

Chasing Hard-to-Get Cases in Panel Surveys: Is it Worth it?

Nicole Watson & Mark Wooden

*Melbourne Institute of Applied Economic and Social Research,
University of Melbourne*

Abstract

In many population surveys, fieldwork effort tends to be disproportionately concentrated on a relatively small proportion of hard-to-get cases. This article examines whether this effort is justified within a panel survey setting. It considers three questions: (i) are hard-to-get cases that are interviewed different from other interviewed cases? (ii) do cases that require a lot of effort in one survey wave require a lot of effort in all waves? and (iii) can easy-to-get cases be re-weighted to eliminate biases arising from not interviewing hard-to-get cases? Using data from a large nationally representative household panel survey, we find that hard-to-get cases are distinctly different from easy-to-get cases, suggesting that failure to obtain interviews with them would likely introduce biases into the sample. Further, being hard-to-get is mostly not a persistent state, meaning these high cost cases are not high cost every year. Simulations confirm that removing hard-to-get cases introduces biases, and these biases lead to an understatement of the extent of change experienced by the population. However, we also find that under one of five fieldwork curtailment strategies considered, the bias in population estimates that would arise if the hard-to-get cases were not pursued can be corrected by applying weights. Nevertheless, this conclusion only applies to the curtailment strategy involving the smallest decline in sample size. Biases associated with curtailment strategies involving larger sample size reductions, and hence greatest cost savings, are not so easily corrected.

Keywords: HILDA Survey, sample representativeness, longitudinal surveys, fieldwork curtailment strategies, fieldwork efficiency



© The Author(s) 2019. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

The return to additional survey fieldwork effort, as measured by additional survey respondents, invariably declines with the rate of response. Obtaining very high response rates to population surveys thus typically requires concentrating fieldwork effort, especially towards the end of the fieldwork period, on a relatively small proportion of cases who are hard-to-get. But is the extra effort and cost spent on achieving high response rates justified?

Most previous research on the fieldwork effort involved in following up hard-to-get cases has been undertaken within the context of cross-sectional surveys (e.g., Billiet et al., 2007; Fitzgerald & Fuller, 1982; Hall et al., 2013; Heerwegh et al., 2007; Lin & Schaeffer, 1995; Lynn et al., 2002; Stoop, 2005). In these studies, the analysis usually revolves around a comparison of the characteristics of two groups of respondents: those defined as 'hard-to-get' and the remainder. This, of course, ignores the group that objectively are the hardest to get – the non-responders – reflecting the fact that often little is known about non-respondents in cross-section surveys. Of course, some limited information about non-respondents can be garnered from frame characteristics (Etter & Perneger, 1997), though the range of variables available is usually quite limited, or from non-response follow-up studies that investigate the reasons for non-response (Stoop, 2005, pp. 146-156), though these studies also suffer from non-response issues. In contrast, a large amount is usually known about panel survey respondents who subsequently do not respond in a later wave. As a result, considerable research has been undertaken into the causes and consequences of panel attrition (for example, Behr et al., 2005; Lepkowski & Couper, 2002; Lugtig et al., 2014; Watson & Wooden, 2009). Panel surveys also provide a rich setting for examining the effectiveness of fieldwork effort in following up hard-to-get cases each wave. Previous studies on this issue that have involved panel survey data, however, have mostly focused on identifying hard-to-get cases in just one survey wave (e.g., Haring et al., 2009; Larroque et al., 1999; Ullman & Newcomb, 1998). This is surprising given the ramifications for pursuing

Acknowledgments

The research reported on in this paper was, in part, supported by an Australian Research Council Discovery Grant (#DP1095497). This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the authors and should not be attributed to either DSS or the Melbourne Institute.

Direct correspondence to

Nicole Watson, Melbourne Institute of Applied Economic and Social Research,
Level 5, 111 Barry Street, University of Melbourne, Victoria 3010, Australia
E-mail: n.watson@unimelb.edu.au

or not pursuing cases extend well beyond a single wave. In this article, we examine whether the fieldwork effort devoted to obtaining hard-to-get interviews across six annual survey waves is justified.

Another feature of previous research using either cross-sectional or panel data is the wide variation across studies in how a 'hard-to-get' case is defined. The most common types of definitions employed include any case that: requires a large number of number of visits or calls (Cottler et al., 1987; Hall et al., 2013; Heerwegh et al., 2007; Kennickell, 2000; Lin & Schaeffer, 1995; Lynn et al., 2002; Yan et al., 2004); has refused earlier in the fieldwork period (Billiet et al., 2007; Hall et al., 2013; Fitzgerald & Fuller, 1982; Lin & Schaeffer, 1995; Lynn et al., 2002; Woodruff et al., 2000; Yan et al., 2004); or was interviewed late in the fieldwork period (Etter & Perneger, 1997; Haring et al., 2009; Kennickell, 2000; Lahaut et al., 2003; Larroque et al., 1999; Studer et al., 2013; Ullman & Newcomb, 1998; Yan et al., 2004).

Most studies find that hard-to-get cases are different from easy-to-get cases. These differences extend from socio-demographic variables such as age (Cottler et al., 1987; Hall et al., 2013; Kennickell, 2000; Larroque et al., 1999), sex (Cottler et al., 1987), race (Cottler et al., 1987; Hall et al., 2013), and education (Cottler et al., 1987; Etter & Perneger, 1997; Kennickell, 2000; Larroque et al., 1999), to more substantive variables such as employment (Hall et al., 2013), occupation (Larroque et al., 1999), income (Etter & Perneger, 1997; Kennickell, 2000), wealth (Kennickell, 2000), smoking (Woodruff et al., 2000), substance use (Studer et al., 2013), and physical health (Etter & Perneger, 1997). Obtaining interviews with these hard-to-get cases is expected to reduce biases in survey estimates. How important this reduction in bias is, however, depends on how similar the interviewed hard-to-get cases are to the non-respondents.

For longitudinal surveys, decisions about how much effort to devote to pursuing hard-to-get cases should be influenced, at least in part, by expectations about the likelihood of retaining such sample members in subsequent waves. Being a hard-to-get respondent in one wave, for example, has been found to be predictive of attrition in the next (Haring et al., 2009; Watson & Wooden, 2009). More generally, does the extra effort (and cost) required to interview the hard-to-get cases fall persistently on the same cases from wave to wave? As far as we are aware, this is an issue not considered in any previous research.

Further, and perhaps most importantly, relatively few studies have tested in a simulation setting whether re-weighting the easy-to-get cases can reduce the potential biases introduced from not pursuing interviews with the hard-to-get cases. And those studies that have been conducted (e.g., Billiet et al., 2007; Hall et al., 2013) have used cross-sectional data.

This article uses the Household, Income and Labour Dynamics in Australia (HILDA) Survey, a household-based panel study, to examine three related questions.

1. Are hard-to-get cases that are ultimately interviewed different from other interviewed cases?
2. Do cases that require a lot of effort in one survey wave require a lot of effort in all waves?
3. Can easy-to-get cases be re-weighted to eliminate biases potentially arising from not interviewing hard-to-get cases?

