

Dead or Alive? Dealing with Unknown Eligibility in Longitudinal Surveys

Nicole Watson¹

Longitudinal surveys follow people over time and some of these people will die during the life of the panel. Through fieldwork effort, some deaths will be reported or known, but others will be unobserved due to sample members no longer being issued to field or having inconclusive fieldwork outcomes (such as a noncontact that is not followed by a contact at a later wave). The coverage of deaths identified among sample members has flow-on implications to nonresponse correction. Using the Household, Income and Labour Dynamics in Australia (HILDA) Survey, four methods are used to examine the extent of missing death reports. The first method matches the sample to the national death register. The second method uses life-expectancy tables to extrapolate the expected number of deaths among the sample with unknown eligibility. The third method is similar but models deaths from data internal to the survey. The fourth method models deaths as part of the attrition process of a longitudinal survey. The last three methods are compared to the first method and the implications for the construction of balanced panel weights and subsequent population inference are explored.

Key words: HILDA Survey; death register; life-expectancy tables; survival model; weighting methods.

1. Introduction

Both cross-sectional and longitudinal surveys have some sample units with unknown eligibility that need to be addressed in some way. In the population frame used to select the sample in a cross-sectional survey, or the first wave of a longitudinal survey, some units will be out of scope. For example, in an area-based frame some dwellings may be empty or may contain people that are not in scope of the survey. Sometimes it may not be possible to make contact with the occupants of a dwelling to ascertain whether they are in scope or not, or indeed if there is anyone resident at that address at all. Survey practitioners tend to focus on high contact rates to ensure the group of units with unknown eligibility is reasonably small.

In a longitudinal survey, the situation is more complex. A sample is selected at one point in time and then is periodically interviewed over a number of years or even decades. While

¹ Melbourne Institute of Applied Economic and Social Research, University of Melbourne, Level 5, 111 Barry St, University of Melbourne Vic 3010, Australia. Email: n.watson@unimelb.edu.au

Acknowledgments: This research makes use of unit record data from the HILDA Survey (Release 13). 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 article, however, are those of the author and should not be attributed to either DSS or the Melbourne Institute. The author thanks John Henstridge, Robert Breunig, an Associate Editor of the Journal and three anonymous referees for their helpful comments on an earlier version of this article.

eligible to be interviewed at the start of the panel, some sample members will die or move abroad during the life of the panel and thus become ineligible to be interviewed at a later wave. Through fieldwork effort, the eligibility of some sample members will be reported or known while others will be unobserved due to sample members no longer being issued to field or having inconclusive fieldwork outcomes (such as a noncontact not followed by a contact in a later wave). The total number of noncontacts grows each wave, mainly due to cases being lost or not being issued to field, and some of these sample members will be out of scope. To illustrate, the Household, Income and Labour Dynamics in Australia (HILDA) Survey is a household-based longitudinal survey that began in 2001. By Wave 13 (conducted in 2013), 25% of the 13,969 Wave 1 respondents were not contacted: four percent were lost following a move (the majority of whom were lost by Wave 3), 20% were not issued to field as a result of a prior refusal or permanent health condition that makes them unable to participate, and 0.4% were issued to field but the interviewer could not make contact. Figure 1 shows the development of these different noncontact categories over time.

The coverage of deaths and moves abroad identified among sample members has important implications for nonresponse correction in constructing sample weights. Many longitudinal surveys provide data users with a range of weights, both cross-sectional and longitudinal, reflecting the different populations to be represented for different analytic objectives and the sample cases that can contribute to these estimates (Lynn 2009). For example, many longitudinal analyses for the period from Wave 1 to Wave t are concerned with individuals living in the population throughout that time, so would exclude people who have died or moved abroad since Wave 1. Some analyses may be concerned with individuals living in the population in Wave 1 who subsequently died by Wave t . Other analyses may focus on individuals that move abroad after Wave 1. Without appropriately accounting for whether sample members with unknown eligibility have died or moved abroad after Wave 1, longitudinal estimates (and cross-sectional estimates post Wave 1) may be biased.

This article examines what impact unknown deaths have on estimates relating to the population in Wave 1 who die by Wave t as well as estimates for individuals living in the

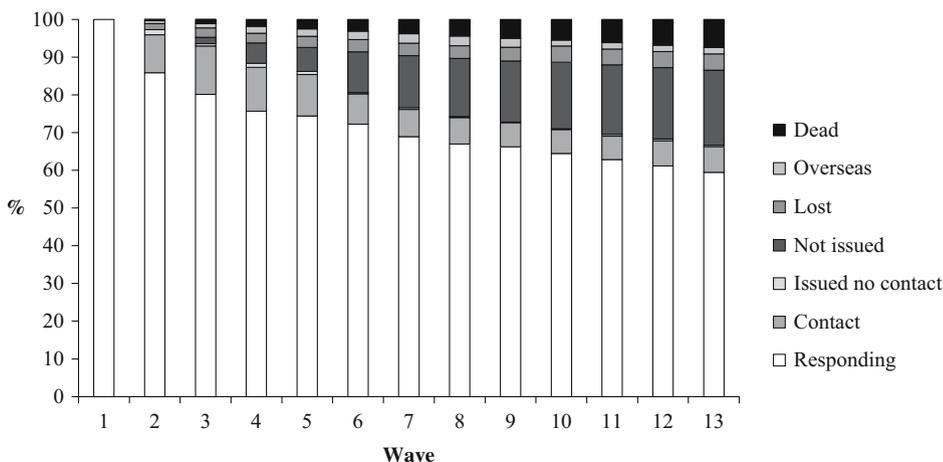


Fig. 1. Outcomes of Wave 1 respondents ($n = 13,969$).

population in Wave 1 through to Wave t . The focus is on deaths, rather than moves abroad, for several reasons:

- i) it is the larger of the two groups (there were 243 Wave 1 respondents reported to have moved overseas by Wave 13 compared to 1,038 Wave 1 respondents who had died),
- ii) it is an absorbing state so can only grow over time, and
- iii) linked survey and administrative death data can provide a benchmark to which alternative weighting methods can be compared.

This analysis of unknown deaths will help inform how other types of unknown eligibility can best be handled in the weighting process.

In the construction of weights for responding sample members in cross-sectional surveys, cases with unknown eligibility may be addressed by a separate (usually simple) step in the weighting process to adjust for different eligibility rates among certain classes of the sample (Valliant et al. 2013, 315). Alternatively they may be handled in the response-propensity adjustments, or implicitly by benchmarking the responding sample to known population totals relating to eligible members in the population. In a similar way, multiple methods exist for calculating response rates in the presence of unknown eligibility. The proportion of eligible cases among those with unknown eligibility is assumed to be 0, 1 or some fraction based on known cases or external data sources. As yet there are no recommendations as to the preferred approach for dealing with unknown eligibility (Smith 2009).

