey response, though only rarely do researchers distinguish between effects at the contact stage and at the cooperation stage. Many researchers, for example, point to the benefits of interviewer continuity in enhancing survey response (more on this below), but 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.

Finally, Nicoletti and Buck (2003) have also examined whether differences in interviewer workloads might affect contact probabilities. The estimated coefficients are mostly negatively signed when using the BHPS data but positively signed when using data from the GSOEP. The magnitude of the coefficients, however, are always quite small, and of only marginal significance at best.

2.3 Obtaining the Cooperation of the Sample Member

The factors that affect response once contact has been made are highly varied and often out of the interviewer's control. For example, willingness to participate in further waves of interviews can be expected to vary with individual personality traits and with how busy sample members are. Nevertheless, the way surveys are designed and administered can also have a marked effect on response. For example, cooperation with surveys has been found to be influenced by the use of incentives, interviewer experience and continuity, respondent identification with the study, the survey topic, and the interview experience in prior waves.

Incentives

Panel studies often use incentives, both in the form of cash payments and gifts, to improve the initial response rate and reduce subsequent attrition. The experience with incentives is reviewed elsewhere in this book (see Lynn and Laurie, Chapter X in this volume) and so we do not spend much space on that issue here. Suffice to say that incentives can enhance response rates, though their effectiveness varies widely depending on how they are designed and administered.

Interviewer Effects

It is widely recognized that interviewers play a vital role in gaining cooperation from respondents. In cross-sectional surveys, for example, response rates have been found to improve with interviewer age and experience and are also influenced by interviewer attitudes and confidence (e.g., Couper and Groves, 1992; Groves and Couper, 1998; Japec and Lundqvist, 1999; Martin and Beerten, 1999). In longitudinal studies the role of the interviewer is arguably even more important given the task of maintaining respondent interest is likely to become more difficult as respondents start to lose interest in the study or even question its salience. Interestingly, Nicoletti and Buck (2003) reported evidence that high interviewer workloads (households per interviewer) were negatively associated with response probabilities (conditional on contact being made). Such findings suggest that concentrating workloads on the most experienced interviewers may not necessarily bring any response gains if those workloads are too high.

As previously noted, one factor that is regularly reported as being beneficial for response in longitudinal studies, at least those administered on a face-to-face basis, is interviewer continuity (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; Zabel, 1998). The size of the effect, however, is highly variable across studies, and in some cases contradictory. For example, both Behr *et al*. (2005) and Nicoletti and Peracchi (2005) analysed the ECHP experience, with the latter finding small insignificant effects and the former highly significant and large relationships. More importantly, the coefficient on a simple dummy variable identifying interviewer continuity will almost certainly be biased away from zero. This is because attrition also occurs among interviewers, and interviewer attrition can be expected to be higher in areas where survey response rates are lower. In other

words, the interviewer continuity variable will often capture part of any systematic difference across areas in response. Nevertheless, there is still disagreement on just how large this bias is. Campanelli and O'Muircheartaigh (2002), using data from the first four waves of the BHPS, the same data used by Laurie *et al*. (1999), distinguished between regions with and without interviewer attrition and reported that nonrandom interviewer attrition explained all of the interviewer continuity effect reported by Laurie *et al*. (1999). In contrast, Nicoletti and Buck (2003), also using BHPS data (as well as data from the GSOEP), reported that interviewer continuity remained an important and valid predictor of response even after controlling for the presence of a random component linked to interviewers.

Very differently, Olsen (2005) dealt with the endogeneity issue by including, in his analysis of response to the NLSY79, two variables that measured the ability of the fieldwork company to exploit interviewer continuity in the current round, and not whether interviewer continuity actually held. While he reported evidence of a response advantage to interviewer continuity, the magnitude of the effect was quite small, indicating an improvement in the attrition rate of only 0.7 percentage points, which was only apparent after the respondent had been interviewed twice by that interviewer.

