Survey Research Methods (2020)
Vol. 14, No. 1, pp. 1-18
doi:10.18148/srm/2020.v14i1.7422

ISSN 1864-3361
http://www.surveymethods.org

Measuring geographic mobility: Comparison of estimates from longitudinal and cross-sectional data

Nicole Watson
University of Melbourne, and
University of Queensland
Australia

Attrition of sample members from a longitudinal survey can undermine the quality of the data and its research potential, especially when the sample members who drop out are different from those who do not. People who move house are more likely to drop out of the survey as they are harder to locate, and once located, may be harder to interview in the remaining fieldwork time available. Moving coincides with many other life events (such as changes in marital status, the birth of a child, buying a home, changes in employment, or retirement) and if movers are not adequately interviewed, this may result in the study under-representing these changes and the events that occur after a move. This paper examines the weighted estimates of the rate of moving by age in a long running household panel study, the Household, Income and Labour Dynamics in Australia (HILDA) Survey, compared to official cross-sectional data sources and probabilistically linked Census data. Geographic mobility is examined over one-, five- and 10-year periods. Some of the differences that occur in the mobility estimates is a result of item non-response or recall error in the cross-sectional sources but little evidence is found of the differential impact of attrition in the HILDA Survey. There is, however, some indication that the longitudinal survey data may underrepresent long distance moves. Other differences between the data sources are investigated by fitting logistic regression models of mobility to estimate the effect of housing tenure and education levels over the life course. These models show similar overall trends, but there is some evidence of differential effects for renters with lower education levels which may be due, at least in part, to the differential role recall error plays in these measures. Overall, these findings reassure longitudinal data users of the quality of geographic mobility estimates from the HILDA Survey and encourage similar comparisons to be made for other longitudinal data sources.

Keywords: residential mobility, internal migration, HILDA Survey, recall error, item non-response, attrition

1 Introduction

Longitudinal data has become increasingly important in understanding social and economic change. Attrition of sample members from a longitudinal survey over time has the potential to undermine the quality of the data and therefore limit its research potential, especially when the sample members who drop out are different from those who do not. A contributing factor to attrition is the inability to locate sample members when they change address. Much effort can be spent on tracking sample members who move without providing forwarding contact details. Inevitably some movers will not be located. Further, this tracking effort can take a long time, leaving little time to contact and interview the sample member once they have been located before the fieldwork ends. In a cross-national study of attrition in household panel studies in Australia, Britain and Germany, Watson and Wooden (2014) found that, all else being equal, the probability of re-interviewing sample members following a move falls 2 to 12 percentage points compared to when they do not move. Over time, this pool of untraceable sample members grows along with the number of people who are tracked but not interviewed, leading to concerns that, apart from the reduction in sample size, the remaining sample is different from (i.e., unrepresentative of) the underlying population from which the sample was initially selected.

There are many factors that can influence the decision to move house, including family, education, employment, housing and health factors. Short-distance moves (often referred to as residential mobility in the literature) are typically related to family and housing factors whereas long-distance

Contact information: Nicole Watson, Melbourne Institute of Applied Economic and Social Research, University of Melbourne VIC 3010, Australia (email: n.watson@unimelb.edu.au).

moves (often referred to as migration or internal migration in the literature) tend to be related to education and employment reasons (Buck, 2000; Geist & McManus, 2008). People who are in their twenties and thirties move much more than other age groups and there is a gradual decline in geographic mobility in the older age groups. Changes in marital status and birth of a child are common life course triggers for a move (Clark & Huang, 2003; Spallek, Haynes, & Jones, 2014). Dissatisfaction with the present dwelling (occupancy to bedroom ratio, cost, neighbourhood, etc.) or a desire to transition from renting to owning may result in the occupants looking for alternative housing (Baker, Bentley, Lester, & Beer, 2016; Clark & Onaka, 1983). Changes in employment, retirement and onset of disability can trigger a decision to move house (Bell & Ward, 2000). In short, moves coincide with many other life events and if the sample members of a longitudinal study who move are not adequately followed and interviewed, this may result in the study underrepresenting these changes and the events that occur after a move.

Not only is there evidence that mobility is associated with change surrounding the move, there is also evidence that it is associated with higher post-move change. In the two years following a short or long distance move, there is an increased rate of income instability, employment changes and family changes (Geist & McManus, 2008). Again, not adequately following movers in a longitudinal study would result in these changes going unobserved and thus introduce biases into estimates.

The extent of geographic mobility varies substantially across countries, though Australia tends to have high rates of mobility that is somewhat similar to the US, Canada and New Zealand (Greenwood, 1997). The proportion of the Australian population moving between 2000 and 2001 (as measured from Census data) was 18% (Australian Bureau of Statistics, 2009) though this annual mobility rate has declined to 15% and held steady for the subsequent three Censuses undertaken in 2006, 2011 and 2016.[1] The proportion of the US population that moved each year in the 1980s was around 18% but this has steadily fallen to 11% by 2016, which is the lowest it has been since 1948 (Greenwood, 1997; Ihrke, 2017). The annual mobility rates in 1997 for European countries tend to be much lower, with rates in the UK, the Netherlands, France and Denmark being between 8% and 11% and Austria, Greece, Ireland, Italy, Portugal and Spain being less than 5% (Gregg, Machin, & Manning, 2004). Research by Caldera Sánchez and Andrews (2011) into housing market factors that affect the rate of mobility indicates that mobility rates are higher in countries with lower property transaction costs, more responsive housing supply, lower rent controls and lower tenant protection. Australia meets all of these criteria with the exception of having moderate rather than low transaction costs in buying and selling houses (Caldera Sánchez & Andrews, 2011). In terms of the rate of home ownership in Australia, around 70% of all households are owner occupied which quite similar to the UK and the US (OECD, 2016). Also in large countries like Australia, the proportion of moves linked to job-related reasons is much higher than in smaller countries where a new job further away may mean a longer commute rather than a move (Caldera Sánchez & Andrews, 2011). An example of this is the mining boom (starting in the early 2000s and led by growth in demand in Asian markets) which drew people to the resource-rich states. Some workers moved directly to the mining areas while others moved to major population centres and began long distance commuting as fly-in, fly-out workers (Productivity Commission, 2014).

That said, it is easier to track some moves compared to others in a longitudinal survey. Some moves are planned well ahead of time (such as those associated with marriage, studying, and retiring) and respondents can provide details of their planned move in the interview prior to their move. Other moves are acted on very quickly (such as a separation from a marriage or de facto relationship or eviction from a rental property) and the fieldwork company then needs to rely on the contact details the respondent has provided (mobile and work telephone numbers, email address, contact details of friends or family members who might know where they are if they move). The degree of family or community connectedness may not only affect the likelihood of a move, but also the ability to locate an individual once they move. It may also be easier to find short-distance moves rather than long-distance moves as the workplace and friend network remain largely unchanged. In Australia, for example, in 2006 37% of the individuals who moved from their address 12 months ago were still within the same local area (i.e. Statistical Local Area), 41% were within the same regional area (i.e., Statistical Division) but different local area, 12% were within the same state but not the same region, and 10% were interstate (Australian Bureau of Statistics, 2009). Similarly, people who move multiple times between waves may also be harder to track as they likely become less connected to the people associated with their last known location. To track movers in a longitudinal survey, a range of proactive and reactive methods are adopted and are operationalised at the interviewer or office level. For a summary of tracking methods used in longitudinal surveys, see Couper and Ofstedal (2009), and Laurie, Smith, and Scott (1999).

Further, when a respondent reports information about a move, the answer may be subject to recall error. Few people can recall the exact date of an event and therefore they will instead use various heuristics to obtain a likely answer. The most common method to date an event refers to the connected stream of events that occur over time, followed by

[1] Obtained via ABS Census TableBuilder using data from 2006, 2011 and 2016 Census of Population and Housing.

use of landmark events (i.e. significant events specific to the respondent's life such as a wedding, birth of a child, a death of a friend or family member, holidays or a birthday) and guessing (Tourangeau, 2000). When a respondent can narrow down the event to occurring within a certain window, they are more likely to report that it occurred in the middle of the window rather than at the beginning or end. It has also been demonstrated that the longer the recall period, the lower the accuracy of the data reported (Menard, 2002). As the recall period gets longer, respondents rely less on recall and more on inference to determine an answer (Tourangeau, 2000). This occurs for two reasons: it is harder to retrieve information about events that occurred in the more distant past and the reference boundary is less distinct. When the respondent is uncertain about the boundary, they may mistakenly report events that actually occurred before the boundary or omit events which actually occurred after the boundary (this is known as forward telescoping and backward telescoping). Pleasant events are recalled with more accuracy than unpleasant events, but both of these types of events are recalled more accurately than neutral events (Skowronski, Betz, Thompson, & Shannon, 1991; Thompson, 1985). Further, well-rehearsed events (those thought about or talked about) are remembered better than those that are not rehearsed (Thompson, 1985). Research specific to recall issues regarding residential moves found that compared to other less significant life events, recall errors are less likely for residential moves (Hall, 2015) and that the more salient the move – such as an interstate move or a move associated with marriage, birth of a child or change in employment – the better the move will be recalled (Smith & Thomas, 2003).

This paper compares the rate of moving by age in the Household, Income and Labour Dynamics in Australia (HILDA) Survey, a long running household panel study, with estimates from official cross-sectional data sources, the Australian Census of Population and Housing and the General Social Survey, along with probabilistically linked Census data. This will shed light on whether movers are underrepresented in the HILDA Survey sample as a result of attrition and the extent to which the weights adequately correct for this. Geographic mobility is examined over one-, five- and 10-year periods. In addition, models of mobility which estimate the effect of housing tenure and education levels over the life course are compared across the different data sources.

