

HILDA Project Discussion Paper Series

Wave 21 data quality

Nicole Watson
Mossamet Kamrun Nesa
Michelle Summerfield

HILDA Project
Discussion Paper No. 1/22
November 2022



Contents

Introduction	1
Restrictions during fieldwork	1
Impact on response rates	2
Individual interview.....	2
Self-Completion Questionnaire	4
Impact on mode	6
Individual interview.....	6
Self-Completion Questionnaire	8
Impact on response timing.....	10
Impact on item responses	13
Item non-response	14
Completeness of multi-item responses	17
Primacy versus recency in list-style questions	19
Straight lining	20
Rounding	20
Length of open-ended responses	21
Conclusions	22
References	23
Appendix	24

List of Tables

Table 1: Re-interview rates by region	2
Table 2: Showcard usage (%) during individual interview	7
Table 3: Interviewer report of respondent situation (%)	8
Table 4: Average item non-response (%) by section.....	14
Table 5: Percentage of item non-response for specific variables	17
Table 6: Number of options chosen at multi-item response questions	18
Table 7: Percentage of respondents selecting first option to list-style questions with four or more response options where showcards had been used in waves 18 and 19	19
Table 8: Percentage of respondents providing exactly the same response across all items in multi-item question (with four or more items)	20
Table 9: Percentage of respondents rounding amounts reported for current wages and salaries or Financial Year age pension	21
Table 10: Number of characters recorded in occupation and industry descriptions.....	21

List of Figures

Figure 1: Response rates for individuals associated with previous wave responding households	3
Figure 2: SCQ response rate.....	4
Figure 3: SCQ response rate by age group	5
Figure 4: Proportion of PQs completed via telephone, by region	6
Figure 5: Proportion of SCQs completed online by age group	9
Figure 6: Number of PQ interviews completed by rolling week.....	10
Figure 7: Cumulative number of PQ interviews in Sydney and Melbourne	11
Figure 8: Cumulative number of PQ interviews in rural New South Wales and rural Victoria	11
Figure 9: Distribution of gap between completion of PQ and SCQ	12
Figure 10: Average item non-response rate in each SCQ section, by mode and wave	15
Figure 11: Cumulative percentage of breakoffs when completing the SCQ	16

Introduction

The fieldwork for wave 21 of the Household, Income and Labour Dynamics in Australia (HILDA) Survey, conducted between July 2021 and March 2022, was affected by various lockdowns and restrictions that were in place due to the COVID-19 pandemic. A previous HILDA discussion paper (Watson, Jin and Summerfield, 2021) detailed the changes made to the wave 20 fieldwork procedures in the first year of the pandemic and examined the impact this had on the data quality. For the most part, the data quality in wave 20 was very similar to previous waves. There was a relatively small reduction in the re-interview rate in wave 20 and a small improvement in the amount of missingness in the Self-Completion Questionnaire (SCQ). More concerning was the reduction in response rates for new entrants and children turning 15.

This paper focuses on the data quality for wave 21. While it was initially planned to return to face-to-face interviewing this wave, these plans were changed with the escalation of COVID-19 cases in New South Wales, Victoria and the Australian Capital Territory just prior to fieldwork commencing. Many of the fieldwork changes implemented in wave 20 were therefore continued in wave 21. Chief among these was interviewing primary by telephone rather than face-to-face. As a result, the following features were also retained: i) showcards (in the form of an A5 sized booklet) were mailed to all households and made available online; ii) the SCQ was made available online as well as in hardcopy format; and iii) the incentive offered was \$40 per completed individual interview and \$20 for a completed SCQ.

Further, experienced interviewers were trained via live webinars. New and less experienced interviewers were trained face-to-face in Queensland, Western Australia and South Australia and via live webinars in the remaining states. All interviewers had self-paced learning to complete prior to group training. The interviewer training was also extended by an additional two weeks to allow for smaller webinar groups.

The aspects of data quality examined in this paper include response rates, response timing, missing data, use of multi-item response options, responses chosen for questions with long response lists, straight lining, rounding, and the length of responses provided at open-ended questions. These data quality measures for wave 21 are typically compared to those for waves 18 to 20.

Restrictions during fieldwork

The fieldwork for wave 21 was conducted from 27 July 2021 to 14 March 2022. This overlapped with extended lockdown periods in Victoria, New South Wales and the Australian Capital Territory:

- Melbourne lockdown (#5) from 16 July to 27 July 2021 (overlapping with day 1 of fieldwork).
- Melbourne lockdown (#6) from 5 August to 21 October 2021 (days 10 to 87 of fieldwork).
- Rural Victoria lockdown from 21 August to 9 September 2021 (days 26 to 45 of fieldwork).
- Sydney lockdown from 26 June to 10 October 2021 (overlapping with day 1 to 76 of fieldwork).
- Parts of Rural New South Wales were in lockdown for various durations over the period from 21 July to 10 September 2021 (overlapping with days 1 to 46 of fieldwork).
- Canberra lockdown from 12 August to 14 October 2021 (days 17 to 80 of fieldwork).

During these lockdowns people were only allowed to leave their home to shop for food and other essentials, to go to medical appointments, to get vaccinated, to exercise, and for essential work. Even outside of lockdown periods, a number of restrictions were still in place, such as limits on the number of visitors to a household, density limits at various venues, and wearing masks in certain situations.

Impact on response rates

Individual interview

Telephone interviewing tends to result in lower response rates than face-to-face interviewing (Holbrook, Green and Krosnick, 2003). However, with the COVID-19 pandemic, many sample members would expect to be contacted by telephone, at least during the lockdown periods.

The response rate for previous wave respondents in wave 21 was 94.0%.¹ This is lower than that achieved in wave 20 (by 1.2 percentage points) and lower than what was achieved on average in waves 18 and 19 (by 2.5 percentage points). This is somewhat disappointing given it is the lowest response rate achieved for over 10 years but there are a number of reasons for these differences:

1. It was very unclear how long the 2021 lockdowns would last or under what conditions the restrictions would be eased. It was initially thought that the lockdowns would be relatively short (as occurred earlier in 2021), so some sample members preferred to wait until lockdowns were lifted so that they could be interviewed face-to-face. Unfortunately, when the lockdowns were lifted a portion of these sample members were then too busy with other activities to do an interview.
2. The response rates for Melbourne and rural Victoria (which were very good for 2020) were more impacted by the lockdown in 2021 (see Table 1). The enthusiasm for telephone interviews early in the fieldwork period evident during the lockdowns in Victoria in 2020 (Watson, Jin and Summerfield, 2021) was lacking in 2021 after multiple extended lockdowns.
3. New interviewers were used during wave 21, as is the usual practice, though no new interviewers were used for wave 20. Based on the wave 9 to 19 experience, households approached by a new interviewer in the initial fieldwork period have, on average, response rates that are 1.8 percentage points lower for previous wave respondents compared to where experienced interviewers are used. The actual difference between the response rates for previous wave respondents initially approached by experienced vs new interviewers in wave 21 was similar (1.7 percentage points lower for new interviewers). Nevertheless, this could only explain about 0.2 percentage points of the wave 21 response rate given the fraction of previous wave respondents approached by new interviewers is relatively small (14.7%).

Table 1: Re-interview rates by region

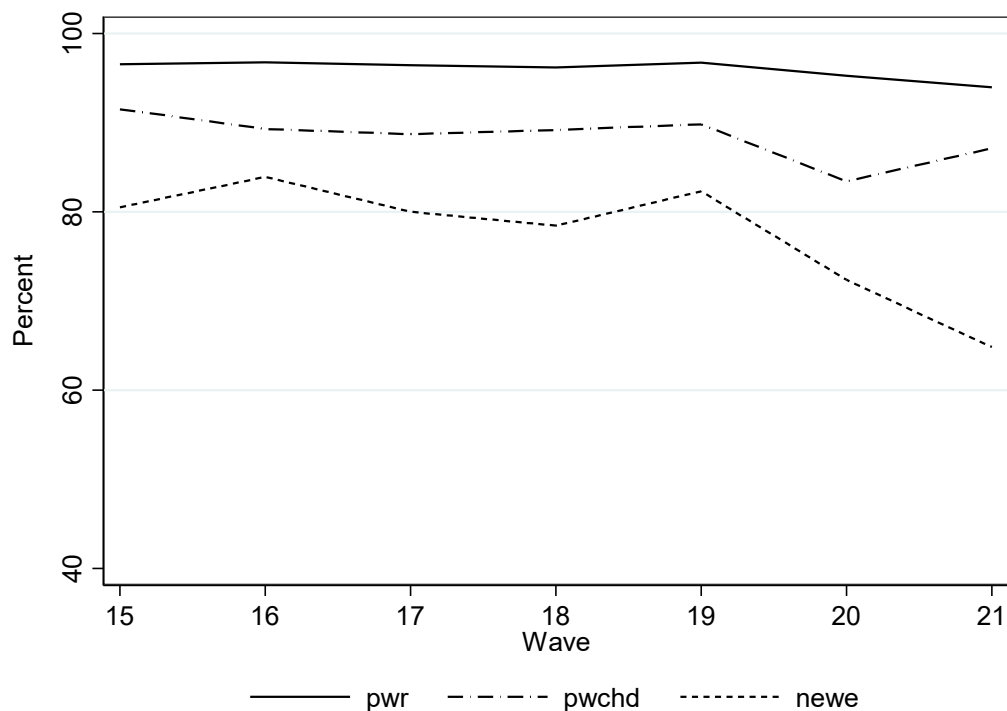
	Wave 18	Wave 19	Wave 20	Wave 21	W20 less ave(W18&19)	W21 less ave(W18&19)
Sydney	95.5	96.4	94.6	93.8	-1.3	-2.1
Melbourne	97.1	97.5	96.9	94.2	-0.3	-3.1
Brisbane	95.9	97.0	95.3	94.7	-1.1	-1.7
Canberra	95.9	96.9	93.8	95.2	-2.6	-1.2
Other major capitals (Adelaide, Perth)	96.8	95.8	95.7	93.2	-0.6	-3.1
Rural NSW	96.2	98.1	94.8	94.8	-2.3	-2.3
Rural Vic	96.0	94.9	95.1	92.9	-0.4	-2.5
Rural Qld	95.1	96.1	95.4	94.0	-0.2	-1.6
Other	96.5	97.5	92.9	93.4	-4.1	-3.6
Total	96.2	96.7	95.2	94.0	-1.2	-2.5

¹ This is calculated as the percentage of respondents in wave 20 that were re-interviewed in wave 21, excluding those that have moved overseas or died.