We build on previous research in a number of ways. First, we define hard-to-get cases in five different ways and so can assess how sensitive conclusions are to the choice of measure. Second, we analyze the extent to which being hard-to-get is a state that persists over time. Third, we examine whether the biases that may result if fieldwork is curtailed over an extended period (six annual survey waves) can be eliminated by re-weighting the remaining (i.e., easy-to-get) cases.

Data

The HILDA Survey is a panel that began in 2001 with a three-stage stratified clustered nationally representative sample of households (Watson & Wooden, 2012). There were 19,914 people living in the 7682 responding households in wave 1. These people are followed over time and the sample is extended to include all people living with these original sample members at the time of the subsequent interviews. Interviews are conducted annually with all sample members aged 15 years or older. The vast majority (over 90 percent) of these interviews are undertaken face-to-face, with the remainder by telephone.

The initial responding sample was achieved from a total of 11,693 households identified as in-scope, giving a wave 1 household-level response rate of 66 percent (AAPOR RR1). Annual re-interview rates of individuals are high, rising from 87 percent in wave 2 to over 94 percent by wave 5, and remaining above that level in all subsequent waves.

Within each wave of fieldwork there are three distinct phases, with each successive phase increasingly focusing on sample members that are hardest to locate, contact and interview. The initial fieldwork phase is concentrated in August to September. Non-responding and partially responding households are reviewed and re-issued for follow-up fieldwork in the next phase (October to December). The third fieldwork phase (in January to February) is used to contact households that were difficult to trace or where it is believed further contact attempts may be successful.

Methods

Defining Hard-to-Get Cases

We examine a range of different definitions of ‘hard-to-get’, based on the length of time since commencement of fieldwork, whether an initial refusal was received, and the number of calls made. The most natural delineations of time for the HILDA Survey are the fieldwork stages described earlier, with the survey manager deciding at the end of the first and second fieldwork stages who among the non-respondents should be re-approached. The two time-based definitions of hard-to-get cases used here are:

- Definition A: The individual was interviewed during a follow-up stage of fieldwork.
- Definition B: The interview was completed after the New Year.
- Alternatively, the survey manager may choose not to re-issue to field anyone who initially refused. This suggests a third definition:
- Definition C: The individual initially refused before being interviewed.

Finally, we create binary variables based on the number of calls exceeding some threshold. A call is counted if it was a face-to-face visit or if it was a telephone call that resulted in an appointment, an interview, or other information to finalize the outcome of an individual. From 2009 (wave 9), a change from pen-and-paper interviewing to computer-assisted personal interviewing facilitated the collection of detailed call records. Using these records, we can determine the number of calls made to the household before a particular individual is interviewed. The distribution of these calls, based on data pooled from waves 9 to 14, is shown in Figure 1. For this analysis, we focus on two specific thresholds – 7 or more calls, and 13 or more calls required to obtain an interview. Obviously, a number of different thresholds could have been selected due to the greater granularity of call-based measures compared to those used in the first three definitions. The choice of the particular thresholds used here reflects the operational requirements imposed on the company engaged to undertake the fieldwork for the HILDA Survey. Specifically, an interviewer must make at least six calls to a household in a particular fieldwork period before they can return the household to the office with an inconclusive outcome (such as a non-contact), and then up to a further 6 calls after making contact to interview sample members. This provides two further definitions.

- Definition D: 7 or more calls were made to the household by the time the individual was interviewed.
- Definition E: 13 or more calls were made to the household by the time the individual was interviewed.

Figure 1 Distribution of calls made before interviewing sample member, waves 9 to 14 combined

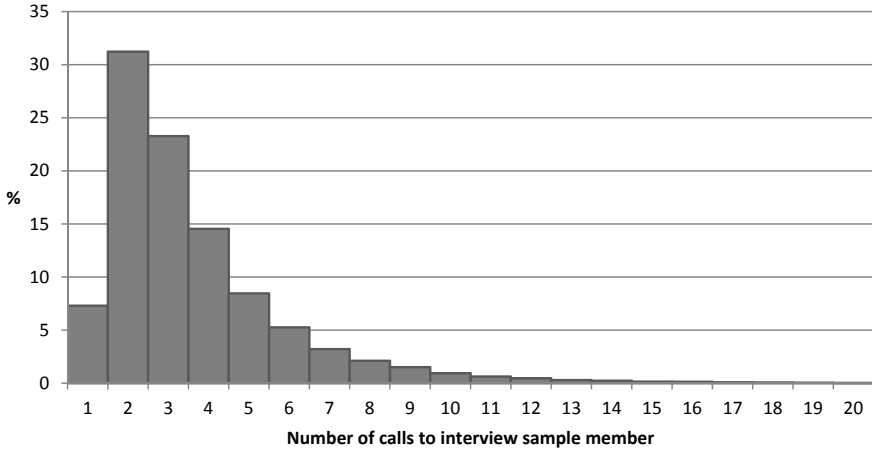


Figure 2 Percentage of Interviewed Cases That Were 'Hard-to-Get', by Wave

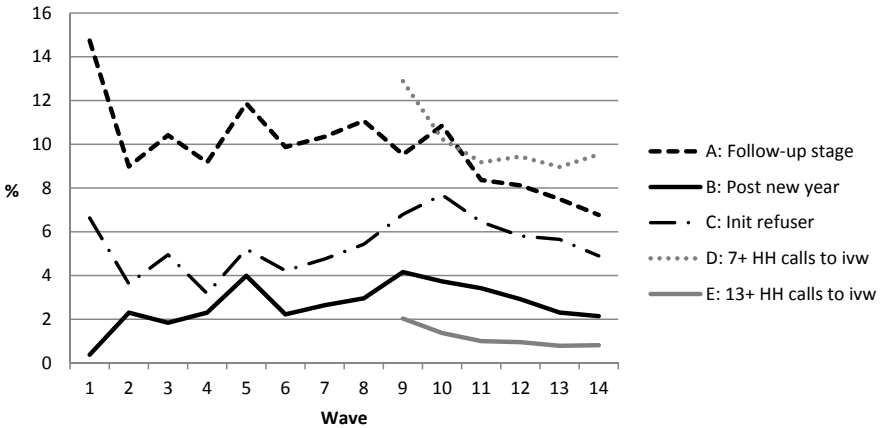


Figure 2 shows the proportion of interviews that were hard-to-get according to each of these five definitions, and how this has varied over time. Approximately 10 percent of interviews required a follow-up period of fieldwork to achieve the interview (definition A); though there is a noticeable decline in this proportion in later waves (waves 11 to 14). Only about half of these follow-up cases were due to an initial refusal (definition C) in the early waves, but this rises to around 70 percent in waves 9 to 14. This shift coincides with a change in fieldwork provider, which occurred

after wave 8, suggesting either a change in re-issuing practice or a greater ability on the part of the new provider to convert initial refusals to interviews. The proportion of cases interviewed after the New Year (definition B) each wave is relatively small (2 to 4 percent) and varies somewhat wave to wave. The proportion of cases defined as hard-to-get when using call counts varies substantially depending on the particular call threshold applied. Using a cut-off of 7 or more calls to define a hard-to-get case (definition D) results in 9 to 13 percent of the interviewed cases being classified as hard-to-get. When the higher cut-off of 13 or more calls is used (definition E), the proportion of interviewed cases defined as hard-to-get declines to just 1 to 2 percent.