A variety of practices have also been adopted by panel studies to adjust for unknown deaths in the construction of weights for the balanced panel from Wave 1 to Wave t . Typically general population studies make no adjustment for unknown deaths, as appears to be the case for the German Socio-Economic Panel (Kroh 2013), the British Household Panel Survey (BHPS) (Taylor 2010), and the Swiss Household Panel (Weaver 2011). There are three notable exceptions, though each have taken a different approach. First, the US Panel Study of Income Dynamics (up to 1993) used life tables from the US Census Bureau to predict the deaths of noncontacted sample members (Gouskova et al. 2008). From 1993, they assumed that all new nonrespondents were alive, presumably due to very high contact rates. Second, the Canadian Survey of Labour and Income Dynamics applies a nonresponse adjustment to known deaths, although they are not included in the nonresponse modelling to calculate the size of the adjustment (LaRoche 2007). Third, up until 2013 the HILDA Survey counted known deaths as a response in the nonresponse modelling process in constructing the weights (Watson 2012), an approach which is suggested by Kalton and Brick (2000). In 2014, the HILDA Survey sample was matched to the national death register (Watson and Summerfield 2014). Longitudinal studies of older individuals have more to gain from matching their samples to national death registers, as this provides the opportunity to gather additional information about the cause of death along with the exact timing. This is done for the US Health and Retirement Study (Hayward undated; HRS undated). Some other ageing studies take a different approach. The Survey of Health, Ageing and Retirement in Europe benchmarks the longitudinal sample to mortality-adjusted benchmarks (SHARE 2013). The English Longitudinal Study of Ageing does not appear to make any correction for unknown deaths (Scholes et al.

2008). Longitudinal studies of children are likely to ignore the issue of unknown deaths as the number of missing deaths is likely to be very small (see, for example, [Hawkes and Plewis 2006](#), for an analysis of the UK National Child Development Study).

There appears to only be one study that has looked explicitly at the impact of adjusting for unknown eligibility on weighted estimates in a longitudinal survey. [Sadig \(2014\)](#) estimated and applied an adjustment factor to BHPS response-propensity-corrected weights. This adjustment factor was calculated as the ratio of the survival rate in the population (obtained from official statistics) to the survival rate in the (unweighted) sample for each sex and age group (with nine age categories). He found that the adjusted weights reduced both the mean and standard deviation of the weights (as older respondents were receiving higher weights than they ought prior to the adjustment). The adjustment also resulted in sizeable differences to population estimates of subjective health status and some differences to regression coefficients of a model of self-reported health. This approach makes two assumptions that may be violated to some degree. The first is that all individuals in the sample are equal. Despite the initial sample having an approximately equal probability of selection design, there were certainly differences in the probabilities of response in the first wave ([Taylor 2010](#)). The second assumption is that the sample has the same survival probabilities as the population, which may not be accurate in early waves because the sample was selected from private dwellings (and excludes people living in institutions such as nursing homes). It will be more accurate the longer the timeframe involved ([Sadig's](#) paper reported on the 18-year panel).

How many deaths are missing from the part of the sample that is no longer contacted? What is the best way to adjust for these deaths in constructing weights for the balanced panel from Wave 1 to Wave t ? How do these missing deaths and weighting methods affect estimates for the population that die between Wave 1 and Wave t ? How do they affect the estimates for individuals living in the population in Waves 1 through to Wave t ? To answer these questions, data from the HILDA Survey is used. The HILDA Survey began with an approximately equal probability sample selected from private dwellings across Australia. The sample is clustered to allow for face-to-face interviewing. The Wave 1 household-level response rate was 66% (96% contact rate and 68% response rate given contact). Sample members are followed and those aged 15 and older are interviewed annually, collecting a broad range of socioeconomic data. While people living in nonprivate dwellings (i.e., institutions such as nursing homes) are excluded from the Wave 1 sample, sample members who move into nonprivate dwellings in Wave 2 or later are followed and interviewed. The sample is extended to include people who join a sample member's household for as long as they are co-resident, though some of these people are converted to permanent sample members (such as new births to permanent sample members). Four different methods to adjust for unknown deaths in a longitudinal survey are examined in this article. The first method matches the sample to the national death registry, and as a high match rate is achieved, this method serves as the gold standard to which the other methods are compared. The second method uses life-expectancy tables to extrapolate the expected number of deaths in the original sample over time. The third method models survival time from the observed sample and applies this to the nonresponding sample. The fourth method includes known deaths as responses in the nonresponse adjustment process over time. To assess the performance of each method

against the death-register matching method, the number of deaths, the timing of these deaths, and the sociodemographic characteristics of those who die are compared. Various population estimates and model estimates are also compared for individuals living in the population from Wave 1 to 13, as these estimates may be impacted by differences in the portion of the Wave 1 population that is removed due to death by Wave 13.

The article proceeds as follows. Section 2 describes the four methods in detail. Section 3 compares the number and composition of actual and estimated deaths under each method. Section 4 examines the balanced panel weights produced and the implications for estimation are outlined in Section 5. Section 6 concludes the article.

2. Methods to Account for Deaths

This section describes four alternative methods for estimating the number of unknown deaths in the sample. The first three methods seek to identify or impute deaths in the part of the sample that is no longer contacted. The fourth method weights up the known deaths to allow for unknown deaths.

2.1. Method 1: Matching to Death Register

Following the twelfth wave of data collection, the HILDA Survey sample was matched to the national death register. A name-based probabilistic linking strategy was used which matched on name (first name and surname), date of birth and sex ([Australian Institute of Health and Welfare 2011](#)). A series of passes were made over the data to allow for variations in the recorded data. Match weights were assigned to linked records based on how closely the records matched: high weights indicate the matched pairs have exact or very similar data and low weights indicate they have very different data. The main contribution to the match weight is made by the names, based on how common each name is and how close the names are in the matched pair. An agreement in day, month, year, or sex increases the weight by 1 for each component and a disagreement in day, month, or sex decreases the weight by 1 for each component. The weight is decreased by 1 for each year the date of birth disagrees. The weights are used to decide between multiple matches and the higher-weighted match is taken. The weights are also used to decide when a match is likely. To determine the cutoff point to use, the distribution of the match weights for the sample known to be alive in Wave 12 (i.e., they were part of a responding household) is compared to the known dead sample (i.e., reported to be dead in Wave 12 or earlier). [Figure 2](#) shows the distribution of the match weights for the known live sample (grey line), the known dead sample (black line), and also the entire sample (dashed grey line). As an example, there were 1,191 matched records with a match weight of 11. This is made up of 863 cases among the known live sample at Wave 12, four cases among the known dead sample, and 324 cases where it was unknown whether they were dead or alive (this last group is not shown as a separate line on the graph). Almost all (98%) sample members whose status is known that have a match weight of 29 or more are known deaths. By comparison, for those sample members with a known status that have a match weight of 21 or less, only 0.5% are known deaths. As a result, the matches with a match weight above 21 were clerically reviewed for consistency with the date of last contact. Overall, 95% of the known (or reported) deaths could be matched to a death-register record. Included in

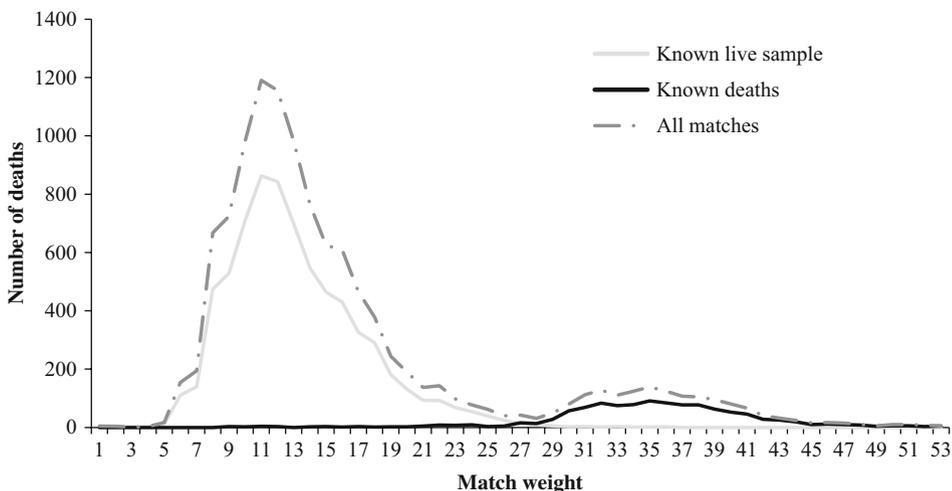


Fig. 2. Death-registry match weights, all match pairs.

this number were eight known deaths that had a match weight below 22 (amounting to 0.7% of all known deaths). Of the five percent of known deaths that could not be matched, the main reasons were as follows:

- they had died overseas,
- the surname did not match (and was not close), and
- the date of birth was not close.

