



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Wooden, M;WATSON, N

Title:

The Household, Income and Labour Dynamics in Australia (HILDA) Survey

Date:

2021

Citation:

Wooden, M. & WATSON, N. (2021). The Household, Income and Labour Dynamics in Australia (HILDA) Survey. *Jahrbücher für Nationalökonomie und Statistik*, 241 (1), pp.131-141. <https://doi.org/10.1515/jbnst-2020-0029>.

Persistent Link:

<https://hdl.handle.net/11343/288854>

# The Household, Income and Labour Dynamics in Australia (HILDA) Survey

Nicole Watson and Mark Wooden\*

Melbourne Institute of Applied Economic and Research, University of Melbourne, Australia

**JEL classification:** C81, R20

**Keywords:** Australia, HILDA Survey, household panel surveys, longitudinal data

## Abstract

This paper provides a brief summary of the Household, Income and Labour Dynamics in Australia (HILDA) Survey, a nationally-representative household panel survey. It describes the survey's key design features, provides an overview of its content, and reports on response rates and sample sizes. It also highlights a few examples of research utilising the data, discusses two challenges currently facing the study, and provides details on how to access the data.

## 1 Introduction

The Household, Income and Labour Dynamics in Australia (HILDA) Survey is a nationally-representative household panel survey that commenced in 2001 with a sample of 13,969 persons from 7682 Australian households. Its design was based on other household panel studies, and especially the British Household Panel Survey (BHPS) and the German Socio-Economic Panel (SOEP). Thus, like them, the HILDA Survey seeks to collect data once a year from all adult members of the originally selected households, any children of those original sample members (once they turn 15), and any other adults who subsequently join those households. And like the older studies, the HILDA Survey was designed with the broad objective of providing researchers and policy-makers a tool for examining a wide range of

---

\* **Corresponding author: Mark Wooden**, Melbourne Institute of Applied Economic and Research, University of Melbourne, Victoria 3010, Australia. Email: m.wooden@unimelb.edu.au.

**Nicole Watson:** Melbourne Institute of Applied Economic and Research, University of Melbourne, Victoria 3010, Australia. Email: n.watson@unimelb.edu.au

economic, demographic, and social policy issues, but with a strong focus on work, income and family. These similarities with other household panel studies, both in terms of design and in content, mean the HILDA Survey data is potentially of great value for cross-national research purposes, and indeed there are already a number of examples of research where the HILDA Survey data has been used in tandem with data collected in other countries.

The study is funded by the Australian Government, currently through its Department of Social Services (DSS), but responsibility for the design and administration of the survey and for the preparation of data for release rests with the Melbourne Institute of Applied Economic and Social Research, a research-only department within the University of Melbourne. It, in turn, subcontracts fieldwork services out to a private market research company – since 2009 that role has been filled by Roy Morgan Research.

This paper provides a short descriptive summary of the study. It: (i) briefly describes the key features of the survey; (ii) provides a broad overview of the subject areas covered; (iii) reports on the achieved sample sizes and response rates; (iv) highlights a few examples of research utilising the data; (v) discusses challenges currently facing the study; (vi) provides details on how to access the data; and (vii) directs readers to other sources where more detailed documentation can be found.

## **2 Key Features**

### **2.1 Design**

Like other long-running household panel studies, such as the Panel Study of Income Dynamics (PSID), SOEP and the BHPS (and its successor, the UK Household Longitudinal Study [UKHLS]), the HILDA Survey is a household panel survey with an indefinite life design. Thus, unlike the more common longitudinal cohort study, the sample for the HILDA Survey is continually replenished over time through the recruitment of children born to existing sample members.

Indeed, all persons who move into a household where an original sample resides are added to the sample each wave. These new sample members, however, are only added to the sample on a permanent basis if they have a child with an original sample member or if they are the child of an original sample member. The only exception to this is recent immigrants: from wave 9, new household members who arrived in Australia for the first time after 2001 were also added to the sample on a continuing basis. This rule was

subsequently amended, in wave 16, to only cover immigrants arriving after 2011 (reflecting the timing of the introduction of a refreshment sample).