2 Data sources

2.1 Summary of the four data sources

A summary of the design features of these four different sources of geographic mobility estimates is provided in Table 1. Each of these data sources are described in turn.

The HILDA Survey is a nationwide household panel survey which began in 2001 (DSS and MI, 2018a, 2018b; Summerfield et al., 2017). [2] Interviews are conducted annually and the core content asked every year includes questions on employment, income and families. Rotating modular content includes questions on wealth, retirement, fertility, education and health. Interviews are conducted with all household members aged 15 and older, and as such, the HILDA Survey, like other household panel studies, are very useful for examining the dynamics of mobility decisions within households and exploring 'linked lives' along this dimension over time (Coulter, van Ham, & Findlay, 2016). The HILDA Survey sample has a multi-stage, stratified and clustered design with 7682 responding households in the initial wave from a total of 11,693 in-scope households, resulting in a household response rate of 66%. The sample is restricted to usual residents in private dwellings and excludes people living in very remote parts of Australia. The wave 1 responding households contain 19,914 individuals, which encompass 13,969 respondents, 1158 non-responding adults and 4787 children aged under 15. Individuals from these responding households are followed over time and, if aged 15 years and older, interviewed. The sample is extended in subsequent waves to include other members of their household that were not part of the initial household. The vast majority of interviews (over 90%) are conducted face-to-face with the remainder conducted by telephone. No proxy interviews are permitted. In wave 6 (2006) the proportion of wave 1 respondents that were re-interviewed, excluding people who had died or moved abroad, is 77%. This re-interview rate in 2011 and 2016 is 70% and 65% respectively. There are two measures of whether a person moves in the HILDA Survey. The first measure is obtained by comparing the address from the previous wave to the current wave address.[3] For most waves, the proportion of interviews conducted within 30 days of the anniversary of previous wave interview is between 77% to 83%, however in waves 2, 11 and 12 this rate was between 67% to 69%. The second measure is determined by the respondent's answer to the questions "Since we last interviewed you on *[date of last interview]*, have you changed your address?", and if they have moved "In what month (and year) did you move to your current address?". If the timing of the interviews in subsequent waves is less than 12 months, then the

[2] This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Survey 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 author and should not be attributed to the Australian Government, DSS or the Melbourne Institute. The data are available through Dataverse at the Australian Data Archive (dataverse.ada.edu.au).

[3] Address changes that are purely corrections to the address are identified by the interviewer and are not counted as a move.

equivalent question in the prior wave is used to determine if the move to the current location occurred in the last 12 months. To obtain the five- and 10-year measures of mobility, these two approaches were used to look back over the answers provided in prior waves.

The Australian Census of Population and Housing (referred to subsequently as the "Census") occurs every five years by the Australian Bureau of Statistics (ABS). The HILDA estimates are compared to estimates from the 2006 and 2011 Censuses. The Census is conducted on the second Tuesday in August (called "Census night") and one or more persons in the household fills in the Census form on behalf of the other members of the household. An online form has been progressively introduced since 2006, with 11% of the households completing the Census online in 2006, increasing to 34% in 2011 (Australian Bureau of Statistics, 2017). The remaining households completed the Census on paper forms. The net undercount for the 2006 and 2011 Censuses is 2.7% and 1.7% respectively, and the Census counts have been adjusted accordingly by the Australian Bureau of Statistics (Australian Bureau of Statistics, 2018). The Census question from which the mobility estimates are derived from asks "Where did the person usually live one year ago (at *[specific date exactly one year ago from Census night]*)". A similar question is asked of where the person was living five years ago. The mobility estimates are extracted via TableBuilder, an online interactive interface that the ABS use to enable registered users to specify customised aggregated tables.[4]

The ABS has also produced the Australian Census Longitudinal Dataset (ACLD) which is a 5% random sample of the 2006 Census that has been linked to the 2011 Census data without name and address information (Australian Bureau of Statistics, 2013). A combination of deterministic and probabilistic linkages were used, with 44% of the dataset linked through a deterministic link that included age, sex, day and month of birth, and a very small geographic area (meshblock) of the 2006 address (in 2011 this was the address reported for where the individual was living five years ago). The remainder of the linkages were made on a probabilistic basis through 10 different passes of the data picking up potential matches on a range of other characteristics. Weights are created to adjust for the sample design, Census undercoverage, and missed links with the final weights calibrated to population benchmarks derived from the 2011 Census for the population in scope of both the 2006 and 2011 Censuses (Chipperfield, Brown, & Watson, 2017). Estimates of geographic mobility over a 10-year period are obtained from the ACLD (via TableBuilder) using the Census question about where the person was living five years ago as described earlier but now measured at two time points (2006 and 2011).

The General Social Survey (GSS) is a repeated cross-sectional survey conducted by the ABS that collects data on a range of social dimensions similar to those included in the HILDA Survey. The GSS sample is representative of people living in private dwellings excluding those living in very remote areas of Australia. Smaller states and low socio-economic areas were oversampled. The mobility estimates are drawn from the 2014 GSS which achieved a response rate of 80% from in-scope dwellings (Australian Bureau of Statistics, 2015). One person aged 15 or over was randomly selected in each dwelling to be interviewed, resulting in 12,932 interviews. People aged 15 to 24 are asked about moving out of their parental home, whereas people aged 25 and older are asked "How long have you lived in this dwelling?" and if the response is less than five years then they are asked "How many times have you moved in the last 5 years?". The mobility estimates (and associated standard errors) were extracted via TableBuilder. As the data available is limited to an indicator of whether a person reported "less than 1 year" or "one year or more" for the first question and whether or not a person was asked the second question, these two indicators were used to estimate mobility in the last 12 months and in the last 5 years for people aged 25 and older.

2.2 Sources of error

The potential sources of error are different for the four different data sources. In terms of non-response, the HILDA Survey has non-response in the initial wave (unit non-response), as occurs with cross-sectional surveys like the GSS, and then is subject to the impacts of attrition (wave non-response) that occurs in subsequent waves. Further, the amount of non-response in the initial wave of the HILDA Survey is higher than that for the GSS. The Census is also subject to non-response and the Census Post Enumeration Survey quantifies the level of net undercount which is usually very low. These sources also have item non-response where a sample member was interviewed or enumerated but a response to the particular question was not provided either because the respondent did not know the answer or refused to provide it. The level of item non-response is very low for the interviewer administered questionnaires (less than 1% in the HILDA Survey for the question on changing addresses since the last interview and 1.7% in the GSS) but is much higher in the household-completed questionnaire used in the Census (5.2% to 8.5%).

A particular source of error that is unique to the ACLD among these four data sources is errors that come from the data linkage process. These include missed links where two records should have been matched but were not and incorrect links where two records from different people have been linked when they should not have been. The ABS estimates that between 5 and 10% of records linked between the 2006 and 2011 Census were incorrect links (Australian Bureau of

[4]TableBuilder is available on the ABS website (www.abs.gov.au).

Table 1
Design features of different sources of geographic mobility

	HILDA Survey	Census 2011	ACLD 2011	GSS 2014
Type	Longitudinal survey	Census	Probabilistically linked Census file	Cross-sectional survey
Mode of data collection	Interviewer administered. 90% face-to-face, 10% telephone	Household-complete (hardcopy or online)	Household-complete (hardcopy or online)	Interviewer administered. Face-to-face
Who is interviewed	Each person aged 15+ in the household	One or more persons answers the questions for each person in the household so have some proxy reports	One or more persons answers the questions for each person in the household so have some proxy reports (could be different people answering at each Census)	Randomly selected person aged 15+ from the household
Move flag	Moved within last 1 year, 5 years, or 10 years	Moved within last 1 year, 5 years	Moved within last 10 years	Moved within last 1 year, 5 years
How move status is determined	Flag 1: Comparison of previous and current address each wave, Flag 2: Respondent report of whether changed address since last interview	Asks address of where living 1 year ago and 5 years ago	Asks address of where living 1 year ago and 5 years ago. Combining question across two Censuses gives move in last 10 years	Asks how long lived in dwelling, and how many times moved in the last 5 years[a]
Population to which estimates relate[b]	Residents of private dwellings in Australia in 2001, and living in Australia until 2011 (or 2014 if comparing to GSS)	All residents in Australia in 2011. Those who arrived in 2002 or later have been excluded for the purposes of this comparison.	All residents in Australia in 2001, 2006 and 2011	All residents of private dwellings in Australia in 2014. Those who arrived in last 10 years have been excluded for the purposes of this comparison.
Primary sources of error	Initial wave non-response; Attrition; Flag 1: Error due to inexact timing of interviews. Move out and back to same address as last wave would be recorded as no move; Flag 2: Recall error 1 year	Census undercount; Proxy reporting error; Move out and back to same address as 1 year ago (or 5 years ago) would be recorded as no move; Recall error 1 year and 5 years	Census undercount; Linkage error (missing links, false links); Proxy reporting error; Move out and back to same address as 1 year ago (or 5 years ago) would be recorded as no move at each Census; Recall error 1 year and 5 years	Non-response; Duration of two or more spells at the same address may be added together by the respondent when determining answer; Recall error 1 year and 5 years

[a] Data on length of time resident in dwelling is not available in Table Builder (other than to indicate whether the respondent reported less than 1 year or one year or more). As a result, whether the subsequent question on the number of times moved in the last five years was asked or not was used to indicate whether a person had moved in the last five years.
[b] Different minimum age cut-offs are used depending on the comparison being made.