These geographical differences are eliminated once a range of respondent characteristics are controlled for (see Table A1 in the Appendix for details of the logistic regression model predicting re-interview). Table A1 also shows that people less likely to be re-interviewed are those who are young or old, did not complete the SCQ last wave, have more adults in their household and, for waves 18 and 19, completed the previous wave interview by telephone. Interestingly, in wave 20, those who completed the previous wave interview by telephone were actually more likely to be interviewed, perhaps because they were accustomed to this type of approach. This also occurred in wave 21, but this telephone group from wave 20 is now the vast majority of the sample. Also, in waves 20 and 21 and in difference to earlier waves, people working long hours (55 or more hours per week) and those with long-term health conditions were less likely to participate.

The response rates for two other sample groups are also of particular interest. These are children turning 15 and new entrants. Figure 1 shows these response rates together with those for previous wave respondents (discussed above). We restrict the sample to individuals associated with households responding in the previous wave as these are the most comparable over time. The response rate for children turning 15 in wave 21 is 87.1%, which is an improvement on wave 20 and closer to what is usually achieved. The response rate for new entrants in wave 21, however, is even lower than what was achieved in wave 20 (64.8% compared to 72.4% in wave 20 and 82.3% in wave 19). New entrants, in particular, are harder to interview by telephone as there are fewer opportunities to build rapport with them when interviewing others in the household compared to face-to-face interviewing. The children turning 15 would likely have seen their parents being interviewed face-to-face in wave 19 or earlier.

Figure 1: Response rates for individuals associated with previous wave responding households



Note: pwr=previous wave respondent, pwnr=previous wave non-respondent, pwchd=previous wave child, newe=new entrant.

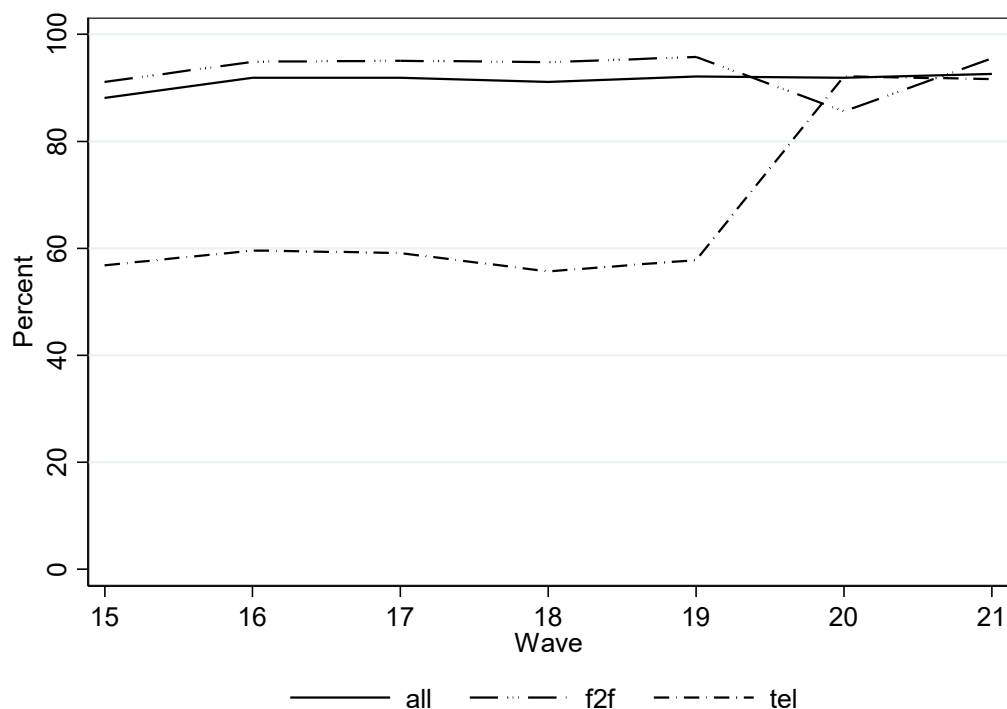
Self-Completion Questionnaire

Of the individuals interviewed in wave 21, 92.6% also completed a Self-Completion Questionnaire. This is the highest response rate achieved since wave 2 when it was 93.0%. Figure 2 shows that those interviewed face to face have a higher SCQ response rate (95.5%) compared to those interviewed by telephone (91.6%). The SCQ incentive and mix of respondents interviewed by telephone in waves 20 and 21 resulted in vastly better response rates than for those interviewed by telephone in earlier waves.²

Figure 3 shows the SCQ response rate by 5-year age group. The introduction of the online SCQ option (and possibly also the SCQ incentive) improved the response rates for people aged 25 to 49.

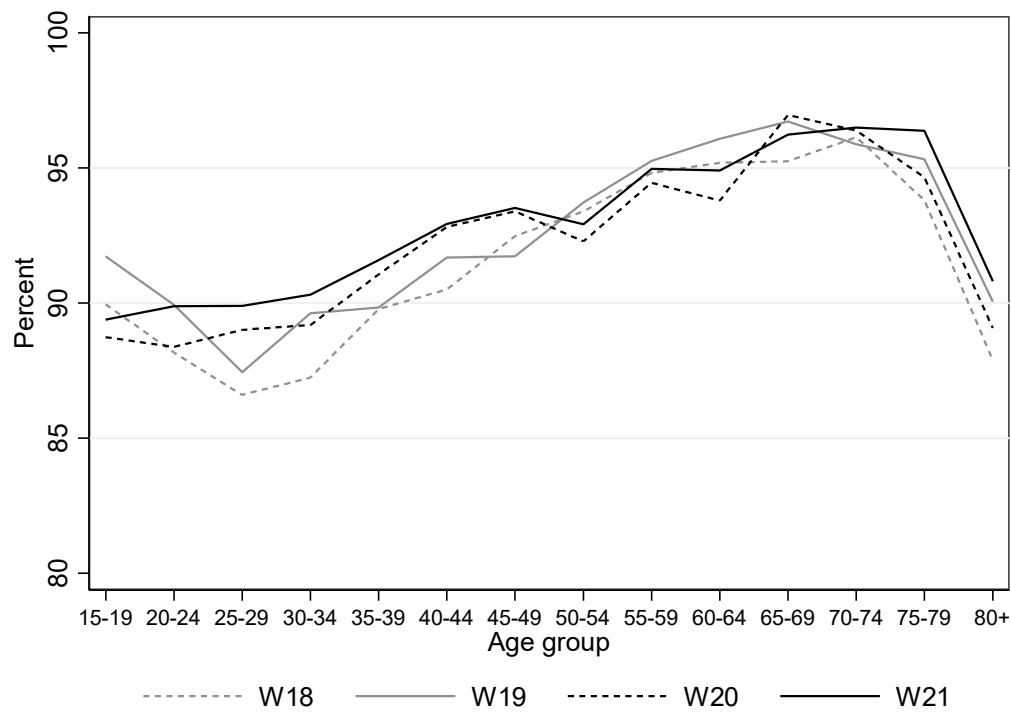
For the most part, the characteristics associated with response to the SCQ in waves 20 and 21 were similar to those observed in waves 18 and 19 (coefficients for a logistic regression model predicting response to the SCQ for waves 18 to 21 are provided in Table A2 in the Appendix). Non-respondents to the SCQ are more likely to be male, born overseas in a non-English speaking country, not highly educated, employed full-time, or living with a long-term health condition. In terms of their household characteristics, SCQ non-respondents are more likely to be renting, living with children under 15, living in Sydney, Melbourne or remote areas, or part of a partially responding household. In waves 20 and 21, there was no response differential for the SCQ in terms of age, but in other waves the young were less likely to complete the SCQ. Also, in difference to earlier waves, we find that in wave 20, those who were separated, divorced or widowed and the unemployed were less likely to complete the SCQ. Of these four factors, those who were separated continued to be less likely to complete the SCQ in wave 21.

Figure 2: SCQ response rate



² Face to face interviews were only undertaken with 4% of respondents in wave 20 and 24% of respondents in wave 21 (as discussed later) whereas in earlier waves the rate was 90% or more.

Figure 3: SCQ response rate by age group



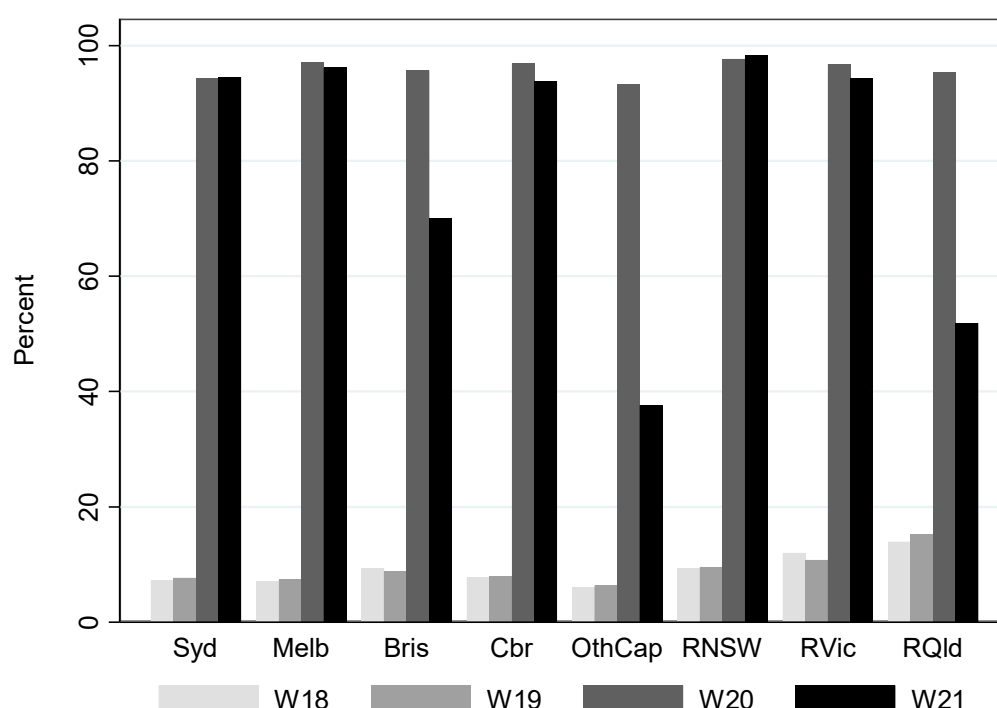
Impact on mode

Individual interview

Use of mode

During the lockdown periods in 2021, interviews were conducted exclusively by telephone. In other areas (and when COVID-19 restrictions permitted in the states affected by the lockdowns), interviewers were directed to make initial contact attempts by telephone and offer respondents the choice of a face-to-face or telephone interview. On average, 75.7% of interviews were undertaken by telephone.³ The rate of telephone interviews was obviously much higher in the states affected by lockdowns (i.e., New South Wales, Victoria and the Australian Capital Territory) as shown in Figure 4. The rates in these areas are similar to the wave 20 experience. In the other parts of the country, the telephone use in wave 21 varied from 38 to 70% of the interviews.

