HILDA PROJECT DISCUSSION PAPER SERIES
No. 2/07, June 2007

Recall Bias in Reported Events: HILDA, Waves 1 - 5

Alison Goode

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

Acknowledgements

This paper was presented at the ACSPRI Social Science Methodology Conference, University of Sydney, Sydney, Australia, 10 -13 December 2006.

Contents

Introduction

The primary purpose of longitudinal surveys is to collect data about individual change over time. For many subject areas this can be achieved by asking survey respondents to answer identical questions at every interview which provide information concerning their state at the specific interview point in time. Asking questions about today is clearly perceived to generate accurate information as it taxes the memory of the respondent as little as possible. However, relying on this type of question alone would miss out on what happens to the individual between interviews. With interviews often as far apart as one year, this would leave many gaps in the information contained in the data. In order to get around this problem, another method used to collect information, especially common in longitudinal/panel surveys, is to ask respondents to recall their behaviour over a specified period of time retrospectively, in order to fill in the gaps between interviews.

These recall questions generally take two forms. The first is where respondents are asked to recall their behaviour over a past specified period of time, typically for yearly interviews stretching from the time of the current interview to a year back. The year is generally split month by month thus producing a calendar of events. This information is repeatedly collected at every interview, thus building up overlapping information and a continuous history of an individual over time, aiming to capture any changes that may have occurred between interviews.

The second form of recall question is where respondents are asked to recollect a given status at some specific point of time in the past (e.g. at the time of last interview, were you….?). The focus of this paper is to investigate the extent of, and characteristics associated with, recall error in this second form of recall question in the Household, Income and Labour Dynamics in Austrlia Survey (HILDA).

Collecting retrospective information is especially valuable when investigating the dynamics of economic outcomes such as labour force participation. Both forms of questions are frequently used within the same survey. One of the main reasons for using the second form of recall question is to avoid asking respondents unnecessary questions about events which are not relevant to what has happened to them. For example, if a respondent is currently employed and was employed continuously from the time of the last interview, there is little point in asking the respondent unnecessary questions about job search. By contrast, if a respondent is currently employed and was not employed at the time of the last interview, asking job search questions makes sense. Hence, by targeting questions to those respondents who can provide useful information, a survey can reduce the number of questions asked of respondents and reduce respondent fatigue.

As recall questions require respondents to remember events that happened in the past, the accuracy of such retrospective data depends on two main factors. Firstly, the accuracy of recall data will depend on the *ability* of respondents to remember past events accurately. Certain individual characteristics, such as age and gender, may play a role in the accuracy of responses. Studies, using German and UK data, show that there are age and gender differences in how males and females recall their employment status in the previous year (Jũrges, 2005 and Paull, 2002). Certain busy lifestyles or life circumstances may also influence the ability to recall accurately. The more events there are to remember, the harder it may be to remember all of them accurately. For example, for individuals who have irregular employment patterns, it may be harder to remember what state they were in at every point in

time, or at the time of their last interview, as opposed to an individual who has been in steady employment.

Another major factor in this context will be the length of time elapsed between any two interviews and the events that need to be recalled. In the case of people with irregular employment patterns, it is reasonable to assume that it would be easier to recall their precise employment status of last week rather than of 11 months ago. This will be even more the case for individuals who have to remember further than a year back (i.e. respondents who have had more than a year between interviews or have not responded in one or more interviews but have responded in a later year). Research suggests that there are significant negative effects on recall accuracy, as time elapsed lengthens between interviews (Jũrges, 2005; Paull, 2002; Martini, 1989).

Secondly, the accuracy of recall data will depend on the *willingness* of respondents to remember correctly. The willingness of respondents to remember correctly will largely depend on two factors: the importance that respondents attach to the event they are being asked to recall and/or its social desirability. On the one hand, if the event has little or no significance for the individual he/she may simply not remember the event at all. On the other hand, social desirability of the whole question or of specific answers to a question may influence how a respondent answers. Using the British Household Panel Survey data, Paull (2002) finds that shorter spells of unemployment are less likely to be recalled than other types of spells. Evidence also suggests that people are less willing to recall "unpleasant" events or events that present them under a socially unfavourable light. For example, in the Health Economics literature self-reported health is often said to suffer from "justification bias" in that persons who are unemployed tend to overstate how bad their health is in order to "justify" their being unemployed (Crossley & Kennedy, 2000). This source of inaccuracy in retrospective data should always be borne in mind but is not the focus of this paper.

It is clear that the collection of retrospective data adds a new layer of complexity over and above that normally encountered in the collection of survey data. Firstly, given that longitudinal surveys are typically carried out annually for reasons of cost, the length of time between interviews may, in itself, lead to inaccuracies (Lynn *et al*, 2005). Secondly, the panel aspect of longitudinal surveys means that if there are patterns in the type of responses individuals give, these responses will impact on the degree to which the data will be an accurate representation of the population over time. Lastly there is the issue of there being no guarantee that respondents will respond in the same way for the same event in subsequent interviews despite respondents being asked to recall the exact same previous event(s) in every interview. For example, a respondent may give one answer in t and another answer for the same event in $t+1$(Horvath, 1982).

It is therefore important that researchers using longitudinal studies, are aware of the extent to which people make 'mistakes' and whether these 'mistakes' are made randomly or not. The very nature of survey data collection entails a fair amount of statistical noise. It is well understood that a sample has never exactly the same characteristics as the population from which it is derived. If any 'mistake' (that is, any inaccurately reported information) is due to random error, this is not an insurmountable problem when analysing the data. It just means that any inference made will be simply less precise. In surveys with large numbers of observations this statistical noise may not be a major problem as the larger the number of observations, the more precise the estimates will be. However, if the same or same types of people are consistently recalling the same thing 'wrongly' over time, this will introduce systematic recall bias in the data, which will be harder to deal with.

In order to guarantee the quality of a sample, efforts are always made at the survey design and the data collection stage to generate samples that are representative of the population. The Household, Income and Labour Dynamics in Australia is, of course, no exception to this. As part of determining the quality of the sample, it is important to ascertain whether recall bias is present in the data.[1] If recall bias exists it will have important implications on the measurement of change over time of individuals as it would lead to over or under estimation of change.

Watson and Wooden (2004) found that between waves 1 and 2 of the HILDA Survey there was a considerable difference in the proportion of respondents who recalled in wave 2 that they were not employed in wave 1, when, according to information collected in wave, they had been employed in wave 1 (4.6%) and some 7% recalled in wave 2 that they had been employed in wave 1 when they had in actual fact not been employed.

It is this respondent recalled data on their employment status in the HILDA Survey which is of specific interest to this paper. If a respondent recalls his/her employment status incorrectly, he/she will then be asked a set of further questions, within the Previous Employment section of the survey, which either he/she should not be answering, or indeed not answering questions he/she should be answering. This will lead to 'spurious' or 'missing' data for these individuals.

The aim of this paper is to investigate this issue, concentrating on the following three questions:

(i) To what extent do respondents remember past events incorrectly;

(ii) If particular types of events are recalled incorrectly; and,

(iii) Do certain types of people respond in different ways?

The paper will identify the characteristics of respondents that may be associated with the generation of inaccurate retrospective data and will investigate characteristics associated with a higher probability to give a wrong answer. It will do this by looking specifically at one employment question asked annually in the HILDA Survey which asks respondents to recall their employment status *at the date of their last interview*. This will be compared with the employment status of respondents, recorded with a high degree of accuracy during their previous interview, based on a series of questions which ask about their current employment status.

The paper will begin by looking at how many respondents "get it wrong" and "how they get it wrong". After defining the possible different types of mistakes, the paper will present them descriptively and analyse their probability of occurrence using multivariate regression analysis. Conclusions will be drawn as to the implications the results may have on the use of this data.

The HILDA Survey data

The Household, Income and Labour Dynamics in Australia Survey is an annual survey of a nationally representative sample of around 7000 households, comprising some 15000 adults. Sample members are re-interviewed in successive waves. Any new adults entering

[1] While it is not easy to differentiate between whether respondents are not able rather than not willing to recall correctly, it is possible to investigate whether any 'incorrect' reported behaviour is random or systematic.

a sample household are also interviewed, as long as they remain in the sample household. Interviews are conducted over a period of six months, starting in August every year. The first wave of interviews began in 2001. A more detailed description on the modeling of the survey can be found in Watson and Wooden (2004). The main focus of the survey is employment, income and the family.