The Interview Experience

One influence about which there is little disagreement is respondents' previous interview experience. Interviewer assessments of how cooperative the respondent was, how much they enjoyed the interview and / or the ease with which the survey questions could be answered, have invariably been found to be highly predictive of cooperation at the next survey wave (e.g., Kalton *et al*., 1990; Branden *et al*., 1995; Laurie *et al*., 1999; Hill and Willis, 2001; Lepkowski and Couper, 2002; Olsen, 2005).

It is also generally assumed that missing data, or item non-response, is indicative of an unpleasant or negative interview experience (Loosveldt *et al*., 2002). It thus follows that item non-response at one wave should also be predictive of unit non-response at the next, and again the weight of evidence suggests this is so (Burkam and Lee, 1998; Zabel, 1998; Laurie *et al*., 1999; Loosveldt *et al*., 2002; Schräpler, 2002; Lee *et al*., 2004; Hawke and Plewis, 2006), though there are exceptions. Serfling (2004), for example, in contrast to Schräpler (2002), could find no evidence of a monotonic relationship between item nonresponse on income-related items and unit nonresponse in data from three waves of the GSOEP. He did, however, report weak evidence of an inverse U-shaped relationship. The proposed explanation here is that some respondents will only participate in the survey because they know they do not actually have to answer questions on sensitive or difficult to answer topics.

Another aspect of the interview experience often asserted to be an important influence on sample attrition is the time taken to administer the interview, with the likelihood of respondent cooperation expected to be inversely related to expected interview length. A number of longitudinal studies have included interview length as a predictor of response in the next wave, but have found little evidence of a significant negative effect. Indeed, the estimated coefficient is often positively signed (Branden *et al*., 1995; Hill and Willis, 1998; Zabel, 1998). With hindsight such findings are easily explained. 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 the respondents who are most engaged by the study and find it most enjoyable will typically have longer interviews. Testing the effect of interview length thus requires experimental evidence, and there are few examples of such experiments being conducted on longitudinal data. Nevertheless, Zabel (1998) reported on the experience of the PSID which made an explicit attempt to reduce the interview length in 1972 and which appeared to have had the effect of reducing attrition, though the magnitude of the effect was

not large. This finding is consistent with the weight of evidence from experimental research conducted on cross-section surveys which, according to the review by Bogen (1996, p. 1024), leads to the conclusion that the "relationship between interview length and nonresponse is more often positive than not, but it is surprisingly weak and inconsistent".

It is also widely accepted that survey delivery mode has a significant impact on response rates, with methods involving greater levels of personal contact between the survey administrator and the respondent usually found to have higher rates of respondent cooperation, but for a greater cost (see Yu and Cooper, 1983). This thus suggests that changing modes of collection during the administration of a panel could affect sample attrition. Unfortunately, while a number of large panel studies (including, for example, the PSID, the UK National Child Development Study, the 1992-97 British Election Panel Study and the 1984-91 Australian Longitudinal Survey) have changed survey modes, rarely is the impact of this change on response rate reported on. Zabel (1998), however, reported that the switch to telephone interviewing in the PSID following five successive years of personal interviewing had a positive, but statistically insignificant, impact on attrition.

Very differently, providing respondents with multiple modes for responding, as is the practice in the GSOEP, is also expected to reduce respondent reluctance, a hypothesis that has received support from experiments conducted in conjunction with cross-section surveys (e.g., Voogt and Saris, 2005). We, however, are unaware of any experimental evidence gathered from a longitudinal design. Indeed, in many longitudinal studies employing predominantly face-to-face survey methodologies, 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 is likely to 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 support this hypothesis.

2.4 The Role of Respondent Characteristics

As mentioned in the introduction, associations between respondent characteristics and survey response are not the main focus of this chapter. This stands in marked contrast to the many studies of the correlates of sample attrition where personal and household characteristics represent most, if not all, of the explanatory variables included (e.g., Becketti *et al*., 1988; Behr *et al*., 2005; Burkam and Lee, 1998; Fitzgerald *et al*, 1998; Gray *et al*., 1996; Hawkes and Plewis, 2006; Kalsbeek *et al*., 2002; Watson, 2003). The emphasis placed on respondent characteristics in these studies should not be surprising. First, and most obviously, detailed information about non-respondents is something that is readily available for all persons who exit longitudinal surveys – it is only non-respondents at wave 1 about whom relatively little is usually known. Second, differences in attrition propensities may be suggestive of possible attrition bias, and thus can be used to help correct for such bias. The importance of respondent characteristics thus cannot be ignored here. In any case, the effect of survey design features may be conditional upon respondent characteristics, so it is appropriate to control for respondent characteristics when attempting to identify the effects of survey design. This is the approach taken in Section 3.