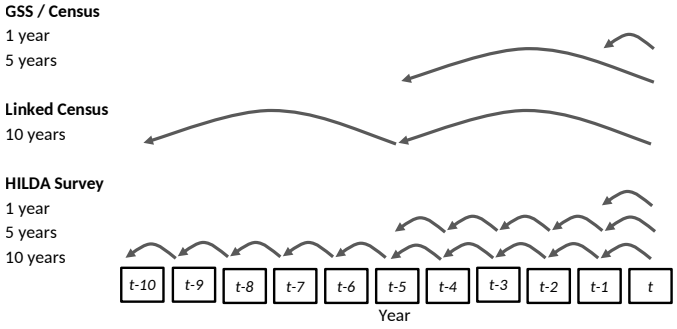


Figure 1. Comparison of recall tasks for respondents
Note: GSS=General Social Survey; HILDA Survey = Household, Income and Labour Dynamics in Australia Survey.

Statistics, 2013). A portion of these incorrect links could have been avoided by linking the whole of the 2006 Census with the 2011 Census rather than a sub-sample but this is a very resource intensive exercise so was not viable at the time. Alternatively, a lower tolerance for incorrect links could have been applied though this would result in a greater proportion of missing links.

There are some small differences in the populations to which the different estimates relate, even after they have been made as similar as possible given the data available. The HILDA Survey sample used to create the mobility estimates is restricted to people living in Australia since 2001. To be consistent with this, the Census estimates are restricted to exclude people arriving in Australia in 2002 or later. The GSS estimates are restricted to exclude people arriving in the last 10 years (individual year of arrival is not available, so it is not possible to restrict the sample to exclude arrivals in 2002 and 2003). Note also that the HILDA Survey sample in 2001 was restricted to people living in private dwellings, but people are followed into both private and non-private dwellings when they subsequently move. The Census includes people in both dwelling types, so the difference here is likely to be small. The GSS is restricted to people living in private dwellings in 2014. It is not expected that this difference would impact the estimates much as the number of people living in non-private dwellings is very small.[5] Further, where there are differences in the mobility estimates, the estimates that include people living in non-private dwellings (such as student accommodation, hostels, prisons, staff quarters and nursing homes) will be higher than those that are restricted to people in private dwellings as non-private dwelling residents are generally more mobile than the rest of the population. This will likely affect the estimates for the elderly more so than younger age groups due to a greater concentration of the elderly living in nursing homes.

With regards to the cognitive task asked of the respondents, the sources again differ. Figure 1 shows the time period over which the respondents are asked to recall the event of moving. In the Census and GSS, respondents are asked to recall over a one- and five-year period. In the HILDA Survey, the recall task is restricted to the time since the last interview a year ago. Five of these responses are combined together to determine if someone moved within the last five years. Ten of these responses are used to determine if there is a move in the last 10 years. As noted earlier, the accuracy of respondent's answers are expected to decline the longer the reference period involved. Nevertheless, a different type of problem can occur in panel surveys when respondent histories are stitched together across multiple interviews leading to a disproportionate amount of change occurring at the boundary (or 'seam') between two waves. These seam effects are a result of combining data that are subject to measurement and processing errors such as omission, misplacing events in time, misclassifying or re-defining past information, and keying or coding errors (Lynn, Buck, Burton, Jäckle, & Laurie, 2005). It is expected that as the focus of this analysis is on whether a move occurred or not (rather than how many moves occurred) in a specific timeframe, the impact of seam effects for the HILDA measure should be very small, particularly where the timeframe is large (i.e., 5 years or 10 years).

Further, and aside from the timeframe involved, the actual information that respondents have been asked to recall is different: the HILDA Survey asks whether they have changed their address since a specific date; the GSS asks how long the respondent has lived at the current address; and the Census asked if the person was living at the same address one (or five) year(s) ago. Respondents may use different heuristics to obtain a response to these different questions, possibly resulting in different answers. In addition, there may be some moves that are not counted depending on the question asked. And where a person moves away from a certain address and later moves back (for example, leaving and then returning to the parental home), these moves may be captured in some of

[5]In the 2011 Census, for example, 2.5 percent of the Australian population were enumerated in non-private dwellings on Census Night and probably only half of these are likely to be permanent residents (based on the type of dwelling reported).

these questions but not in others. The Census question (related to 1 year or 5 years ago) and the first HILDA measure (related to 1 year ago) compare addresses recorded for two points in time, so a move out after the first point in time and a return before the second point in time would not be recorded as a move. Over 5 years, the first HILDA measure compares addresses recorded for five points in time, so is less susceptible to this problem. The GSS question asks for how long someone has lived at an address and some respondents may conceptualise this as how long (in total) they have lived at a particular address and may add all of their spells of living at that address together and provide their answer based on that. The question that appears least susceptible to this problem is the question asked of the HILDA respondents about whether they have changed their address since the last interview. Regarding the two HILDA measures, there may be some differences between them simply due to timing. The first HILDA measure compares addresses of the household at two consecutive household interviews whereas the second measures asks the respondent to compare address they were at last wave on the date of their interview with their current address. The individual interviews are done on the same day as the household interview or afterwards. Over a five or ten year period these differences should have minimal impact.

3 Methods

Estimates of the rate of moving by 5-year age groups are generated from the HILDA Survey for 2011 for the proportion of people who have moved within the last year, last 5 years and last 10 years. Longitudinal weights are applied and standard errors are calculated using Jackknife estimation with 45 replicates. These longitudinal weights adjust for initial wave non-response and attrition from the panel in waves 2 to 11 (i.e. 2002 to 2011).[6] These estimates are used to compare to those from the 1-year and 5-year mobility rates from the 2011 Census and to the 10-year mobility estimates from the 2006-2011 ACLD. To compare to the 2014 GSS estimates of mobility over a 1-year and 5-year period, the HILDA estimates relate to 2014 and apply the longitudinal weights that adjust for initial wave non-response and attrition in 2002 to 2014.

As there are two mobility measures in the HILDA Survey – reported by the respondent since last interview (reported measure) and observed by comparing addresses at subsequent time points (observed measure) – these can be compared to understand what types of people under-report or over-report a move with the intention of choosing one of these measures to compare to ABS data. Focusing on the proportion of people moving in the last five years, the reported measure, weighted by the longitudinal responding person weight (applicable to people interviewed each wave so the reported measure is obtained),[7] is compared to the observed measure weighted in two ways: i) by applying the longitudinal responding person weight, and ii) by applying the longitudinal enumerated person weight (applicable to people who are part of a responding household where the observed measure can be ascertained). To understand the person characteristics associated with the differences between these two measures a new variable representing the discrepancies in the measures is derived with three categories: i) no difference between reported and observed measures for 5-year mobility, ii) respondent did not report a move, but a move was observed, iii) respondent reported a move, but a move was not observed. A multinomial logit model was fitted to this "error" variable with selected explanatory variables. The sample is restricted to original sample members (i.e., those that are part of wave 1 responding households) and observations are included from three time points: 2006, 2011 and 2016. The standard errors allow for the correlation of errors using the clustered sandwich estimator as the observations on the same individuals over time are not independent. Characteristics that may be associated with cognition and recall errors (for a discussion, see Dex (1995), Smith and Thomas (2003) are included in the model, these being: age (defined here in 10-year age bands for parsimony of the model), sex, whether respondent has a bachelor degree (or higher), and broad country of birth (i.e., Australia, main English speaking country, or not main English speaking country). Also included are several control variables. The first two controls identify situations where it may be unclear for respondents who spent time at multiple residences which is their main or primary residence, such as fly-in fly-out workers and people with a holiday house.[8] The first of these controls is an indicator variable for people who work in the mining industry as it is common for employers to fly in workers to the mining site to work a number days and then fly them back to their hometown to rest. The second control is an indicator of whether the respondent belongs to a household that owns a holiday house. The third control is the average rate of mobility for the 5-year age group to which the respondent be-

[6] For details of how this weight is constructed, see Watson (2012).

[7] The longitudinal responding person weight provided in the HILDA datasets are for persons aged 15 and older in wave 1 (2001) who have been interviewed in subsequent waves. To obtain estimates for persons aged 15-24 in 2011, the longitudinal responding person weight has been amended to include the longitudinal responding person weights for subsets of the sample that age into the responding sample. That is, for people who are aged 24 in 2011, the longitudinal responding person weight used is for wave 2 to 11 (i.e. they would have been first eligible to be interviewed in wave 2), and for those aged 23 in 2011 the weight used is for wave 3 to 11, and so on.

[8] While it would also have been ideal to include indicators for people who are seasonal workers, those living apart together, or others who potentially have multiple residences, these situations could not be identified sufficiently in the HILDA Survey data.

longs as errors may be more prevalent for people who move more often. The fourth control provides for a quadratic effect of the difference between the 5-year anniversary date of the respondent's interview at the start of the window and the interview date at the end of the 5-year window. This measure aims to pick up differences in the two measures that may be caused by different timing of the interviews. Finally, wave dummies are included to allow for any differences in recall error that may occur over time, for example, as the respondent becomes more accustomed to being interviewed. Based on this analysis (which is reported in the next section), the observed mobility measure (rather than the reported mobility measure) from the HILDA Survey is used for comparing to the ABS estimates.

The ABS estimates for geographic mobility are obtained via TableBuilder for the 2011 Census, the linked 2006-2011 ACLD and the 2014 GSS. Standard errors are extracted for the GSS estimates. While standard errors are relevant to the ACLD as it is a 5% sample from Census records, they are not available in TableBuilder. Nevertheless, they will be very minimal given the particularly large sample involved.