Figure 4: Proportion of PQs completed via telephone, by region



Note: Syd=Sydney; Melb=Melbourne; Bris=Brisbane; Cbr=Canberra (ACT); OthCap=Other capital city (Adelaide; Perth); RNSW=Rural New South Wales; RQld=Rural Queensland; OthR=Other rural areas.

To understand the characteristics of the respondents associated with completing the PQ by telephone (versus a face-to-face interview), a logistic regression model was fitted to the data for waves 18 to 21. The coefficients for this model are provided in Table A3 in the Appendix.

In waves 18 and 19, respondents were more likely to complete the interview by telephone if they were in their 20s and 30s, living in a flat, more highly educated, living in regional or rural areas, living at a new address since the last wave, or a member of a partially responding household. They were less likely to complete the interview by telephone if they were born in a country where the main language spoken was not English, working part-time, or living in large households.

³ In wave 19 and earlier, telephone interviews were used where the respondent lived outside the range of our interviewer network, where the respondent express a strong desire for a telephone interview, or as a method of last resort.

The majority of these factors continued to be associated with telephone interviews in waves 20 and 21 even though telephone became the primary mode of interview. However, there were some differences. Respondents were less likely to be interviewed by telephone in wave 20 or 21 if they rented or had a serious long-term health condition. Also, there was no association between being interviewed by telephone and working part-time hours or moving between waves. In wave 21, respondents were more likely to be interviewed by telephone if they lived in Melbourne or Sydney, as expected given the restrictions. Further, and just for wave 20, there was no association between being interviewed by telephone and working long hours, living in households with a larger number of adults, or being a member of a partially responding household.

Showcards

The way the showcards were made available to the respondents in wave 20 and 21 was different to earlier waves. Previously, the showcard booklet was provided by the face to face interviewer for the interview and the telephone interviews did not use showcards. For waves 20 and 21, a smaller showcard booklet was mailed to the sample households or provided online. Table 2 shows the percentage of individual interviews where the showcards were used by mode of interview. There was greater use of the online showcards in wave 21 than in wave 20, but the proportion of telephone interviews using showcards declined from 68.8% in wave 20 to 64.1% in wave 21. When showcards were not used, the question text that the interviewer read out was automatically modified for this context.

To examine the respondent characteristics associated with using showcards, a logistic regression model was fitted to the PQ respondents in wave 20 and 21 (see Table A4 in the Appendix). In general, respondents not using showcards were more likely to be male, born in a non-English speaking country, not married, working long hours, with lower education, or living at a new address since the last wave. In terms of their household characteristics, respondents not using showcards were more likely to have children in their household or be part of a partially responding household.

Table 2: Showcard usage (%) during individual interview

	Wave 20		Wave 21	
	Face to face	Telephone	Face to face	Telephone
Paper showcards	91.2	61.3	94.1	38.6
Online showcards	0.1	7.5	1.3	25.4
No showcards	8.6	31.2	4.6	35.9
Number of interviews	730	16,340	4,025	12,524

Interview situation

After the individual interview, interviewers provide some information about the interview situation (see Table 3). In wave 21, interviewers reported there was another person present during the interview for 23.9% of the interviews (this excludes 4.0% of cases where the interviewer did not know). This is a little higher than in wave 20 but both are substantially lower than in waves 18 and 19.

In terms of the other characteristics of the interview situation, the situation for waves 20 and 21 were very similar to waves 18 and 19.

Table 3: Interviewer report of respondent situation (%)

	Wave 18	Wave 19	Wave 20	Wave 21
Others present during interview	35.1	32.8	22.1	23.9
Whether others influenced interview a fair amount or a great deal (where others were present during the interview)	4.0	3.7	4.5	5.0
Excellent or good understanding	95.7	96.4	95.8	96.5
Excellent or good cooperation	98.4	98.9	98.9	99.0
Not suspicious	98.5	99.0	98.5	98.9
Referred to documents	39.5	36.8	39.7	43.7

Self-Completion Questionnaire

The process for the Self-Completion Questionnaire (SCQ) in wave 21 was the same as used in wave 20:

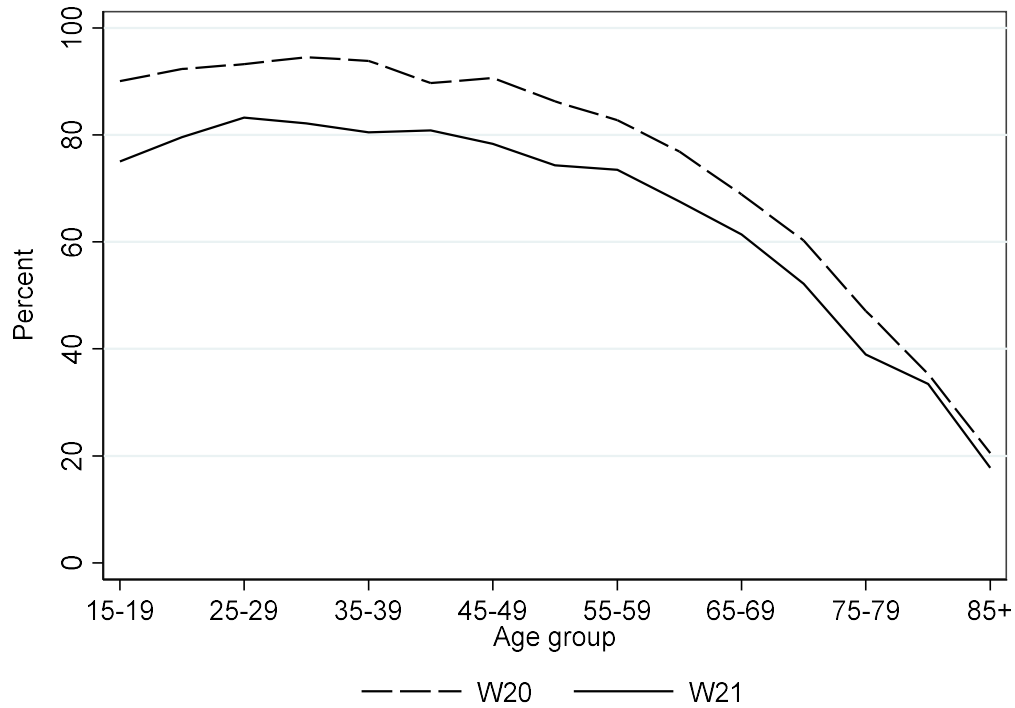
1. A link to the online option was included in the primary approach letter sent to respondents two weeks before the interviewer attempted contact with the household.
2. After contact was made with a household, the interviewer could discuss with the respondent how they would like to complete the SCQ (online or paper) if they had not already done so. Alternatively, the interviewer could have this discussion with the respondent at the end of their individual interview.
3. If the respondent chose the online option, an email was sent to their email address and an SMS was sent to their mobile. Where the SCQ had not been completed, reminders were sent by email and text message to complete the online SCQ on days 3, 10 and 17 (after the PQ interview or after the respondent activated the SCQ link in the primary approach letter if this was before the PQ interview).
4. If the respondent chose the paper option, the paper SCQ was mailed out from the office (if the interview was done by telephone) or handed to the respondent (if the interview was completed face-to-face). The follow up of the SCQ occurred as follows:
 - a. Where the SCQ was mailed to the respondent, they received a call on day 9 (after the PQ interview) to see that the respondent received the SCQ package in the mail.⁴ Another call is made on day 21 to remind the respondent to return the SCQ by mail if they had not already done so.
 - b. Where the interviewer gave the SCQ to the respondent, the interviewer would seek to pick up the SCQ when they were in the area so there was no official reminder protocol in these circumstances other than reminders to the interviewer from the office at 21 days to return the SCQs to the office.

Overall, 71.2% of SCQ respondents completed the online version (compared to 82.0% online in wave 20). Figure 5 shows the proportion of SCQs completed online by 5-year age groups. The reduction in online use in wave 21 compared to wave 20, particularly for the younger respondents, is due to the greater use of face to face interviewing where the interviewer could hand over the paper version of the SCQ and provide the SCQ incentive in cash when it is picked up. We are working on a way for the interviewers to provide a cash incentive for the SCQs completed online.

⁴ Due to longer postal times to non-metropolitan regions and to Western Australia, the follow up calls to households in these areas are made on day 11 rather than day 9.

To further understand the profile of SCQ respondents who completed their SCQ online rather than hardcopy, a logistic regression model was fitted to wave 20 and 21 data. The coefficients for this model are provided in Table A5 in the Appendix. In general, respondents more likely to complete the SCQ online were younger (as previously mentioned), married, highly educated, employed, interviewed by telephone, or have internet access at home. They were less likely to be born in a non-English speaking country, living with a long-term health condition, or renting.