Before that, however, we summarise what the literature tells about the characteristics of non-respondents to longitudinal surveys and in what ways they differ from respondents. Note that the scope of this review is limited to studies that have analysed sample attrition, though we draw on insights from the much larger body of evidence based on single-wave cross-section surveys where appropriate. Further, we restrict our attention to those studies that have employed multivariate methods. Where possible we attempt to distinguish between contact and cooperation but as observed earlier, we are limited by the small number of studies that

have explicitly separated out these two stages in the response process. Finally, we limit our attention to the major demographic and socio-economic variables that are regularly included in attrition models.

Sex

Studies of survey response nearly always find that response rates are higher among women than among men, and this is no less true of analyses of survey attrition. The main reason usually cited for this is that women are at home more frequently. Nevertheless, there is limited evidence that even conditional on contact, men may be slightly more likely to discontinue survey participation (Lepkowski and Couper, 2002; Nicoletti and Buck, 2003).

Age

A widely reported result in cross-section surveys is that response rates are lowest for both the youngest and oldest members of the population. The evidence from longitudinal surveys confirms that survey response rates tend to be relatively low among young people. At the other end of the age spectrum the results are less clear. We would expect the older population to be less mobile and thus easier to reach, and a number of studies have reported evidence consistent with this hypothesis (e.g., Gray *et al.*, 1996; Lepkowski and Cooper, 2002; Thomas *et al.*, 2001). In aggregate, however, the evidence is more mixed. Some studies have found that overall attrition propensities are rising in old age (e.g., Becketti *et al.*, 1988; Fitzgerald *et al.*, 1988), others have reported the reverse (e.g., Hill and Willis, 2001), while still others have reported no clear evidence in either direction (e.g., Behr *et al.*, 2005; Nicoletti and Buck, 2003; Nicoletti and Peracchi, 2005). Perhaps it is not age *per se* that matters for survey cooperation but the salience of the survey content. Hill and Willis (2001), for example, speculated that the positive relationship found between age and survey response in their study of attrition from wave 3 of the US Health and Retirement Study (HRS) might reflect the greater relevance of the HRS to the eldest members of their sample.

Race / Ethnicity

Studies that have included controls identifying racial and ethnic minority groups have generally found negative relationships with response probabilities (e.g., Burkam and Lee, 1998; Zabel, 1998). Those studies that distinguish between contact and cooperation, however, find that this relationship is mainly due to lower rates of contact and not higher rates of refusals (Gray *et al.*, 1996; Lepkowski and Cooper, 2002). That said, an important mediating factor that appears to have been largely ignored (but see Burkam and Lee, 1998) is language speaking ability. In English-speaking countries, for example, cross-section surveys almost always report higher rates of survey non-response among non-English speakers.

Marital Status

Marital status is another variable where there is widespread consensus; single people have a higher propensity to attrit than do married people. Whether this higher rate of attrition is due to lower rates of contact or higher rates of refusal is less clear. The results reported by Gray *et al.* (1996) suggest it is mainly a function of lower contact probabilities, but supportive evidence for this from other studies of survey attrition is lacking.