Information on employment and unemployment is collected in the following ways:

- All eligible respondents are asked questions about their work activity in the last seven days from the date of current interview. Respondents are defined as either being *in paid employment* (employee, self employed) or *not in paid employment* (not in the labour force, unemployed).
- Dependent on the *defined* employment status of respondents, i.e. in paid employment or not in paid employment, they are then asked different sets of questions.
- Those in paid employment are asked about their current employment: hours of work; reason for working part time; preference for hours worked; how many jobs they have; length of tenure in current job; occupation; sick leave entitlements; paid leave entitlements; on-the-job training; and firm size.
- Those not in paid employment answer questions on: looking for work; reasons for not actively seeking work; reasons for having trouble getting a job; and, if could work, minimum acceptable hourly wage.

All respondents (not in paid employment and in paid employment) are also asked questions on previous employment, of which the first question in this section is identical for all respondents and asks "Were you employed at the date of your last interview?" As respondents are asked this same question, either in the *in paid employment* section of the questionnaire or in the *not in paid employment* section, there are two variables (*_pjemply, _pjemp* respectively) generated for this one question which are, by definition, mutually exclusive of one another.[2]

For both 'in paid employment' and 'not in paid employment' respondents, the response given for this same question will then direct the respondent to a targeted set of further questions within the Previous Employment section. Therefore the questions asked will depend on whether the respondent recalls being employed or not being employed (see Figure 1 below).

For individuals in paid work at current interview, those who recall being employed at the time of their last interview answer questions on: how many jobs they held at that time; whether working for same employer; if different employer; reason for leaving that job; and whether has held any other jobs in between job at last interview and job at current interview (Figure 1: question group **1**).

For individuals in paid work at current interview, those who recall not being employed at the time of their last interview answer questions on: other jobs since last interview; hours worked in last job prior to current job; terms of employment; occupation and reason for leaving that job (Figure 1: question group **2**).

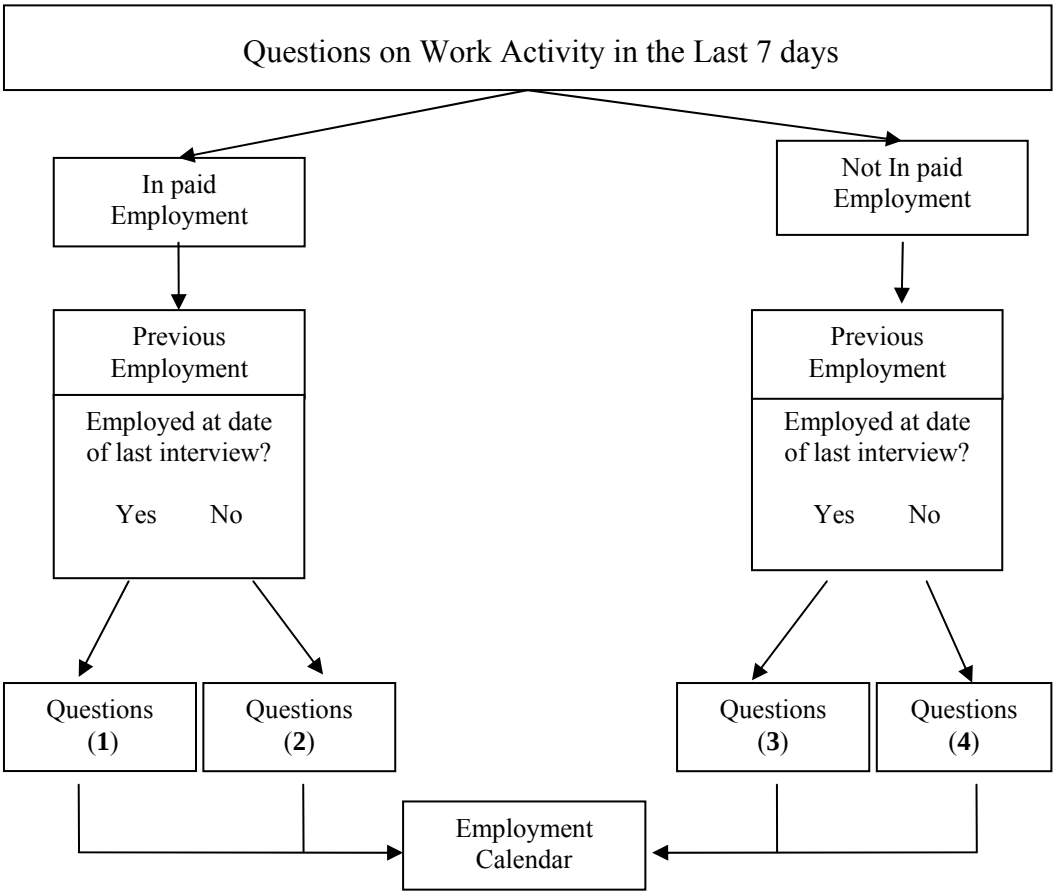
[2] Though it is asking exactly the same question, only those *in paid employment* will be asked *_pjemply*, only those not in *paid employment* will be asked *_pjemp*.

For individuals not in paid work at current interview, those who recall being employed at the time of their last interview answer questions on: reason for leaving that job; whether has held any other jobs in between job at last interview and job at current interview; main reason for leaving most recent job; hours worked in that job: industry of that job; and, terms of employment (Figure1: question group **3**).

For individuals not in paid work at current interview, those who recall not being employed at the time of their last interview answer questions on: other jobs since last interview; hours worked in last job prior to current job; terms of employment; occupation and reason for leaving that job. If the individual reports having had a job in the last 12 months, he/she answers questions on education/training schemes in the last 12 months (Figure1: question group **4**).

Finally, all respondents, both those in paid employment and those not in paid employment are asked to fill in the same activity calendar, where information on spells of employment, unemployment, and education are collected. The calendar covers the entire preceding financial year (year ended 30 June). For each month, there are three spells, each spell covering a third of a month. The starting month and year of each spell is collected where respondents have had multiple jobs in the last financial year. This question is repeated every year thus creating the overlapping activity information from the first interview to the next interview.[3]

Figure 1: Structure of HILDA Survey questions on employment



[3] Obviously, this overlapping information is only created where individuals have responded in consecutive years.

Thus, once respondents have completed their first interview, for all those respondents with subsequent interviews, there is data on their employment status (*_esbrd*), as defined at the time of each interview, and, respondent recalled data on their employment status at the date of their last interview in the current (i.e. subsequent) interview (*_pjemp, _pjemply*) as well as information on what respondents have done in the intervening time between interviews.

Constructing the data for analysis

For the purposes of this paper an unbalanced dataset is created using five waves of HILDA data and includes all eligible adults, aged 15 years and above, who provide two or more interviews over the 5 years of data and where there is complete information for the question "*at the date of the last interview were you employed*' for each interview- to- interview (see Table 1). A total of 315 interview-to-interview observations were dropped from the sample due to incomplete information for the question "*at the date of the last interview were you employed*'.[4]

Table 1: Number of individual interviews for each interview-to-interview period

Interview	*Total Number Interviewed*	*Incomplete Information*	*Total Number . Interview-to-Interview*
1st to 2nd	14,407	84	14,323
2nd to 3rd	12,675	70	12,605
3rd to 4th	11,137	61	11,076
4th to 5th	9,311	100	9,211
Total	47,530	315	47,215

Table 2 shows the distribution of individuals by gender, the mean age for each of the four possible interview-to-interview periods and their employment status in the first interview of each interview-to-interview period.

Table 2: Interview-to-interview by gender, age, and employment status

	Interview-to-Interview				
	1st to 2nd	*2nd to 3rd*	*3rd to 4th*	*4th to 5th*	*All*
Female	7451	6,599	5,848	4,906	
Male	6872	6,006	5,228	4,305	
Mean Age	42.3	44.2	46.0	48.0	45.1
In Paid Employment	8836	7,921	7,055	5,862	
Not in Paid Employment	5487	4,684	4,021	3,349	
Total Number of Individuals	14323	12605	11076	9211	47,215

Construction of the dependent variable

Over the five HILDA waves, the data contain a maximum of four possible interview-to-interview periods that an individual can have completed. In turn, there are four possible combinations of employment states each individual can be in from one interview to the next, each of which is mutually exclusive.

[4] These are observations where the question *"at the date of last interview were you employed?"* has not been asked due to a respondent's answer to a preceding question.