Assessing the Impact of Pursuing Hard-to-Get Cases

Multinomial logistic models of the three interview outcomes at wave t – easy-to-get interview, hard-to-get interview, and not interviewed – are used to assess whether the hard-to-get cases are appreciably different from the easy-to-get cases (research question 1). We include a range of personal and household characteristics, all measured at wave $t-1$, that are often found to be associated with non-response (see Watson & Wooden, 2009). These include: age (in 10-year bands), sex, marital status (6 categories), number of adults living in the household, number of children (aged less than 15) living in the household, education level (6 categories), country / region of birth (3 categories), whether the sample member has a restrictive long-term health condition, area of residence (9 categories), employment status (6 categories), real equivalized (i.e., household size adjusted) gross annual (financial year) household income (with missing values imputed; see Hayes & Watson, 2009), whether an owner-occupier of a home, whether the household moved between waves $t-1$ and t , and a set of wave indicators.

Missing data on covariates resulted in the loss of just 520 observations (0.7 percent) from the models employing the first three definitions of hard-to-get, leaving a total of 77,315 person-wave observations. For the last two hard-to-get definitions, a further 54 person-wave observations were dropped due to missing call record information. To allow repeated observations on the same individuals, the multinomial logistic models are fitted as two-level hierarchical models where level 1 is the wave observation and level 2 is the individual. Two random effects, which were allowed to be correlated, were assumed for the different interview outcomes.

To assess whether individuals are hard-to-get repeatedly over time simply because of their particular socio-demographic characteristics (research question 2), we rerun the above set of multinomial logit models and include an indicator variable for whether the individual was hard-to-get in wave $t-1$.

Finally, we test whether excluding the hard-to-get cases materially affects key estimates from the study (research question 3). We examine whether the different

sample curtailment strategies are associated with significant differences in selected personal and household characteristics, and assess whether these differences can be eliminated through the application of survey weights constructed for the reduced sample under each of the five curtailment strategies that only contains the easy-to-get cases. We then similarly test for differences in responses to 15 selected estimates of change over time. The weights used relate to a “balanced” panel of respondents from wave 1 to 14 where the hard-to-get cases have been dropped from wave 9 onwards. The balanced panel weights were constructed by adjusting the wave 1 cross-sectional weights for attrition from wave 1 to wave 14 (by multiplying by the inverse of the response propensity that is modelled on a range of wave 1 socio-economic characteristics and some post-wave 1 mobility information where available). The weights are then calibrated to a set of external wave 1 totals. This follows the same methodology employed to construct the regular HILDA Survey weights (Watson, 2012). Standard errors of the difference between the full sample and the truncated sample (i.e., after excluding the hard-to-get cases) for each definition of hard-to-get, were calculated using jackknife estimation with 45 replicates.

To ensure that all definitions are examined across the same timeframe, all analyses that follow are restricted to the outcomes observed in waves 9 to 14.

Results

Are the Hard-to-Get Cases Different From Other Cases?

The coefficients from the estimation of multinomial logit regression models with random effects predicting interview outcomes are shown in Table 1. Separate estimates are provided for each of the five definitions of hard-to-get.

Regardless of the definition used, the hard-to-get group is distinctly different from both the easy-to-get and the non-respondents. Compared to easy-to-get respondents, hard-to-get respondents tend to be younger, single, live in a household with three or more adults, less educated, born in a non-English-speaking country, have higher incomes, not have a restrictive long-term health condition and live in households that have moved. The likelihood of being a hard-to-get case also increases with household income (though at a declining rate) and hours worked. Non-respondents when compared to easy-to-get cases, tend to be relatively young, live in larger households, have not completed high school and likely to live in households that have moved since the previous interview.

Table 1 Multinomial Logit of Interview Outcome, with Random Effects: waves 9 to 14

Characteristics from interview in wave <i>t-1</i>	A: Follow-up stage		B: Post New Year		C: Initial refuser		D: 7+ calls to interview		E: 13+ calls to interview	
	Hard	Non-resp	Hard	Non-resp	Hard	Non-resp	Hard	Non-resp	Hard	Non-resp
<i>Age group (base=15-24)</i>										
25-34	-0.056	-0.087	-0.044	-0.096	-0.034	-0.084	0.035	-0.085	0.017	-0.101
35-44	-0.052	-0.003	-0.183	-0.030	-0.050	-0.007	-0.027	-0.017	0.014	-0.009
45-54	-0.117	-0.312**	-0.250*	-0.333**	-0.195*	-0.331**	-0.094	-0.340**	0.094	-0.334**
55-64	-0.279**	-0.809**	-0.487**	-0.830**	-0.494**	-0.841**	-0.430**	-0.899**	-0.373	-0.840**
65+	-0.303**	-0.254#	-0.304#	-0.220	-0.563**	-0.288#	-0.734**	-0.333*	-0.382	-0.243#
Female	0.019	0.025	0.071	0.035	0.071	0.028	-0.043	0.025	-0.070	0.030
<i>Marital status (base=married)</i>										
De facto	0.091	0.142	0.185#	0.145	-0.013	0.130	0.209**	0.152#	0.316*	0.110
Separated	0.092	0.091	0.381*	0.113	0.042	0.062	0.259*	0.152	0.391	0.149
Divorced	-0.142	0.003	0.109	0.040	-0.227*	-0.021	0.168#	0.070	0.511*	0.047
Widowed	-0.160	0.410*	0.059	0.401*	-0.118	0.410*	0.223#	0.522**	-0.210	0.447**
Never married & not living with partner	0.247**	0.347**	0.404**	0.327**	0.042	0.300**	0.361**	0.369**	0.721**	0.298**
<i>Number of adults (base=1 adult)</i>										
2 adults	0.099	0.118	0.106	0.114	0.037	0.111	0.247**	0.161	0.188	0.101
3 adults	0.227**	0.174#	0.285*	0.173#	0.207*	0.160	0.505**	0.242*	0.622**	0.165
4 or more adults	0.150#	0.217#	0.325**	0.243*	0.146#	0.211#	0.716**	0.337**	0.724**	0.213#
<i>Number of children (base=0 children)</i>										
1 child	-0.036	-0.093	-0.060	-0.097	0.023	-0.090	0.024	-0.108	0.203	-0.095
2 children	-0.022	-0.178#	-0.092	-0.197*	0.096	-0.155	-0.023	-0.215*	0.213	-0.193*
3 or more children	-0.016	-0.355**	-0.009	-0.315*	0.041	-0.339*	0.037	-0.382**	0.497**	-0.331*