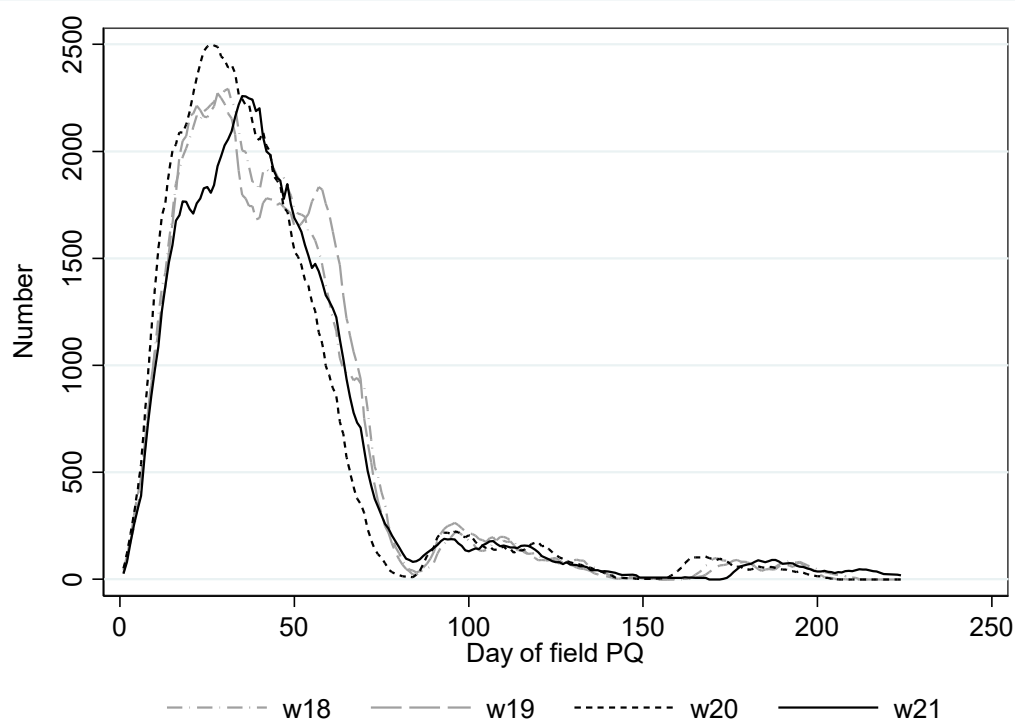
Figure 5: Proportion of SCQs completed online by age group



Impact on response timing

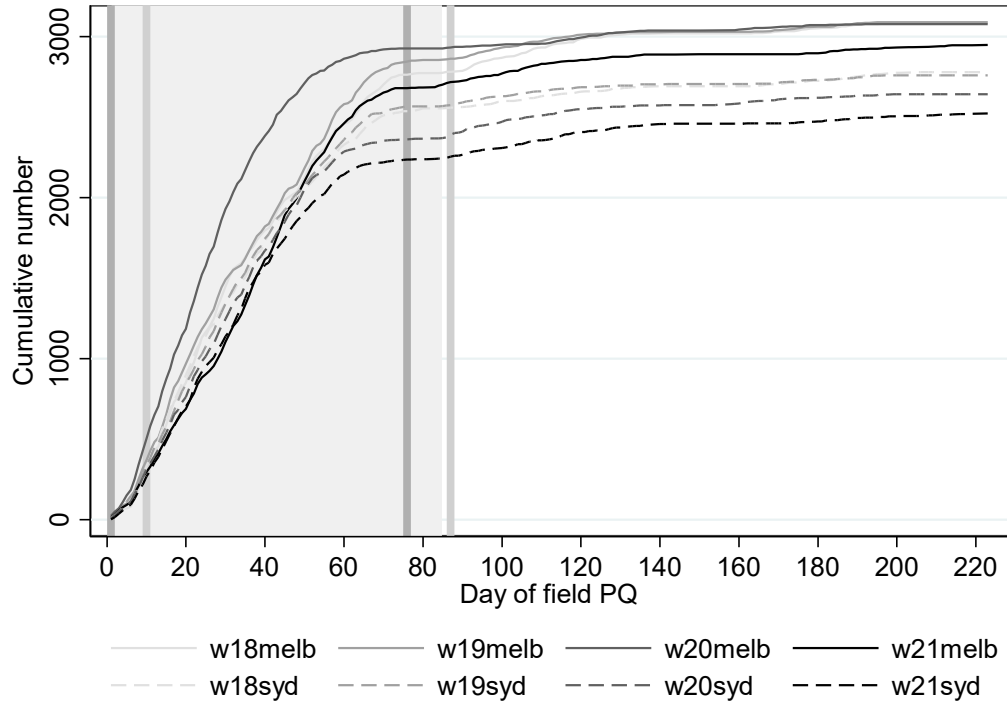
Compared to the three previous waves, the PQ interviews in wave 21 were completed somewhat later in the fieldwork period. This is due in part to the hesitancy of some respondents in lockdown areas to complete a telephone interview early in the fieldwork period when they would rather do a face-to-face interview after the lockdowns lifted (as mentioned earlier). It is also due to the interviewer training being extended by two weeks to accommodate webinar training of some of the new and less experienced interviewers in Victoria and South Australia. Figure 6 shows the number of PQs completed in a rolling week throughout the fieldwork period. For example, the number of interviews completed in a rolling week on day 50 is the number of interviews completed on days 54 through to 50. The slowing of work completed in wave 21 is evident around days 15 to 30. Note that in wave 20, the interviews were completed slightly earlier than usual as with telephone interviewing interviewers were able to progress through their allocated sample quicker without having to factor in travel time to get to their areas in person.

Figure 6: Number of PQ interviews completed by rolling week



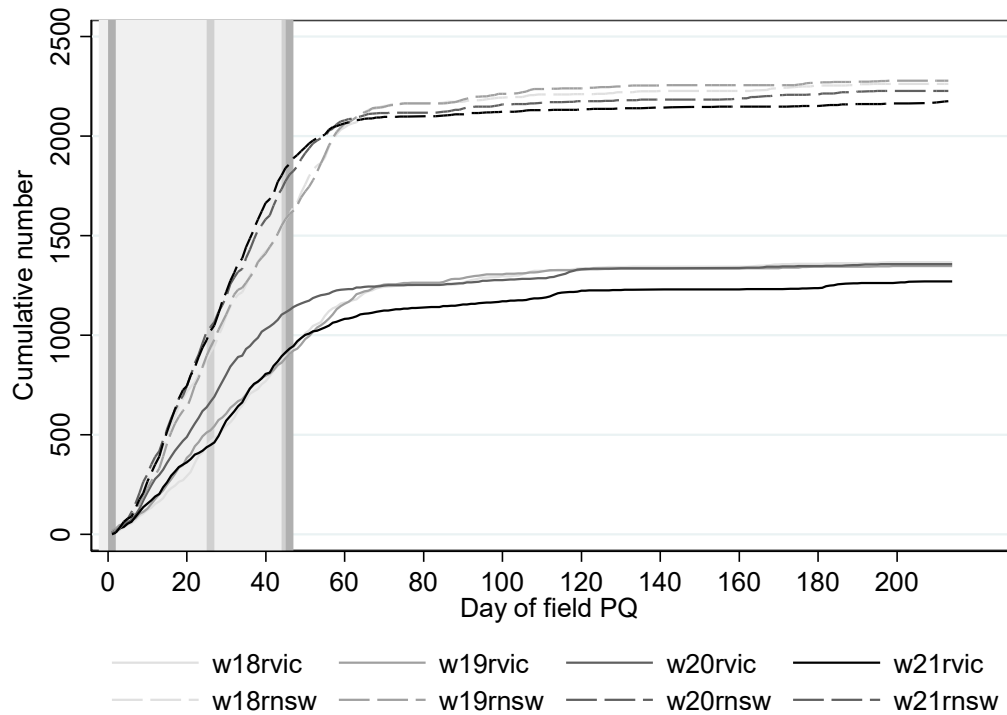
A different way of looking at the work completed is the cumulative number of interviews completed by each day of field. Figure 7 shows the cumulative count for Sydney and Melbourne and Figure 8 show the cumulative count for rural New South Wales and rural Victoria. Also shown in these figures are the lockdown periods for wave 20 (shaded grey) and wave 21 (between the vertical lines). Both the Melbourne and Sydney fieldwork progressed slower in wave 21 than previous waves. The fieldwork in Melbourne, in particular, was a lot slower than in wave 20, despite being in lockdown both waves. This is partially due to the new and less experienced interviewers in Victoria not being trained until week 5 (normally they would be trained in week 1 and 2). The fieldwork in rural New South Wales is on par with wave 20 in the initial fieldwork period, but did not seek the same improvements later in the fieldwork period (October to March). In rural Victoria, the fieldwork was on par with waves 18 and 19 initially, but this was not sustained past day 60 (beginning of October). This is due to an increase in refusals in wave 21, particularly in areas affected by lockdowns.

Figure 7: Cumulative number of PQ interviews in Sydney and Melbourne



Note: In wave 21, lockdown in Sydney from day 1 to 76 of fieldwork (bounded by dark vertical bars) and lockdown in Melbourne on day 1 and 10 to 87 (the latter bounded by light vertical bars). In wave 20, lockdown in Melbourne occurred from day 1 to 85 of fieldwork (shaded grey).

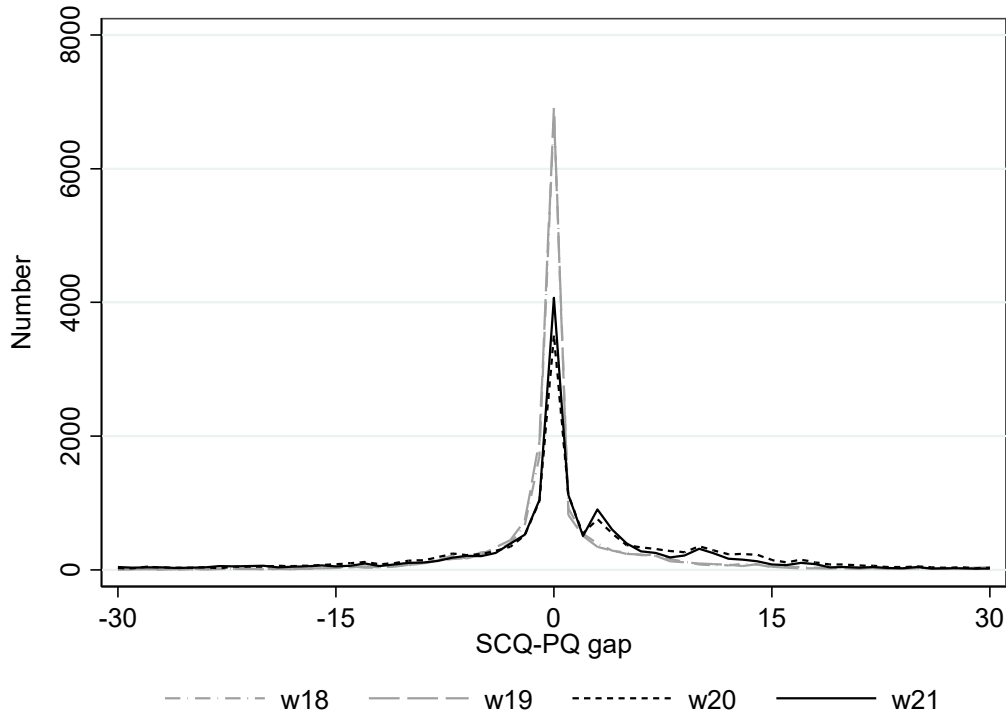
Figure 8: Cumulative number of PQ interviews in rural New South Wales and rural Victoria



Note: In wave 21, lockdown in parts of rural New South Wales from day 1 to 46 of fieldwork (bounded by dark vertical bars) and lockdown in rural Victoria on day 26 to 46 (bounded by light vertical bars). In wave 20, lockdown in rural Victoria occurred from day 1 to 45 of fieldwork (shaded grey).