There are two possible answers an individual can give to the question "Were you employed at the date of your last interview?", asked in *t+1*: (i) *Yes, I was employed* or (ii) *No, I was not employed.*

A derived employment status variable (*_esbrd*) defines respondents, at the time of a specific interview date, as being employed, unemployed or not in the labour force. For the purposes of this paper, an assumption is made that the employment status variable (*_esbrd*), which is derived from questions asking respondents about their work activity in the last seven days from the date of current interview, is correct. This assumption can be justified by the fact that this variable is derived from a set of questions which ask respondents to only remember a very short and very current period of time. For the purposes of all analysis, this variable is the reference point to define respondents as having recalled incorrectly or not.[5] The three employment status categories of the variable *_esbrd* (employed, unemployed and not in the labour force) are collapsed into two categories, in paid employment, not in paid employment (unemployed and not in the labour force) in order to correspond directly with the other variables of interest: *_pjemply* and *_pjemp*.

Sixteen binary variables are first constructed from the question "Were you employed at the date of your last interview" (*_pjemply, _pjemp*) and the previous interview employment status (*_esbrd*), one for each of the four possible combinations of employment states each individual can be in from the four interview-to-interview periods that the data contains. A further four binary variables are then constructed from these sixteen variables where each same type of mistake for each of the four interview- to- interview periods is pooled together. If the respondent recalled his/her employment status incorrectly (when matched with their employment status in their previous interview, as defined by the variable *_esbrd*) the value is one and the answer is defined to be a *Mistake*. Otherwise it is defined to be a *Correct* answer and takes the value of zero. A respondent can only be in one of the four possible combinations of employment states from one interview to the next. This distinction divides the data into four sub-samples, each sub-sample containing individuals who give either the specific *Correct* answer or make the specific *Mistake* that is possible in this sub-sample. From these four variables, a further single binary variable is created which pools all these mistakes together, where the value is one for anyone who makes any kind of mistake, and zero for all others (that is, those that recalled correctly for any of the four possible employment states).

As said above, respondents in each sub-sample can only be in one employment state combination from interview-to-interview and make only one mistake and each mistake will be different from that in other sub-samples.[6] Figure 2 shows the four possible employment states sub-samples and the four possible mistakes, one for each of the four sub-samples. As each sub-sample and therefore the recall mistake that an individual can make is mutually exclusive, the sub-samples, denoted by **E1, E2, E3** and **E4** in Figure 2, can also denote the type of mistake an individual makes.

[5] It should be noted that not all inconsistencies that may exist between respondents recalled employment status and employment status as defined in the previous interview can be definitively attributed to recall error. Employment status is based on International Labour Office (ILO) definitions, and the ILO definition of paid employment will not necessarily accord with what individual respondents think of as paid employment (Wooden & Watson, 2004). However, it cannot be known how people think of paid employment and in order to carry out this analysis a reference point is necessary.

[6] Individuals can make more than one type of mistake but only for another interview-to-interview period.

Figure 2: Definitions of the four possible recall mistakes

	Sub-sample *Employment State* *Interview-to-Interview*	*Recall Mistake*
E1 (Mistake 1)	In paid employment at first interview (*t*) and in paid employment at next interview (*t+1*)	Reported (at *t+1*) to be not in paid employment at *t*
E2 (Mistake2)	In paid employment at first interview (*t*) and not in paid employment at next interview (*t+1*)	Reported (at *t+1*) to be not in paid employment at *t*
E3 (Mistake 3)	Not in paid employment at first interview (*t*) and in paid employment at next interview (*t+1*)	Reported (at *t+1*) to be in paid employment at *t*
E4 (Mistake 4)	Not in paid employment at first interview (*t*) and not in paid employment at next interview (*t+1*):	Reported (at *t+1*) to be in paid employment at *t*

Explanatory variables

Included in the analysis are those factors which have been suggested by the existing literature (as set out in the introduction), to influence respondents' recall of events. These include: age (split into six age groups); gender; elapsed time between interviews (duration); life satisfaction; number of jobs held in the last financial year; and, employment states for interview-to-interview period.

Other socio-demographic variables include: marital status; geographical region (state); English as first language; whether in full-time education; education level; number of children ever had; occupation (manual/non-manual); Index of relative socio-economic disadvantage (quintiles); and, general health. A detailed description of the explanatory variables can be found with descriptive statistics in the Appendix (Table A1a and Table A1b, respectively).

Descriptive Analysis

Table 3 shows the number of respondents for all possible interview-to-interview employment status combinations. This table is generated from the variable which is derived from questions asking respondents about their work activity in the last seven days from the date of each (current) interview and which defines respondents' employment status at each interview (*_esbrd).*[7]

[7] This is interview- to- interview, not consecutive wave on wave interviews.

Table 3: Employment status – interview-to-interview

Interview -to- interview	***E1*** *(Stable)*	***E2*** *(Changed)*	***E3*** *(Changed)*	***E4*** *(Stable)*	*Total*
1 – 2	8065	771	951	4536	14,323
2 - 3	7279	642	756	3928	12605
3 - 4	6517	538	568	3453	11076
4 - 5	5481	381	465	2884	9211
Total	**27342**	**2332**	**2740**	**14801**	**47,215**
% Stable	58%			31%	89%
% Changed		5%	6%		11%
Total %	58%	5%	6%	31%	100%

The majority of respondents (58%) appear to be in a stable employment status. A large minority (31%) are in a stable non-employment status. Only 11% of the sample has changed employment status between interviews, 5% leaving paid employment and 6% returning to it.

Appendix Table A2 divides these four sub-samples by gender (%) and mean age. More males are in the category paid employment/paid employment from interview-to-interview than females. In the other three categories there are more females. This reflects the overall distribution of employment by gender. More women stay at home or move in and out of the labour force. In terms of mean age, the oldest group is those not in paid employment in both interviews. This reflects that the retired and the long-term unemployed, who are generally older, are included in this group.[8]

Table 4 shows each of the four sub-samples and how respondents recalled their previous employment status in *t* at *t+1*. This table shows that very few respondents who are in the same employment status from one interview to the next recall their previous employment status incorrectly (percentage figures in bold in the table). For the four interview periods 2.6%, 1.6%, 2.1% and 1.8% of those in paid employment in both interview *t* and interview *t+1* recall their employment state incorrectly. Similarly low percentages of those not in paid employment in both interview *t* and interview *t+1* recall their employment state incorrectly (2.7%, 2.5%, 2.4% and 1.9%, for the four interview periods respectively).

However, the picture is different for those respondents whose employment status changes from one interview to the next (percentage figures in italics in the table). For the four interview periods, 29.2%, 26.3%, 25.5% and 20.2% of those respondents who are not in paid employment at interview *t+1* and were in paid employment at interview *t* recalled their

[8] As there is no mandatory retirement age in Australia, there is no obvious age cutoff point for this data. In the multivariate section regressions were carried out with different age cutoff points to look into this issue. Given that no substantial differences were found in the results, it was decided, on balance, to not truncate the sample by age.

employment status incorrectly. That is, they recalled themselves as being not employed at the time of interview *t*.

For the four interview periods where respondents are in paid employment at interview *t+1* and were not in paid employment at interview *t*, 33.8%, 36.5%, 35.2% and 42.8% of respondents recall themselves as being employed at the time of interview *t*. Clearly, respondents who have changed employment status from one interview to the next make many more recall errors than their counterparts who are in the same employment status in *t* and *t+1*.

Table 4: Recall Mistakes by interview and employment status

	Employment status in Interview(t+1)							
	2		3		4		5	
	In Paid *Empl*	*Not in Paid Empl*	*In Paid Empl*	*Not in Paid Empl*	*In Paid Empl*	*Not in Paid Empl*	*In Paid Empl*	*Not in Paid Empl*
In paid employment at Intervierw(*t*)								
Recalled employed	7,853	546	7,160	473	6,378	401	5,385	304
Recalled Not employed	212	225	119	169	139	137	96	77
Total	8,065	771	7,279	642	6,517	538	5,481	381
% Incorrectly recalled	**2.6**	*29.2*	**1.6**	*26.3*	**2.1**	*25.5*	**1.8**	*20.2*
Not In paid employment at Interview (*t*)								
Recalled not employed	630	4,415	480	3,831	368	3,370	266	2,827
Recalled employed	321	121	276	97	200	83	199	57
Total	951	4,536	756	3,928	568	3,453	465	2,884
% Incorrectly recalled	*33.8*	**2.7**	*36.5*	**2.5**	*35.2*	**2.4**	*42.8*	**1.9**
Total in t+1	9,016	5,307	8,035	4,570	7,085	3,991	5,946	3,265

Note: Total sample size 47,215

This next section looks at the All Mistakes and the four types of Mistakes in more detail. Table 5a shows the frequency of the total number of occurrences of each type of recall mistake (Mistake 1, Mistake 2, Mistake 3 and Mistake 4) and for the pooled data (All Mistakes). The percentage figure in brackets for *All Mistakes* is the proportion of respondents who make a mistake as a percentage of the *total sample* over the five years of data. For all 4 possible interview-to-interview periods, 5.4% of respondents recall their employment status

incorrectly. The percentage figures in the last column of Table 5a are the proportion of respondents within each particular interview-to interview employment states (and the recall mistake this opens up to them) and shows the big differences in the probability of making a mistake depending on the employment states of individuals from interview-to-interview.