## **2.2 Samples**

The initial sample comprised private households residing in private dwellings in 2001. These were selected using a multi-stage approach. First, a sample of 488 Census Collection Districts (CDs) was selected from across Australia (each of which consists of approximately 200 to 250 households). Second, within each of these CDs, all dwellings were enumerated and a sample of 22 to 34 dwellings selected (the precise number depending on the expected response and occupancy rates of the area). Finally, within each dwelling, up to three households were selected to be part of the sample. After excluding addresses subsequently discovered during the fieldwork to be out of scope (because the dwelling was vacant, not a primary private residence, or none of the occupants met the selection criteria – most importantly, occupants had to have resided, or expected to reside, in Australia for at least 12 months), this process resulted in a total of 11,693 households being approached in field

A top-up (or refreshment) sample was added in 2011, with a further 125 CDs selected. This resulted in an additional 3117 in-scope households approached in field.

## **2.3 Data collection**

Data are collected once a year via interview and a self-administered paper form – the self-completion questionnaire (or SCQ).

The interviews involve two components – one component administered to just one person in the household, which in a normal year averages around 12 to 13 minutes in duration, and the other administered to all household members aged 15 years and older, and normally averages around 35 minutes. Most interviews are conducted face-to-face, usually within the respondent's home. The telephone, however, is used for persons who reside outside the range of our interviewer network, where respondents express a strong desire for a telephone interview, or as a method of last resort. In recent years, telephone interviews have represented around 8 to 10% of all person interviews conducted.

The SCQ is a 20-page form that is given to all persons completing a personal interview. It is often completed and collected at the time of interview. Where this is not possible, the

interviewer is required to return to every household at least once to collect the completed forms. In the case of telephone respondents, the SCQ is administered via mail.

The fieldwork period spans late July to early February the following year, but with about 95% of interviews completed between August and November.

## **2.4 Incentives**

Participation is incentivized. Since wave 5 this has taken the form of a payment for each interview completed plus a bonus payment if everyone in the household provides an interview.<sup>1</sup> The amount of the incentive has increased over time, from AU\$25 in wave 5 to AU\$40 by wave 19. The form of the payment has also changed. Prior to wave 9 the payment was delivered well after interview as a cheque in the mail. From wave 9 the payment for those interviewed in person has been provided as cash at the time of interview.

## **3 Content**

### **3.1 Household interview**

This component involves two separate instruments – the Household Form (HF) and the Household Questionnaire (HQ).

The HF establishes the composition of the household at the time of interview and thus is where both household leavers and joiners are identified, and where relationships between household members are recorded. It also asks a small number of questions identifying some basic characteristics of all household members (e.g., sex, date of birth, and employment status).

The HQ is focused primarily on the use of childcare services and on housing and housing costs. Additionally, rotating content is included every four years on household wealth, material deprivation, the health of children in the household, and children's education.

### **3.2 Person interview**

This component also involves two components – the New Person Questionnaire (NPQ) and the Continuing Person Questionnaire (CPQ).

---

<sup>1</sup> In waves 1 to 4, an incentive was provided at the household level only: AU\$20 for a part household response (i.e., at least one but not all adults responded) and AU\$50 for a full household response.

The NPQ, as its name suggests, is administered to those participating in the survey for the first time. It includes all the content included in the CPQ, as well as additional questions about aspects of a person's history. This includes sequences on (among other things): selected demographic characteristics (e.g., country of birth, year of arrival, whether an Indigenous Australian, and whether English is the first language); family background; educational attainment; employment history; and marital history.

The CPQ comprises a stable core of questions administered every year and a series of rotating modules that are now repeated every four years. The core includes sections on:

- education, where we seek to update each respondent's educational attainment;
- employment status, which provides the key questions critical for identifying both labour force status (employed, unemployed or not in the labour force) and employment status (employee or self-employed);
- current employment, where details about a range of job characteristics are collected;
- persons not in employment, which is focused on the job search activity of those not in work;
- other labour market activities, which includes a calendar recording basic labour market activity since the 1<sup>st</sup> July of the preceding year, as well questions about work-related training, and use of employment-related leave;
- income, which seeks details on each source of individual income over the preceding financial year, as well as current income from wages and government benefits and pensions;
- family formation, which is predominantly concerned with the amount of contact children have with non-resident parents, the amount of income support received or paid in respect of children, and fertility intentions;
- partner relationships, including both marital and de facto partners; and
- "living in Australia", which ranges over a number of issues including long-term health conditions and disability, caring, and life satisfaction.