Changes in the mode of delivery of the PQ and SCQ in waves 20 and 21 has modified the distribution of the gap between the PQ and SCQ. A similar proportion of respondents completed the SCQ before their PQ interview but the proportion completing both on the same day was 22.5% in wave 20 and 26.7% in wave 21 (compared to 44% in both waves 18 and 19). Figure 9 shows the distribution of the gap between the PQ and SCQ within 30 days of the PQ. More SCQs were completed within two weeks after the PQ in wave 20 and 21 with peaks relating to the days the online SCQ reminders were sent to the respondents.

Figure 9: Distribution of gap between completion of PQ and SCQ



Impact on item responses

Changing the survey mode (or the relative mixture of modes) can result in differences in responses if it affects how the respondent arrives at an answer. The quality of the response is a result of how well the respondent understands the question, retrieves relevant information, integrates that information to form an overall judgement, and then formulates a response (Tourangeau, Rips, and Rasinski 2000, p.7). How well a respondent performs this task depends on their level of effort to obtain the relevant information and their willingness to disclose that information.

In a face-to-face interview, the interviewer can respond to both visual and aural cues from the respondent to allow more time for the respondent to formulate a response, to repeat the question or provide additional information or clarifying statements (as may be provided with optional question text). The interviewer can slow down or speed up the pace of the interview based on their perception of the respondent's level of comfort with the interview. The respondent is typically fully engaged in the face-to-face interview process as they are often sitting down with the interviewer and tend not to be engaged with other tasks at the same time. The interviewer can observe interruptions from others and react accordingly. Showcards can also be provided to the respondent to help them select an appropriate response for certain questions.

In a telephone interview, the interviewer can only respond to aural cues from the respondent and there may be a temptation for the interviewer or the respondent to fill in any awkward silences. It is harder for the interviewer to pace the interview exactly right for the respondent in a telephone interview compared to a face-to-face interview. This can result in shorter telephone interviews compared to equivalently scripted face-to-face interviews (Jäckle et al., 2006). Telephone respondents may also be multitasking during the interview or they may be in a public place which may affect their ability to concentrate on the interview or, in the latter case, their willingness to disclose answers. In the HILDA Survey the vast majority of interviews (95%) were conducted while the respondent was at home in wave 20 and 21, as in earlier waves. We also continue to use showcards for 68.8% and 64.1% of the telephone interviews in waves 20 and 21 respectively.

In general, face-to-face surveys tend to have slightly less overall item non-response (de Leeuw 1992) and slightly longer responses to open ended questions than telephone surveys (de Leeuw 1992, Jäckle et al. 2006) but do not tend to differ on socially desirable responses (de Leeuw 1992).⁵ Some studies have also found more straight lining (also known as non-differentiation) on battery questions in telephone surveys compared to face-to-face surveys where the same response option is chosen across multiple items in a question set (Green et al. 2001, Holbrook et al. 2003, Jäckle et al. 2006).

For self-administered forms, such as our Self-Completion Questionnaire, there may also be some differences by mode. The main differences are that an online questionnaire routes the respondent through the questionnaire, reducing skip errors, and avoids the need for the respondent to mail back (or have an interviewer pick up) a hardcopy questionnaire. In an online questionnaire, the questions are typically displayed one at a time whereas with a hardcopy questionnaire the respondent can see the overall size of the questionnaire and all questions on each two-page opening of the questionnaire booklet. Mode differences may also occur if the way the questionnaire is laid out on the screen is different to how it is displayed on the hardcopy questionnaire. For example, the hardcopy questionnaire displays a multi-item question in a matrix form (e.g. SCQ questions A3, A9) but the online version of the questionnaire may use a different style.

Online (web) surveys typically have lower item non-response than surveys with hardcopy questionnaires (Kwak and Radler 2002; Shin et al. 2012). There are fewer skip errors in web surveys compared to mail surveys (Fricker and Schonlau 2012). Straight lining behaviour has been found to be similar between online and mail surveys (Kim et al. 2019).

⁵ While some studies have found some small differences in socially desirable responses between modes, it has generally not been in the expected direction (Green et al. 2001, Holbrook et al. 2003, Jäckle et al. 2006). They have found that telephone respondents are slightly more likely to give socially desirable responses even though there is a greater social distance between the respondent and the interviewer than in face-to-face interviews. Nevertheless, the type of questions where an effect has been found would typically be placed in our self-completion questionnaire (such as attitudes, beliefs, risky behaviours, religion) or not asked (e.g., voting behaviour).

In the following sections, various aspects of data quality are considered for waves 18 to 21. These are item non-response, number of items chosen at questions with multi-item responses, proportion of the first response option chosen at questions with a long list of responses, straight lining (choosing the same response option in multi-item questions), rounding, and the length of open-ended questions.

Item non-response

The average item non-response rate for each section of the survey instruments is provided in Table 4. It is calculated from the variables common to all four waves. The item non-response rates for wave 20 and 21 are very similar to waves 18 and 19 in the interview components (HF, HQ and PQ). In the SCQ, the item non-response rate in wave 20 is lower than that seen in waves 18 and 19 for most sections (Sections B, D and E) though wave 21 rate is similar to these earlier waves. This is due to a higher rate of breakoffs in the online SCQ in wave 21 than wave 20 (see Figure 11 discussed below).

Table 4: Average item non-response (%) by section

	Wave 18	Wave 19	Wave 20	Wave 21
Household Form	0.18	0.19	0.20	0.27
Household Questionnaire				
Section Q: Childcare	0.22	0.24	0.23	0.27
Section R: Housing	1.04	0.98	1.01	1.17
Person Questionnaire				
Section AA: Country of birth, migration, language (NPQ only)	0.09	0.24	0.36	0.20
Section BB: Family background (NPQ only)	3.76	4.34	4.56	4.33
Section A: Education	0.05	0.06	0.05	0.08
Section B: Employment status	0.01	0.02	0.03	0.03
Section C: Employment	0.17	0.17	0.20	0.34
Section D: Not Employed	0.80	0.78	0.92	0.98
Section E: Calendar and job training	0.02	0.02	0.02	0.02
Section F: Income	0.91	0.94	1.16	1.24
Section G: Family formation	0.49	0.61	0.49	0.53
Section H: Relationships	0.29	0.33	0.23	0.29
Section K: Health, caring, mobility	0.02	0.02	0.02	0.11
Section T: Tracking	0.03	0.03	0.04	0.04
Self-Completion Questionnaire				
Section A: Health and wellbeing (SF36)	1.63	0.87	1.30	1.04
Section B: Lifestyle and living situation	0.92	0.89	0.75	0.86
Section C: Finances	3.43	1.92	2.14	2.91
Section D: Attitudes towards job	4.32	4.33	3.41	4.55
Section E: Attitudes towards parenting	2.20	2.26	1.91	2.32

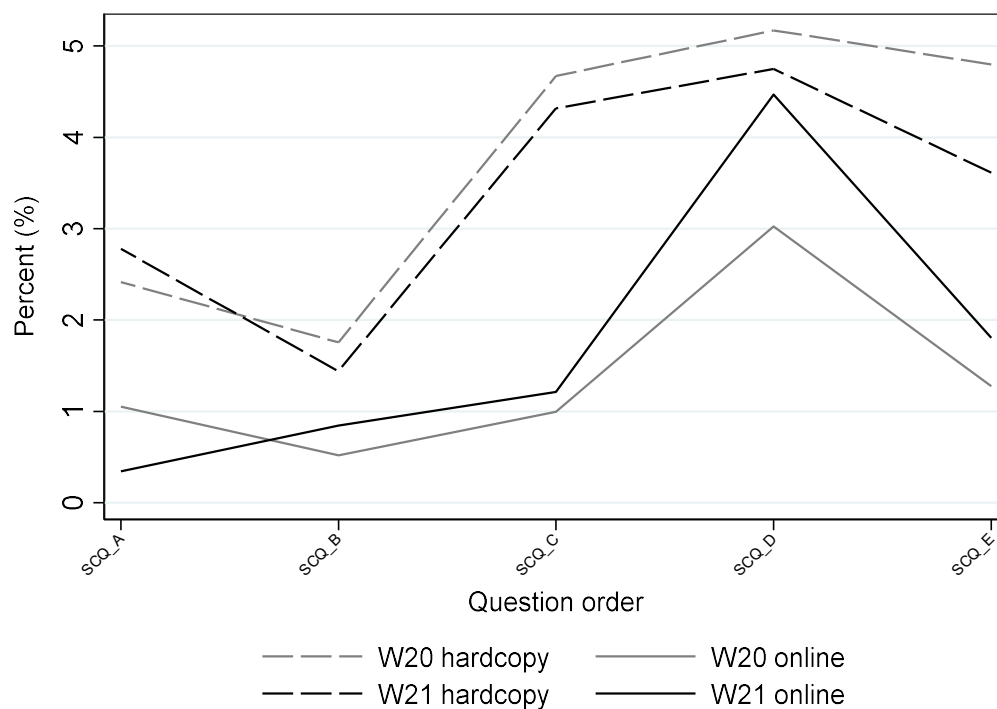
Note: The item non-response rate for a section is the number of questions with a “don’t know” or “refused” response divided by the number of questions asked, restricted to questions common to waves 18-21. The item non-response rate is calculated from the General Release, with the exception of Section B in the PQ where it is calculated from the Restricted Release. The average is calculated across the section rates of all individuals.