Table 5a: Total number of Recall Mistakes by sub-sample

Employment state t-t+1	*MistakeType*	*Make Mistake*		*Total*	*% of Mistakes for each sub-sample*
		Yes	*No*		
(E1)	M*istake 1*	566	26777	27342	2.1%
(E2)	*Mistake 2*	608	1724	2332	26.1%
(E3)	*Mistake 3*	996	1744	2740	36.4%
(E4)	*Mistake 4*	358	14443	14801	2.5%
	All Mistakes	2528 (5.4%)	44687 (94.6%)	47215 (100%)	

Table 5b shows the total number for each type of Mistake as a percentage of all recall mistakes. Those belonging to Mistake 2 and Mistake 3 represent over 60% of all those respondents who make any mistake. Again this suggests that respondents who have changed employment status from one interview to the next make more recall errors than their counterparts who are in the same employment status in *t* and *t+1*.

Table 5b: The four Recall Mistakes as a percentage of All Mistakes

	Total	*%*
E1: Mistake 1	566	22.4
E2: Mistake 2	608	24.1
E3: Mistake 3	996	39.4
E4: Mistake 4	358	14.2
Total No. of Mistakes Made	2528	100.0

It is interesting to note that for those individuals in employment states E1 and E2, a very similar percentage of individuals recall incorrectly. (Mistakes 1 (22.4%) and Mistake 2 (24.0%)), the recall being 'unfavourable'[9] for both groups. For those individuals in employment states E3 and E4, double the number of individuals make Mistake 3 (39.4%) than Mistake 4 (14.2%) although both groups are making the same 'favourable' recall.[10]

[9] That is, individuals recall being not in paid employment at interview t when they had been defined as being in paid employment This can be viewed as an 'unfavourable' recall.

[10] That is, individuals recall being in paid employment at interview t when they had been defined as not being in paid employment This can be viewed as an 'favourable' recall.

This would suggest again that what seems to matter is the employment status at the time of recall (and not the employment status that is being recalled) and whether the employment status changed between interviews rather than what that status was in *t* or *t+1*.

Table 5c shows the frequency of the number of all recall mistakes for the pooled data by select characteristics. More females than males make a mistake and the mean age of those that make a mistake is 39.2 years. Those that are married or are cohabiting make more mistakes than single, separated, divorced or widowed people. Almost 89% of those that make a mistake are either born in Australia or have English as their first language and almost 15% of them are in full time education. Over half of those that make mistakes report having had one job in the last financial year, some 21% report having had no job, and 23.9% report having had two or more jobs.

Table 5c: All Recall Mistakes by select characteristics

	All Mistakes		
Characteristic	*Number*	*Total*	*%*
Male	1,133	2528	44.8
Female	1,395		55.2
Mean Age	39.2		
Mean Duration b/n Interviews (Days)	396		
Married/Cohabit	1,453	2528	57.5
English as first language	2,243	2528	88.8
In FT Education	373	2528	14.8
No.of Jobs in last year			
None	537	2528	21.2
1	1387	2528	54.9
2	432	2528	17.1
3 or more	172	2528	6.8

Table 5d breaks Table 5c down further by select characteristics for those that make Mistakes 1, 2, 3 and 4. The percentage figures for All Mistakes characteristics are given in the final column. This highlights the differences in characteristics of respondents depending on what type of employment state they are in from interview *t* to interview *t+1*. For example, the mean age of respondents who make any kind of mistake is 39.2 years. However, the mean age of respondents who make Mistake 1 is a lot lower, at 33.9 years. Likewise, there are more males who make Mistake 1, although for those who make any kind of mistake, it is females who make more mistakes. It is interesting to note that in terms of the number of jobs that respondents have had in the last financial year (that is, leading up to the time of interview *t+1*), some 70% of those that make Mistake 2, which is to recall being not in paid employment at interview *t* when in actual fact they were employed, had no jobs in the last financial year, whereas respondents who make either Mistake 1, Mistake 3 or Mistake 4, over 80% in all three types of mistakes recall having had either 1 or 2 jobs.

Table 5d: The four Recall Mistakes by select characteristics

	Mistake 1		*Mistake 2*		*Mistake 3*		*Mistake 4*		*All Mistakes*
	N	%	N	%	N	%	N	%	
Characteristic									
Male	304	53.7	269	44.2	427	42.9	133	37.2	44.8%
Female	262	46.3	339	55.8	569	57.1	225	62.8	55.2%
Married/Cohabit	308	54.4	364	59.9	576	57.8	185	51.7	57.5%
English	501	88.5	545	89.6	880	88.4	317	88.5	88.8%
In FT Education	73	12.9	97	16.0	150	15.1	53	14.8	14.8%
Number of jobs in last financial year									
None	51	9.0	431	70.9	8	0.8	47	13.1	21.2%
1	354	62.5	132	21.7	649	65.2	258	72.1	54.9%
2	115	20.3	29	4.8	243	24.4	45	12.6	17.1%
3 or more	46	8.1	16	2.6	96	9.6	14	3.9	6.8%
Mean Age	33.9		43.2		37.2		38.9		39.2
Mean Duration (Days)	394.7		385.2		396.0		415.1		396
All Mistakes	566		608		996		358		2528

To conclude this section, the number of times individuals make a mistake within all 5 waves of data are summarized in Table 6 below. Table 6 shows that 70% of those individuals who make All Mistakes make only 1 mistake and 27.6% make 2 mistakes. Very few individuals make a mistake more than twice (2.4%). Of the 70% that make one mistake only, some 57% of individuals' employment state has changed from interview-to-interview (Mistakes 2 and 3). Similarly, for those that have made 2 mistakes (27.6%), some 78% of individuals' employment state has changed from interview-to-interview and for those making 3 mistakes (2.1%) just over 60% of individuals' employment state has changed from interview-to-interview.

Table 6: Number of mistakes by individuals by Mistake Type and All Mistakes (pooled)

	Number of Mistakes Made by Individuals						
	1 only		*2*		*3*		*Total*
Mistake Type	*N*	*%*	*N*	*%*	*N*	*%*	*N*
Mistake1	486	27.4	73	10.5	7	11.7	566
Mistake2	347	19.6	241	34.5	20	33.3	608
Mistake3	672	38.0	306	43.8	18	30.0	996
Mistake4	265	15.0	78	11.2	15	25.0	358
Total		100		100		100	
All Mistakes (Pooled)	1770	70.0	698	27.6	60	2.4	2528

Multivariate analysis

Estimation method

This section investigates the data in a multivariate context using regression analysis. The main advantage of multivariate regression is that it allows the researcher to examine how the variable under investigation (the dependent or left-hand-side variable) may be associated with several factors (the independent, explained or right-hand-side variables) simultaneously, where these factors may be (and typically are) to a degree correlated between themselves.

This section follows the previous analysis in the way the sub-samples have been defined. The main contribution of this section is that it estimates the probability that a person belonging to each of these sub-samples may make a recall mistake or not. The method of estimation is the PROBIT estimation which is appropriate for cases where the dependent variable takes the values of either 1 or 0. To recap the types of mistakes made, as defined in Table 5, there are four dependent variables (one for each sub-sample and mistake) where if the respondent recalls incorrectly the dependent variable takes the value of 1, if not the dependent variable takes the value of 0. STATA/SE 9.0 was used for all estimations. STATA estimation command *dprobit* was used in order to produce marginal effects at the mean values of the right-hand-side variables. To account for the multiple observations of respondents within the pooled sample, observations were clustered[11] by respondent identity and t-ratios are reported (based on heteroskedasticity corrected standard errors).

Five estimations are carried out and each estimation is presented separately below. The first estimation is for **all types** of mistakes pooled together (All Mistakes) and the next four estimations are for each sub-sample and its associated mistake. The first model estimates the probability that a person may make any recall mistake or not. The next four models estimate the probability that a person belonging to any of the four sub-samples may make the specific recall mistake that is possible for that sub-sample.