Table 1 continued

Characteristics from interview in wave t-1	A: Follow-up stage		B: Post New Year		C: Initial refuser		D: 7+ calls to interview		E: 13+ calls to interview	
	Hard	Non- resp	Hard	Non- resp	Hard	Non- resp	Hard	Non- resp	Hard	Non- resp
<i>Highest level of education (base=Year 11 or below)</i>										
Year 12	0.103#	-0.134	-0.036	-0.176*	0.025	-0.144	0.078	-0.153#	-0.124	-0.173*
Cert III or IV	0.083	-0.072	0.094	-0.085	0.039	-0.077	0.120*	-0.062	-0.136	-0.096
Diploma	0.025	-0.268*	-0.163	-0.311*	-0.051	-0.291*	0.048	-0.260*	0.079	-0.275*
Graduate	-0.154*	-0.449**	-0.290*	-0.491**	-0.375**	-0.492**	-0.127*	-0.435**	-0.464**	-0.457**
Post graduate	-0.182*	-0.625**	-0.226#	-0.640**	-0.365**	-0.659**	-0.220**	-0.629**	-0.411*	-0.606**
<i>Country of birth (base=Australia)</i>										
Main English-speaking country	-0.034	-0.130	-0.130	-0.152	-0.143	-0.157	-0.017	-0.140	-0.034	-0.143
Not main English-speaking country	0.416**	0.377**	0.432**	0.350**	0.271**	0.338**	0.431**	0.382**	0.537**	0.305**
Long term health condition	-0.114*	-0.010	-0.269**	-0.009	-0.139**	-0.023	-0.127**	-0.015	-0.215#	-0.011
<i>Area of residence (base=Major city: Sydney)</i>										
Major city: Melbourne	-0.164*	-0.187#	-0.257*	-0.170	-0.168*	-0.185#	-0.131*	-0.197#	-0.305*	-0.169#
Major city: Brisbane	-0.086	-0.161	-0.204	-0.160	-0.185#	-0.187	-0.667**	-0.277*	-1.105**	-0.179
Major city: Adelaide	-0.062	-0.082	-0.197	-0.095	-0.186#	-0.099	-0.521**	-0.175	-1.563**	-0.121
Major city: Perth	-0.108	0.072	-0.215	0.071	-0.271*	0.022	-0.296**	0.016	-0.527**	0.058
Major city: other	-0.423**	-0.478**	-0.416**	-0.470**	-0.545**	-0.487**	-0.855**	-0.580**	-1.134**	-0.459**
Inner regional	-0.048	-0.109	-0.326**	-0.137	-0.165*	-0.131	-0.680**	-0.250**	-1.041**	-0.151
Outer regional	0.182*	0.301**	0.202#	0.264*	0.026	0.270*	-0.722**	0.131	-0.727**	0.262*
Remote	0.629**	0.686**	1.004**	0.699**	0.608**	0.694**	-0.205#	0.498**	-0.308	0.536**
<i>Employment status (base=not in labour force)</i>										
Employed, working <=34 hrs	-0.057	-0.127	-0.100	-0.114	-0.006	-0.133	0.100*	-0.125	0.121	-0.128
Employed, working 35-44 hrs	0.087	0.264**	0.143	0.270**	0.208**	0.252**	0.134*	0.273**	0.365*	0.269**

Table 1 continued

Characteristics from interview in wave $t-1$	A: Follow-up stage		B: Post New Year		C: Initial refuser		D: 7+ calls to interview		E: 13+ calls to interview	
	Hard	Non- resp	Hard	Non- resp	Hard	Non- resp	Hard	Non- resp	Hard	Non- resp
Employed, working 45-54 hrs	0.204**	0.321**	0.241*	0.313**	0.308**	0.319**	0.255**	0.342**	0.424*	0.331**
Employed, working 55+ hrs	0.480**	0.609**	0.381**	0.564**	0.542**	0.588**	0.441**	0.587**	0.568**	0.570**
Unemployed	0.018	0.238#	0.066	0.246*	0.143	0.235#	0.152#	0.257*	0.186	0.261*
Equivalentized HH income (10^5)	0.435**	0.151	0.402**	0.181	0.382**	0.168	0.443**	0.157	0.317	0.125
Equivalentized HH income squared (10^{10})	-0.049**	-0.003	-0.046	-0.008	-0.038#	-0.005	-0.071**	-0.004	-0.091	-0.002
Owner occupier	0.015	-0.103	-0.014	-0.097	0.033	-0.094	-0.148**	-0.132*	-0.267*	-0.121#
Household moved (btw $t-1$ and t)	0.533**	0.419**	0.839**	0.418**	0.605**	0.426**	0.502**	0.398**	0.575**	0.322**
<i>Next wave (base=wave 9)</i>										
Wave 10	0.303**	0.180*	0.065	0.113	0.286**	0.168*	-0.218**	0.058	-0.420**	0.096
Wave 11	0.020	0.170*	0.067	0.168*	0.155*	0.201*	-0.307**	0.104	-0.485**	0.159*
Wave 12	0.053	0.339**	-0.061	0.319**	0.111#	0.358**	-0.198**	0.289**	-0.518**	0.320**
Wave 13	-0.019	0.373**	-0.188*	0.362**	0.124*	0.410**	-0.222**	0.328**	-0.582**	0.367**
Wave 14	-0.115*	0.344**	-0.302**	0.339**	-0.022	0.377**	-0.143**	0.331**	-0.585**	0.354**
Constant	-3.489**	-4.703**	-4.939**	-4.732**	-3.766**	-4.719**	-2.734**	-4.526**	-5.675**	-4.594**
Var (random effects)	2.155**		3.202**		2.411**		1.547**		3.403**	
Cov (random effects for hard and non- resp)	2.728**		3.496**		3.003**		2.181**		3.041**	
Log-likelihood	-29,694		-18,890		-25,832		-31,852		-14,753	
N(person-year observations)	77,315		77,315		77,315		77,261		77,261	

Note: # $p < 0.10$; * $p < 0.05$; ** $p < 0.01$.

How Persistent are Hard-to-Get Cases?

Do cases that require a lot of work in one wave require a lot of work in all waves? Figure 3 shows the proportion of hard-to-get cases at one wave that are interviewed in subsequent waves but were hard-to-get. It shows that the level of reoccurrence is relatively low, with 9 to 24 percent of hard-to-get cases in one wave classified as hard-to-get in the next wave, and the rate of persistence in being classified as hard-to-get declines over time, with 5 to 17 percent classified as hard-to-get four waves later. The large majority (75 to 90 percent, depending on the hard-to-get definition used) of hard-to-get cases are classified as easy-to-get in the next wave.

Does the relatively small amount of persistence observed in the hard-to-get cases remain after controlling for respondent characteristics? To test this, we modify the model presented in Table 1 (which predicts whether a case will be easy-to-get, hard-to-get or a non-respondent) and include an indicator of whether the individual was hard-to-get in the prior wave (when the other characteristics included in the model were measured). Table 2 reports the estimated coefficients and mean predicted probabilities for this variable. We find a strong negative association between being hard-to-get in one wave and being easy-to-get in the next. The predicted probability, holding all else constant, of being an easy-to-get case (using definition A; i.e., whether they require follow-up work or not) for those who were easy-to-get in the previous wave is 89.1 percent. This compares with 80.9 percent of those who were previously hard-to-get. The differences in the predicted probabilities are similar for the other four definitions; i.e., 7.2 percentage points for definition B, 8.4 percentage points for definition C, 8.3 percentage points for definition D, and 6.5 percentage points for definition E.