When the SCQ is completed online (in wave 20 and 21), the respondent is moved to the next relevant question and appropriate skips are followed. It is still possible for the respondent to not answer a question, but a warning message is displayed on the screen to indicate a question has not been answered and asks the respondent to click on the forward button again if they meant to skip the question. In comparison, a respondent using the hardcopy SCQ can skip answering some questions, miss questions placed in the right-hand column on the page, or skip two entire pages of the SCQ if they turn over two pages rather than one. The respondent can also look ahead and see what is involved with a series of question and decide not to answer them.

Graphs of the item non-response for the common questions in each section of the SCQ are provided in the Appendix (Figure A1 to Figure A5). Two questions styles were used for multi-item questions in wave 20 which have a common stem to introduce the topic, a common response scale, and a set of items that ask about different aspects (e.g., A3 asks how the respondent's health limits their ability with certain).⁶ The first question style used was a carousel question which was probably less intuitive for the respondent, did not display warning messages if an item was skipped, and resulted in higher item non-response (particularly the first time it is used at A3). The second question style used was the accordion question which performed better than the carousel question. In wave 21, the carousel style questions were replaced with the accordion style. As a result, the proportion of item non-response at A3, in particular, was improved.

The item non-response rates in the SCQ also differ by mode. Figure 10 shows that the hardcopy SCQs have higher item non-response. Of course, some of the difference can be explained by the different age profiles of the respondents (see Figure 5).

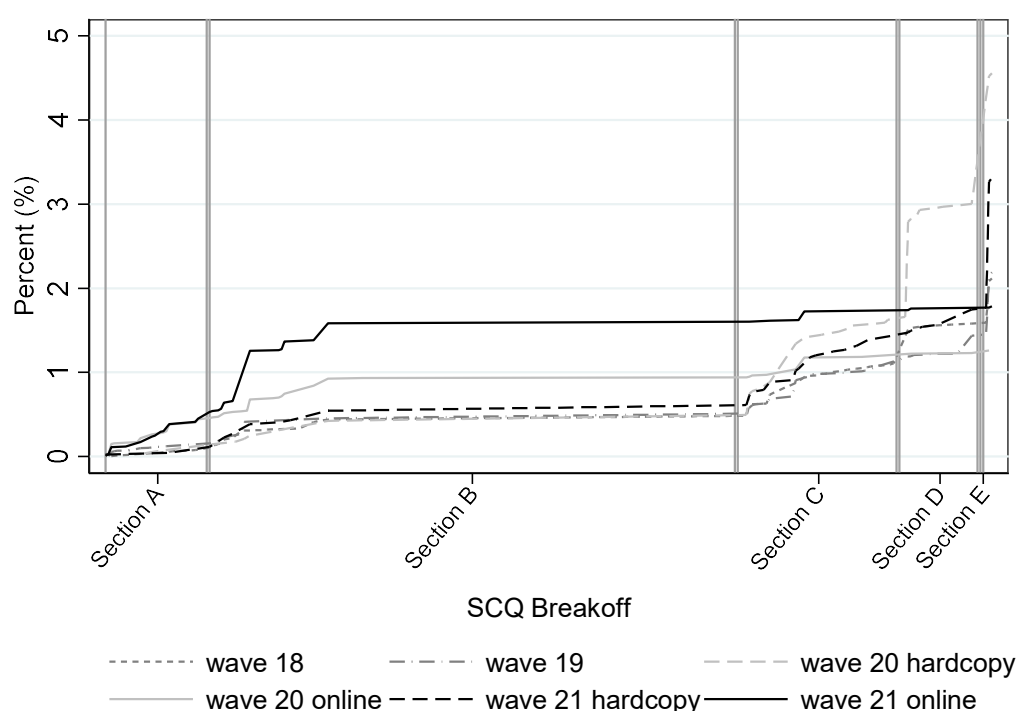
Figure 10: Average item non-response rate in each SCQ section, by mode and wave



⁶ A detailed description of these two question styles with screen shots is provided in the wave 20 data quality paper (Watson, Jin, Summerfield, 2021).

We can also compare the percentage of breakoffs that occur throughout the SCQ. A break-off occurs when the respondent stops responding to questions in the SCQ. We might expect this to occur more when using the online SCQ as respondents are called away from the online form to other things on their computer or in their home. People completing the paper version of the SCQ may be more likely to return to complete it at a later point in time. After restricting the variables to those in common between the four waves, the proportion of breakoffs in the SCQ in wave 21 is 2.2%, compared to 2.1% in wave 18, 2.2% in wave 19 and 1.8% in wave 20. Figure 11 separates the percentage of breakoffs in the SCQ by mode in waves 20 and 21. Breakoffs are more likely to occur when the SCQ is completed online rather than on hardcopy. They also occur in different places for the different modes. The breakoffs on the hardcopy form tend to occur in Section C (finances) but the breakoffs online tend to occur in Section B (lifestyle and living situation).⁷ The increase in the breakoff rate in wave 21 compared to wave 20 occurs in the online version of the SCQ (after B11 and before B19, mainly at the questions on waist measurement – where the respondents are encouraged to use the tape measure provided with their primary approach letter – and psychological distress).⁸

Figure 11: Cumulative percentage of breakoffs when completing the SCQ



We turn now to variables that generally have higher rates of non-response, being those that ask for dollar values. Table 5 shows the item non-response for a selected set of derived variables that are subsequently imputed in the HILDA Survey datasets. The item non-response rates in wave 20 and 21 are generally similar to previous waves, however wages and salaries, government allowances, and parenting payments (at least for the last FY) having marginally higher rates of item non-response in these waves. Item non-response for house value declined in wave 20, likely due to homeowners reconsidering their housing needs during 2020.

⁷ The breakoffs in Section D (asked of people in paid work) and E (asked of parents of children under 18) in hardcopy form in waves 20 and 21 are typically a result of respondents missing the screening question at the start of these sections. This is more likely to occur when the section is not relevant to them which will be more frequent among the hardcopy respondents as they are older (the mean age of respondents completing the SCQ via hard copy in wave 21 is 55 compared to 41 for the online respondents).

⁸ In wave 17, when this question on waist measurement was last included in the SCQ, the tape measure was provided to the respondent in the SCQ pack, and so would be more likely to be nearby when the respondent was completing the SCQ.

Table 5: Percentage of item non-response for specific variables

Variable	Wave 18	Wave 19	Wave 20	Wave 21
Responding Persons¹				
Current Income				
Wages and salaries	3.1	2.7	3.3	3.5
Government pensions	4.0	4.0	5.1	4.2
Government allowances	5.8	5.9	4.6	6.0
Parenting payments	2.5	2.4	3.7	2.2
Financial Year Income				
Wages and salaries	4.2	4.7	5.0	5.7
Business income	12.4	13.6	12.9	12.7
Investment income	15.8	15.5	16.1	14.9
Private pension	4.9	5.5	3.9	4.4
Private transfer	6.8	8.5	9.2	6.9
Government pensions	3.1	3.0	3.7	2.9
Government allowances	3.3	3.7	6.1	7.3
Parenting payments	4.2	4.5	7.0	5.0
Total Financial Year income	10.6	10.9	11.2	11.6
Households²				
Total Financial Year income	23.0	23.1	24.9	26.7
House value	3.7	2.7	2.4	3.2
House debt	6.2	5.2	5.1	6.4
Mortgage repayments	2.4	2.3	2.4	2.4
Rent	1.4	1.6	1.6	1.5

Notes: 1. The percentage is calculated for non-zero cases for the income components and total income.

2. The percentage is calculated for non-zero cases for the housing variables and for zero and non-zero cases for household total income.

Completeness of multi-item responses

Another way the data quality of responses could be affected by a change in mode or the removal of some showcards is that multi-item response questions could have a different number of items selected by the respondent. There could be less probing by the interviewer over the telephone or the respondent may not consider some options as valid without seeing a list of possibilities on a showcard. We would then expect to see a lower number of options selected at multi-item response questions in wave 20 and 21. What we see (in Table 6) is that the number of options selected at multi-item response questions in wave 20 and 21 is very similar to waves 18 and 19. There is one question, E26 on requirements from Centrelink or employment services, where the number of responses selected has dropped. However, this is likely a real change given the COVID-19 pandemic.

Table 6: Number of options chosen at multi-item response questions

Question	Wave 18	Wave 19	Wave 20	Wave 21
Household Questionnaire				
Q28: Family Tax Benefit payments	1.01	1.02	1.03	1.02
Personal Questionnaire				
A9a: Qualifications studying for since last interview	1.06	1.05	1.05	1.06
A11: Qualifications completed since last interview	1.04	1.02	1.02	1.03
A7a: Qualifications ever completed (NPQ)	1.38	1.42	1.46	1.43
C9b: Days usually worked	4.14	4.17	4.19	4.24
D2: Activities to look for work	2.94	2.88	2.86	2.71
D6: Difficulties getting a job	2.46	2.58	2.39	2.43
D11: Reasons not looking for work	1.36	1.37	1.32	1.34
E12: Aims of work-related training	2.77	2.67	2.74	2.70
E26: Requirements from Centrelink or an employment services provider	1.38	1.37	1.19	1.19
F7: Salary sacrifice received – main job	1.21	1.22	1.20	1.20
F11: Non-cash benefits received – main job	1.57	1.58	1.58	1.60
F19: Salary sacrifice received – other jobs	1.07	1.20	1.29	1.11
F22b: Non-cash benefits received – other jobs	1.31	1.26	1.43	1.51
F28: Current pensions and allowances received	1.03	1.03	1.02	1.03
F30a: Other current pensions and allowances received	1.01	1.01	1.01	1.01
F36: Financial year salary sacrifice received	1.23	1.22	1.19	1.20
F40: Financial year non-cash benefits received	1.57	1.57	1.54	1.58
F46: Financial year salary sacrifice received with incorporated businesses	1.29	1.30	1.57	1.17
F50: Financial year non-cash benefits received with incorporated businesses	2.30	2.28	2.36	2.37
F61a: Financial year pensions and allowances received	1.14	1.14	1.22	1.17
F67: Use of lump sum payments	1.26	1.32	1.22	1.21
F69a: Other sources of income	1.05	1.05	1.05	1.07
H1a: Marital status changes since last interview	1.01	1.00	1.01	1.00
K10b: Long-term health condition	2.69	2.70	2.62	2.59
K8: Care for household member with long-term health condition	1.03	1.04	1.05	1.05
K66: Care for people living elsewhere	1.03	1.03	1.04	1.02
K82: Reasons for moving since last interview	1.27	1.30	1.21	1.21
Self-Completion Questionnaire				
C3b: Ways to get money for an emergency	1.74	1.73	1.68	1.67

Note: Question numbers relate to the wave 21 questionnaire.