After assessing the age profile of movers, two other comparisons are made using summary statistics. The first compares the age-sex profile of movers. The second compares the type of move with respect to how far away the move takes the individual from their initial location. As the ABS data available on this measure is limited and seems to only be available in an ABS publication using the 2006 Census (Australian Bureau of Statistics, 2009, Table 6.1), the comparisons are made for the 5-year mobility rates with the 2006 HILDA estimates. Further, rather than measuring distance, the type of move is restricted to four categories: i) the same Statistical Local Area (SLA), ii) within the same Statistical Division (SD) but not the same SLA, iii) within the same state but not the same SD, and iv) interstate. SLAs are based on the administrative areas of local governments and the SDs are a collection of SLAs that define a broad city (such as Melbourne, Sydney, etc.) or in rural areas include areas surrounding one or more major towns with well-defined social and economic links (Australian Bureau of Statistics, 1999).

Finally, the predicted probability of moving based on HILDA and ABS data is estimated to determine if the relationships between variables are comparable. The data from the ABS are obtained as population weighted aggregated cross-classified tables so the frequencies are rescaled to sum to the effective sample size of the relevant sample.[9] These frequencies are then used as frequency weights when fitting logistic regression models to the binary variable indicating whether the individual has moved or not. The number of covariates needs to be kept small to avoid the impact on the overall integrity of the model of the small random adjustments the ABS makes (for confidentiality reasons) to tables with cells containing very small counts.[10] Three covariates included: i) 5 year age group, iii) indicator of whether the household own their own home; and ii) indicator of having a bachelor degree or higher. Other covariates were considered but discarded as they did not have as much of an impact on predicting the probability of moving as the covariates retained (the other covariates included sex, marital status, labour force status, and broad geographic region). The fitted model takes the form

$$\begin{aligned}\Pr(\text{move} = 1) = F(&b_0 + b_1 I(H)\, I(\text{own})\, \text{bs1} + b_2 I(H)\, \text{bs1} + \\ &+ b_3 I(G)\, I(\text{own})\, \text{bs1} + b_4 I(G)\, \text{bs1} + \ldots + b_{13} I(H)\, I(\text{own})\, \text{bs4} \\ &+ b_{14} I(H)\, \text{bs4} + b_{15} I(G)\, I(\text{own})\, \text{bs4} + b_{16} I(G)\, \text{bs4} + \\ &+ b_{17} I(H) \text{bs5} + b_{18} I(G)\, \text{bs5} + b_{19} I(H)\, I(\text{bach}) + \\ &b_{20} I(G)\, I(\text{bach}) + + b_{21} I(G)) \quad ,\end{aligned}$$

where $F(z) = e^z/(1 + e^z)$ is the cumulative logistic distribution, $I(\text{ds})$ is an indicator if the data source is the HILDA Survey (denoted by H) or the ABS source (in this case, the GSS, denoted by G), $I(\text{own})$ is an indicator of whether the person lives in a dwelling owned by one (or more) of the household members, $I(\text{bach})$ is an indicator of whether the person has a bachelor degree or higher, and bs1 to bs5 are the B-splines created for age based on knots at ages 25–29, 45–49 and 85 and older (using the `bspline` Stata command (Newson, 2012). This particular combination of knots for the splines were chosen as best compromise between high Receiver Operator Characteristic (ROC), low Akaike information criterion (AIC), low Bayesian information criterion (BIC) for the GSS data and avoids overfitting the mobility curve. For the comparison between the HILDA Survey and the ACLD, an additional knot was added at age group 15–19 to allow for the larger number of age groups for which data is available. This model allows separate curves across the age distribution to be fitted for each dataset and for home ownership status and then provides a shift effect for people with a bachelor degree (or higher).

4 Results

4.1 Differences in observed and reported mobility in the HILDA Survey

Focusing first on the differences in the observed and reported mobility estimates in the HILDA Survey, Figure 2

[9] The effective sample sized used for the GSS is 10,000 (this is an approximation as there is no information available on the relative sample sizes by state, urban/rural areas, and low/high income families. For the ACLD it is simply the size of the sample (1.44m) as the ACLD is a simple random sample. For the HILDA Survey the effective sample size is 6490.

[10] The ABS advise against aggregating more finely classified data where possible due to the impact the small random adjustment of small cells can have on the higher level aggregates (Australian Bureau of Statistics, 2011).

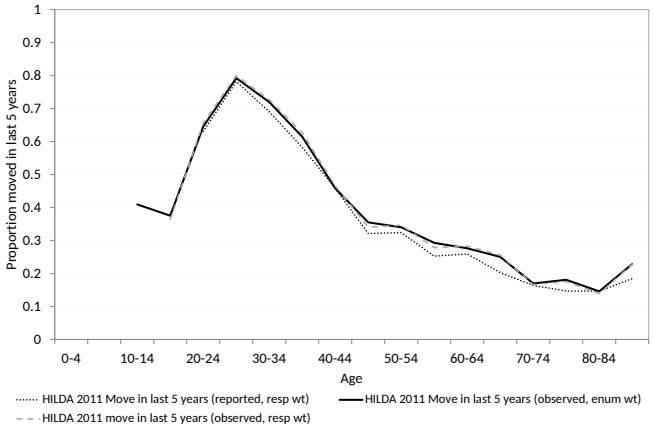


Figure 2. Proportion moving in the last 5 years (reported vs observed), HILDA 2011
Note: HILDA = Household, Income and Labour Dynamics in Australia Survey; resp wt = longitudinal responding person weight; enum wt = longitudinal enumerated person weight.

Table 2
Multinomial logit of differences in reported and observed 5-year mobility, waves 6, 11 and 16

	Did not report move, but address changed		Report move, but did not change address	
Characteristics in wave t	Coeff.	S.E.	Coeff.	S.E.
Age (base=19-29)				
30-39	0.378***	0.128	0.153	0.223
40-49	0.581**	0.263	0.115	0.473
50-59	0.744**	0.333	−0.069	0.589
60-69	0.920**	0.371	0.070	0.655
70-79	0.876**	0.422	−0.216	0.755
80+	0.768*	0.451	0.072	0.778
Female	−0.037	0.070	−0.283**	0.126
Has bachelor degree (or higher)	−0.236***	0.087	−0.066	0.151
Country of birth (base=Australia)				
Main English Speaking	−0.056	0.120	0.499*	0.272
Not Main English Speaking	−0.017	0.113	0.107	0.187
Worked in mining industry in last 5 years	−0.138	0.257	0.396	0.363
Has holiday house	0.307**	0.126	0.293	0.230
Mean mobility rate for 5-year age group	1.882**	0.756	0.583	1.310
Distance from 5-year anniversary date/10^2	0.134*	0.001	−0.880***	0.002
Distance from 5-year anniversary date squared/10^4	0.230***	0.000	0.290*	0.000
Wave (base=Wave 6)				
Wave 11	0.058	0.082	−0.222	0.162
Wave 16	−0.367***	0.084	−0.334**	0.156
Constant	−4.683***	0.547	−4.912***	0.971

Significance: *** p<0.01, ** p<0.05, * p<0.1

compares these estimates for the proportion of people in 2011 who have moved in the last 5 years. The black line shows the proportion based on the observed measure (weighted by the longitudinal enumerated person weight). The black dotted line shows the proportion moving based on the respondent reported measure (weighted by the longitudinal responding person weight). To demonstrate that the differences in these two estimates is not caused by differences in the weights, the grey dashed line shows the proportion based on the observed measure but weighted by the longitudinal responding person weight. The choice of weight changes the estimates very little. However, the estimates based on the reported mobility measure are typically below the observed mobility rates, suggesting recall errors are affecting the reported measure in a way that moves are being omitted or pushed further back in time outside the 5-year window (backward telescoping), particularly for those aged 45 and older which may be related to declining cognitive functioning. The unweighted HILDA estimates (not shown) are on average 2 percentage points lower than the weighted estimates and up to 5 percentage point lower on some estimates. This correction is due to the response propensity models used in the construction of the weights which include covariates related to the propensity to move (such as age, sex, education, employment, marital status and housing tenure) as well as the number of times moved in the 10 years prior to initial interview, the likelihood of moving, and observed moves (Watson, 2012).

Considering now the three separate 5-year windows available in the HILDA data (i.e., ending in 2006, 2011 and 2016), the rate of consistency between the observed and reported mobility indicators is actually very high (96%). However, there are 3% of respondents who do not report a move but a move was observed based on address changes, and 1% of respondents who reported a move but a move was not observed. Fitting a multinomial regression model predicting the type of error made results in the coefficients and standard errors shown in Table 2. The characteristics of the respondents who make these two types of errors are quite different. Respondents who do not report a move but a move was observed based on address change are more likely to be older and not have a bachelor degree (or higher). Sex and country of birth have no effect on this particular type of error. People who had a holiday house were more likely to have a change in address but not report a move than those who did not have a holiday house, suggesting that residency may be a more fluid concept for these people. People working in the mining industry were no more or less likely to have differences in their reported and observed mobility. In addition, and as expected, the further the interview date at the end of the 5-year window is from the 5-year anniversary of the initial interview at the start of the window, the greater the chance of this mismatch occurring. Interestingly, the rate of this type of error reduces in wave 16 (2016) compared to the earlier two waves. This may be due to the respondent getting better at answering these questions (a form of panel conditioning). It may also be due to the introduction of computer-assisted personal interviewing which started in wave 9 (documented by Watson, 2010) as the date of last interview is dynamically inserted into the question presented to the interviewer to ask of the respondent. Prior to 2009, the interview was conducted with a paper questionnaire and the date of last interview was written on the front cover of the questionnaire by the interviewer. The interviewers were instructed to incorporate this date into the questions as required during the interview. Administration of the interview via a paper questionnaire could have resulted in higher error rates if the interviewer was not always diligent in copying the date onto the questionnaire, using the precise date of the last interview (and instead used the month or year) or using only the phrase "when we last interviewed you". Turning now to the alternative type of error, where the respondent reported a move but no change in address was observed, this type of error is more likely to occur for men, but age, education level, and country of birth do not have an effect (at least not at the 5 percent significance level). As expected, this type of error is more likely for those interviewed in wave 6 and the error rate reduces further in wave 16 (consistent with both panel conditioning and the move to computer-assisted personal interviewing with the greater accuracy this would afford).