In summary, the large majority of hard-to-get cases (over 80 percent under all definitions) are easy-to-get come the next survey wave. This is not to say, however, that there is no state persistence; a hard-to-get case is still much more likely (around twice as likely) to be hard-to-get next wave than an otherwise comparable case classified as easy-to-get.

Can the Differences in Hard-to-Get Cases be Corrected by Weighting?

Finally, we examine whether the differences between estimates obtained using only the easy-to-get cases and those obtained using both the easy-to-get and hard-to-get cases can be eliminated by applying weights generated specifically for each truncated sample. We first consider the impact fieldwork effort has on the personal and household characteristics included in Table 1 (with the exception of residential mobility, which is included later in Table 4). Table 3 reports the unweighted and weighted estimates for these variables, measured as of 2014 (i.e., in wave 14).

Figure 3 Average Percentage of Hard-to-Get Interviewed Cases in Future Waves Conditional on Being Hard-to-Get in Wave *t*

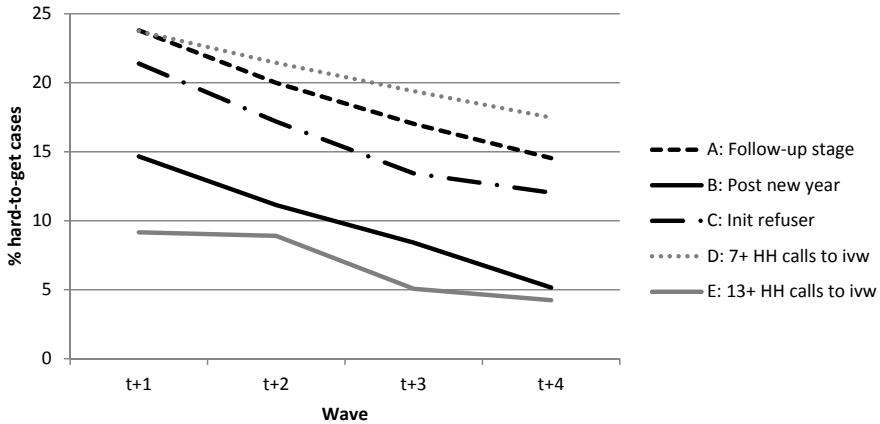


Table 2 Coefficient and Predicted Probabilities for Hard-to-Get in Prior Wave in Multinomial Logit Model of Interview Outcome with Random Effects

	Easy-to-get at <i>t</i>		Non-respondent at <i>t</i>			
	Coeff	Mean predicted probability	Coeff	Mean predicted probability		
		Hard at <i>t-1</i>		Easy at <i>t-1</i>	Hard at <i>t-1</i>	Easy at <i>t-1</i>
Hard-to-get in wave <i>t-1</i>						
Definition A: Follow-up stage	-0.800**	80.9	89.1	0.059	6.9	4.0
Definition B: Post New Year	-1.205**	86.2	93.4	-0.238*	7.9	4.2
Definition C: Initial refuser	-0.936**	82.4	90.8	0.003	7.3	4.1
Definition D: 7+ calls to interview	-0.714**	80.2	88.4	0.181*	6.9	3.8
Definition E: 13+ calls to interview	-1.315**	88.5	95.0	-0.188	9.2	4.2

Note: Models include controls for all the covariates shown in Table 1. # $p < 0.10$; * $p < 0.05$; ** $p < 0.01$.

The weights used to calculate these estimates are for the sample of individuals interviewed in both wave 1 and wave 14, and interviewed in every intervening wave. We describe this as the full balanced panel, though strictly speaking the sample is not completely balanced – respondents that moved abroad and subsequently returned to Australia were also retained. For the curtailed samples, cases that were hard-to-get in any wave between 9 and 14 are dropped from the balanced

panel. The full balanced panel from wave 1 to 14 includes 6707 individuals. This declines to 6572 when cases requiring 13 or more calls (definition E) are excluded (a 2 percent reduction), or 6245 cases if the post New Year fieldwork (definition B) is dropped (a 7 percent reduction). Greater reductions in the sample occur when the broader definitions of hard-to-get are used. The balanced panel contains 5661 cases if all initial refusers (definition C) are dropped (a 16 percent reduction), 5225 cases if all follow-up fieldwork (definition A) is abandoned (a 22 percent reduction), or 5046 cases if cases requiring 7 or more calls (definition D) are dropped (a 25 percent reduction).

The unweighted and weighted estimates for the personal and household variables for the full balanced panel are presented in the first two columns of Table 3. The weighted estimates are constructed by weighting the responses provided by both easy- and hard-to-get cases by the wave 1 to 14 balanced panel weight available in the HILDA Survey dataset. The unweighted estimates are similarly restricted to cases that have a positive balanced panel weight to aid comparison of the weighted and unweighted estimates. The following columns in the table provide (for each definition of hard-to-get): i) the difference between the unweighted estimate for the full balanced panel and the unweighted estimate obtained after dropping the relevant hard-to-get cases from waves 9 to 14; and ii) the difference between the weighted estimate for the full balanced panel and the estimate obtained by applying the recalculated balanced panel weight after dropping the relevant hard-to-get cases from waves 9 to 14. The estimates are marked to indicate the p-value for the two-sided z-test for whether this difference is statistically different from zero (# $p < 0.10$; * $p < 0.05$; ** $p < 0.01$).

We find that the definition of hard-to-get that shows the largest number of differences in the unweighted estimates is the curtailment strategy that drops the most cases (definition D which drops cases requiring 7 or more calls) and is least able to be corrected by the weights. The curtailment strategy affecting the personal and household estimates the least is definition E, which drops people requiring 13 or more calls. Further, these estimates are most amenable to correction by the application of weights (while one estimate is not corrected, this is expected by chance alone). Nevertheless, this strategy involves a very small decline in the number of cases followed, and hence the potential for costs savings is commensurately small. Arguably, our results suggest that the best curtailment strategy in terms of maximising sample reduction (and thus saving fieldwork effort) while minimising the effect on estimates is strategy A (not pursuing persons into the follow-up fieldwork phase). However, application of weights is still unable to correct for differences observed on at least three variables (age, country of birth and income).

Next we focus on a subset of variables that relate to change over time, some of which have been much analyzed by users of the HILDA Survey data. The first five estimates relate to changes in the family: the proportion who got married in the

last five years; the proportion who separated from a marriage or were widowed in the last 5 years; the proportion that began a de facto relationship in the last year; the proportion who had a new birth in the last year; and the proportion who had a new birth in the last 5 years. There is one measure relating to income: the increase in the 5-year average income between the start and end of the panel (i.e., 2001-05 versus 2010-14). There are four estimates related to employment: whether a new job was started in the last year; whether retired in the last year; for those self-employed in 2009, the proportion that switched to being an employee by 2014; and for those who were employees in 2009, the proportion that transitioned to self-employment by 2014. In terms of health, we include the proportion of people who experienced the onset of a long-term health condition between 2009 and 2014. The final group of four estimates relate to housing: the proportion who moved house in the last year; the proportion who moved house in the last five years; the proportion who transitioned from living in a home that was not owned (i.e., was rented or provided rent-free) to one that was owned between 2009 and 2014; and the proportion who transitioned from living in a home that was owned to one that was not between 2009 and 2014.