This method adds a total of 265 (unweighted) deaths of Wave 1 respondents to the datasets by Wave 13, representing a 26% increase in the number of deaths recorded. This also increases the response rate of Wave 1 respondents re-interviewed in Wave 13 (excluding those considered out of scope in Wave 13) from 65% to 67%.

Matching to the death register will be superior to the other methods if the match rate is high. As the match rate of reported deaths to the death register is high for the HILDA sample, this method is used as the gold standard against which the other methods are compared. This method is more costly and time-consuming (in terms of the clerical work to review matches) than the other methods and may not be able to be undertaken every wave.

2.2. Method 2: Using Life Tables

The second method employs age- and sex-specific death rates from life tables produced annually by the Australian Bureau of Statistics (ABS 2014). These death rates are the proportion of deaths that occur between exact age x and $x + 1$, and is denoted as p_x . To apply these death rates to the HILDA sample where a sample member is aged x years and m months at the date of last contact, a weighted combination of p_x and p_{x+1} is used. That is, the probability that a person aged x years and m months at Wave t dies between Wave t and $t + 1$ is:

$$p_{txm} = \frac{(12 - m)}{12} p_x + \frac{m}{12} p_{x+1}$$

These rates are used to calculate the probability of a death between Wave t and $t + n$ for a person aged x years and m months at Wave t as:

$$p_{mxm} = p_{txm} + \sum_{j=1}^{n-1} p_{t+j,x+j,m} \prod_{k=0}^{j-1} (1 - p_{t+k,x+k,m})$$

That is, the probability of a death between Wave t and $t + n$ is the sum of the probability of death between Wave t and $t + 1$, the probability of death between Wave $t + 1$ to $t + 2$ for those who survive to Wave $t + 1$, the probability of death between Wave $t + 2$ to $t + 3$ for those who survive to Wave $t + 2$, and so on.

These rates of death are applied from the last wave in which a sample member is part of a *responding* household (where the household roster has been completed and at least one person has been interviewed). Each person is assigned a random number from the uniform distribution and is flagged as an imputed death in Wave n if the random number is below p_{mxm} . An alternative would be to assign death from the wave of last *contact* with a sample member (i.e., those lost to tracking, not issued, and issued but not contacted), but the interviewer may not speak with the sample member directly as this contact could have been with another member of the household or with some other person and the eligibility of the sample member may not have been ascertained. This method adds 279 deaths of Wave 1 respondents to the datasets by Wave 13, representing a 27% increase in the number of deaths recorded.

The life-table method is a relatively simple method to apply; however, life tables usually reflect survival rates in the population as a whole irrespective of whether the individuals live in institutions or not. If the sample has been selected in such a way as to exclude people living in institutions (as is often the case) or if the following rules preclude following people into institutions, then the general population survival rates may be too low (and the death rates too high) when applied to the sample in the short term.

2.3. Method 3: Model Deaths Explicitly

The third method uses the deaths observed through the fieldwork to model the survival curve for sample members. By using the available survey variables, it may be possible to predict deaths more accurately than by just using age and sex alone as in Method 2.

Survival time of Wave 1 sample members is modelled as a Weibull proportional hazards model. Sample members are designated to be alive, dead, or right censored (when the sample is not known to be either alive or dead). The survival time t for individual j is modelled as:

$$S_j(t) = \exp\left\{-\exp(\mathbf{X}_j\boldsymbol{\beta})t_j^p\right\}$$

where \mathbf{X} is the matrix of covariates, $\boldsymbol{\beta}$ is the vector of fitted coefficients and p is the Weibull shape parameter estimated from the data. Models for males and females are built separately. For each, a model for those belonging to a Wave 1 responding household (where only some basic information is available) is constructed and another for Wave 1 respondents (for whom more detailed information is available). The covariates considered for sample members belonging to a responding household are age, presence of long-term health condition, whether married/de facto married, and employment status. The covariates for Wave 1

respondents includes similar variables from the individual interview together with education and occupation. The models are unweighted, as the aim is to predict deaths in the sample rather than the population. If certain segments of the population had been oversampled (for example, immigrants), then variables could be included in the model or separate models could be developed to reflect this. This is not the case for the HILDA Survey and including weights make little difference to the results.

The fitted models are used to estimate the time to death for sample members in the last wave in which they are part of a responding household and were not known to have died at a later wave. The characteristics used for these sample members are those collected in the last responding wave. Each person is given a random number from the uniform distribution and when that number exceeds the predicted probability of survival in Wave $t + n$, the sample member is assumed to have died in Wave $t + n$. This method adds 262 deaths of Wave 1 respondents to the datasets by Wave 13, representing a 24% increase in the number of deaths recorded.

It is expected that by modelling deaths from the data internal to the survey, the limitations of the life-table method (Method 2) will be overcome. It should more accurately predict deaths in samples that do not align with the overall population (to which the life tables relate). It should also more accurately reflect the sociodemographic characteristics of those that die as it uses other characteristics besides age and sex to model the probability of death. This method assumes that there is reasonably good coverage of reported deaths within the portion of the sample recently contacted.

2.4. Method 4: Model Deaths Implicitly

The fourth method implicitly models deaths when adjusting for attrition in the weighting process. In the ideal world of 100% response rates, the eligibility of all sample members would be known in all waves. In reality, with response and contact rates less than 100%, sample members are known to be alive (as they are part of a responding household), known to be dead, or are not known to be either dead or alive (as they are part of a nonresponding household). When adjusting for nonresponse in the construction of the balanced panel of responding persons in Waves 1 through 13, unknown deaths can *implicitly* be taken account of by classifying the sample into two groups:

- Response – interviewed in Wave 1, and then interviewed, overseas, or dead in Waves 2 to 13, or
- Nonresponse – all other individuals interviewed in Wave 1.

This method of counting sample members who die or are not eligible to be interviewed (because they move overseas or, in some surveys, move into institutions) as responses was suggested by [Kalton and Brick \(2000\)](#).

Note that there are a number of different ways that weights can be developed for longitudinal surveys and [Kalton and Brick \(2000\)](#) provide an overview of the various choices that can be made. Different methods can be used to calculate nonresponse adjustments (via weighting classes, classification trees, logistic regression, or generalised raking), or how these adjustments are applied over time (wave by wave, in a single step, or in segments of waves). There are also different choices regarding poststratification

adjustments, including whether or not they are made, and if made, to which population they refer (the population as at Wave 1, or some adjustment to this population to exclude deaths, emigration and immigration to reflect the live longitudinal population from Wave 1 to Wave t). It is beyond the scope of this article to consider how all of these choices interact with the ways in which unknown eligibility may be handled. Rather, the method for dealing with unknown eligibility has been varied in this article while holding the remaining aspects of the weighting strategy constant. It is not expected that these choices in other aspects of the weighting strategy would greatly affect the results.