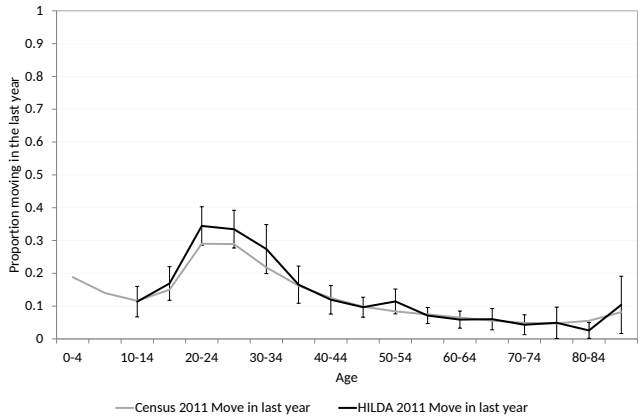
As these results from the multinomial logistic model for the first type of error (which is the largest of the two error types) are consistent with recall errors associated with cognitive functioning, the HILDA estimates used in the comparisons to ABS sources are those derived from the observed mobility measure (rather than the reported mobility measure).

4.2 Age profile of movers

Turning now to the comparison of HILDA and ABS mobility rates, Figure 3 shows the comparison of the one-year mobility rates from the HILDA Survey to the 2011 Census (in the left panel) and the 2014 GSS (in the right panel). While the Census mobility rate is lower than the HILDA mobility rate for people aged 20–34, they are not statistically significant. The only significant difference evident between the HILDA and the Census for the one-year mobility rate is for those aged 80–84 where the HILDA rate is lower than the Census rate. Similarly, the 2014 GSS and HILDA one-year mobility rates are very similar across the age distribution. The only significant difference between the two sets of estimates is for the 65–69 year olds where the HILDA mobility rate is higher than that from the GSS.

Greater differences between the data sources are evident when considering the five-year mobility rate. The left panel of Figure 4 provides the five-year mobility rates for 2011

(a) HILDA and Census 2011



(b) HILDA and GSS 2014

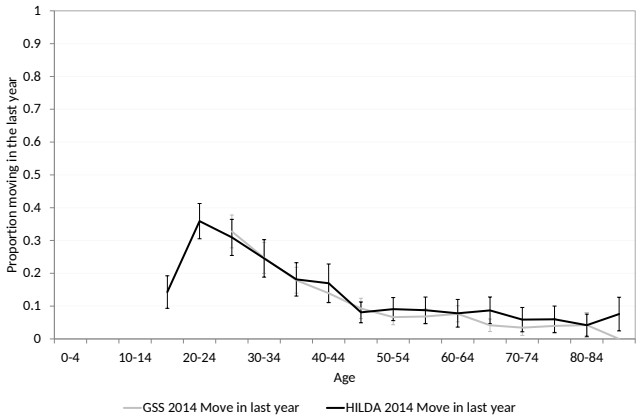
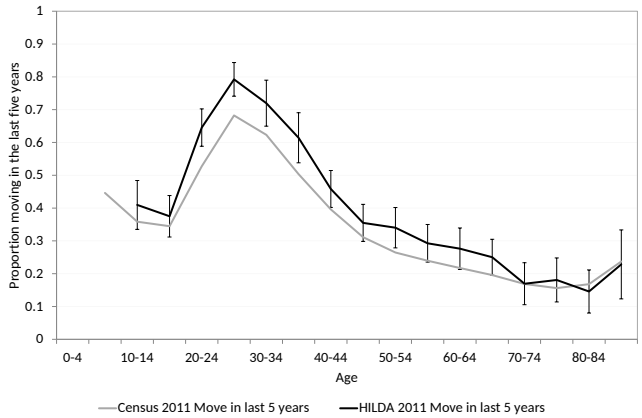


Figure 3. Proportion moving in the last year, HILDA, Census 2011 and GSS 2014
Note: HILDA = Household, Income and Labour Dynamics in Australia Survey; GSS = General Social Survey.

(a) HILDA and Census 2011



(b) HILDA and GSS 2014

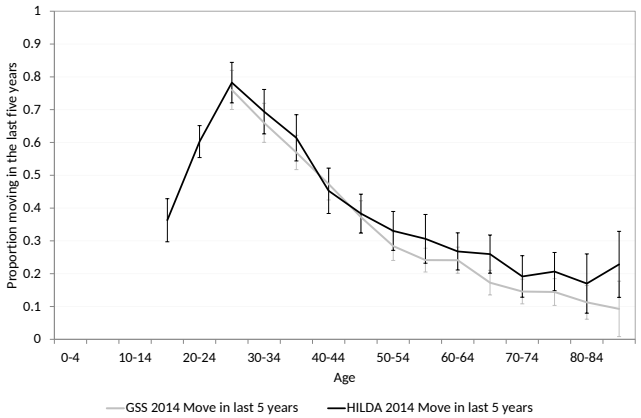


Figure 4. Proportion moving in the last years, HILDA, Census 2011 and GSS 2014
Note: HILDA = Household, Income and Labour Dynamics in Australia Survey; GSS = General Social Survey.

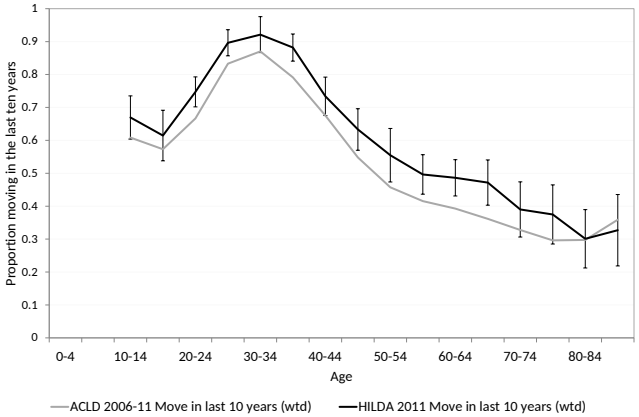


Figure 5. Proportion moving in the last ten years, HILDA and ACLD 2006–2011
Note: ACLD = Australian Census Longitudinal Dataset; HILDA = Household, Income and Labour Dynamics in Australia Survey.

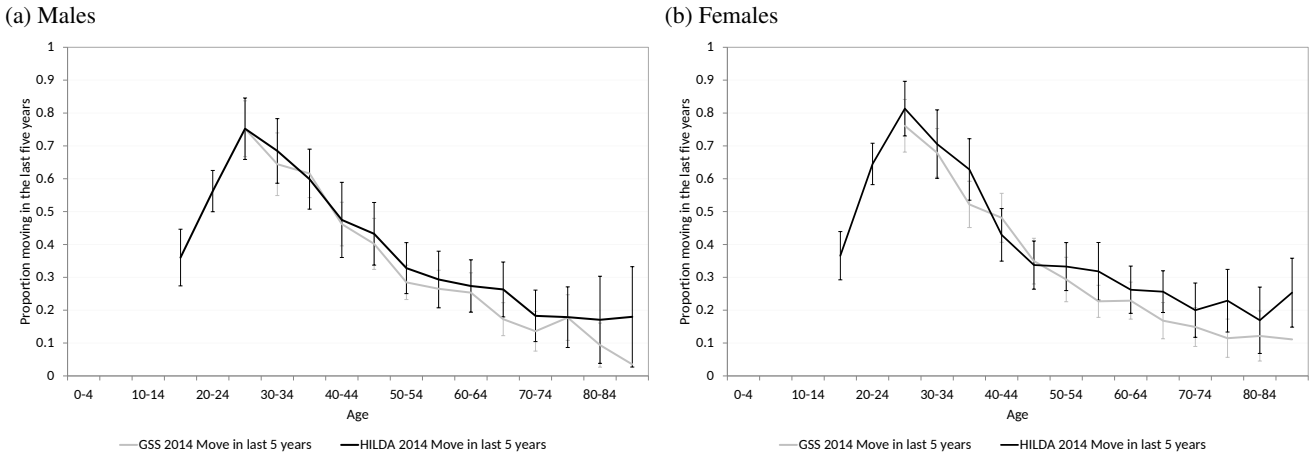


Figure 6. Proportion moving in the last five years, by sex, HILDA and GSS 2014
Note: GSS = General Social Survey; HILDA = Household, Income and Labour Dynamics in Australia Survey.