The four-yearly rotating modules or sequences are concerned with: household wealth; health; education, skills and abilities; and a multi-topic cluster focused on family formation, non-coresidential relationships, and retirement. A noteworthy feature of the education rotation is the inclusion of three short tests of cognitive ability (see Wooden 2013).

### 3.3 SCQ

The SCQ consists mainly of questions which are difficult to administer in a time-effective manner in a personal interview, or which respondents may feel slightly uncomfortable answering in a face-to-face interview. The types of topics covered each year include: health status (the SF36 health survey); lifestyle behaviours and outcomes, such as smoking, exercise, alcohol consumption, and height and weight; relationship satisfaction; social interaction and support; time use; life events; financial stress; and, since wave 5, household expenditure. Other topics appear on a less frequent basis, examples of which include: psychological distress (Kessler 10); religion; neighbourhood characteristics; locus of control; participation in community activities; gambling; illicit drug use; and personality traits.

## 4 Response and Sample Sizes

Of the 11,693 in-scope household approached in wave 1, interviews were obtained with 7682 households, resulting in an initial wave response rate of 66%. A higher response rate of 69% was achieved for the top-up sample added in 2011, with interviews obtained from 2153 households of the 3117 in-scope households approached.

Figure 1 shows the re-interview rates of the previous wave respondents (dashed lines) along with the re-interview rates of respondents to the first, or initial, wave (solid lines) for each wave for both the original, or main, sample and the top-up sample.<sup>2</sup> The re-interview rate of previous wave respondents in the original sample rises from 87% at wave 2 to over 96% for wave 9 onwards, rates which compare very favourably with those reported for other longitudinal studies (see Watson et al. 2019). However, despite the relatively low rates of attrition, sample losses cumulate over time: at wave 18, 62% of the original wave 1 sample members were interviewed. For the top-up sample, re-interview rates of previous wave respondents have reached 95% by wave 18. Also 76% of the wave 11 top-up sample respondents were interviewed in wave 18 (compared to 72% in the main sample at wave 8 – the equivalent point of sample development).

---

<sup>2</sup> Sample members found to be out of scope (due to death, moving abroad, or temporary sample members moving out of sampled households) are not included in the denominator when calculating these re-interview rates.

Turning now to the overall sample size, a total of 270,611 individual interviews and 243,289 SCQs have been completed over the first 18 waves of the HILDA Survey. Figure 2 shows the number of interviews completed and number of SCQs returned. With the introduction of the top-up sample, the total number of individuals interviewed increased from around 13,500 interviews to 17,600 interviews per wave. Note also that while the number of interviews achieved each wave fell until wave 5, since then gains to the sample from new sample members have outweighed the losses due to attrition. The SCQ response rate in wave 1 was 94%, falling to a low of 87% in wave 8 but, following the introduction of various initiatives designed to improve the response rate, increased to around 91% to 92% in later waves.

## **5 Research Uses**

With 18 waves of data collection behind it, it is impossible to summarise in this short article the contribution the HILDA Survey data has been making to research; the volume of output is just too great. The number of persons who are approved to use the unit-record data each year now exceeds 850, and the total number of known papers using the data that have been published in academic journals exceeds 1200. And on top of that is the countless (but almost certainly, much larger) number of book chapters, reports, policy briefs, discussion papers, conference papers and the like.

Like its German cousin, the SOEP (Goebel et al. 2019), major areas of investigation have included the life course, inequality and mobility, and psychological outcomes (and especially subjective well-being). Labour market outcomes, such as labour force status, hours of work and wages, have also been much studied. In some cases, data use reflects the peculiar nature of Australian institutions (e.g., the many papers on casual employment), whereas in other cases it is a function of the HILDA Survey filling significant data gaps (e.g., household wealth, personality traits).

As might be expected, the majority of researchers are based in Australia. Nevertheless, a significant fraction of licensed users are based overseas. Further, while most research relies solely on the HILDA Survey data, there are numerous papers where the HILDA Survey data provides the Australian input into an international comparison. Examples here include research into the stress cost of children (Buddelmeyer et al. 2018), life satisfaction (Headey and Muffels 2017), health inequality (Schurer et al. 2014), the correlates of personality

(Hakulinen et al. 2015a, 2015b), non-standard employment (OECD 2015), and the association between hours of work and mental health (Otterbach et al. forthcoming). That said, in our view the HILDA Survey data is still relatively under-exploited in comparative cross-country research.