Before results are presented the reader is reminded of an underlying assumption in the analysis of the four sub-samples, as it relates to the way results are interpreted. There are four sub-samples in this analysis depending on the employment status of the survey respondents at any two neighbouring interview points *t* and *t+1*. These four sub-samples were defined in Figure 3.

When each member of each sub-sample is interviewed at *t+1*, he/she can either mis-report his/her employment status at t or not. Note that the member of each sub-sample can only make one type of recall mistake, by way of their belonging to one and not any of the other sub-samples (see Figure 3). To make this clear, someone who is in Paid Employment at *t+1* and was in Paid Employment at *t* can only make the mistake of reporting that he/she was Not in Paid Employment at *t*. The same applies for the other three sub-groups.

The following section begins by analysing the total sample, that is, for those who have made any type of mistake. The four specific mistake sub-samples are then analysed as if they are completely independent, treating the way they were selected into their respective employment

[11] Clustering by respondent identity specifies that the observations are independent across groups (clusters) but not necessarily independent within groups

status as a secondary issue.[12] In statistical terms this means that all results will be conditional on individuals belonging to each of the sub-samples. The advantage of this approach is that, where there may be any systematic differences between making a recall mistake between the sub-samples, these will be shown in the separate estimations for the four recall mistake categories. The results of all five estimations are presented and, in the light of the results from the analysis of the four sub-samples, the first analysis of the total sample will be re-examined.

Estimation results

Table 7 presents the *dprobit* estimation results for *ALL MISTAKES*. Marginal effects can be interpreted as the association between the probability that a mistake may happen with the variable in question. Where a single variable is concerned, such as In Full Time Education, the marginal effect of 0.0103 in the estimation (Table 7) implies that those in Full Time Education have a 1.03 percentage point higher probability than those not in full time education of recalling incorrectly their past employment status. The t-ratio of 2.93 and p-value of 0.003 suggest that this estimate is very precise and that the hypothesis that this marginal effect is equal to zero is rejected by the data. Marginal effects of variables that belong to a group (such as the group of health status variables) are interpreted in the same way with the only difference being that there is a common reference category for the whole group (E.g. in the case of the health status variables this is health category "excellent health"). Looking at marginal effects of different variables in Table 7, age, whether respondents are in full time education or not, the number of jobs respondents have had in the last financial year, the number of time periods answered in the employment calendar, children and employment status at interview *t* and interview *t+1* are all associated with the probability of making a recall mistake. The association of age is modest as only the 15 to 24 year olds have a 1.9 percentage point higher probability of recalling incorrectly than the reference category which is the 65 year old and over. No other age group is statistically significant. Those in full time education have a 1.0 percentage point higher probability of recalling incorrectly than those not in full time education. The more children respondents have ever had increases the likelihood that respondents will recall incorrectly. Respondents with one or two children have a 0.6 percentage point higher probability of recalling incorrectly, while those with three or more have a 1.2 percentage points higher probability of recalling incorrectly.

As the number of jobs that respondents have had in the last financial year from the time of interview *t+1* increases, the more likely respondents are to recall incorrectly. Those reporting one job, two jobs and three or more jobs have a 1.9, 3.7 and 6.2 percentage point(s) higher probability of recalling incorrectly than those who have reported having had no jobs, respectively.

Lastly, the employment status of respondents in interview *t* and in interview *t+1* is highly statistically significant. For those who are in paid employment at interview *t* and not in paid employment in interview *t+1* and those not in paid employment at interview *t* and in paid employment in interview *t+1*, respondents have a 28.8 and 35.9 percentage points (respectively) higher probability of recalling incorrectly than those who are in paid employment in both interview *t* and interview *t+1*. For those who are not in paid

[12] It is understood that selection of respondents into one of the four employment status categories is not random (there is substantial evidence in labour economics research on this). It should be noted that the non-randomness of this selection is not the subject of this paper. This paper takes respondents in their present employment status and then looks at what they do in terms of making a recall mistake or not.

employment at interview *t* and not in paid employment in interview *t+1* (type Mistake 4), the probability drops to 2.5 percentage points.

Clearly, the employment status of individuals from one interview to the next strongly influences whether individuals recall their employment status at the time of their last interview correctly or not. There is a clear indication that those who have a different employment status in interview *t*+1 to that at interview *t* are far more likely to recall incorrectly than those individuals who are in the same employment state in both interviews.[13] Concluding the discussion of Table 7 results is the suggestion that the different mistakes should also be looked at separately.

[13] Note that these coefficients reflect the raw (unconditional) percentages for the different types of mistakes in the split sub-samples.

Table 7: Probit results for All Mistakes

ALL MISTAKES			
Characteristic	*Marg Effect*	*t-ratio*	*P-value*
Male	0.002	1.14	0.2530
Age15-24 (age 65 and above)[1]	0.019	3.74	0.0000**
Age25-34	0.004	1.10	0.2700
Age35-44	0.003	0.81	0.4180
Age45-54	0.002	0.56	0.5770
Age55-64	0.004	1.16	0.2470
Married or cohabiting (separated,divorced,widowed)[1]	0.001	0.55	0.5850
Very good health (excellent health)[1]	0.001	0.29	0.7700
Good health	0.003	1.18	0.2370
Fair/poor health	0.004	1.34	0.1790
Life Satisfaction	0.000	0.25	0.8000
English as first language or born in English speaking country	-0.004	-1.58	0.1130
Victoria (New South Wales)[1]	-0.003	-1.59	0.1130
Queensland	-0.003	-1.40	0.1600
South Australia	-0.004	-1.52	0.1290
Western Australia	-0.002	-0.77	0.4420
Tasmania/ Northern Territory, ACT	0.003	0.86	0.3890
Full time education	0.010	2.93	0.0030**
Non-Manual worker	0.001	0.41	0.6800
University educated (Educated up to Year 11 or less)	0.001	0.29	0.7730
Year 12 plus educated	-0.001	-0.27	0.7890
Duration between interviews in days	-0.003	-1.01	0.3130
DurationSq	0.004	2.25	0.0250**
1 job in last financial year (0 jobs in last financial year)	0.019	4.95	0.0000**
2 jobs in last financial year	0.037	6.62	0.0000**
3+ jobs in last financial year	0.062	7.61	0.0000**
Number of time periods answered in employment calender	0.001	3.72	0.0000**
Number of children:1-2	0.006	2.18	0.0290**
Number of children:3+	0.012	3.97	0.0000**
Index of Relative socio-economic disadvantage			
2nd quintile (1st quintile)	-0.005	-1.70	0.0900*
3rd quintile	0.001	0.41	0.6790
4th quintile	0.001	0.45	0.6510
5th quintile	-0.002	-0.71	0.4790
Employed to unemployed (employed to employed)	0.288	40.32	0.0000**
Unemployed to employed	0.359	50.07	0.0000**
Unemployed to Unemployed	0.025	4.75	0.0000**
log-L (restr)		-9848.33	
log-L (unrestr)		-7436.52	
Pseudo R Square		0.2449	
Chi Square		5252.54	
Observations		47173	

[1] ( ) denotes the reference category. For Tables 8a, 8b and Table 9 the characteristics have been abbreviated for space reasons. The reference categories are as in Table 7.

*denotes statistical significance at the 10% level **denotes statistical significance at the 5% level

Tables 8a and 8b present the *dprobit* estimation results for Mistakes 1 and 2, and Mistakes 3 and 4, respectively. Before an interpretation of the results from the estimations of the four sub-samples, two main points need to be stressed. Firstly, Mistakes 1 and 2 are about misreporting past Paid Employment status and Mistakes 3 and 4 are about misreporting past Not in Paid Employment status. Hence if there is any justification bias it will probably run in the same direction in these two categories. Secondly, Mistakes 1 and 4 are made by respondents who were in the same employment status in both interviews and Mistakes 2 and 3 are made by respondents who were in a different employment status in the two interviews. Mistakes 1 and 4 are low probability events and Mistakes 2 and 3 are much higher probability events. Bearing in mind these similarities and differences between the four mistake categories a brief interpretation of the results follows.