The particular weighting strategy adopted in the HILDA Survey is as follows. The probability of response is modelled on numerous Wave 1 characteristics including age, sex, marital status, ability to speak English, employment status, hours worked, education level, health status, number of times moved in last ten years, interviewer-rated cooperation, whether the self-completion component was returned, geographical location, remoteness area, number of people in the household, household type, condition of dwelling, number of calls made, and whether the household was partially responding. Some post-Wave 1 mobility information was included when available. The balanced panel weight w_{r1_13} for responding persons in Waves 1 through 13 is taken to be the Wave 1 responding person weight w_{r1} divided by this probability of response in Waves 2 through 13 p_{r2_13} :

$$w_{r1_13} = \frac{w_{r1}}{p_{r2_13}}$$

As a result, those respondents who share similar characteristics with nonrespondents receive a higher adjusted weight than those who do not. These weights are then calibrated to the Wave 1 benchmarks for age, sex, labour-force status, social marital status, household composition, and geographic area.

To obtain a weighted estimate of the sample deaths, this balanced panel population weight is rescaled to the Wave 1 responding sample size. This method estimates an additional 317 deaths by Wave 13, representing a 30% increase in the weighted number of known deaths (weighted by the Wave 1 responding person weight rescaled to the sample size).

It is expected that this method would produce similar estimates to the death-modelling method (Method 3) as they both use data internal to the survey.

3. Comparison of Methods

A comparison of the number of deaths reported, matched, or estimated by the different methods is given in [Table 1](#). Part 1 of the table provides the weighted estimate of the number of deaths cumulated from Wave 2 to each subsequent wave. The weight used to create the estimate for all methods except Method 4 (weighting) is the Wave 1 responding person weight rescaled to the sample size. The weight used for Method 4 is the balanced panel weight for Waves 1 through to Wave t calculated via that method (as described in Subsection 2.4). Part 2 of the table shows the difference in the number of deaths estimated via Method 1 to the other methods. Part 3 provides the standard errors for these differences; these are calculated using jackknife estimation from the 45 replicates

Table 1. Cumulative number of deaths by wave and method.

	Wave												
	2	3	4	5	6	7	8	9	10	11	12	13	
1. Estimated number of deaths													
Reported deaths	56	159	270	368	455	542	629	710	782	867	963	1036	
Method 1: Death register	71	178	291	403	506	610	728	829	940	1060	1204	1325	
Method 2: Life tables	66	179	309	433	544	656	757	874	971	1089	1232	1342	
Method 3: Model deaths	63	170	294	402	516	630	745	861	966	1078	1210	1333	
Method 4: Weighting	64	166	284	377	483	597	703	831	953	1077	1231	1353	
2. Difference from Method 1													
Reported deaths	-15	-18	-21	-35	-51	-68	-100	-119	-158	-193	-242	-289	
Method 2: Life tables	-5	1	17	30	38	46	29	44	30	29	27	17	
Method 3: Model deaths	-8	-7	3	-1	10	19	17	31	25	18	6	8	
Method 4: Weighting	-8	-11	-8	-26	-23	-14	-25	2	13	16	27	28	
3. Standard error of difference													
Reported deaths	4.2	4.5	4.4	5.8	6.8	9.0	9.8	12.5	12.4	15.5	19.0	21.0	
Method 2: Life tables	5.0	7.0	8.8	11.1	13.2	16.3	15.4	18.3	16.4	18.2	19.6	22.4	
Method 3: Model deaths	6.2	6.0	6.9	7.3	9.4	11.3	12.4	12.1	15.4	17.5	18.4	19.2	
Method 4: Weighting	3.9	10.0	12.4	14.2	16.1	18.6	19.6	22.3	24.0	27.1	35.7	33.8	
4. P(difference = 0)													
Reported deaths	0.000												
Method 2: Life tables	0.310	0.903	0.048	0.008	0.004	0.005	0.062	0.015	0.066	0.116	0.163	0.441	
Method 3: Model deaths	0.201	0.218	0.692	0.883	0.287	0.084	0.171	0.009	0.102	0.305	0.754	0.672	
Method 4: Weighting	0.051	0.254	0.532	0.069	0.155	0.469	0.200	0.933	0.601	0.543	0.452	0.404	

Note: Numbers are weighted by the population weight rescaled to number of Wave 1 respondents ($n = 13,969$). For all except Method 4, the weight is the Wave 1 responding person population weight. For Method 4, the weight is the balanced panel weight for Wave 1 to Wave t , where t is the wave number for the column. P-values in bold and italics indicate the estimate is significantly different from Method 1 at the 5% and 10% level respectively.

routinely used to construct the HILDA weights. Part 4 of the table reports the probability that these differences are equal to zero.

The death-register matching approach (Method 1) adds 68 new deaths (weighted) by Wave 7 and 289 by Wave 13 to the number of deaths already reported in the sample. As not all of the known deaths were matched (five percent were unmatched), it would be reasonable to expect there to be a similar portion of deaths not matched to the death register among sample members with unknown eligibility. This would translate to about four further deaths by Wave 7 and 15 by Wave 13, suggesting the true number of missing deaths is about 72 in Wave 7 and 304 in Wave 13. Nevertheless, as the death-register matching was of reasonably high quality, the other methods are compared to this method. The life-table method (Method 2) significantly overestimates the number of missing deaths by Wave 4 and does not come back into line with the death-register results until Wave 10. This is likely to be a result of selecting the initial sample from private dwellings (thus excluding nursing homes and other institutions), whereas the life tables relate to all people living in Australia. This healthy sample bias disappears after a decade as the next cohort of sample members age and are followed into these institutions prior to death. Modelling the deaths observed in the data (Method 3) or implicitly modelling the deaths through the attrition adjustment (Method 4) produce very similar results to those obtained by matching to the death register. Only by Wave 9 is the cumulative number of deaths higher when modelling the deaths (Method 3) compared to the death-register matching method (at the 5% significance level).