The population estimates for the subset of variables are presented in the first column of Table 4. These estimates are constructed by weighting the responses provided by easy- and hard-to-get cases (the full balanced panel) by the wave 1 to 14 balanced panel weight available in the HILDA Survey dataset. Subsequent columns in the table provide (for each definition of hard-to-get): i) the difference between this first population estimate and the one obtained by applying the recalculated balanced panel weight after dropping the relevant hard-to-get cases from waves 9 to 14; and ii) the p-value for the two-sided z-test for whether this difference is statistically equivalent to zero.

We find that the impact of dropping the hard-to-get cases on the selected 15 population estimates is minimal when using the definition involving the loss of fewest cases (definition E); the estimated differences are both very small and a long way from statistically significant. Use of any of the other four definitions, which all involve larger sample losses, results in larger changes in the population estimates. Interestingly, curtailment strategy A, which above we suggested was the best strategy in terms of maximising the reduction in fieldwork effort while having the least impact on the estimates, now appears to be the one that results in the most harm to these estimates of change over time. In general, there is evidence of biases in favour of stability rather than change. For example, the increase in 5-year average income is understated by 6 percent under definition B, 7 percent under definition D, and 10 percent under both definitions A and C. Similarly, under all four of these curtailment strategies the easy-to-get cases are significantly less likely to move house and to separate from marriages.

Table 3 continued

Characteristic	Difference in estimate when hard-to-get cases excluded											
	Estimate based on full balanced panel		A: Follow-up stage		B: Post New Year		C: Initial refuser		D: 7+ calls to interview		E: 13+ calls to interview	
	Unwtd	Wtd	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b
<i>Number of children (base=0 children)</i>												
1 child	9.8	10.7	-0.2	-0.2	-0.1	-0.1	-0.4*	-0.2	-0.5*	0.0	-0.1	0.0
2 children	10.0	11.3	0.2	0.4	0.1	0.5**	-0.2	0.2	-0.8**	-0.2	0.1	0.1*
3 or more children	4.5	5.2	0.0	0.1	-0.1	-0.1	-0.2	0.0	0.2	0.5**	-0.1	-0.1
<i>Highest level of education (base=Year 11 or below)</i>												
Year 12	10.9	12.4	-0.1	-0.1	-0.1	-0.1	-0.3	-0.1	-0.6*	-0.5#	0.0	0.0
Cert III or IV	22.4	22.9	-0.1	0.0	0.0	0.0	-0.3	-0.2	0.0	0.2	0.0	0.0
Diploma	10.6	10.4	-0.2	-0.1	0.1	0.2	-0.1	0.1	-0.3	0.1	0.0	0.1
Graduate	13.9	14.3	0.0	-0.2	0.1	0.3	0.1	0.0	-0.8*	-0.4	0.0	0.0
Post graduate	12.7	11.7	-0.2	-0.1	-0.3*	-0.4*	0.2	0.1	-0.6*	0.1	0.0	0.0
<i>Country of birth (base=Australia)</i>												
Main English-speaking country	11.1	10.3	-0.1	0.0	0.0	-0.1	0.1	-0.1	0.1	0.0	0.0	0.0
Not main English-speaking country	11.2	14.8	-0.5#	-1.0**	-0.2#	0.0	-0.2	-0.6*	-1.0**	-1.4**	0.0	-0.1
Long term health condition	36.4	35.3	1.1**	0.7#	0.4*	0.2	0.8**	0.3	2.5**	1.1**	0.3**	0.1
<i>Area of residence (base=Major city: Sydney)</i>												
Major city: Melbourne	16.1	17.1	1.0**	0.3#	0.2	0.1	0.5*	0.3	-1.6**	-0.2	-0.2	0.0
Major city: Brisbane	8.0	8.3	-0.2	0.2	-0.1	0.0	-0.2	0.2	0.1	0.0	0.0	0.0
Major city: Adelaide	5.9	5.2	-0.3	0.0	0.0	0.0	-0.1	0.0	0.1	0.0	0.1*	0.0
Major city: Perth	6.8	6.8	-0.2	-0.1	-0.1	-0.2	0.0	-0.2	-0.6**	-0.5*	0.0	0.0

Table 3 continued

Characteristic	Difference in estimate when hard-to-get cases excluded											
	Estimate based on full balanced panel		A: Follow-up stage		B: Post New Year		C: Initial refuser		D: 7+ calls to interview		E: 13+ calls to interview	
	Unwtd	Wtd	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b	Unwtd ^a	Wtd ^b
Major city: other	8.7	9.1	0.5*	0.3	0.0	0.1	0.4*	0.2	0.9**	0.4#	0.0	0.1
Inner regional	27.3	23.8	-0.2	-0.2	0.2	0.1	-0.1	-0.2	1.8**	0.6#	0.2**	0.1
Outer regional	11.3	10.0	-0.6#	-0.2	0.0	0.1	-0.3	0.0	1.1**	0.5*	0.1	0.0
Remote	1.4	1.4	-0.2#	-0.1	-0.1	0.0	-0.1	-0.1	0.0	0.0	0.0	0.0
<i>Employment status (base=not in labour force)</i>												
Employed, working <=34 hrs	19.1	18.2	0.4*	0.5#	0.0	0.1	0.1	0.4#	-0.3	0.5	0.0	0.0
Employed, working 35-44 hrs	23.5	25.4	-0.5#	-0.4	-0.2	-0.1	-0.7**	-0.6	-1.8**	-1.0*	0.0	0.0
Employed, working 45-54 hrs	10.2	11.1	-0.6**	-0.4	-0.1	0.0	-0.4**	-0.2	-0.8**	0.2	-0.1#	0.0
Employed, working 55+ hrs	5.6	5.8	-0.4#	-0.2	-0.1	-0.1	-0.2#	-0.1	-0.7**	-0.5#	-0.1	-0.1
Unemployed	1.7	1.8	-0.1	-0.1	0.0	0.1	-0.1#	0.0	-0.3**	-0.2#	0.0	0.0
Equalized HH income (/10 ⁵)	44,382	45,057	-1601**	-1627**	-791#	-917*	-1183**	-1264*	-2017**	-1337*	-280*	-153
Owner occupier	77.1	74.1	0.2	0.7#	0.4*	0.1	0.3	0.2	1.3**	1.4**	0.1	-0.1

Notes: a – The difference between the unweighted estimate using easy- and hard-to-get cases and the unweighted estimate that would be obtained if no interviews had been completed with the hard-to-get cases. b – The difference between the weighted (population) estimate using easy- and hard-to-get cases and the weighted (population) estimate that would be obtained if no interviews had been completed with the hard-to-get cases. # p<0.10; * p<0.05; ** p<0.01.