Gender and marital status matter only in the case of Mistake 3: male and married or cohabiting respondents have a higher probability of mis-reporting their past Not in Paid Employment status once they have moved into paid employment by 4.2 and 1.3 percentage points, respectively. Age matters only when employment status has changed. Both mistakes 2 and 3 are more likely to be made by older people. Being in Full Time Education influences the probability of making Mistakes 1, 3, and 4. The probability of an individual in full time education making Mistakes 1 and 3 are 0.7 and 6.1 percentage points higher (respectively) than those who are not in full time education. Interestingly, those in full time education who make Mistake 4 are 0.2 percentage points *less* likely to make a mistake and recall their past not in employment as being in employment at interview *t*. The difference between those that make Mistakes 1 and 3 and those that make Mistake 4 could be explained by the possibility that those in full time education who are not in paid employment at either point in time may not work at all while studying.

Respondents with higher levels of education are less likely to report not employed when actually employed and more likely to misreport employed when actually not employed. In both cases the bias is in favour of reporting being in paid employment in the previous year. The time elapsed between interviews matters only in misreporting past employment (Mistakes 1 and 2) and not past Not in Employment (Mistakes 3 and 4). The association with Mistakes 1 and 2 takes a rather flat U-shaped form where the shorter breaks between interviews are associated with a lower probability of misreporting past employment.[14]

The number of jobs in the last financial year (in the employment calendar) is negatively associated with Mistakes 1 and 2 and positively associated with Mistakes 3 and 4. That is, the more jobs in the last financial year, the *less* likely that past paid employment status will be misreported (Mistakes 1 and 2) and the more likely that not being in paid employment will be misreported (Mistakes 3 and 4). Again here one can see a bias in favour of reporting being in paid employment by remembering more accurately having been in paid employment rather than not in paid employment. The number of children respondents have had is positively associated with Mistake 4 only. Those with three or more children have a higher probability of misreporting their past not in paid employment.

[14] The turning point for Mistake 1 is 620 days and 1002 days for Mistake 2.

Table 8a: Probit results for Mistakes 1 and 2

	Mistake 1			*Mistake 2*		
True status at *t:*	In Paid Employment at *t*			In Paid Employment at *t*		
True status at *t+1*:	In Paid Employment at *t+1*			Not in Paid Employment at *t+1*		
Recall Mistake at *t+1*	Not in paid employment at *t*			Not in paid employment at *t*		
Characteristic	Marg Effect	t-ratio	P-value	Marg Effect	t-ratio	P-value
Male	0.000	0.29	0.7700	0.007	0.32	0.7480
Age15-24	0.013	1.86	0.0630*	-0.084	-1.65	0.1000*
Age25-34	-0.002	-0.35	0.7260	-0.135	-3.31	0.0010**
Age35-44	-0.004	-0.79	0.4270	-0.087	-2.14	0.0320**
Age45-54	-0.006	-1.15	0.2520	-0.111	-2.78	0.0050**
Age55-64	-0.006	-1.31	0.1890	-0.091	-2.37	0.0180**
MarrCohab	-0.003	-1.71	0.0880*	-0.043	-1.65	0.1000*
VGhealth	0.003	1.52	0.1300	-0.050	-1.62	0.1050
Ghealth	0.004	1.94	0.0520*	-0.031	-0.99	0.3220
FPhealth	0.005	1.57	0.1170	-0.034	-0.99	0.3230
LifeSatisf	0.000	0.38	0.7070	0.004	0.61	0.5440
EnglishLang	-0.008	-2.61	0.0090**	-0.057	-1.53	0.1250
VIC	-0.005	-2.31	0.0210**	-0.008	-0.27	0.7900
QLD	-0.002	-0.87	0.3840	-0.002	-0.08	0.9380
SA	0.001	0.37	0.7100	-0.024	-0.66	0.5110
WA	-0.005	-1.97	0.0490**	-0.041	-1.11	0.2690
TASNTACT	0.003	0.76	0.4450	0.025	0.47	0.6420
FTEducation	0.007	1.89	0.0590*	0.051	1.27	0.2030
Non-Manual	-0.001	-0.27	0.7900	-0.039	-1.51	0.1320
UnivEd	-0.005	-2.34	0.0190**	-0.009	-0.30	0.7640
Yr12plusEd	-0.006	-3.44	0.0010**	-0.004	-0.16	0.8720
Duration	-0.005	-1.73	0.0830*	-0.178	-4.17	0.0000**
DurationSq	0.004	2.53	0.0110**	0.090	3.19	0.0010**
1Job	-0.415	-15.37	0.0000**	-0.766	-26.11	0.0000**
2Jobs	-0.044	-14.31	0.0000**	-0.305	-18.26	0.0000**
3+Jobs	-0.022	-12.84	0.0000**	-0.232	-11.57	0.0000**
EmplCalend	0.000	0.73	0.4660	0.002	0.47	0.6390
1-2Children	0.003	1.34	0.1820	-0.009	-0.25	0.8000
3+Children	0.004	1.28	0.2020	0.019	0.51	0.6130
SeconQuint2	0.000	0.13	0.8950	-0.063	-1.82	0.0690*
SeconQuint3	0.002	0.65	0.5150	-0.086	-2.50	0.0130**
SeconQuint4	-0.001	-0.43	0.6660	-0.058	-1.67	0.0940*
SeconQuint5	-0.002	-0.81	0.4210	-0.012	-0.35	0.7270
log-L (restr)	-2750.52			-1338.12		
log-L (unrestr)	-2517.45			-713.79		
Pseudo RSq	0.0847			0.4626		
ChiSq	417.55			891.01		
Observations	27321			2332		

*denotes statistical significance at the 10% level **denotes statistical significance at the 5% level

Table 8b: Probit results for Mistakes 3 and 4

	Mistake 3			*Mistake 4*		
True status at *t*:	Not in Paid Employment at *t*			Not in Paid Employment at *t*		
True status at *t+1*:	In Paid Employment at *t+1*			Not in Paid Employment at *t+1*		
Recall Mistake at *t+1*:	In Paid Employment at *t*			In Paid Employment at *t*		
Characteristic	Marg Effect	t-ratio	P-value	Marg Effect	t-ratio	P-value
Male	0.042	2.03	0.0430**	-0.001	-1.41	0.1590
Age15-24	-0.439	-7.83	0.0000**	0.005	1.98	0.0480**
Age25-34	-0.361	-7.90	0.0000**	0.002	1.23	0.2170
Age35-44	-0.355	-7.86	0.0000**	0.001	0.77	0.4410
Age45-54	-0.325	-7.60	0.0000**	0.007	3.06	0.0020**
Age55-64	-0.276	-5.96	0.0000**	0.001	0.85	0.3940
MarrCohab	0.128	5.23	0.0000**	0.000	-0.44	0.6620
VGhealth	-0.024	-0.93	0.3540	-0.001	-0.84	0.4010
Ghealth	-0.010	-0.36	0.7190	0.000	0.16	0.8730
FPhealth	0.002	0.05	0.9570	0.002	1.38	0.1680
LifeSatisf	0.006	0.83	0.4050	0.000	1.07	0.2840
EnglishLang	-0.034	-1.08	0.2820	0.001	0.62	0.5330
VIC	0.011	0.42	0.6750	-0.001	-0.74	0.4610
QLD	0.001	0.02	0.9830	-0.001	-1.39	0.1640
SA	-0.038	-1.01	0.3120	-0.002	-1.76	0.0780*
WA	-0.014	-0.38	0.7020	0.000	0.08	0.9370
TASNTACT	0.024	0.57	0.5710	0.000	-0.07	0.9450
FTEducation	0.061	2.01	0.0450**	-0.002	-1.86	0.0630*
Non-Manual	-0.021	-0.90	0.3680	0.001	0.91	0.3610
UnivEd	0.036	1.32	0.1880	0.002	1.65	0.0980*
Yr12plusEd	0.013	0.55	0.5800	0.002	2.07	0.0390**
Duration	0.046	1.32	0.1870	0.002	1.98	0.0470**
DurationSq	-0.021	-0.87	0.3850	0.000	-0.56	0.5760
1Job	0.512	10.89	0.0000**	0.273	26.40	0.0000**
2Jobs	0.684	12.23	0.0000**	0.382	16.48	0.0000**
3+Jobs	0.681	12.33	0.0000**	0.368	10.75	0.0000**
EmplCalend	0.010	2.80	0.0050**	0.000	2.06	0.0400**
1-2Children	0.048	1.48	0.1400	0.002	1.35	0.1770
3+Children	0.045	1.22	0.2240	0.003	1.72	0.0850*
SeconQuint2	-0.079	-2.54	0.0110**	0.000	-0.07	0.9450
SeconQuint3	0.012	0.34	0.7340	-0.001	-0.57	0.5650
SeconQuint4	0.005	0.16	0.8740	0.000	0.25	0.8030
SeconQuint5	0.020	0.62	0.5380	0.001	0.52	0.6020
log-L (restr)		-1794.80			-1681.89	
log-L (unrestr)		-1533.46			-900.84	
Pseudo RSq		0.1456			0.4644	
ChiSq		353.45			1270.92	
Observations		2739			14781	

*denotes statistical significance at the 10% level **denotes statistical significance at the 5% level

Finally, Table 9 provides an overview of the results from all five probit estimations. This shows more clearly how different characteristics may have different associations with the four different types of mistakes which respondents can make. Some of these characteristics are also associated with making any mistake, that is, the pooled sample (All Mistakes), while other characteristics which are associated with a specific type of mistake show no association for those in the pooled sample or are associated but the sign changes. For example, of the age groups, there is a positive association between making any mistake and being in the Age 15 to 24 year group yet all age groups are negatively associated with making Mistake 2 or Mistake 3. It would appear that due to small sample sizes for both Mistakes 2 and 3 and large sample sizes for Mistakes 1 and 4 which are both positively associated with the 15 to 24 year age group, the result found in the first estimation for the pooled sample (All Mistakes) is being driven primarily by Mistakes 1 and 4.