The four methods differ somewhat in the distribution of deaths over time. [Figure 3](#) shows the year of death for reported deaths and the total number of deaths for the four methods. As more of the sample is withheld from fieldwork over time, a greater portion of deaths are missed in later waves. This is seen in the light grey dotted line, which has the highest number of deaths in 2003 and then generally falls over time after that. The number of deaths in 2002 is lower than in 2003 as the initial sample is selected from private dwellings and the respondents are well enough to take part in an interview in 2001. Matching to the death register helps correct the number of deaths to a generally more stable number over time from 2004. Note that the death-register matching results in a lower number of deaths recorded for 2003 than for the reported deaths. This is because the *exact* timing of these deaths is not always reported. For about one third of the deaths identified in the field, all that is known about the timing of the death is that the death was not reported at the last wave and is reported at the current wave. A quarter of those deaths are assumed to have occurred in the previous year and three quarters in the current year (as approximately 50% of the interviews are completed by September). Of the cases where a year of death was reported, only three percent did not match the year of death recorded in the death register. In terms of the other methods, the life-table method (Method 2) and the death-modelling method (Method 3) distributes the deaths in a fairly similar manner to the death-register matching, though there are significantly more deaths in 2003 with the life-table method. The weighting method (Method 4), where the balanced panel weight for Waves 1 through to 13 is applied to the known deaths, shows substantially more deaths in early years and fewer in the later years. This is a result of greater coverage of deaths (i.e., known deaths compared to all deaths) in the early waves and less in the later waves. This may be ameliorated by modelling attrition on a wave-by-wave basis rather than in one step

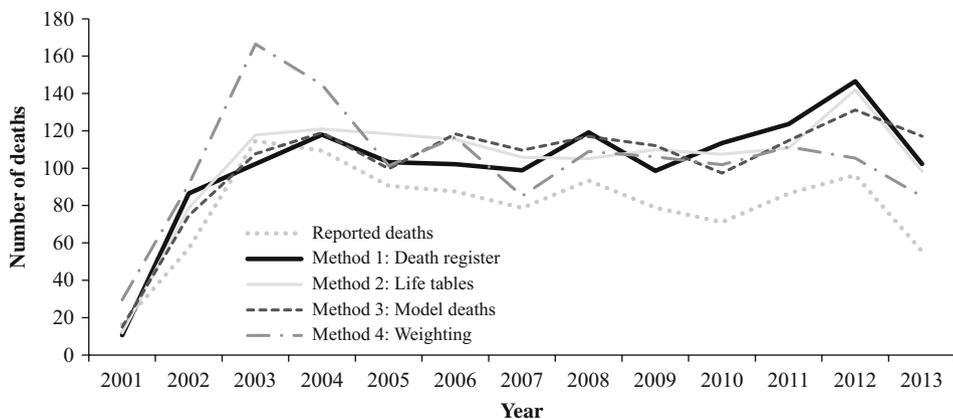


Fig. 3. Year of death of Wave 1 respondents (weighted).

from Wave 1, however this may also increase the variability of the weights. Alternatively, a benchmark for the number of additional deaths estimated via each subsequent balanced panel could be included in the poststratification adjustment step of the weighting process as the cumulative number of deaths reported in Table 1 are consistent with the number of deaths identified from matching to the death register.

Aside from the total number and distribution of the deaths over time, how well do the methods reflect the sociodemographic characteristics and health conditions (as reported at Wave 1) of those who subsequently die? Table 2 provides a summary of the (weighted) Wave 1 characteristics of the reported deaths compared to the deaths identified via the four methods by Wave 13. Compared to the death-register matching (which is expected to be the closest to the truth), the people reported to have died are more likely to be younger, male, live in a two-adult household, have a long-term health condition, or have a higher mental health score. These characteristics to some extent reflect the household situations in which a death is more likely to be reported, as another sample member can report on the death. Deaths of older females living alone are less likely to be reported. For the two methods that impute missing deaths based on respondent characteristics (Methods 2 and 3), the additional deaths are likely to be (at Wave 1) younger, born in countries where the main language spoken is not English, residing in a capital city, have higher physical functioning scores, and higher mental health scores, and be less likely to live in single-person households than those identified via the death-register matching. This suggests that, while these two methods achieve the same number of deaths by Wave 13 as the death-register matching, they are not sufficiently discriminating in the types of people imputed as deaths over the life of the panel. For the twelve characteristics considered in Table 2, Method 4 (weighting) is indistinguishable from Method 1.

4. Weights

To explore the implications of these different methods of dealing with unknown deaths in a longitudinal survey, a set of weights for the balanced panel from Wave 1 through to Wave 13 is produced for each method. The weighting process for Method 4 is described

Table 2. Mean characteristics of deaths.

Wave 1 characteristic	Method 1: Death register		Method 2: Life tables		Method 3: Model deaths		Method 4: Weighting	
	Est	P(Rpt-M1 = 0)	Est	P(M2-M1 = 0)	Est	P(M3-M1 = 0)	Est	P(M4-M1 = 0)
Age (years)	69.4	0.006	70.0	0.000	68.8	0.000	69.6	0.266
Female (%)	43.2	0.008	45.6	0.777	44.4	0.229	44.9	0.537
Born abroad in main English-speaking country (%)	13.5	0.645	13.3	0.923	13.5	0.720	13.2	0.848
Born abroad not in main English-speaking country (%)	13.3	0.109	14.5	0.082	17.1	0.003	14.2	0.848
Single-adult household (%)	30.3	0.351	31.1	0.025	29.0	0.002	31.1	0.967
Two-adult household (%)	55.0	0.003	52.4	0.968	53.7	0.151	52.4	0.989
Three or more adult household (%)	14.6	0.033	16.5	0.174	17.3	0.397	16.6	0.990
Married/defacto (%)	60.2	0.295	59.2	0.215	61.0	0.051	58.1	0.459
Lives in capital city (%)	56.3	0.602	55.9	0.006	58.6	0.002	55.8	0.953
Has long-term health condition (%)	58.8	0.009	56.6	0.216	55.7	0.366	58.5	0.180
SF-36 Physical functioning score (0–100)	56.4	0.411	56.0	0.002	57.7	0.001	56.3	0.716
SF-36 Mental health score (0–100)	73.7	0.040	73.1	0.007	73.8	0.013	74.0	0.084

Note: Bold and italicised *p*-values indicate the estimate is significantly different from the Method 1 estimate at 1% level and 5% level respectively.

Table 3. Summary of weights by method.

Weight description	All respondents (aged 15+ in Wave 1)			All respondents (aged 65+ in Wave 1)		
	Mean	StdDev	CV	Mean	StdDev	CV
Method 0: Known death as out of scope, unknown death as nonresponse	1898.2	1226.2	0.646	2014.0	1294.8	0.643
Method 1: Death register	1852.9	1199.3	0.647	1731.4	1113.8	0.643
Method 2: Life table	1851.3	1196.1	0.646	1704.6	1081.5	0.634
Method 3: Model death	1856.3	1201.9	0.647	1760.4	1132.4	0.643
Method 4: Known death as response	1848.2	1191.9	0.645	1699.7	1097.2	0.646

earlier in Subsection 2.4 and the other three methods follow the same process, with the exception of how deaths are treated. For these three methods, deaths that are reported, matched (as in Method 1), or imputed (as in Methods 2 and 3) do not have an attrition adjustment applied. They do, however, have a separate benchmark applied and this is calculated as the weighted sum of deaths (using the Wave 1 cross-sectional weight). The small portion of reported deaths (five percent) not matched to the death register in Method 1 are weighted according to the procedure adopted for Method 4. Another method is also considered here which treats known deaths as out of scope and unknown deaths as being alive nonrespondents (denoted as Method 0). This method is included as it is explicitly or implicitly adopted by a number of longitudinal studies, as mentioned earlier.

Table 3 provides the mean, standard deviation, and coefficient of variation (= standard deviation/mean) of the weights for the individuals living in the population from Wave 1 to 13. Note that the mean of the weights is not the same across all methods, as different numbers of deaths are assumed under each method. The coefficient of variation is a useful measure to compare across weights with different means as it is not unit specific. Weights with a lower coefficient of variation are preferred over those with a higher one (assuming no difference in the bias of resulting estimates) as the estimates will be more efficient. These summary statistics are provided for all adults and for those aged 65 and over.