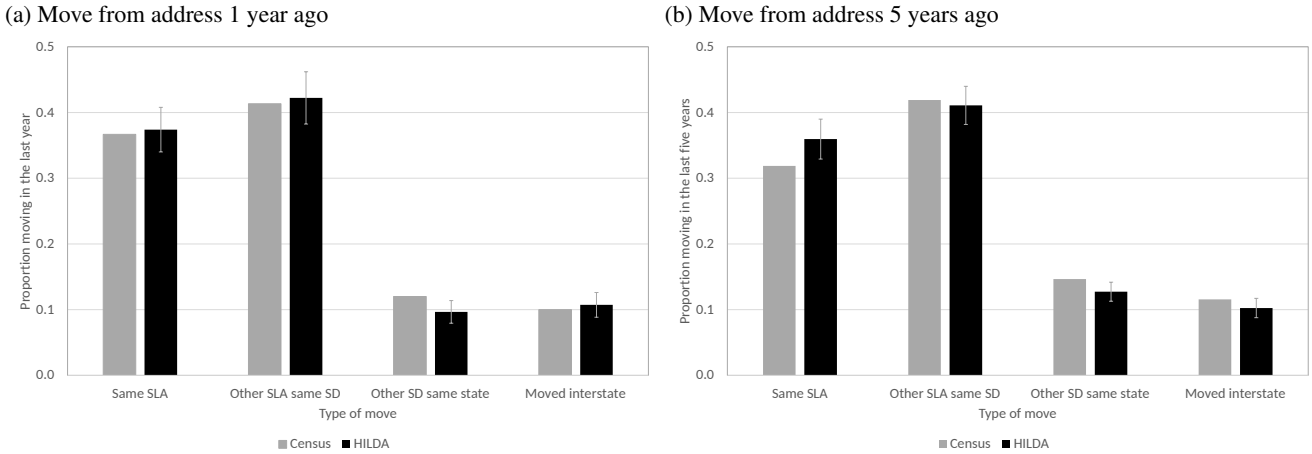


Figure 7. Proportion of each type of move, Census and HILDA, 2006
Note: HILDA = Household, Income and Labour Dynamics in Australia Survey; SLA = Statistical Local Area; SD = Statistical Division.

based on the HILDA Survey and the Census. For almost all age groups, the Census estimate is below the HILDA estimate with significant differences for those aged 20 to 29, 40 to 44 and 65 to 69 at the 5% level (and differences for those aged 30 to 39 and 50 to 54 at the 10% level). It is likely that the household-completion nature of the Census is having an effect for people aged 20 to 39 as the rate of missingness to the mobility question is approximately 10% for this age group and more of these are likely to be movers than non-movers (for those who move it might seem simpler to skip the entire question rather than provide only partial information about their address five years ago if they do not remember or do not want to provide the full address). The right panel of Figure 4 provides the rates for 2014 from the HILDA Survey and the GSS. The HILDA estimates are closely aligned with the GSS estimates, though the five-year GSS mobility estimates are generally lower for people aged 45 and older (but only significantly different at the 10% level for a couple of age categories, being those aged 65–69 or 85+). The lower mobility rates for the older people in the population for the GSS is likely a result of recall error, with more distant moves being omitted or pushed further back in time to outside the timeframe of interest. Had the reported mobility estimates from the HILDA Survey been used here, the gap between the two sets of estimates is reduced for some estimates, though the significant differences remain.

Compared to the 10-year mobility estimates from the ACLD (as shown in Figure 5), the HILDA estimates are generally between 5 to 10 percentage points higher across most of the age distribution. The estimates are significantly different for those aged 25–29 or 35–69. The lower estimates from the ACLD may simply be the combination of the lower estimates in the 5-year mobility rates in the 2006 Census (not shown) and the 2011 Census (as reported earlier). However,

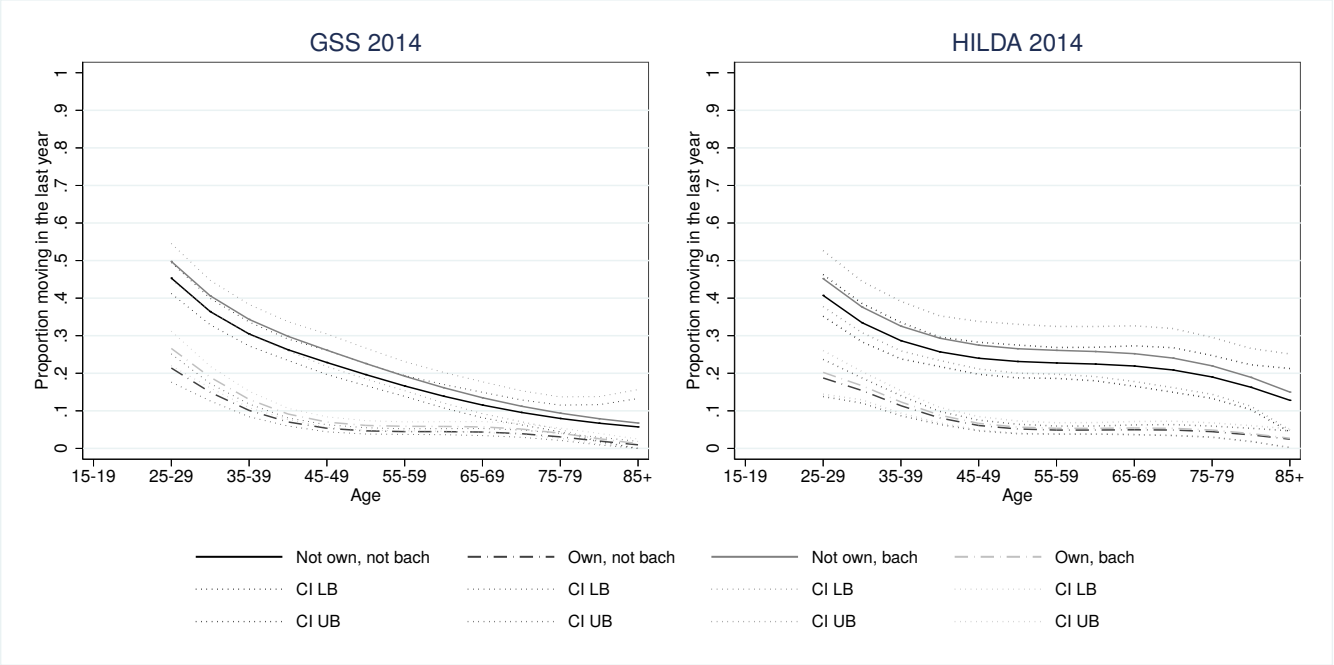


Figure 8. Predicted probability of 1-year geographic mobility by home ownership and bachelor education, GSS and HILDA, 2014

Note: GSS = General Social Survey; HILDA = Household, Income and Labour Dynamics in Australia Survey; Not own = Does not own home; Own = Owns home; Not bach = Does not have bachelor degree or higher; Bach = Has bachelor degree or higher; CI LB = Lower bound of confidence interval; CI UB = Upper bound of confidence interval.

linkage errors could also result in higher estimated mobility rates if the likelihood of moving is not independent across the two five-year periods that are being combined (some people are more likely to move than others over time, so if a highly mobile person were mismatched to a person not likely to move then the mobility rate would be higher than if the linkage error were lower).

As the GSS estimates are more reliable than the Census estimates, the subsequent analysis of one-year and five-year mobility rates use the GSS data. The 10-year mobility rates are only available from the ACLD (it is not available in the GSS data in TableBuilder) so the Census data is still used in this case.

4.3 Other characteristics of movers

Turning now to other characteristics of movers in addition to age, Figure 6 provides the age-sex distribution of movers within a 5-year window for the GSS and HILDA Survey. The expected pattern according to other studies of mobility is that women typically have higher mobility rates than men in their 20s, and men typically have higher mobility rates than women in their 30s or 40s, around the time most would be raising children (Bell & Ward, 2000; Rees, Bell, Duke-Williams, & Blake, 2010). This pattern is seen in the HILDA data but is not as evident in the GSS data. Nevertheless, the HILDA and GSS estimates are not significantly different across the age range for men and only one difference is apparent for women aged 75–79 where the HILDA mobility rate is higher than the GSS rate.

Figure 7 restricts the focus to just those that move and shows how far the move takes individuals from their initial location one year ago (in the left panel) and 5 years ago (in the right panel). For the most part, the proportion of each type of move from the address a year ago agrees between the Census and HILDA data, however the HILDA estimate for the proportion of moves that are to a different Statistical Division but within the same state is significantly lower than the Census estimate (0.097 vs 0.120). Over a 5-year window, the HILDA Survey appears to overstate short distance moves compared to the Census (0.360 vs 0.319) and understate long distance moves (0.127 vs 0.147) for moves to a different Statistical Division but within the same state, and 0.103 vs 0.116 for moves interstate though the latter is significant only at the 10% level). One plausible explanation for this finding is that it is harder to locate people who move long distances as they often also leave their work and social networks. An alternative explanation, at least in part, is that long distance moves