Table 9: Overview of estimation results

		Mis-reporting Past Paid Employment Status		*Mis-reporting Past Not in Paid Employment Status*	
	All Mis-reporting	*Mistake 1*	*Mistake 2*	*Mistake 3*	*Mistake 4*
Male	Not Sig	Not Sig	Not Sig	++	Not Sig
Age15-24	++[1]	+	-	--	++
Age25-34	Not Sig	Not Sig	--	--	Not Sig
Age35-44	Not Sig	Not Sig	--	--	Not Sig
Age45-54	Not Sig	Not Sig	--	--	++
Age55-64	Not Sig	Not Sig	--	--	Not Sig
MarrCohab	Not Sig	-	-	++	Not Sig
VGhealth	Not Sig	Not Sig	Not Sig	Not Sig	Not Sig
Ghealth	Not Sig	+	Not Sig	Not Sig	Not Sig
FPhealth	Not Sig	Not Sig	Not Sig	Not Sig	Not Sig
LifeSatisf	Not Sig	Not Sig	Not Sig	Not Sig	Not Sig
EnglishLang	Not Sig	--	Not Sig	Not Sig	Not Sig
VIC	Not Sig	--	Not Sig	Not Sig	Not Sig
QLD	Not Sig	Not Sig	Not Sig	Not Sig	Not Sig
SA	Not Sig	Not Sig	Not Sig	Not Sig	-
WA	Not Sig	--	Not Sig	Not Sig	Not Sig
TASNTACT	Not Sig	Not Sig	Not Sig	Not Sig	Not Sig
FTEducation	++	+	Not Sig	++	--
Non-Manual	Not Sig	Not Sig	Not Sig	Not Sig	Not Sig
UnivEd	Not Sig	--	Not Sig	Not Sig	+
Yr12plusEd	Not Sig	--	Not Sig	Not Sig	++
Duration	Not Sig	-	--	Not Sig	Not Sig
DurationSq	++	++	++	Not Sig	Not Sig
1Job	++	--	--	++	++
2Jobs	++	--	--	++	++
3+Jobs	++	--	--	++	++
EmplCalend	++	Not Sig	Not Sig	++	++
1-2Children	++	Not Sig	Not Sig	Not Sig	Not Sig
3+Children	++	Not Sig	Not Sig	Not Sig	+
SeconQuint2	-	Not Sig	-	--	Not Sig
SeconQuint3	Not Sig	Not Sig	--	Not Sig	Not Sig
SeconQuint4	Not Sig	Not Sig	-	Not Sig	Not Sig
SeconQuint5	Not Sig	Not Sig	Not Sig	Not Sig	Not Sig
Empl to Unempl	++				
Unempl to Empl	++				
Unempl to Unempl	++				

[1] ++/--denotes statistical significance at the 5% level. +/- denotes statistical significance at the 10% level.

Conclusion

The aim of this paper was to investigate whether there is recall bias present in the HILDA data, with particular reference to recalling past employment status correctly or not, concentrating on the following three questions: (i) to what extent do respondents remember past events incorrectly; (ii) are particular types of events recalled incorrectly; and, (iii) do certain types of people respond in different ways? Results show that 5.4% of respondents of the total sample, over 5 years, incorrectly recall their employment status at the date of their last interview when asked in their 'current' interview.

Descriptive analysis shows that there appear to be clear differences in which respondents are making a mistake and this seems to depend on whether respondents' employment state has changed or remained stable between interviews. Of the four possible types of mistakes respondents could make (see Figure 2), the majority of mistakes are made by those whose employment state has changed between interviews.

The multivariate analysis of the probability that a respondent makes any mistake shows that being in full time education, the number of children, the number of jobs in the last financial year, possibly the time elapsed between interviews (this is not entirely clear) and the number of jobs reported in the employment calendar, are all factors statistically significantly associated with the probability of making a mistake. Above all, the combination of employment states at interview *t* and interview *t+1* are, by far the most significant factors regarding the probability of making a mistake. This last result motivated estimating all the four possible different types of mistakes separately.

The differences in the results between the different types of mistakes vindicates this strategy and highlights that respondents have a much higher probability of making a mistake when being in particular employment states at interview *t* and interview *t+1*. It also finds that the four sub-samples (one for each of the four possible employment states individuals can be in from one interview to the next) have two similarities. Firstly, Mistakes 1 and 4 concern individuals who are in the same employment state in both interview *t* and interview *t+1* (that is, in paid employment in both interviews or not in paid employment in both interviews, respectively), while Mistakes 2 and 3 concern individuals who are in different employment states at interview *t* and interview *t+1*. The two groups in stable employment states (Mistakes 1 and 4) comprise, by far, the largest group and had a very low probability of making a mistake. The two groups of changed employment state (Mistakes 2 and 3) are much smaller but have a very high probability of making a mistake. Furthermore, the factors that influence these mistakes are different. The patterns of association between the probability of making a specific mistake (that is, Mistake1, 2, 3, or 4) is found to differ.

It would appear that with a few exceptions (most notably for males and those who are married/cohabitating who make Mistake 3) the results of the first estimation (All mistakes) are being driven by the magnitude of the effect of a factor within one or a combination of the four sub-samples. For example, there is a negative association between number of jobs and making Mistake 1 or Mistake 2. There is a positive association between number of jobs and making Mistake 3 or Mistake 4. The magnitude of the positive associations is far greater than that of the negative associations in Mistake 1 and Mistake 2. Hence the net effect of the association between number of jobs and All Mistakes is positive and a weaker effect.

In conclusion, this paper finds that there is systematic employment recall bias present in the HILDA data, the most important factor associated with this bias being the employment state that individuals are in at interview *t* and interview *t+1*. Further, this paper finds that the factors associated with the probability of an individual making a mistake will change, both in magnitude and direction, depending on their exact employment state at interview *t* and interview *t+1*.

Some further research on employment recall bias is needed in two areas. Firstly, out of all respondents, some 30% make more than one mistake. The nature of repeated mistakes is an issue that needs to be investigated further. However, the small number of individuals who make more than one mistake contained in the four interview-to-interview periods limits any robust investigation. Secondly, further research into matching individuals who make a mistake with how those individuals recall employment spells in the employment calendar would be useful. Findings from this piece of research could then be combined with the results of this paper, thus providing a more accurate picture of employment recall bias in the data.

References

Akerlof, G. A and Yellan, J. (1985), 'Unemployment through the Filter of Memory', *Quarterly Journal of Economics* 100: 747-773.

Brewer, M. and Paull, G. (2005), 'The Consistency and Reliability of the Activity History Data in the Families and Children Study (FACS)', Working Paper No 25, Department for Work and Pensions,UK.

Crossley, T. F. and Kennedy, S. (2000), 'The Stability of Self Assessed Health Status', Discussion Paper No. 421, Economics Program, Australian National University.

Elias, P. (1987), 'Who forgot they were unemployed', ISER Working Paper 97-19, Institute for Social and Economic Research, Essex, UK

Horvath, F. (1982), 'Forgotten Employment: recall bias in retrospective data'. *Monthly Labour Review* Vol. 105:40-43.

Jũrges, H. (2005), 'Unemployment, Retrospective Error, and Life Satisfaction', MEA Discussion Paper Series, No. 05089, Mannheim Research Institute for the Economics of Aging, University of Mannheim.