Under Method 0, the mean of the weights for the older members of the population are substantially higher as it is assumed that all unknown deaths are alive nonrespondents and respondents with similar characteristics (such as being of a similar age) receive higher weights. All methods are very similar in terms of their coefficient of variation.

5. Impact on Analysis

In this section, various population estimates and model estimates are compared across the five weighting schemes. Table 4 provides weighted estimates for Wave 13 (2013) of the proportion of people with a long-term health condition, self-rated health status (“excellent”, “very good”, “good”, “fair”, and “poor”), health satisfaction, life satisfaction, and a rating of how much their health limits the work they can do (including paid work, work around the house or garden). The weights used here are the balanced panel weights for the

Table 4. Weighted estimates for various health measures at Wave 13, by age.

	Method 0	Method 1	Method 2	Method 3	Method 4
Aged 15+					
Has long-term health condition (%)	37.3	36.5	36.5	36.6	36.4
Health status (%)					
Excellent	8.5	8.6	8.6	8.6	8.6
Very good	33.8	34.1	34.1	34.1	34.2
Good	37.0	37.0	36.9	37.0	37.0
Fair	16.7	16.3	16.3	16.4	16.2
Poor	4.1	4.0	4.0	4.1	4.0
Health satisfaction (0–10)	7.0	7.0	7.0	7.0	7.0
Life satisfaction (0–10)	7.9	7.9	7.8	7.9	7.9
Health limitation (0–10)	1.5	1.5	1.5	1.5	1.5
Aged 65+					
Has long-term health condition (%)	70.9	70.4	70.3	70.5	70.3
Health status (%)					
Excellent	2.2	2.0	2.0	2.1	2.0
Very good	20.3	20.7	20.8	20.6	20.9
Good	38.8	38.7	38.6	38.7	38.7
Fair	32.1	31.8	31.9	32.0	31.8
Poor	6.6	6.7	6.7	6.7	6.6
Health satisfaction (0–10)	6.8	6.8	6.8	6.8	6.8
Life satisfaction (0–10)	8.2	8.2	8.2	8.2	8.2
Health limitation (0–10)	3.4	3.4	3.4	3.4	3.4

Note: Method 0 weights unknown death as nonresponse. Method 1 obtains extra deaths from death register. Method 2 imputes extra deaths based on life tables. Method 3 imputes extra deaths from model of deaths in sample. Method 4 weights known deaths as response.

13-wave panel. Estimates are provided for all people aged 15 and over, and for those aged 65 and over.

There are no significant differences between the means calculated under the death-register method and any of the other methods. The estimate that is closest to being significantly different is the proportion of people aged 65 and over reporting ‘very good’ health for weights calculated with no adjustment versus those with the death-register matching (Method 0 vs Method 1; difference=0.4%, p -value=0.149).

Next the effect that the various death-adjustment methods have on two weighted random-effect logistic models is considered; one model is focused on life satisfaction and the other model examines the extent of health limitations. It is possible for the marginal distribution of these outcome variables (i.e., conditional on covariates in the model) to be affected by the weighting method even though the unconditional means are unaffected. Both models considered here have been restricted to older adults (aged 65+) as the impact of how unknown deaths are handled in the weights (if there is any) is likely to be greatest among this group due to their higher rate of death. The models were fitted in Stata using the *gllamm* command which allows for probability weights in the calculation of the random effects logistic model. The dependent variable for the first model is high life satisfaction (i.e., a

score of eight or higher out of ten), and the dependent variable for the second model is a rating of eight or higher out of ten in terms of how limiting the respondent's health condition is on the work they can do. The covariates include age, sex, broad country of birth (indicators of whether born overseas in a mainly English-speaking country or not), real equivalised household income, marital status, education level, region, whether another adult was present during the interview, and (for the first model only) presence of a long-term health condition. These variables were included following other studies of life satisfaction and health limitations (Green 2011; Au and Johnston 2014). The weights used here are the balanced panel weights for the 13-wave panel. The standard errors are calculated using the replicate weights and thus take account of the complex sample design. Table 5 presents the regression coefficients for the life-satisfaction model. The significant coefficients differ by at most seven percent from the baseline model (Model 0). Which weighting method is adopted is irrelevant to this model: the coefficients for the significant variables for Models 2, 3, and 4 differ by less than two percent compared to those for the death-register method (Method 1). None of these differences are significant. The coefficient closest to being significantly different from the Method 1 coefficients is the indicator variable for remote Australia for Method 0 (diff = 0.187, p -value = 0.119).

In the model of health limitations presented in Table 6, the significant coefficients differ by at most 13% from the baseline model (Method 0). Again, there is little difference resulting from the method used to adjust for unknown deaths, with significant coefficients changing by no more than five percent. For the covariate sex, the coefficient becomes marginally significant under Method 4 (weighting), reflecting only a small change in the p -value for the coefficient which is close to the ten-percent significance level in the other models. None of the differences in the coefficients of these models are significantly different from zero. The coefficient closest to being significantly different from Method 1 is for age in Method 0 (diff = -0.079, p -value = 0.180).

An alternative multilevel modelling approach proposed by Skinner and Holmes (2003) could have been used here. This method incorporates all balanced panels from Wave 1, not just the balanced panel from Wave 1 to 13, and thus increases the number of person-year observations used in the models by nearly 80%. Under this modelling approach, the Wave 1 weight is used as the individual (Level 2) weight and a rescaled version of the balanced panel weights is used for the wave observations (Level 1). The wave-observation (Level 1) weights are rescaled so that the sum of these weights equals the number of wave observations for an individual (Scaling Method 2 in Rabe-Hesketh and Skrondal 2006). The model parameters under this approach are more robust to the particular choice of weighting method (Method 0–4) adopted. The parameter estimates differ by at most two percent in the model of life satisfaction from the baseline method (Method 0) and at most five percent in the model of health limitations (results not shown).

6. Summary and Conclusions

This article examined four different methods for adjusting for unknown deaths in the construction of weights for a longitudinal sample. Matching the sample to the death register will produce the most accurate results assuming the matching process is of high quality. It will also provide the opportunity to add further information to the dataset, such

Table 5. Random effects logistic regression coefficients for high life satisfaction (>=8 out of 10), aged 65+.

	Method 0 b (se)	Method 1 b (se)	Method 2 b (se)	Method 3 b (se)	Method 4 b (se)
Age/10	1.309 (1.857)	0.627 (1.848)	0.860 (1.959)	0.886 (1.890)	0.633 (1.817)
Age squared/100	-0.097 (0.121)	-0.053 (0.121)	-0.069 (0.129)	-0.070 (0.124)	-0.054 (0.119)
Sex	0.153 (0.258)	0.120 (0.261)	0.119 (0.266)	0.127 (0.267)	0.102 (0.224)
Born overseas, main English-speaking country	-0.099 (0.340)	-0.087 (0.360)	-0.088 (0.349)	-0.094 (0.362)	-0.070 (0.313)
Born overseas, not main English-speaking country	-1.146*** (0.337)	-1.192*** (0.348)	-1.196*** (0.342)	-1.187*** (0.360)	-1.183*** (0.334)
Real equiv. household income/100,000	1.740** (0.774)	1.820** (0.785)	1.840** (0.791)	1.800** (0.783)	1.810** (0.774)
Real equiv. household income squared/100,000 ²	-0.695 (0.622)	-0.728 (0.624)	-0.744 (0.629)	-0.711 (0.623)	-0.701 (0.616)
Separated / divorced	-0.622** (0.280)	-0.661** (0.304)	-0.658** (0.328)	-0.641** (0.313)	-0.666** (0.279)
Widowed	-0.420** (0.181)	-0.421** (0.196)	-0.428** (0.195)	-0.419** (0.197)	-0.415** (0.174)
Single	-0.748 (0.654)	-0.675 (0.661)	-0.690 (0.672)	-0.717 (0.677)	-0.712 (0.550)
Graduate	-0.905 (0.774)	-0.903 (0.811)	-0.889 (0.828)	-0.901 (0.833)	-0.887 (0.728)
Year 11 and below	-0.292 (0.209)	-0.265 (0.224)	-0.271 (0.228)	-0.276 (0.233)	-0.275 (0.209)
Has long-term health condition	-0.601*** (0.140)	-0.591*** (0.153)	-0.594*** (0.146)	-0.593*** (0.148)	-0.593*** (0.125)

Table 5. Continued.