Table 4 Population Estimates of Selected Wave 14 (2014) Characteristics Specific to Change and the Effect of Excluding Hard-to-Get Cases

Population characteristic	Pop. estimate based on full balanced panel	Difference in population estimate when hard-to-get cases excluded									
		A: Follow-up stage		B: Post New Year		C: Initial refuser		D: 7+ calls to interview		E: 13+ calls to interview	
		Diff. ^a	p-value	Diff. ^a	p-value	Diff. ^a	p-value	Diff. ^a	p-value	Diff. ^a	p-value
<i>Family</i>											
Got married in last 5 years (%)	6.4	-0.3	0.246	-0.2	0.367	-0.2	0.351	-0.4	0.199	-0.1	0.367
Separated / widowed from marriage in last 5 years (%)	5.5	-0.8	0.000	-0.4	0.011	-0.8	0.000	-0.7	0.003	-0.1	0.427
Began de facto relationship in last year (%)	1.3	-0.2	0.133	-0.1	0.178	-0.1	0.575	-0.1	0.327	0.0	0.778
Had child in last year (%)	2.3	0.0	0.797	-0.1	0.451	0.0	0.876	0.1	0.643	0.0	0.450
Had child in last 5 years (%)	11.8	0.0	0.934	-0.1	0.397	-0.2	0.290	0.1	0.873	-0.1	0.378
<i>Income</i>											
Increase in 5-year average financial year income 2001-05 to 2010-14 (\$) ^b	8123	-792	0.013	-510	0.021	-790	0.017	-587	0.048	-226	0.177
<i>Employment</i>											
Started new job in last year (%) ^c	12.4	-0.4	0.306	-0.3	0.231	-0.2	0.514	-0.4	0.516	-0.1	0.348
Retired in last year (%)	1.3	-0.2	0.048	0.0	0.370	-0.1	0.233	-0.2	0.102	0.0	0.324
Transition from self-employed in 2009 to employee in 2014 (%) ^d	30.7	0.8	0.656	-0.4	0.568	0.9	0.441	-1.1	0.548	-0.6	0.298
Transition from employee in 2009 to self-employed in 2014 (%) ^e	3.9	-0.1	0.579	-0.2	0.200	-0.3	0.244	-0.1	0.725	-0.1	0.315

Table 4 continued

Population characteristic	Pop. estimate based on full balanced panel	Difference in population estimate when hard-to-get cases excluded											
		A: Follow-up stage		B: Post New Year		C: Initial refuser		D: 7+ calls to interview		E: 13+ calls to interview			
		Diff. ^a	p-value	Diff. ^a	p-value	Diff. ^a	p-value	Diff. ^a	p-value	Diff. ^a	p-value		
<i>Health</i>													
Onset of long term health condition b/w 2009 and 2014 (%)	17.7	0.4	0.417	0.0	0.892	0.2	0.525	0.6	0.202	0.2	0.143		
<i>Housing</i>													
Moved b/w 2009 and 2014 (%)	39.7	-1.7	0.000	-0.5	0.026	-1.3	0.000	-1.7	0.001	-0.2	0.091		
Moved b/w 2013 and 2014 (%)	12.3	-0.7	0.041	-0.2	0.173	-0.1	0.734	-0.5	0.145	0.0	0.825		
Transition from not living in owned home in 2009 to living in owned home in 2014 (%) ^f	28.7	1.7	0.028	0.5	0.251	0.9	0.189	2.6	0.005	-0.1	0.809		
Transition from living in owned home in 2009 to not living in owned home in 2014 (%) ^g	8.3	-0.4	0.293	-0.3	0.118	-0.4	0.257	-0.5	0.189	0.1	0.187		

Notes: a – The difference between the population estimate using easy- and hard-to-get cases and the population estimate that would be obtained if no interviews had been completed with the hard-to-get cases. b – Gross (i.e., before tax) annual (measured over a financial year; i.e., 1 July to 30 June) real (measured in 2001 prices) personal income. c – Employed persons only. d – Self-employed persons only. e – Employees only. f – Non-home owners only. g – Home owners only.

Discussion

This paper has examined the effect of pursuing hard-to-get cases in a panel setting. We used data from waves 9 to 14 of the HILDA Survey and applied five different definitions of being hard-to-get (based on time in field, whether a refusal was initially obtained, or the number of calls required to achieve the interview). Using different definitions provides a test of the sensitivity of the findings to different possibilities of curtailing the fieldwork effort. Our results suggest three key findings.

First, survey respondents who are hard-to-get, regardless of the definition used, are distinctly different from those who are easy-to-get. This means that in pursuing the hard-to-get cases, we are not simply bringing into the sample more of the same and thus replicating the biases that exist in the sub-sample of early-to-get cases.

Second, being hard-to-get is mostly not a persistent state. The vast majority of sample members who are hard-to-get in one wave (80 to 90 percent) will be easy-to-get in the next wave. This suggests that difficulty obtaining interviews with a case in one wave is largely situational and such cases will not routinely be difficult to interview over a longer time span.

Third, we have uncovered evidence that it is possible to curtail some elements of fieldwork – notably capping the number of call attempts to no more than 12 – without noticeably affecting population estimates. That is, any biases that might arise can be largely rectified through the use of appropriate sample weights. This conclusion, as might be expected, applies to the definition of hard-to-get involving the smallest decline in sample size. When we consider other more significant curtailment strategies involving greater sample losses, and hence greater cost savings, however, the effects on population estimates are more serious. The sample that is lost through these more expansive curtailment strategies tends to be those who have experienced greater change in their lives. Even with the curtailment strategy involving the second smallest decline in the sample size (via dropping the post New Year fieldwork) where the wave-specific estimates can be corrected by weighting, the estimates relating to change over time could not be. Of course, it is not just the number of cases that are dropped that is important, but also what type of cases are dropped. A limitation inherent in examining different definitions of ‘hard-to-get’ is that they will result in different numbers of cases being dropped. However, for the two curtailment strategies that did involve a similar decrease in the number of cases (A and D), we find evidence of different impacts. That is, the curtailment strategy that restricts the number of calls to 6 resulted in substantially more differences in the unweighted and weighted wave-specific estimates but fewer differences in estimates of change over time than the strategy that involved no follow-up fieldwork.

So has devoting effort to chasing hard-to-get cases in the HILDA Survey been worth it? Our answer is a qualified yes. Hard-to-get cases have characteristics that

are, on average, quite different from other respondents, suggesting that failure to obtain interviews with them would likely introduce biases into the sample. At the same time, most of these more costly cases are not high cost every year. One qualification is that our simulations suggest that the number of calls to a household could be limited to 12 without significant losses to the sample integrity. This strategy, however, results in a relatively modest reduction in overall sample size (just 2 percent). That said, it is also important to bear in mind that we have only examined the effect of curtailment on a limited set of population estimates; it may be that even very modest curtailment strategies could have significant effects on other estimates.

We expect that these findings are relevant to other longitudinal surveys that employ face-to-face, telephone and possibly even online methodologies. While the definitions of hard-to-get versus easy-to-get may need to change (especially for online surveys), this study has shown that the findings are similar across definitions. Not pursuing the hard-to-get cases could cause biases in estimates that are not able to be eliminated through weighting, and these biases tend to favour stability rather than change over time. We encourage researchers to replicate this analysis with other longitudinal studies. We also encourage use of other definitions of 'hard-to-get', such as the number of calls to first contact and the use of reminder emails or texts (in online surveys).