Household Size and Composition

The evidence from cross-section surveys indicates that single-person households will typically have lower contact probabilities than larger households and may even have higher refusal rates, though the evidence here is more mixed. The results of both Gray *et al.* (1996) and Thomas (2001) suggest that it is only the impact on contact propensity that matters for attrition. A small number of studies also distinguish between households not just on the basis

of the number of family members, but whether those family members are children or adults. More often than not it is the presence of children that seems to be most strongly, and negatively, associated with attrition probabilities (e.g., Fitzgerald *et al*., 1988; Kalton *et al*., 1990; Nicoletti and Peracchi, 2005; Zabel, 1998). Again, the most plausible explanation for this finding is that the presence of children increases the likelihood of the respondent being home. Somewhat perplexing are the results reported by Nicoletti and Buck (2003). They found in both their BHPS and GSOEP samples that the number of adults in the household was associated with both lower contact rates and higher refusal rates. Watson (2003) also reported higher attrition rates in larger households, but only after single-person households and the number of children had been controlled for.

Education

Education is usually thought to be positively associated with survey response, mainly because those with higher educational attainment are likely to better appreciate the utility of research and information gathering activities (Groves and Couper, 1998, p. 128). The evidence from attrition studies is mostly in line with this hypothesis (e.g., Behr *et al.*, 2005, Fitzgerald *et al.*, 1998; Gray *et al.*, 1996; Lepkowski and Couper, 2002; Lillard and Panis, 1998; Watson, 2003), though the magnitude of the relationship is arguably quite small.

Home Ownership

Home ownership is frequently included in models of survey attrition and in nearly all cases attracts a negative coefficient (Fitzgerald *et al.*, 1998; Kalton *et al.*, 1990; Lepkowski and Couper, 2002; Watson, 2003; Zabel, 1998). The rationale for the inclusion of this variable offered by Lepkowski and Couper (2002) is that it is an indicator of community attachment, and so is expected to be positively associated with contact propensity. In fact, their results suggest that cooperation may also be positively affected, but other studies have found that the effect mainly works through enhanced contact probabilities (Gray *et al*., 1996; Thomas *et al.*, 2001; Nicoletti and Peracchi, 2005).

Income

A widely held view is that response rates, at least for cross-section surveys, tend to be lowest in both tails of the income distribution. The evidence from longitudinal surveys, however, is mixed. While a number of studies have reported evidence of a quadratic (inverted u-shaped) relationship between income measures and re-interview rates (Becketti *et al.*, 1988; Hill and Willis, 1998), others have reported no significant decline at high income levels (e.g., Fitzgerald *et al*., 1998; Kalton et al., 1990; Watson, 2005). The magnitude of the estimated effects, however, are typically small and given the number of studies that have found no evidence of any significant relationship (e.g., Gray *et al.* 1996; Lepkowski and Couper, 2002; Nicoletti and Peracchi, 2005; Zabel, 1998), it can probably be concluded that income is relatively unimportant for attrition.

Labour Force Status

It might be expected that, compared with the unemployed and the economically inactive, employed persons will be both harder to contact, since they spend more time away from the home, and more likely to refuse, given the greater opportunity cost of their time. The analyses of pooled data from the ECHP by Nicoletti and Peracchi (2005) and Watson (2003) both report evidence supportive of this hypothesis with survey re-interview rates highest among the economically inactive. The country-specific analysis of these same data by Behr et al. (2005), however, found labour market inactivity significantly and positive associated with survey response in only four of the 14 participant countries. Very differently, analysis of PSID data suggests no strong association between employment status and attrition. The

analysis of Fitzgerald *et al.* (1998) initially revealed higher attrition rates among non-workers than among workers, but this effect became insignificant in the presence of other controls. Zabel (1998) also reported an insignificant relationship. In contrast, and counter to expectations, Gray *et al.* (1996) found attrition rates to be lowest among the employed, the result of both lower refusal rates and a greater likelihood of tracing employed respondents. Similarly, Lepkowski and Couper (2002) found employment to be positively and strongly associated with contact propensity and positively but weakly associated with cooperation propensity in one of their samples (in the other the relationships were insignificant). To further confuse matters, Nicoletti and Buck (2003) reported a significantly higher cooperation propensity among the economically inactive in BHPS data, but significantly lower contact probabilities for both the unemployed and the inactive in GSOEP data.