Finally, as the panel continues to mature, one area where scope for new research should expand is intergenerational mobility and disadvantage (see Murray 2018 and Mooi-Reci et al. 2020 for early examples of research on intergenerational issues using HILDA Survey data).

## 6 Challenges

Like other surveys, the COVID-19 pandemic, has posed major challenges for the HILDA Survey. Social distancing requirements imposed by governments, University directives to cease fieldwork activity involving face-to-face contact with participants, and an expected reduced willingness on both the part of interviewers and interviewees to be involved in face-to-face interviews have led to a change in survey mode. In 2020 we are switching to complete reliance on telephone interviews, at least for phase 1 of fieldwork (with the expectation that it will continue to be the dominant mode in phase 2, which commences in October). This, in turn, will have ramifications for the collection and return of SCQs, which in the normal course of events would now be returned through the mail, with likely serious adverse consequences for return rates.<sup>3</sup> To help ameliorate this, we are thus providing respondents, for the first time, with a web-based version of the SCQ for completion online.<sup>4</sup>

In the longer term, the bigger challenge is the need to recruit immigrants in order to maintain the cross-sectional representativeness of the sample. Immigrants who arrived in Australia since the HILDA Survey began in 2001 were included as part of the 2011 top-up sample, but another refreshment sample is needed to cover those who have arrived since 2011. Australia is a high immigrant country, so this problem cannot simply be ignored. Since 2011 there have been around 200,000 permanent immigrants arriving in Australia each year (which represents 0.8-0.9% of the Australian population), as well as a further 375,000 to 400,000 immigrants arriving on a temporary basis, such as students and skilled

---

<sup>3</sup> On average, only around 61% of respondents interviewed by telephone return the SCQ. In contrast, completed SCQs are received from over 92% of those interviewed face-to-face.

<sup>4</sup> For further details about how the HILDA Survey has responded, and is adapting, to the COVID-10 pandemic, see Wooden (2020).

migrants (Phillips and Simon-Davies 2017). We are currently exploring options for adding quality samples of recent immigrants, preferably every wave.

## **7 Data Access**

Unit record data files are released early December each year. There are two versions of each annual data release: (i) the General Release, which is available to any bona fide researcher (including graduate students), both in Australia and elsewhere; and (ii) the Restricted Release, which is only available to Australian users. The main differences between the two is that the latter contains dates of birth and postcodes of residence, neither of which are included in the General Release.

Prospective users can apply online for access to these datasets through National Centre for Longitudinal Data (NCLD) Dataverse ([dataverse.ada.edu.au/dataverse/nclid](https://dataverse.ada.edu.au/dataverse/nclid)), which, in turn, is managed by the Australian Data Archives (ADA). Obtaining the data is free of charge, but each application is subject to approval by the NCLD, a unit within the Australian Government Department of Social Services. Once approval is granted, users are sent an electronic link that will enable them to download the data directly from the ADA.

Release 18, providing data covering the first 18 waves (2001 to 2018) became available in December 2019.

## **8 More information**

More information about the HILDA Survey can be found in the HILDA Survey User Manual (Summerfield et al. 2019), available from the project website at [melbourneinstitute.unimelb.edu.au/hilda](https://melbourneinstitute.unimelb.edu.au/hilda). This manual provides a much more extensive overview of the survey design, data collection procedures and response metrics than can be provided here. It also provides information about, among other things: matching wave data files to create longitudinal files; missing data conventions; derived variables and coding schemes; imputation procedures for income, wealth and expenditure variables; and the construction and use of weights.

Also available from this site is:

- a set of survey instruments covering every survey wave;

- a searchable online data dictionary, which provides basic information about every variable in the data set (waves in which collected and frequencies); and
- a bibliography of known research publications that have made use of the data.