Finally, we note that we have restricted our attention to potential fieldwork modifications that standardize fieldwork protocols across all cases. An alternative, known as responsive design, is to focus the extended effort only on those cases thought most likely to reduce the bias in key estimates or improve the efficiency of the estimates (Groves & Heeringa 2006; Schouten, Peytchev, & Wagner 2017; Tourangeau et al. 2016). This, however, is far from straightforward in longitudinal surveys or in surveys that cover a wide number of subject domains. Another challenge for all curtailment strategies is that it would require survey funders to shift their focus from response rates as a measure of survey quality to other quality measures (Kreuter 2013).

References

- Behr, A., Bellgardt, E., & Rendtel, U. (2005). Extent and determinants of panel attrition in the European Community Household Panel. *European Sociological Review*, 21(5), 489-512.
- Billiet, J., Philippens, M., Fitzgerald, R., & Stoop, I. (2007). Estimation of nonresponse bias in the European Social Survey: Using information from reluctant respondents. *Journal of Official Statistics*, 23(2), 135-162.
- Cottler, L. B., Zipp, J. F., Robins, L. N., & Spitznagel, E. L. (1987). Difficult-to-recruit respondents and their effect on prevalence estimates in an epidemiologic survey. *American Journal of Epidemiology*, 125(2), 329-139. doi:10.1093/oxfordjournals.aje.a114534

- Etter, J., & Perneger, T. V. (1997). Analysis of non-response bias in a mailed health survey. *Journal of Clinical Epidemiology*, 50(10), 1123-1128. doi:10.1016/S0895-4356(97)00166-2
- Fitzgerald, R., & Fuller, L. (1982). I hear you knocking but you can't come in: The effects of reluctant respondents and refusers on sample survey estimates. *Sociological Methods and Research*, 11(1), 3-32. doi:10.1177/0049124182011001001
- Groves, R. M., & Heeringa, S. G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society, Series A*, 169(3), 439-457. doi:10.1111/j.1467-985X.2006.00423.x
- Hall, J., Brown, V., Nicolaas, G., & Lynn, P. (2013). Extended field efforts to reduce the risk of non-response bias: Have the effects changed over time? Can weighting achieve the same effects? *Bulletin de Méthodologie Sociologique*, 117(1), 5-25. doi:10.1177/0759106312465545
- Haring, R., Alte, D., Völzke, H., Sauer, S., Wallaschofski, H., John, U., & Schmidt, C. O. (2009). Extended recruitment efforts minimize attrition but not necessarily bias. *Journal of Clinical Epidemiology*, 62(3), 252-260. doi:10.1016/j.jclinepi.2008.06.010
- Hayes, C., & Watson, N. (2009). HILDA imputation methods. HILDA Project Technical Paper Series 2/09, Melbourne Institute of Applied Economic and Social Research, University of Melbourne. Available at <http://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-technical-papers/htec209.pdf>
- Heerwegh, D., Abts, Koen., & Loosveldt, G. (2007). Minimizing survey refusal and noncontact rates: Do our efforts pay off? *Survey Research Methods*, 1(1), 3-10. doi:10.18148/srm/2007.v1i1.46
- Kennickell, A. B. (2000). What do the "late" cases tell us? Evidence from the 1998 Survey of Consumer Finances. Presented at the 1999 International Conference on Survey Nonresponse, Portland, Oregon. Available at www.federalreserve.gov/Pubs/OSS/oss2/papers/icsn999.pdf.
- Kreuter, F. (2013). Facing the nonresponse challenge. *The Annals of the American Academy of Political and Social Science*, 645(1), 23-35. doi:10.1177/0002716212456815
- Lahaut, V. M. H. C. J., Jansen, H. A. M., Van de Mheen, D., Garretsen, H. F. L., Verdurmen, J. E. E., & Van Dijk, A. (2003). Estimating non-response bias in a survey on alcohol consumption: Comparison of response waves. *Alcohol and Alcoholism*, 38(2), 128-134. doi:10.1093/alcac/agg044
- Larroque, B., Kaminski, M., Bouvier-Colle, M., & Hollebecque, V. (1999). Participation in a mail survey: Role of repeated mailings and characteristics of nonrespondents among recent mothers. *Paediatric and Perinatal Epidemiology*, 13(2), 218-233. doi:10.1046/j.1365-3016.1999.00176.x
- Lepkowski, J. M., & 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* (pp. 259-272). New York: John Wiley and Sons.
- Lin, I., & Schaeffer, N. C. (1995). Using survey participants to estimate the impact of non-participation. *Public Opinion Quarterly*, 59(2), 236-258. doi:10.1086/269471
- Lugtig, P., Das, M., & Scherpenzeel, A. (2014). Nonresponse and attrition in a probability-based online panel for the general population. In M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick and P. J. Lavrakas (Eds.), *Online Panel Research: A Data Quality Perspective* (pp. 135-153). Chichester: Wiley.

- Lynn, P., Clarke, P. S., Martin, J., & Sturgis, P. (2002). The effects of extended interviewer efforts on nonresponse bias. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey Nonresponse* (pp. 135-147). New York: Wiley.
- Schouten, B., Peytchev, A., & Wagner, J. (2017). *Adaptive Survey Design*. Boca Raton: CRC Press.
- Stoop, I. A. L. (2005). *The Hunt for the Last Respondent: Nonresponse in Sample Surveys*. The Hague: Social and Cultural Planning Office of the Netherlands.
- Studer, J., Baggio, S., Mohler-Kuo, M., Dermota, P., Gaume, J., Bertholet, N., Daeppen, J., & Gmel, G. (2013). Examining non-response bias in substance use research – are late respondents proxies for non-respondents? *Drug and Alcohol Dependence*, 132(1), 316-323. doi:10.1016/j.drugalcdep.2013.02.029
- Tourangeau, R., Brick, J. M., Lohr, S., and Li, J. (2016). Adaptive and responsive survey designs: a review and assessment. *Journal of the Royal Statistical Society, Series A*, 180(1), 203-223. doi:10.1111/rssa.12186
- Ullman, J. B., & Newcomb, M. D. (1998). Eager, reluctant, and nonresponders to a mailed longitudinal survey: Attitudinal and substance use characteristics differentiate respondents. *Journal of Applied Social Psychology*, 28(4), 357-375. doi:10.1111/j.1559-1816.1998.tb01710.x
- Watson, N. (2012). Longitudinal and cross-sectional weighting methodology for the HILDA Survey. HILDA Project Technical Paper Series 2/12, Melbourne Institute of Applied Economic and Social Research, University of Melbourne. Available at <http://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-technical-papers/htec212.pdf>.
- Watson, N., & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys* (pp. 157-181). Chichester: Wiley.
- Watson, N., & Wooden, M. (2012). The HILDA Survey: A case study in the design and development of a successful household panel study. *Longitudinal and Life Course Studies*, 3(3), 369-381. doi:10.14301/lcls.v3i3.208
- Woodruff, S. I., Conway, T. L., & Edwards, C. C. (2000). Increasing response rates to a smoking survey for U.S. Navy enlisted women. *Evaluation and the Health Professions*, 23(2), 172-181. doi:10.1177/016327870002300203
- Yan, T., Tourangeau, R., & Arens, Z. (2004). When less is more: Are reluctant respondents poor reporters? *Proceedings of the Survey Research Methods Section, American Statistical Association*, 4632-4651.