Location

Finally, it is common to allow for regional differences in survey response, with many studies distinguishing between urban and rural localities. The usual expectation is that residents in large cities will be both less available and harder to reach (e.g., due to the security features of their housing) (Groves and Couper, 1998, p. 85). Social isolation explanations for survey participation also suggest that cooperation rates might be superior in smaller rural communities. For the most part the evidence is consistent with these hypotheses, with numerous studies reporting higher survey attrition rates among sample members living in urban locations (Burkam and Lee, 1998; Fitzgerald *et al*., 1998; Gray *et al.*, 1996; Kalton *et al*., 1990; Zabel, 1998). Only Lepkowski and Couper (2002) report any contrary evidence.

3. Predicting Response in the HILDA Survey

We now examine the experience with sample attrition in the HILDA Survey in order to test some of the hypotheses and relationships summarized in the foregoing review. There are at least three reasons why the HILDA Survey is well suited to this task. First, and as noted earlier, it has experienced significant sample attrition over its first four waves. Second, the data available explicitly distinguish between the contact and cooperation stages. Third, and perhaps most significantly, the range of explanatory variables available far exceeds what has been reported in any previous analysis of attrition from a longitudinal survey.

3.1 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 (including the PSID, the GSOEP and the BHPS), the sample is extended each year to include any new household members resulting from changes in the composition of the original households. With the exception of children of original sample members (OSMs) and persons who have a child with an OSM, new sample members only remain in the sample for as long as they live with an OSM. 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.

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.

Details about the evolution of the responding sample over the first four waves are provided in Table 1. This table shows that 10,565, or 76 per cent, of those persons initially interviewed in wave 1 were re-interviewed in wave 4. If deaths and movements out of scope are excluded the four-wave sample retention rate rises to 78 per cent. The wave-on-wave attrition rates (calculated as the proportion of in-scope previous wave respondents not providing an interview) for waves 2, 3 and 4 were 13.2, 9.6 and 8.4 per cent, respectively.

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

Wave first interviewed	*Wave 1*	*Wave 2*	*Wave 3*	*Wave 4*
Wave 1	13969	11993	11190	10565
Wave 2	-	1048	705	594
Wave 3	-	-	833	543
Wave 4	-	-	-	706
TOTAL	13969	13041	12728	12408

3.2 Estimation Approach

Pooled data from the first four waves of the HILDA Survey are now used to estimate a model that employs information about the respondents and their interviews in wave *t-1* to predict response at wave *t*. Earlier we argued that the response process involve three stages. 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 with the sample member; and (ii) obtaining the cooperation of the sample member once contacted. Our model is thus essentially the same as that used by both Nicoletti and Buck (2003) and Nicoletti and Peracchi (2005).

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 an OSM. 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.[5]

3.3 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.

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. In addition, it is 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)[6]; 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.[7] Also included is a measure of the relative socio-economic disadvantage of the neighbourhood in which the respondent lives.[8]

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,[9] being interviewed over the telephone, 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.

---

[5] 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.

[6] 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 74 years and could find no evidence of any significant association with contact probabilities.

[7] 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.

[8] 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.

[9] 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.

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. It is expected 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. 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.[10] In addition, we also include measures of: the length of the personal interview; the length of the household interview; the number of questions asked during the interview (which is a proxy for the relevance of topics covered in the questionnaire to the respondent); the proportion of missing responses in each section of the questionnaire; whether an indigenous Australian; educational attainment; and equivalised household income.

3.4 Results

The results of our model estimation are presented in Table 2. 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), but like Nicoletti and Peracchi (2005) we find that the estimated coefficients 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.[11] 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:

[10] To help identify Model II, dwelling type and housing tenure (in addition to the variable identifying individuals that move) are excluded from the response equations in all models. Neither were significant when included in the response equation in Model I.

[11] 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).

Table 2
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 2 (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: * $0.10 \geq p > 0.05$, ** $0.05 \geq p > 0.01$, *** $p \leq 0.01$.

(i) The optimal interviewer workload for maximizing contact rates is 124 previous wave respondents. 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.[12]

(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 a little 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 2, 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

[12] 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).

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, the interviewer workload does not appear to have had 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. Nevertheless, the magnitude of this effect is relatively small. 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 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 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, 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 relatively low rates of 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).