## References

- Buddelmeyer, H., D.S. Hamermesh, M. Wooden (2018), The Stress Cost of Children on Moms and Dads. *European Economic Review* 109: 148–461.
- Goebel, J., M.M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, J. Schupp (2019), The German Socio-Economic Panel (SOEP). *Journal of Economics and Statistics* 239(2): 345–360.
- Kakulinen, C., M. Elovainio, G.D. Batty, M. Virtanen, M. Kivimäki, M. Jokela (2015a), Personality and Alcohol Consumption: Pooled Analysis of 72,949 Adults from Eight Cohort Studies. *Drug and Alcohol Dependence* 151: 110–114.
- Hakulinen, C., M. Elovainio, L. Pulkki-Råback, M. Virtanen, M. Kivimäki, M. Jokela (2015b), Personality and Depressive Symptoms: Individual Participant Meta-analysis of 10 Cohort Studies. *Depression and Anxiety* 32(7): 461–470.
- Headey, B., R. Muffels (2017), Towards a Theory of Medium Term Life Satisfaction: Similar Results for Australia, Britain and Germany. *Social Indicators Research* 134(1): 359–384.
- Mooi-Reci, I., M. Wooden, M. Curry (2020), The Employment Consequences of Growing Up in a Dual-Parent Jobless Household: A Comparison of Australia and the United States. *Research in Social Stratification and Mobility* 68: e100519.
- Murray, C., R.G. Clark, S. Mendolia, P. Siminski (2018), Direct Measures of Intergenerational Income Mobility for Australia. *Economic Record* 94(307): 445–468.
- Organisation for Economic Co-operation and Development [OECD] (2015), *In It Together: Why Less Inequality Benefits All*. Paris.
- Otterbach, S., A. Charlwood, Y.-K. Fok, M. Wooden, M. (forthcoming), Working-Time Regulation, Long Hours Working, Overemployment and Mental Health. *International Journal of Human Resource Management*.
- Phillips, J., J. Simon-Davies (2017), *Migration to Australia: A Quick Guide to the Statistics*. Parliamentary Library Research Paper Series, 2016-17. Canberra. Can be downloaded at:

[https://www.aph.gov.au/About\\_Parliament/Parliamentary\\_Departments/Parliamentary\\_Library/pubs/rp/rp1617/Quick\\_Guides/MigrationStatistics](https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/rp/rp1617/Quick_Guides/MigrationStatistics).

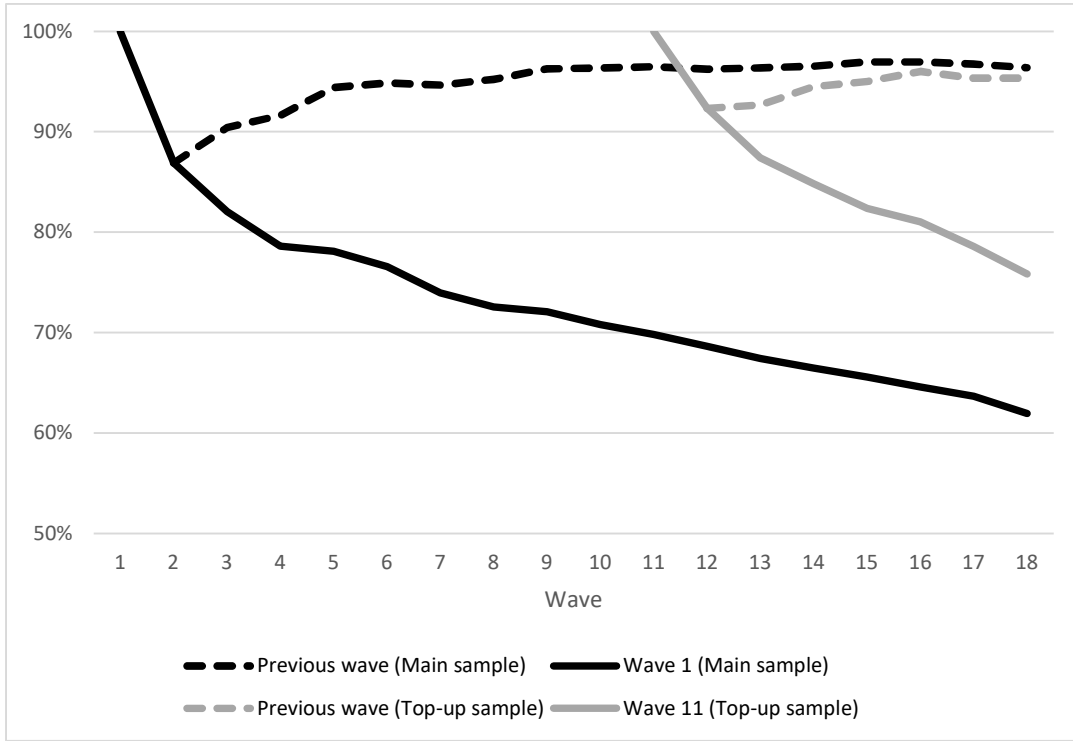
Schurer, S., M.A. Shields, A.M. Jones (2014). Socio-Economic Inequalities in Bodily Pain over the Life Cycle: Longitudinal Evidence from Australia, Britain and Germany. *Journal of the Royal Statistical Society: Series A* 177(4): 783–806.