Primacy versus recency in list-style questions

It has sometimes been found that responses to questions with a long list of response options can change when they are read (such as on a showcard or during an online questionnaire) compared to when they are heard (by being read out by an interviewer on the telephone). Respondents reading through a list may select an earlier option as the earlier options are processed more thoroughly whereas when the interviewer reads out the list the respondent may select an option from later in the list as they can be processed more thoroughly once the interviewer pauses after reading all the options (Krosnick and Alwin 1987). This is known as primacy versus recency.

Table 7 shows the proportion of respondents selecting the first option in list-style questions with four or more (substantive) response options. The mean responses are similar across the waves indicating that primacy versus recency is not an issue for these questions in waves 20 and 21.

Table 7: Percentage of respondents selecting first option to list-style questions with four or more response options where showcards had been used in waves 18 and 19

	Wave 18	Wave 19	Wave 20	Wave 21
Question				
Interviewer reads out list only when respondent does not have showcards in waves 20 & 21				
C10: Current work schedule (8 categories)	48.9	48.8	48.2	49.8
C24: Category of current pay (5 categories)	17.1	17.1	16.9	17.4
C31: Category of working place organisations (6 categories)	45.3	44.5	42.9	44.0
D9: Main activity since last worked or looked for work (8 categories)	18.5	18.6	19.4	19.1
G11b: Current employment status of the other parent whom the youngest children lives elsewhere with (8 categories)	1.0	1.0	1.0	1.0
G24: Current employment of the other parent (7 categories)	2.3	2.5	2.3	2.4
Interviewer reads out list in waves 20 & 21				
G9a: Frequency to see the youngest child living elsewhere (9 categories)	0.3	0.3	0.3	0.2
G10: Contact with the youngest child living elsewhere (5 categories)	1.5	1.4	1.3	1.3
G22a: Contact of the youngest child with the other parent who lives elsewhere (9 categories)	0.3	0.3	0.2	0.3
G23: Feeling about the amount of contact that youngest child has with the other parent (5 categories)	0.8	0.7	0.6	0.6
H4: Current marital status (6 categories)	46.1	45.7	45.4	45.2
H8: Likelihood of marrying current partner (5 categories)	6.6	6.5	7.3	7.4
H9: Likelihood of marrying in the future (if not likely to marry current partner) (5 categories)	0.1	0.1	0.0	0.1
H10: Likelihood of marrying in the future (if no current partner) (5 categories)	5.3	5.0	5.2	5.2

Note: Question numbers relate to the wave 21 questionnaire.

Straight lining

Straight lining occurs when the respondent answers the same response category to all items in a multi-item question. For example, Q5 in the Household Questionnaire asks the respondent to rate on a scale between 0 and 10 how much difficulty they have had with certain aspects of childcare in the last 12 months. If the respondent reports the same number for all 12 items in the list this is considered an instance of straight lining. Table 8 shows the percentage of respondents providing the exact same response to all items in a multi-item question. Again, we are focusing just on those questions asked all four waves. There does not appear to be any changes in straight lining behaviour in the interview components in waves 20 and 21, though the instances of straight lining have potentially decreased for some questions in the SCQ. Differences in the SCQ would be due to how the question is displayed: for the carousel or accordion questions (discussed earlier) the answer for earlier items is hidden from view or folded up under the item text whereas the hardcopy SCQ uses a matrix format which is more susceptible to straight lining.

Table 8: Percentage of respondents providing exactly the same response across all items in multi-item question (with four or more items)

Question	Wave 18	Wave 19	Wave 20	Wave 21
Household Questionnaire				
Q5: Having difficulties with childcare in the last 12 months (12 items)	3.6	2.7	2.2	1.9
Personal Questionnaire				
C35: Job satisfaction (6 items)	4.8	4.7	4.7	4.7
K71: Life satisfaction (8 items)	0.9	0.8	0.8	0.7
Self-Completion Questionnaire				
A9: Feelings during the past 4 weeks (9 items)	0.9	1.0	0.7	0.8
A11: Feelings about own health condition (4 items)	2.7	2.9	2.8	3.3
B19: Satisfaction with family life (8 items)	0.4	0.5	0.4	0.3
B23: Support got from the others (10 items)	1.1	1.1	0.9	1.1
D2: Feelings about current (main) job (21 items)	1.0	1.1	0.9	0.9
E2: Feelings about raising children (4 items)	11.8	11.2	12.3	11.8

Note: Question numbers relate to the wave 21 questionnaire.

Rounding

Another aspect of data quality that can be affected by mode is the level of rounding the respondents may apply at dollar questions. Table 9 shows the level of rounding of responses to two particular dollar value questions in the individual interview, being current wages and salaries and financial year age pension. It shows the distribution of the values that end with zero, one, two, three or four zeros. The level of rounding in wave 20 appears similar to earlier waves, however there does seem to be an improvement in the level of detail provided in wave 21. The latter is likely due to an increase in the occurrence of respondents referring to documents (such as pay slips, tax return, and Centrelink benefit statement) during the interview (43.6% of interviews in wave 21 compared to 39.6% in wave 20).

Table 9: Percentage of respondents rounding amounts reported for current wages and salaries or Financial Year age pension

	Wave 18	Wave 19	Wave 20	Wave 21
Current Wages and Salaries				
Ones	26.3	29.0	27.0	33.0
Tens	18.7	17.7	18.2	16.9
Hundreds	36.0	34.2	37.6	33.8
Thousands	14.2	14.8	13.6	12.7
Ten thousands	4.7	4.3	3.7	3.6
Financial Year Age Pension				
Fortnight amount reported ¹				
Ones	33.6	32.1	35.6	45.8
Tens	36.9	38.2	43.3	34.8
Hundreds	29.5	29.8	21.2	19.5
Annual amount reported ²				
Ones	29.9	28.4	29.0	29.0
Tens	19.9	23.8	19.4	21.5
Hundreds	27.2	25.0	28.0	26.6
Thousands	20.9	20.5	21.2	20.6
Ten thousands	2.2	2.3	2.4	2.3

Notes: 1. The percentage is calculated for non-zero cases for the financial year age pensions reported as a fortnightly amount. Relatively few respondents reported a fortnightly amount: 148 in wave 18, 131 in wave 19, 104 in wave 20 and 118 in wave 21.

2. The percentage is calculated for non-zero cases for the financial year age pensions reported as an annual amount.

Length of open-ended responses

When interviews are conducted on the telephone rather than face-to-face, it may affect the interviewer's level of probing for open ended questions such as occupation and industry. Table 10 shows the number of characters recorded at the two components of the occupation questions and at the industry questions. The number of characters recorded in wave 20 and 21 is consistent with waves 18 and 19.

Table 10: Number of characters recorded in occupation and industry descriptions

	Wave 18	Wave 19	Wave 20	Wave 21
Occupation				
Job title	25.5	27.1	25.0	26.7
Tasks and duties	43.1	46.7	44.8	48.6
Industry	30.0	32.4	31.1	32.8

Conclusions

This paper examined the quality of the wave 21 data following challenges faced during the fieldwork due to the COVID-19 pandemic. We continued to interview mainly by telephone. Face-to-face interviews were possible in some states but not in others for most of the fieldwork. We also continued to use an online option for the SCQ in addition to the paper form.

We had a drop in response rates for those interviewed in the previous wave (94.0% in wave 21, compared to 95.2% in wave 20 and over 96% in waves 18 and 19). This is somewhat disappointing but understandable given the complex environment at the time. The response rates for children turning 15 in wave 21 improved over wave 20 but the response rate for new entrants declined further. It is hard to make connection and build rapport with these individuals over the telephone.

A total of 75.7% of the individual interviews were completed by telephone in wave 21 (compared to 95.7% in wave 20 and less than 10% in earlier waves). Nevertheless, the SCQ response rate was the highest it had been since wave 2 with 92.6% of the PQ respondents completing the SCQ online or via paper. Similar to wave 20, we found that only 27% of the respondents in wave 21 completed the SCQ on the day of their PQ interview (compared to around 44% in earlier waves).

We found similar rates of missingness as in earlier waves for most variables in wave 21. The lower rate of missingness found in the SCQ variables in wave 20 (due to the online SCQ assisting respondents through the correct routing in the questionnaire was not evident in wave 21. This is because the wave 21 online SCQ had a higher rate of breakoffs than in wave 20 (around questions included in wave 21 on waist measurement and depression that were not included in wave 20) which had a counterbalancing effect.

We found no evidence of differences across the waves in the use of multi-item response options, the responses chosen for questions with long response lists, straight lining, rounding, or the length of responses provided at open-ended questions.

Overall, we conclude that the wave 21 data are of a similar quality to previous waves despite the complexities of the fieldwork during the second year of the pandemic.