Lynn, P., Buck N., Burton J., Jäckle, A., and Laurie, H. (2005), 'A Review of Methodological Research Pertinent to Longitudinal Survey Design and Data Collection.' ISER Working Paper No. 2005-29, Institute for Social and Economic Research, Essex, UK.

Martini, A. (1989), 'Seam Effect, Recall Bias, and the Estimation of Labor Force Transition Rates from SIPP, *1989 Proceedings of Survey Research Methods Section of the American Statistical Association,* 387-392. http://www.amstat.org/sections/srms/proceedings

Paull, G. (2002), 'Biases in the Reporting of Labour Market Dynamics', No. Working Paper 02/10, The Institute for Fiscal Studies, UK.

Watson, N. and Wooden, M. (2004), 'Assessing the Quality of the HILDA Survey Wave 2 Data', HILDA Project Technical Paper Series No.5/04, Melbourne Institute for Applied Economic and Social Research, University of Melbourne.

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

Appendix

Table A1a: Explanatory variable descriptions

AGE	Six binary variables for age groups: 15-24, 25-34, 35-44, 45-54, 55-64 and =>65. Min age:15 Max100 years
ELAPSED TIME BETWEEN INTERVIEWS (duration)	Continuous variable: the date of interview for year *(t+1)* subtracted from the date of interview for year *(t)*, in days. A variable for *Days Squared* generated. Minimum and maximum length of time between interviews:183 to 1610 days
GENDER	Binary variable: 1 for males, 0 for females.
MARITAL STATUS	Binary variable: value of 1 for married or cohabitating, and 0 for all others(separated, divorced, widowed and single).
LIFE SATISFACTION	Continuous variable: value of 1 "most dissatisfied" to 10 "most satisfied".
REGION	Six binary variables: Tasmania, Northern Territory and the ACT were collapsed into a single variable due to small cell counts.
ENGLISH AS FIRST LANGUAGE	Binary variable: Value of 1 for those who were born in Australia or those born in an English speaking country, 0 for all others.
FULL TIME EDUCATION	Binary variable:value of 1 if in full time education, 0 for those not in full time education.
HIGHEST EDUCATIONAL ACHIEVEMENT	Three binaryvariables: (i) university educated, (ii) educated to Year 12 or Certificate 1, 11&111 (iii) educated to Year 11 or below. All take the value 1 for those in the category and 0 otherwise.
THE NUMBER OF JOBS HELD IN THE LAST FINANCIAL YEAR	Four binary variables: (i) no jobs, (ii) one job, (iii) two jobs and (iv) more than two jobs. All take the value 1 for those in the category and 0 otherwise.
NUMBER OF CHILDREN RESPONDENTS HAVE	Three binary variables: (i) no children, (ii) one or two children; (iii) three or more children.
OCCUPATION CODING (2-digit ASCO)	Binary variable: value of 1 if non-manual, 0 otherwise
SEIFA 96 DECILE OF INDEX OF RELATIVE SOCIO-ECONOMIC DISADVANTAGE	Collapsed into quintiles due to small cell counts and 5 binary variables created. All take the value 1 for those in the category and 0 otherwise
GENERAL HEALTH	Four 4 binary variables: excellent, (i) very good, (ii) good, (iii) fair and very poor. Fair and very poor were collapsed into one category due to small cell counts. Takes the value 1 if a respondent belongs to that category and 0 otherwise.
NUMBER OF TIME PERIODS FILLED IN EMPLOYMENT CALENDAR	Continuous variable:range from 39 to 54 periods (1/3 month-long)

Table A1b: Explanatory variables – Descriptive Statistics

Variable	*Mean*	*Min*	*Max*	*Total N. Obs.*
Age groups in years				
Age 15 to 24 *(age15to24)* [Mean group age 19.7]	0.146	0	1	47215
Age 25 to 34 *(age25to34)* [Mean group age 29.9]	0.168	0	1	47215
Age 35 to 44 *(age35to44)* [Mean group age 39.6]	0.212	0	1	47215
Age 45 to 54 *(age45to54)* [Mean group age 49.2]	0.181	0	1	47215
Age 55 to 64 *(age55to64)* [Mean group age 59.1]	0.133	0	1	47215
Age 65 or over *(age65over)* [Mean group age 73.8]	0.157	0	1	47215
Time elapsed b/t interviews in days				
Duration/100 (standard deviation)	3.771 (0.97)	1.83	16.10	47215
Duration Square/100,000 (standard deviation)	1.516 (1.33)	0.335	25.92	47215
Gender *(gender)* Male=1	0.475	0	1	47215
Marital Status *(marstat)**	0.636	0	1	47215
Life satisfaction *(lifesat)*	7.917	0	10	47193
Regions				
NSW *(nsw)*	0.306	0	1	47215
Victoria *(vic)*	0.249	0	1	47215
Queensland *(qld)*	0.197	0	1	47215
South Australia *(sa)*	0.095	0	1	47215
Western Australis *(wa)*	0.098	0	1	47215
Tasmania/Northern Territory/ACT *(tasntact)*	0.055	0	1	47215
English as first language *(anengf_1)(yes=1)*	0.892	0	1	47195
In full time education *(edfts) (yes=1)*	0.069	0	1	47215
Educated to level of:				
University *(univ) (yes=1)*	0.283	0	1	47215
Year 12 and above *(yr12plus) (yes=1)*	0.335	0	1	47215
Year 11 or less *(yr11less) (yes=1)*	0.382	0	1	47215

* (*variable name*)

Table A1: Explanatory variables (continued).

Variable	*Mean*	*Min*	*Max*	*Total N. Obs.*
Number of jobs held in last financial year				
None *(cafnj_non) (yes=1)*	0.310	0	1	47215
One *(cafnj_one) (yes=1)*	0.534	0	1	47215
Two *(cafnj_two) (yes=1)*	0.121	0	1	47215
More than two *(cafnj_gttwo) (yes=1)*	0.035	0	1	47215
Number of time periods answered in employment calender *(cantp) (standard deviation)*	45.06 (2.65)	39	54	47215
Number of children				
None *(tchad_0) (yes=1)*	0.311	0	1	47215
One or Two *(tchad_2) (yes=1)*	0.379	0	1	47215
Three or more *(tchad_3plus) (yes=1)*	0.310	0	1	47215
Manual/Non-manual Occupation *(nonman_4)*	0.738	0	1	47215
SEIFA index of socio-economic disadvantage				
Quintile 1: hhda2_1 *(yes=1)*	0.20	0	1	47215
Quintile 2: hhda2_2*(yes=1)*	0.20	0	1	47215
Quintile 3: hhda2_3*(yes=1)*	0.20	0	1	47215
Quintile 4: hhda2_4*(yes=1)*	0.20	0	1	47215
Quintile 5: hhda2_5 *(yes=1)*	0.20	0	1	47215
General Health				
gh1_1: (excellent)	0.198	0	1	47215
gh1_2: (very good)	0.320	0	1	47215
gh1_3: (good)	0.322	0	1	47215
gh1_4: (fair and poor=1)	0.160	0	1	47215
Employment Status in interview *t* and *t+1*				
In paid Employment and In paid employment	0.579	0	1	47215
In paid Employment and Not in paid employment	0.493	0	1	47215
Not in paid Employment and In paid employment	0.580	0	1	47215
Not in paid Employment and Not In paid employment	0.313	0	1	47215

Table A2: Interview-to-interview employment status: by gender and mean age

Interview -to- interview			***E1***		***E2***		***E3***		***E4***		*Total*
			N	%	N	%	N	%	N	%	
1-2	***Gender***	*Male*	4406	54.6	320	41.5	419	44.1	1727	38.1	14323
		Female	3659	45.4	451	58.5	532	55.9	2809	61.9	
	Mean Age		37.2		37.6		30.1		51.3		
2-3	***Gender***	*Male*	3959	54.4	268	41.7	317	41.9	1462	37.2	12605
		Female	3320	45.6	374	58.3	439	58.1	2466	62.8	
	Mean Age		38.3		37.5		31.6		55.1		
3-4	***Gender***	*Male*	3515	53.9	225	41.9	228	40.1	1260	35.9	11076
		Female	3002	46.1	313	58.1	340	59.9	2193	64.1	
	Mean Age		39.3		42.1		34.8		57.8		
4-5	***Gender***	*Male*	2918	53.2	150	39.3	158	34.0	1079	37.1	9211
		Female	2563	46.8	231	60.7	307	66.0	1805	62.9	
	Mean Age		40.8		42.4		38.7		60.8		
	Mean Age All		38.9		39.9		33.8		56.3		