	Method 0 b (se)	Method 1 b (se)	Method 2 b (se)	Method 3 b (se)	Method 4 b (se)
Regional Australia	-0.066 (0.210)	-0.069 (0.232)	-0.064 (0.229)	-0.068 (0.233)	-0.084 (0.227)
Remote Australia	-0.143 (0.412)	-0.331 (0.441)	-0.324 (0.487)	-0.281 (0.446)	-0.304 (0.433)
Other adult present during interview	0.076 (0.140)	0.099 (0.137)	0.104 (0.134)	0.099 (0.138)	0.110 (0.135)
Constant	-1.460 (6.950)	1.143 (6.982)	0.272 (7.405)	0.161 (7.154)	1.173 (6.844)
Number of person-year observations	8796	8796	8796	8796	8796

Notes: 1. Method 0 weights unknown death as nonresponse. Method 1 obtains extra deaths from death register. Method 2 imputes extra deaths based on life tables. Method 3 imputes extra deaths from model of deaths in sample. Method 4 weights known deaths as response.
2. ***, **, and * represent significance at 1% level, 5% level and 10% level respectively.

Table 6. Random effects logistic regression coefficients for work-limiting health condition (≥ 8 out of 10), aged 65+.

	Method 0 b (se)	Method 1 b (se)	Method 2 b (se)	Method 3 b (se)	Method 4 b (se)
Age/10	3.562 (2.999)	4.352 (2.971)	4.321 (3.230)	4.123 (3.177)	4.371 (2.756)
Age squared/100	-0.119 (0.191)	-0.168 (0.189)	-0.167 (0.206)	-0.155 (0.203)	-0.169 (0.174)
Sex	-0.435 (0.288)	-0.434 (0.269)	-0.441 (0.289)	-0.442 (0.273)	-0.414* (0.226)
Born overseas, main English-speaking country	-0.262 (0.333)	-0.270 (0.360)	-0.268 (0.354)	-0.265 (0.348)	-0.283 (0.329)
Born overseas, not main English-speaking country	1.145*** (0.314)	1.193*** (0.333)	1.194*** (0.328)	1.191*** (0.334)	1.201*** (0.300)
Real equiv. household income/100,000	0.519 (1.220)	0.224 (1.190)	0.183 (1.180)	0.302 (1.190)	0.150 (1.180)
Real equiv. household income squared/100,000 ²	-1.050 (1.430)	-0.865 (1.410)	-0.831 (1.400)	-0.919 (1.410)	-0.827 (1.420)
Separated/divorced	-0.668 (0.420)	-0.656 (0.416)	-0.662 (0.434)	-0.660 (0.420)	-0.666 (0.414)
Widowed	0.005 (0.320)	0.021 (0.317)	0.052 (0.312)	0.033 (0.310)	0.026 (0.296)
Single	0.171 (0.788)	0.196 (0.841)	0.211 (0.847)	0.210 (0.834)	0.265 (0.673)
Graduate	0.359 (0.504)	0.393 (0.510)	0.395 (0.486)	0.387 (0.503)	0.404 (0.486)
Year 11 and below	0.771*** (0.253)	0.774*** (0.276)	0.781*** (0.269)	0.785*** (0.283)	0.806*** (0.244)
Regional Australia	0.254 (0.326)	0.279 (0.311)	0.264 (0.328)	0.271 (0.322)	0.292 (0.317)

Table 6. Continued.

	Method 0 b (se)	Method 1 b (se)	Method 2 b (se)	Method 3 b (se)	Method 4 b (se)
Remote Australia	-0.079 (1.065)	0.048 (1.148)	0.060 (1.346)	0.010 (1.215)	0.034 (1.255)
Other adult present during interview	0.174 (0.192)	0.129 (0.209)	0.125 (0.202)	0.141 (0.214)	0.127 (0.202)
Constant	-24.064** (11.92)	-27.140** (11.85)	-26.936** (12.79)	-26.215** (12.56)	-27.290** (10.97)
Number of person-year observations	8784	8784	8784	8784	8784

Notes: 1. Method 0 weights unknown death as nonresponse. Method 1 obtains extra deaths from death register. Method 2 imputes extra deaths based on life tables. Method 3 imputes extra deaths from model of deaths in sample. Method 4 weights known deaths as response.
2. ***, **, and * represent significance at 1% level, 5% level and 10% level respectively.

as a more accurate date of death and cause of death. Nonetheless, there are some difficulties in using the death register. In the HILDA Survey sample, five percent of the known (or reported) deaths were unmatched, suggesting that among the unknown deaths there would be a similar percentage of deaths not identified via this method. By Wave 13, this would amount to approximately 15 missed deaths and these would most likely be recently separated or divorced women who change their surname or people who died after moving overseas. Some adjustment in the construction of the weights would need to be made for these types of people.

Where matching to the death register is not an option or is of low quality, an alternative method will need to be adopted to improve estimates of the number of deaths, the timing of the deaths, and the sociodemographic characteristics of those that die since the start of the panel. The life-table method (Method 2) tends to overestimate deaths in the HILDA Survey sample in the short term (up to around Wave 9) for two reasons. First, the initial sample was selected from private dwellings only, so excludes institutions such as nursing homes. Second, the sample members also needed to be well enough to participate in the Wave 1 interview. For both of these reasons, the sample would be in better health on average than the overall population to which the life tables relate. Modelling the deaths from data internal to the survey (Method 3) proves to be a reasonably good short-term solution, at least in terms of the number of deaths. Unfortunately, the life-table method and the modelling-death method do not adequately reflect the sociodemographic characteristics of people who die according to the death register. Weighting the known deaths as responses provides a good solution for both the number of deaths and the sociodemographic characteristics of those that die, though the standard errors for these estimates are a little higher than the other methods (as a smaller number of cases is being given a higher weight under this method than under the other methods). Method 4, with additional benchmarks to improve the timing of the deaths, is recommended over Methods 2 or 3.

Estimates for the live population from Wave 1 to 13 were unaffected by the method chosen to adjust for the unknown deaths, at least among the variables and the time frame considered in this article. It is plausible that over a longer time frame, or for a survey with a higher rate of unknown eligibility than the HILDA Survey, the estimates for the live population would begin to be affected (as was observed by [Sadig's \(2014\)](#) analysis of an 18-wave panel). This is because the Wave 1 population is split into several population segments by Wave t :

- i) the part of the population who die by Wave t ;
- ii) the part of the population who move overseas in at least one wave by Wave t ; and
- iii) the remaining portion of the Wave 1 population who continuously live in the population to Wave t .