Summerfield, M., S. Bright, M. Hahn M, N. La, N. Macalalad, N. Watson, R. Wilkins, M. Wooden (2019), HILDA User Manual – Release 18. Melbourne. Can be downloaded at: <https://melbourneinstitute.unimelb.edu.au/hilda/for-data-users/user-manuals>.

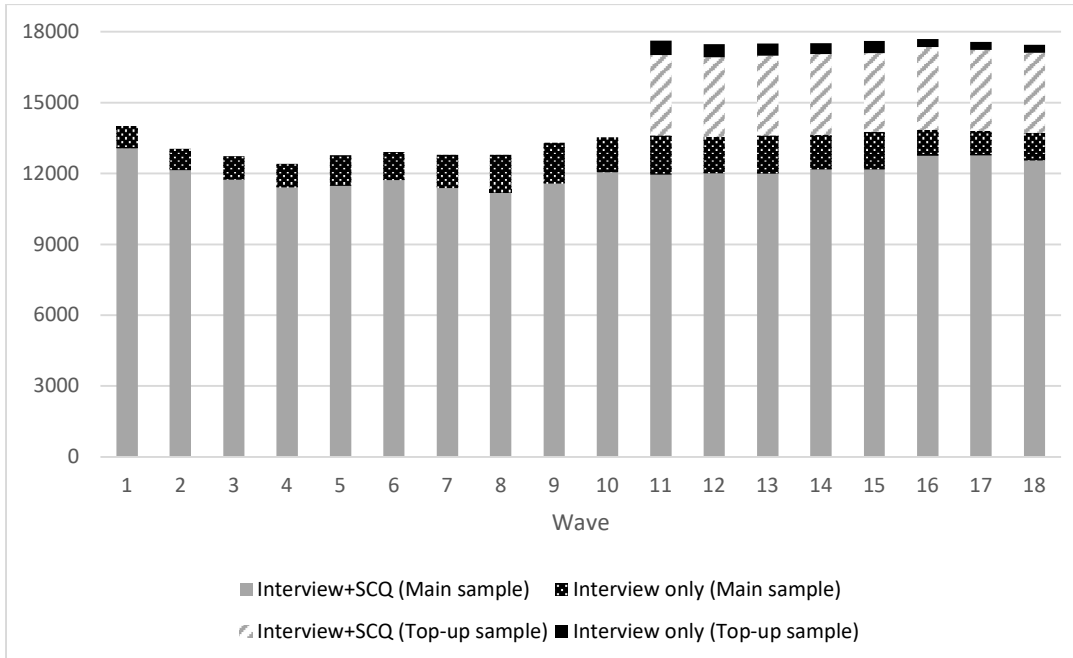
Watson, N., E. Leissou, H. Guyer, M. Wooden (2019), Best Practices for Panel Maintenance and Retention. Pp. 597-622 in: T.P. Johnson, B.-E. Pennell, I.A.L. Stoop, B. Dorer (eds.), *Advances in Comparative Survey Methods: Multicultural, Multinational and Multiregional Contexts (3MC)*. New York.

Wooden, M. (2013). The Measurement of Cognitive Ability in Wave 12 of the HILDA Survey. HILDA Project Discussion Paper Series No. 1/13. Melbourne. Can be downloaded at: <https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/hilda-discussion-papers/hdps113.pdf>.

Wooden, M. (2020). Responding to the COVID-19 Pandemic in the HILDA Survey. HILDA Project Discussion Paper Series No. 1/20. Melbourne. Can be downloaded at: [https://melbourneinstitute.unimelb.edu.au/\\_\\_data/assets/pdf\\_file/0006/3399468/hdps120.pdf](https://melbourneinstitute.unimelb.edu.au/__data/assets/pdf_file/0006/3399468/hdps120.pdf).



**Figure 1:** Re-interview rates of previous wave respondents and initial wave respondents, HILDA Survey, waves 1-18



**Figure 2:** Number of interviews and SCQs completed, HILDA Survey, waves 1-18