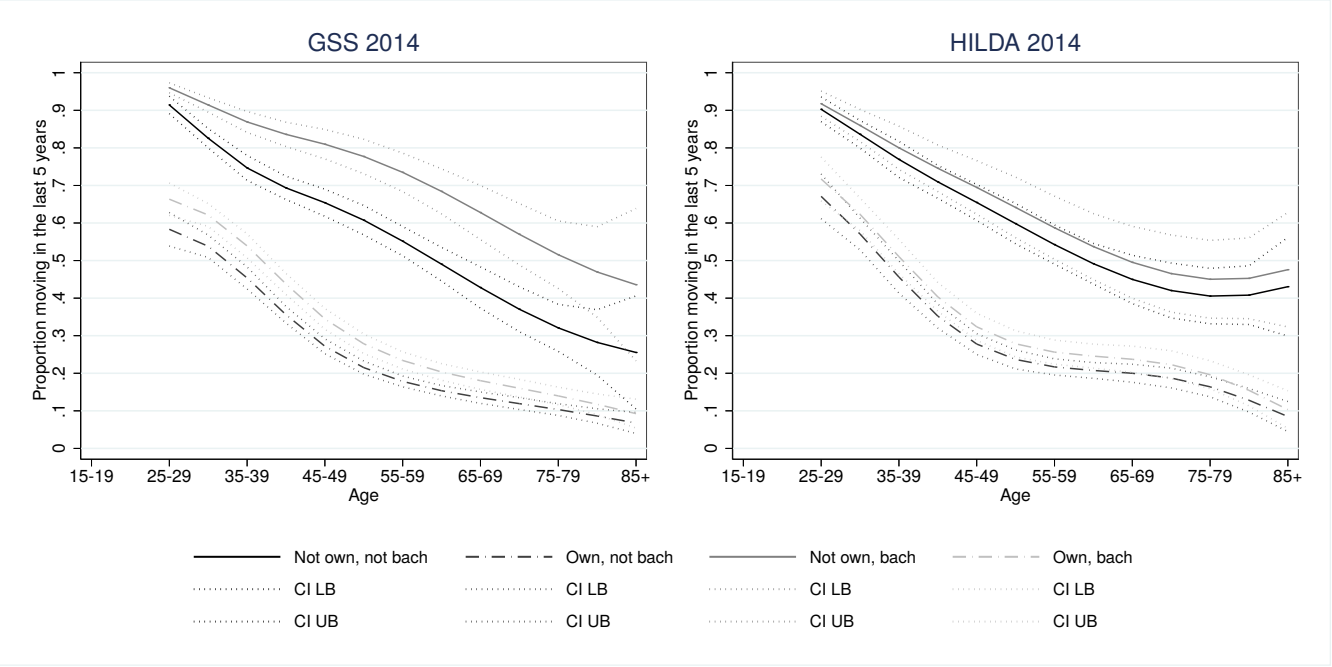


Figure 9. Predicted probability of 5-year geographic mobility by home ownership and bachelor education, GSS and HILDA, 2014

Note: GSS = General Social Survey; HILDA = Household, Income and Labour Dynamics in Australia Survey; Not own = Does not own home; Own = Owns home; Not bach = Does not have bachelor degree or higher; Bach = Has bachelor degree or higher; CI LB = Lower bound of confidence interval; CI UB = Upper bound of confidence interval.

within a state or between states are being recalled more accurately in the Census data as the move is more salient than short distance moves.

4.4 Modelling the probability of moving

The predicted probability of geographical mobility across one-year and five-year periods are estimated for the GSS and HILDA Survey with age, tenure and education covariates and are shown in Figures 8 and 9 respectively. Focusing first on left panel of Figure 8 with the one-year mobility estimates in the GSS, it is apparent that home ownership is a protective factor for mobility across almost all parts of the age distribution. People who rent have 2 to 3 times the rate of mobility of those who own their own home. Having a bachelor degree (or higher) is associated with higher mobility rates. The HILDA estimates are similar, though the effect on the mobility rate for those who do not own their own home is flatter across the middle of the age distribution. The effect of having a bachelor degree (or higher) when either owning or renting is the same between the two datasets. Moving now to Figure 9, there are two aspects of the predicted mobility rates across the five-year window that are noteworthy. First, the five-year mobility rate for people who rent in the GSS are much lower for people without a bachelor degree (or higher) compared to the similar group in the HILDA Survey. Second, the five-year mobility rate for owners is fairly consistent for those aged 55 to 79 whereas the GSS data shows a steady decline with age for these people. It is likely that recall errors are having an effect on the shape of the mobility distribution for the elderly population.

Finally, Figure 10 shows the comparison of 10-year mobility estimates from the ACLD and the HILDA Survey. As expected based on the earlier comparisons by age, the mobility rates in the ACLD are generally compressed compared to the HILDA Survey by between 5 to 10 percentage points. Nevertheless, the overall shape and ordering of these curves is consistent across the two data sources. Interestingly, and as with the five-year mobility rates discussed earlier, the difference in the 10-year mobility rates for renters with and without a bachelor degree or higher is substantially greater in the ACLD data than in the HILDA data.

5 Conclusion

Motivated by concerns about the impact of attrition on geographic mobility in the HILDA Survey, this paper compares the mobility rates from the HILDA Survey to three external

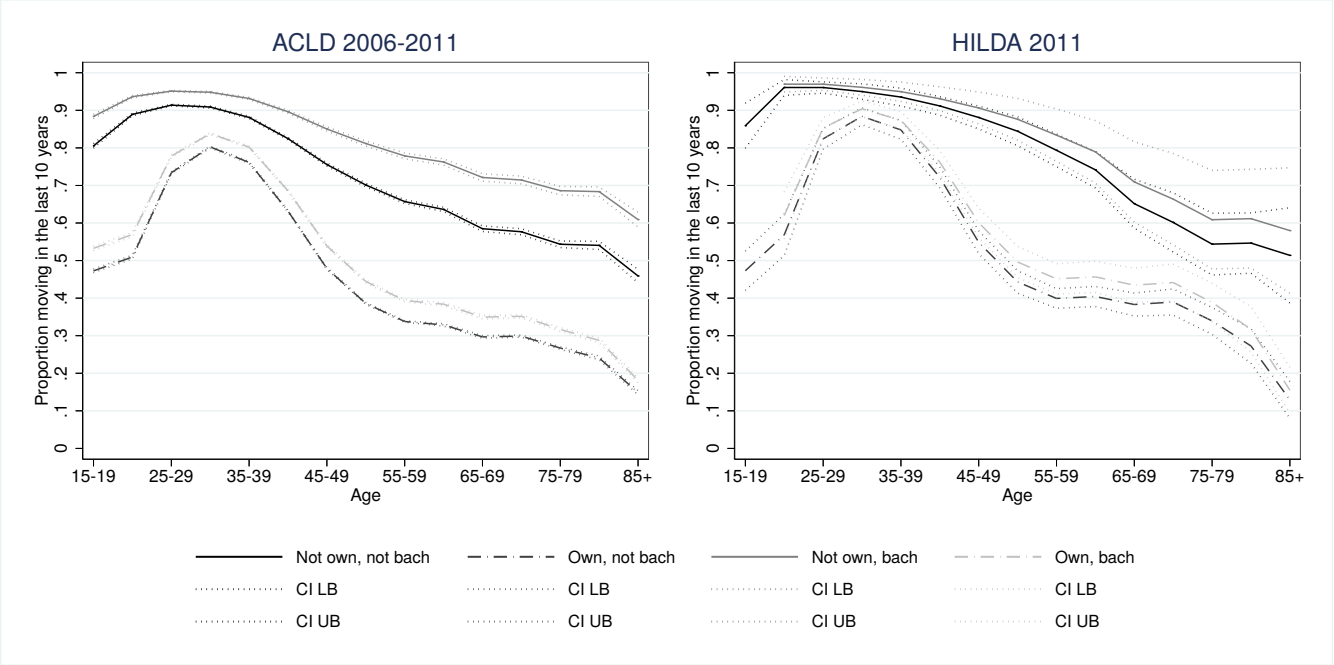


Figure 10. Predicted probability of 10-year geographic mobility by home ownership and bachelor education, ACLD and HILDA, 2011

Note: ACLD = Australian Census Longitudinal Dataset; HILDA = Household, Income and Labour Dynamics in Australia Survey; Not own = Does not own home; Own = Owns home; Not bach = Does not have bachelor degree or higher; Bach = Has bachelor degree or higher; CI LB = Lower bound of confidence interval; CI UB = Upper bound of confidence interval.

data sources: the Census, the Australian Census Longitudinal Dataset, and the General Social Survey. There are four main conclusions from the analysis presented.

First, the HILDA Survey mobility estimates align closely with those from the GSS across the age distribution, and this continues to hold when the estimates are disaggregated by sex. Five-year mobility estimates from the Census are lower than the HILDA estimates at multiple points along the age distribution. This is likely due to higher item non-response by movers than non-movers. As a result, when comparing the HILDA estimates to the linked Census data provided in the ACLD, the two distributions are substantially different, with the ACLD estimates often being 5 to 10 percentage points lower than the HILDA estimates.

Second, cognitive ability and recall errors appear to be having a role in the GSS and Census estimates for people aged 45 and older. This is likely because the recall task requires respondents to report on either how many years they have lived at the current address or where they were living five years ago. This places a much higher cognitive burden on the respondent than reporting only about the last 12 months. Further, even questions about moves in the last 12 months are subject to error and a better measure can be derived from address changes (as long as address corrections can be correctly eliminated).

Third, in tracking moves sample members make in the HILDA Survey, there is evidence to suggest (based on the 2006 Census comparison) that long-distance moves may be harder to trace than short distance moves. An alternative explanation for this finding, at least in part, is that respondents to the Census may be less likely to report details of a local move compared to a longer distance move due to the saliency of the move (consistent with the findings of Smith and Thomas (2003) on the recall of different types of moves).

Fourth, models of the probability of moving in the last five years or 10 years for the various data sources show quite different curves for renters who do not have a bachelor (or higher) degree. The GSS and ACLD show quite a sizeable gap between renters who have a bachelor degree or higher and those who do not (this gap being two to four times the size of the gap observed in the HILDA data). Perhaps this is partly related to recall issues: those who are highly educated might be more motivated to provide more accurate information or they may be more able to recall this sort of information more accurately.

There are a number of limitations of this work. Firstly, the measures used in the different data sources to identify mobility are quite different: the Census asks for the address where

a person was previously, the GSS asks for how long a person has been at their current address, and the HILDA Survey asks in one measure whether they have moved since the last interview (and if so, when that move occurred) and in another measure two addresses are compared. Some differences may have occurred due to these different measures used. It is not possible to disentangle the effects of these measures unless they are asked of the same people at the same time, yet few studies have done this (Bell et al., 2015). The second limitation of this work is the constraint imposed on the number of covariates that could be included in the logistic regression models due to the data extraction method. Nevertheless, multiple variables were considered and the variables having the greatest effect were selected for inclusion in the model.