4. Conclusion

There are perhaps five main conclusions that can be distilled from the review and empirical analysis reported in this analysis. First, it cannot be automatically assumed that the experience from cross-section surveys will always be relevant when considering the correlates with attrition from longitudinal surveys. This is most obvious with respect to the role of the interview experience. In single wave surveys the way the interview is conducted can only adversely affect interview duration and not whether an interview is obtained. In contrast, the evidence presented here suggests that the respondent's perception of the interview experience is possibly the single most important influence on cooperation in future survey waves.

Second, the factors that influence contact are quite distinct from those that influence cooperation, and the empirical modeling process should reflect this. That said, it should still be borne in mind that in most well-resourced surveys contact rates 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.

Third, while this chapter 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. Thus, while a number of potentially important influences were identified in our literature review (e.g., a variety of tracking procedures, the length of the fieldwork period, respondent incentives, multiple modes and interview length), we were unable to say anything directly about their efficacy from the HILDA Survey experience. 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.

Fourth, 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 potentially can be used to assist survey managers 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.

Bogen, K. (1996), 'The Effect of Questionnaire Length on Response Rates – A Review of the Literature', *Proceedings of the American Statistical Association (Survey Research Methods Section)*, pp. 1020-1025.

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.

Buck, N. (2000), 'Using Panel Surveys to Study Migration and Residential Mobility', in D. Rose (ed.), *Researching Social and Economic Change: The Uses of Household Panel Studies*, Routledge, London.

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.

Campanelli, P.A. and O'Muircheartaigh, C. (2002), 'The Importance of Experimental Control in Testing the Impact of Interviewer Continuity on Panel Survey Nonresponse', *Quality and Quantity*, vol. 36, no. 2 (May), pp. 129-144.

Couper, M.P. and Groves, R.M. (1992), 'The Role of the Interviewer in Survey Participation', *Survey Methodology*, vol. 18, no. 2 (December), pp. 263-277.

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.

Foster, K. (1998), *Evaluating Nonresponse on Household Surveys* (GSS Methodology Series no. 8), Government Statistical Service, London.

Gray, R., Campanelli, P., Deepchand, K. 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.

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

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 (August), forthcoming.

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.

Japec, L. and Lundqvist, P. (1999), 'Interviewer Strategies and Attitudes', Paper presented at the International Conference on Survey Non-Response, Portland, Oregon, 28-31 October.

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.

Kalton, G., Lepkowski, J., Montanari, G.E., and Maligalig, D. (1990), 'Characteristics of Second Wave Nonrespondents in a Panel Survey', *Proceedings of the American Statistical Association (Survey Research Methods Section)*, pp. 462-467.

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.

Martin, J. and Beerten, R. (1999), 'The Effect of Interviewer Characteristics on Survey Response Rates', Paper presented at the International Conference on Survey Non-Response, Portland, Oregon, 28-31 October.

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.

Ribisl, K.M., Walton, M.A., Mowbray, C.T., Luke, D.A., Davidson II, W.S. and Bootsmiller, B.J. (1996), 'Minimizing Participant Attrition in Panel Studies Through the Use of Effective Retention and Tracking Strategies: Review and Recommendations', *Evaluation and Program Planning*, vol. 19, no. 1 (February), pp. 1-25.

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.

Serfling, O. (2004), 'The Interaction Between Unit and Item Nonresponse in View of the Reverse Cooperation Continuum: Evidence from the German Socioeconomic Panel (GSOEP)', Paper presented at the 6th International Conference of German Socio Economic Panel Users, Berlin, 24-26 June.

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.

Thomas, D., Frankenberg, E. and Smith, J.P. (2001), 'Lost but Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey', *The Journal of Human Resources*, vol. 36, no. 3 (Summer), pp. 556-592.

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.

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.

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.

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.

Yu, J. and Cooper, H. (1983), 'A Quantitative Review of Research Design Effects on Response Rates to Questionnaires', *Journal of Marketing Research*, vol. 20, no. 1 (February), pp. 36-44.

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.