Differences in the characteristics of one of these groups (as seen for those who die) would eventually result in differences in the other groups as the overall size of the Wave 1 population is fixed.

What is clear from this analysis is that assuming sample members with unknown eligibility are all eligible leads to inaccurate estimates of the total number of people who die, the timing of their death, and their sociodemographic characteristics. Similar findings

are expected for other longitudinal surveys and could be confirmed through replication of these analyses. The impact may be greater in another study if a smaller proportion of deaths is reported or if these deaths are more dissimilar to the unknown deaths than found in the HILDA Survey. This is a problem that only grows over time as more of the sample is withheld from fieldwork for various reasons. Some adjustment to address this issue, either implicitly or explicitly, should be made.

While this article considered deaths specifically, this same issue of unknown eligibility also applies to people who move out of scope in other ways. For the HILDA Survey, this applies to people who move abroad, but for other surveys this could include people who move into institutions. A national registry for these transitions out of scope is unlikely to be available for matching purposes, and the findings presented here may help guide decisions on alternative methods. Aggregate statistics on transitions may be available (such as from Census or migration statistics) which could be used in a similar way to life tables in Method 2. It would certainly be feasible to use the data internal to the survey to model moves abroad or to institutions (as in Method 3) or to weight known out of scopes as responses (as in Method 4). Indeed, for those surveys of specific subgroups in the population it may not be possible to use any external data sources at all, so any adjustment for unknown eligibility would be limited to methods that use internal data (such as Methods 3 and 4). Further work could include a similar assessment for other types of unknown eligibility as presented here for unknown deaths.

7. References

- Au, N. and D.W. Johnston. 2014. "Self-Assessed Health: What Does it Mean and What Does it Hide?" *Social Science and Medicine* 121: 21–28. Doi: <http://dx.doi.org/10.1016/j.socscimed.2014.10.007>.
- Australian Bureau of Statistics (ABS). 2014. *Life Tables, States, Territories and Australia, 2011-2013*. Canberra: ABS. Available at: <http://www.abs.gov.au> (accessed January 2015).
- Australian Institute of Health and Welfare. 2011. *Comparing an SLK-Based and a Name-Based Data Linkage Strategy: An Investigation into the PIAC Linkage*. Canberra: AIHW. Available at: <http://www.aihw.gov.au/publication-detail/?id=10737418279> (accessed November 2015).
- Gouskova, E., S. Heeringa, K. McGonagle, R. Schoeni, and F. Stafford. 2008. "Panel Study of Income Dynamics Revised Longitudinal Weights 1993-2005." PSID Technical Series Paper 08-05. Ann Arbor: University of Michigan. Available at: https://psidonline.isr.umich.edu/publications/Papers/tsp/2008-05_PSID_Revised_Longitudinal_Weights_1993-2005%20.pdf (accessed 30 June 2014).
- Green, F. 2011. "Unpacking the Misery Multiplier: How Employability Modifies the Impacts of Unemployment and Job Insecurity on Life Satisfaction and Mental Health." *Journal of Health Economics* 30: 265–276. Doi: <http://dx.doi.org/10.1016/j.jhealeco.2010.12.005>.
- Hawkes, D. and I. Plewis. 2006. "Modelling Non-Response in the National Child Development Study." *Journal of the Royal Statistical Society, Series A* 169: 479–491. Doi: <http://dx.doi.org/10.1111/j.1467-985X.2006.00401.x>.

- Hayward, M.D. Undated. "Using the Health and Retirement Survey to investigate health discrepancies." Health and Retirement Survey documentation. Ann Arbor: University of Michigan. Available at: http://hrsonline.isr.umich.edu/sitedocs/dmc/hrs_healthdisparities_hayward.pdf (accessed February 2015).
- Health and Retirement Survey (HRS). Undated. "Sampling Weights: Revised for Tracker 2.0 and Beyond." Health and Retirement Survey documentation. Ann Arbor: University of Michigan. Available at: <http://hrsonline.isr.umich.edu/sitedocs/wghtdoc.pdf> (accessed February 2015).
- Kalton, G. and M. Brick. 2000. "Weighting in Household Panel Surveys." In *Researching Social and Economic Change: the Uses of Household Panel Studies*, edited by D. Rose, 96–112. London: Routledge.
- Kroh, M. 2013. *Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984 until 2012)*. DIW Data Documentation 71. Berlin: DIW.
- LaRoche, S. 2007. *Longitudinal and Cross-Sectional Weighting of the Survey of Labour and Income Dynamics*. Income Research Paper Series, 75F0002MIE. Ottawa: Statistics Canada.
- Lynn, P. 2009. "Methods for Longitudinal Surveys." In *Methodology of Longitudinal Surveys*, edited by P. Lynn, 1–19. Chichester: Wiley.
- Rabe-Hesketh, S. and A. Skrondal. 2006. "Multilevel Modelling of Complex Survey Data." *Journal of the Royal Statistical Society, Series A* 169: 805–827. Doi: <http://dx.doi.org/10.1111/j.1467-985X.2006.00426.x>.
- Sadig, H. 2014. *Unknown Eligibility Whilst Weighting for Non-Response: The Puzzle of Who has Died and Who is Still Alive?* ISER Working Paper 2014-35. Colchester: Institute for Social and Economic Research, University of Essex.
- Scholes, S., R. Taylor, H. Cheshire, K. Cox, and C. Lessof. 2008. *Retirement, Health and Relationships of the Older Population in England: The 2004 English Longitudinal Study of Ageing*, NatCen Technical Report P2808. London: National Centre for Social Research.
- SHARE. 2013. *Release Guide 1.1.1 Wave 4*. Munich: Munich Centre for the Economics of Ageing. Available at: http://www.share-project.org/fileadmin/pdf_documentation/SHARE_wave_4_release_guide_1.1.1.pdf (accessed February 2015).
- Skinner, C.J. and D.J. Holmes. 2003. "Random Effects Models for Longitudinal Survey Data." In *Analysis of Survey Data*, edited by R. Chambers and C.J. Skinner, 205–219. Chichester: Wiley.
- Smith, T.W. 2009. *A Revised Review of Methods to Estimate the Status of Cases with Unknown Eligibility*. Report of the Standard Definitions Committee for the American Association for Public Opinion Research. Available at: https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/FindingE.pdf (accessed October 2016).
- Taylor, M.F., ed. 2010. *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex.
- Valliant, R., J.A. Dever, and F. Kreuter. 2013. *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer.
- Watson, N. 2012. *Longitudinal and Cross-Sectional Weighting Methodology for the HILDA Survey*. HILDA Technical Paper Series 2/12. Melbourne: Melbourne Institute of Applied Economic and Social Research, University of Melbourne.

- Watson, N. and M. Summerfield. 2014. *Outcomes from Matching the HILDA Survey Sample to the Death Register*, HILDA Technical Paper Series 2/14. Melbourne: Melbourne Institute of Applied Economic and Social Research, University of Melbourne.
- Weaver, B. 2011. *Combined 2009 and 2010 Technical Weighting Report for the Swiss Household Panel*. Lausanne: Swiss Foundation for Research in Social Sciences. Available at: <http://forscenter.ch/en/our-surveys/swiss-household-panel/documentationfaq-2/methods/weighting/> (accessed February 2015).

Received March 2015

Revised August 2016

Accepted August 2016