The findings of this paper are particularly relevant to longitudinal survey fieldwork managers and users of the HILDA data. Longitudinal survey fieldwork managers are encouraged to review tracking strategies to ensure long-distance movers can be tracked as well as possible. They are also encouraged to replicate this type of work on their own data using relevant official data sources. It is expected that there will be similar findings with other longitudinal studies that have high contact rates, have active follow up of movers within the fieldwork for each wave, and use of indictors of mobility (likely or actual) within the construction of longitudinal weights. Users of the HILDA data are reassured as to the quality of the HILDA data: the geographic mobility estimates from the HILDA Survey compare favourably to several official data sources. That is, this paper has shown that, with the possible exception of the distance moved, the weights successfully take account of the effect that mobility has on attrition.

In terms of future work on this issue, consideration should be given to how best to include mobility and other attrition information, along with relevant sample design information, in substantive analyses (for example, Washbrook, Clarke, & Steele, 2014). It may also be worthwhile focusing on a particular scenario, such as separation of married or defacto couples, which almost certainly leads to mobility of one or both partners and likely to also increase an individual's probability of attrition from the sample. The assistance of linked register data that identified separations (or divorces) would be beneficial to this analysis. And along a different theme, more work could be done in identifying and testing tracking strategies (as discussed in Couper & Ofstedal, 2009) to understand the methods that work well for people who make short distance moves and for those making long distance moves.

Acknowledgement

The author thanks Janeen Baxter, Michelle Haynes, an Associate Editor of this journal and referees for their valuable comments on this paper.

References

Australian Bureau of Statistics. (1999). Australian Standard Geographical Classification (Cat. No. 1216.0). Retrieved from http://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/CA25687100069892CA256889000B0434/$File/12160_final_RGB.pdf

Australian Bureau of Statistics. (2009). Migration, Australia, 2007-08 (Cat. No. 3412.0). Retrieved from http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/3412.0Explanatory%20Notes12007-08?OpenDocument

Australian Bureau of Statistics. (2011). Census Dictionary, 2011 (Cat. No. 2901.0). Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/2901.0Chapter38202011

Australian Bureau of Statistics. (2013). Australian Census Longitudinal Dataset: Methodology and quality assessment, Australia, 2006-2011 (Cat. No. 2080.5). Retrieved from http://www.abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/AF5C5BA7885F6550CA258240000DB036?opendocument

Australian Bureau of Statistics. (2015). General Social Survey: Summary results, Australia, 2014 (Cat. No. 4159.0). Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/mf/4159.0

Australian Bureau of Statistics. (2017). Census of Population and Housing: Understanding the Census and Census Data, Australia 2016 (Cat. No. 2900.0). Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/2900.0~2016~Main%20Features~Understanding%20the%20Census%20and%20Census%20Data~1

Australian Bureau of Statistics. (2018). Census of Population and Housing: Details of Overcount and Undercount, Australia, 2016 (Cat. No. 2940.0). Retrieved from http://www.abs.gov.au/ausstats/abs@.nsf/mf/2940.0

Baker, E., Bentley, R., Lester, L., & Beer, A. (2016). Housing affordability and residential mobility as drivers of locational inequality. *Applied Geography*, *72*, 65–75. doi:10.1016/j.apgeog.2016.05.007

Bell, M., Charles-Edwards, E., Kupiszewska, D., Kupiszewski, M., Stillwell, J., & Zhu, Y. (2015). Internal migration data around the world: Assessing contemporary practice. *Population, Space and Place*, *21*(1), 1–17. doi:10.1002/psp.1848

Bell, M. & Ward, G. (2000). Comparing temporary mobility with permanent migration. *Tourism Geographies*, *2*(1), 87–107. doi:10.1080/146166800363466

Buck, N. (2000). Using panel surveys to study migration and residential mobility. In D. Rose (Ed.), *Researching social and economic change* (pp. 250–272). London: Routledge.

Caldera Sánchez, A. & Andrews, D. (2011). Residential mobility and public policy in OECD countries. *OECD Journal: Economic Studies*, *2011*(1), 185–206. doi:10.1787/19952856

Chipperfield, J., Brown, J. J., & Watson, N. (2017). The Australian Census Longitudinal Dataset: using record linkage to create a longitudinal sample from a series of cross-sections. *Australian & New Zealand Journal of Statistics*, *59*(1), 1–16. doi:10.1111/anzs.12177

Clark, W. A. V. & Huang, Y. (2003). The life course and residential mobility in British housing markets. *Environment and Planning A*, *35*(2), 323–339. doi:10.1068/a3542

Clark, W. A. V. & Onaka, J. L. (1983). Life cycle and housing adjustment as explanations of residential mobility. *Urban Studies*, *20*(1), 47–57. doi:10.1080/00420988320080041

Coulter, R., van Ham, M., & Findlay, A. M. (2016). Rethinking residential mobility: Linking lives through time and space. *Progress in Human Geography*, *40*(3), 352–374. doi:10.1177/0309132515575417

Couper, M. P. & Ofstedal, M. B. (2009). Keeping contact with mobile sample members. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys* (pp. 183–203). Chichester: Wiley.

Dex, S. (1995). The reliability of recall data: A literature review. *Bulletin of Sociological Methodology/Bulletin de Methodologie Sociologique*, *49*(1), 58–89.

DSS and MI. (2018a). The Household, Income and Labour Dynamics in Australia (HILDA) Survey, GENERAL RELEASE 16 (Waves 1-16). ADA Dataverse, V3. doi:10.4225/87/VHRTR5

DSS and MI. (2018b). The Household, Income and Labour Dynamics in Australia (HILDA) Survey, RESTRICTED RELEASE 16 (Waves 1-16). ADA Dataverse, V3. doi:10.4225/87/QFUIBM

Geist, C. & McManus, P. A. (2008). Geographical mobility over the life course: Motivations and implications. *Population, Space and Place*, *14*(4), 283–303. doi:10.1002/psp.508

Greenwood, M. J. (1997). Internal migration in developed countries. In R. M. R. & S. O. (Eds.), *Handbook of population and family economics* (pp. 647–720). Elsevier Science.

Gregg, P., Machin, S., & Manning, A. (2004). Mobility and Joblessness. In D. Card, R. Blundell, & R. B. Freeman (Eds.), *Seeking a premier economy: The economic effects of british economic refors, 1980–2000* (pp. 371–410). University of Chicago Press.

Hall, M. (2015). Measuring internal migration: Retrospective self-report. In F. D. Bean & S. K. Brown (Eds.), *Encyclopedia of Migration*. Dordrecht: Springer.

Ihrke, D. (2017). United States mover rate at a new record low. Retrieved from https://www.census.gov/newsroom/blogs/random-samplings/2017/01/mover-rate.html

Laurie, H., Smith, R. A., & Scott, L. (1999). Strategies for reducing nonresponse in a longitudinal panel survey. *Journal of Official Statistics*, *15*(2), 269–282.

Lynn, P., Buck, N., Burton, J., Jäckle, A., & Laurie, H. (2005). A review of methodological research pertinent to longitudinal survey design and data collection. Retrieved from https://www.iser.essex.ac.uk/files/iser_working_papers/2005-29.pdf

Menard, S. (2002). *Longitudinal Research* (2nd ed.). Beverly Hills, CA: Sage.

Newson, R. B. (2012). Sensible parameters for univariate and multivariate splines. *The Stata Journal*, *12*(3), 479–504.

OECD. (2016). Housing tenures. Retrieved from http://oe.cd/ahd

Productivity Commission. (2014). Geographic Labour Mobility. Retrieved from https://www.pc.gov.au/inquiries/completed/labour-mobility/report

Rees, P., Bell, M., Duke-Williams, O., & Blake, M. (2010). Problems and solutions in the measurement of migration intensities: Australia and Britain compared. *Population Studies*, *54*(2), 207–222. doi:10.1080/713779082

Skowronski, J. J., Betz, A. L., Thompson, C. P., & Shannon, L. (1991). Social memory in everyday life: Recall of self-events and other-events. *Journal of Personality and Social Psychology*, *60*(6), 831–843.

Smith, J. P. & Thomas, D. (2003). Remembrances of things past: Test–retest reliability of retrospective migration histories. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *166*(1), 23–49.

Spallek, M., Haynes, M., & Jones, A. (2014). Holistic housing pathways for Australian families through the childbearing years. *Longitudinal and Life Course Studies*, *5*(2), 205–226. doi:10.14301/llcs.v5i2.276

Summerfield, M., Freidin, S., Hahn, M., Ittak, P., Li, N., Macalalad, N., … Wooden, M. (2017). HILDA User Manual–Release 16. Retrieved from https://melbourneinstitute.unimelb.edu.au/hilda/for-data-users/user-manuals

Thompson, C. P. (1985). Memory for unique personal events: Effects of pleasantness. *Motivation and Emotion*, *9*(3), 277–289.

Tourangeau, R. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.

Washbrook, E., Clarke, P. S., & Steele, F. (2014). Investigating non-ignorable dropout in panel studies of residential mobility. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *63*(2), 239–266.

Watson, N. (2010). The impact of the transition to CAPI and a new fieldwork provider on the HILDA Survey. Retrieved from https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-discussion-papers/hdps210.pdf

Watson, N. (2012). Longitudinal and cross-sectional weighting methodology for the HILDA Survey. Retrieved from https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-technical-papers/htec212.pdf

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