References

- de Leeuw E D (1992). Data Quality in Mail, Telephone, and Face to Face Surveys. Amsterdam: TT Publications.
- Fricker R D, Schonlau M (2012). Advantages and disadvantages of internet research surveys: Evidence from the literature. *Field Methods*, 14, 347-367.
- Green M C, Krosnick J A, and Holbrook A L (2001). The survey response process in telephone and face-to-face surveys: differences in respondent satisficing and social desirability response bias. Report published at <http://users.clas.ufl.edu/kenwald/pos6757/spring02/tch62.pdf>
- Holbrook A L, Green M C, Krosnick J A (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: comparison of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, 67, 79-125.
- Jäckle A, Roberts C, and Lynn P (2006). Telephone versus face-to-face interviewing: mode effects on data quality and likely causes: report on phase II of the ESS-Gallup mixed mode methodology project. ISER Working Paper Series, No. 2006-41. Institute for Social and Economic Research, University of Essex. https://www.iser.essex.ac.uk/files/iser_working_papers/2006-41.pdf
- Kim Y, Dykema J, Stevenson J, Black P, Moberg D P (2019). Straightlining: Overview of measurement, comparison of indicators in mail-web mixed-mode surveys. *Social Science Computer Review*, 37(2), 214-233.
- Krosnick J A, and Alwin D F (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, 51(2), 201-219.
- Kwak N, and Radler B (2002). A comparison between mail and web surveys: response pattern, respondent profile, and data quality. *Journal of Official Statistics*, 18(2), 257-273.
- Shin E, Johnson T P, and Rao K (2012). Survey mode effects on data quality: comparison of web and mail modes in a US National Panel Survey. *Social Science Computer Review*, 30(2), 212-228.
- Tourangeau R, Rips L J, and Rasinski K (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Watson N, Jin Y, and Summerfield M (2021). Wave 20 data quality and the impact of questionnaire and fieldwork changes due to the COVID-19 pandemic. HILDA Project Discussion Paper Series No. 1/21. Melbourne Institute: Applied Economic and Social Research, University of Melbourne. https://melbourneinstitute.unimelb.edu.au/_data/assets/pdf_file/0004/3969310/hdps121.pdf

Appendix

Table A1: Logistic regression coefficients for model predicting response to individual interview (vs non-response) for previous wave respondents

Variable	Waves 18		Wave 19		Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Female	0.160	0.181	0.183	0.209	0.259	0.167	0.402***	0.149
Age (/10)	0.911***	0.303	0.786**	0.380	1.332***	0.279	1.051***	0.264
Age squared (/10 ²)	-0.099***	0.031	-0.069*	0.040	-0.133***	0.029	-0.093***	0.028
Dwelling (base=separate house)								
Semi-detached	-0.697*	0.392	0.629	0.558	-0.331	0.324	-0.155	0.297
Flat	-0.198	0.396	0.624	0.439	0.666*	0.382	-0.216	0.259
Other private dwelling	-	-	-	-	-0.505	1.052	-	-
Non-private dwelling	-0.587	1.065	-	-	0.105	1.078	-0.119	1.021
Renter	0.287	0.238	-0.229	0.282	-0.127	0.200	0.026	0.183
Country of birth (base=Australia)								
Main English-speaking country	-0.413	0.306	-0.531	0.387	0.210	0.356	0.022	0.304
Not main English-speaking country	0.139	0.335	-0.561*	0.302	-0.145	0.283	-0.124	0.258
Part responding HH	-0.070	0.315	-0.346	0.343	0.137	0.356	-0.205	0.242
No SCQ in previous wave	-1.098***	0.294	-1.372***	0.283	-1.386***	0.266	-1.439***	0.168
Telephone interview in previous wave	-0.838***	0.303	-0.583*	0.344	0.809**	0.382	1.406***	0.254
Marital status (base=married)								
De-facto	-0.089	0.315	-0.035	0.335	-0.503**	0.225	-0.126	0.228
Separated	0.156	0.725	-0.218	0.718	0.452	0.749	1.439	0.919
Divorced	0.484	0.633	-0.148	0.556	0.447	0.544	-0.355	0.388
Widowed	1.291	0.795	0.032	0.717	1.272**	0.632	0.254	0.538
Never married	-0.114	0.312	0.342	0.352	0.330	0.298	0.367	0.252
Number of children	-0.103	0.122	0.141	0.144	0.008	0.099	-0.137*	0.072
Number of adults	-0.440***	0.078	-0.399***	0.082	-0.299***	0.072	-0.337***	0.071
Education (base=Year 11 or below)								
Year 12	0.139	0.254	0.124	0.323	0.063	0.249	0.201	0.222
Certificate	-0.135	0.275	0.307	0.295	-0.284	0.229	0.085	0.230
Diploma	0.309	0.397	0.045	0.391	0.428	0.373	-0.023	0.287
Graduate degree	0.405	0.355	0.153	0.366	0.628*	0.354	0.236	0.283
Post-graduate	0.432	0.423	0.885*	0.516	0.610	0.387	0.654*	0.367
Employment and hours (base=not in labour force)								
Unemployed	-0.126	0.433	0.331	0.563	-0.447	0.387	-0.270	0.323
Employed part time	-0.221	0.260	0.186	0.303	0.165	0.303	0.055	0.233
Employed full time (35-54 hours)	0.003	0.291	0.369	0.343	-0.450*	0.273	-0.233	0.232
Employed full time (55 plus hours)	-0.606	0.409	-0.317	0.454	-0.700*	0.416	-0.715**	0.318
Serious long term health condition	-0.164	0.245	-0.396	0.303	-0.615**	0.248	-0.552***	0.204
Location (base=Sydney)								
Melbourne	0.333	0.341	0.049	0.345	0.047	0.284	0.220	0.280
Brisbane	0.634	0.442	0.180	0.413	0.307	0.354	0.281	0.298
Adelaide	-0.006	0.447	0.328	0.533	0.139	0.415	0.411	0.388
Perth	0.662	0.484	0.071	0.463	1.093*	0.558	0.348	0.396
Major city	0.290	0.436	0.626	0.530	0.438	0.384	0.374	0.380
Inner regional	0.270	0.344	0.705*	0.409	0.165	0.289	0.286	0.277

Variable	Waves 18		Wave 19		Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Outer regional	-0.066	0.375	0.432	0.423	-0.275	0.328	0.059	0.297
Remote	1.203	1.018	-0.072	0.652	-0.149	0.827	-0.589	0.458
SEIFA quintile (base=quintile 1 (most disadvantaged))								
Quintile 2	0.316	0.344	-0.426	0.332	-0.380	0.266	-0.359	0.233
Quintile 3	0.037	0.323	-0.091	0.373	0.033	0.291	-0.171	0.254
Quintile 4	0.024	0.334	-0.164	0.365	-0.371	0.272	-0.220	0.258
Quintile 5 (least disadvantaged)	-0.125	0.332	-0.346	0.337	-0.464	0.284	-0.385	0.270
Moved	-0.086	0.252	0.358	0.332	0.467*	0.264	-0.334*	0.202
Constant	4.317***	0.815	4.124***	1.054	2.914***	0.779	1.728**	0.734

Note: * $p \geq 0.05$ & $p < 0.1$, ** $p \geq 0.01$ & $p < 0.05$, *** $p < 0.01$. Standard errors allow for clustering of people in previous wave households.

Table A2: Logistic regression coefficients for model predicting completion of SCQ

Variable	Waves 18		Wave 19		Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Female	0.194***	0.057	0.262***	0.060	0.445***	0.058	0.404***	0.061
Age (/10)	0.426***	0.113	0.358***	0.125	0.126	0.099	0.213*	0.112
Age squared (/10 ²)	-0.031**	0.012	-0.021	0.013	0.001	0.010	-0.008	0.012
Dwelling (base=separate house)								
Semi-detached	0.022	0.151	-0.067	0.169	0.173	0.129	0.287**	0.134
Flat	-0.099	0.119	-0.039	0.130	0.207*	0.107	0.258**	0.116
Other private dwelling	0.749*	0.432	0.145	0.429	-0.038	0.359	-0.203	0.314
Non-private dwelling	-1.307***	0.295	-1.508***	0.309	-0.942***	0.275	-0.643**	0.318
Renter								
Country of birth (base=Australia)	-0.404***	0.095	-0.305***	0.098	-0.229***	0.079	-0.206**	0.083
Main English-speaking country	0.134	0.142	-0.065	0.138	0.004	0.120	-0.111	0.123
Not main English-speaking country	-0.921***	0.104	-0.719***	0.114	-0.431***	0.102	-0.321***	0.109
Part responding HH	-0.714***	0.125	-0.742***	0.131	-0.724***	0.108	-0.768***	0.102
Telephone interview	-2.768***	0.090	-2.928***	0.094	0.361***	0.135	-0.624***	0.103
Marital status (base=married)								
De-facto	-0.061	0.122	-0.142	0.121	-0.278***	0.103	-0.120	0.108
Separated	0.072	0.217	-0.263	0.218	-0.496***	0.179	-0.388**	0.189
Divorced	-0.035	0.159	-0.181	0.168	-0.428***	0.145	-0.244	0.159
Widowed	-0.746***	0.182	-1.017***	0.196	-0.915***	0.162	-0.920***	0.177
Never married	-0.284**	0.119	-0.215*	0.124	-0.579***	0.102	-0.451***	0.109
Number of children	-0.206***	0.041	-0.258***	0.039	-0.157***	0.033	-0.152***	0.038
Number of adults	0.090*	0.052	0.034	0.056	0.072*	0.042	0.120***	0.041
Education (base=Year 11 or below)								
Year 12	0.441***	0.103	0.238**	0.107	0.231**	0.094	0.122	0.098
Certificate	0.198**	0.093	0.218**	0.102	0.186**	0.084	0.018	0.090
Diploma	0.431***	0.127	0.245*	0.137	0.549***	0.131	0.514***	0.137
Graduate degree	0.643***	0.119	0.427***	0.123	0.511***	0.113	0.411***	0.115
Post-graduate	0.737***	0.139	0.617***	0.145	0.584***	0.133	0.506***	0.13
Employment and hours (base=not in labour force)								
Unemployed	0.029	0.160	-0.080	0.175	-0.277**	0.132	-0.274*	0.166

Variable	Waves 18		Wave 19		Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Employed part time	-0.066	0.106	-0.017	0.111	0.018	0.099	-0.155	0.104
Employed full time (35-54 hours)	-0.250**	0.101	-0.238**	0.105	-0.281***	0.091	-0.346***	0.098
Employed full time (55 plus hours)	-0.686***	0.144	-0.483***	0.153	-0.685***	0.140	-0.872***	0.141
Serious long term health condition	-0.299***	0.103	-0.570***	0.105	-0.170*	0.090	-0.353***	0.094
Location (base=Sydney)								
Melbourne	0.122	0.130	0.153	0.138	0.211*	0.115	0.050	0.115
Brisbane	0.381**	0.155	0.406**	0.165	0.396***	0.133	0.252*	0.138
Adelaide	0.683***	0.181	0.375*	0.207	0.311**	0.155	0.192	0.165
Perth	0.968***	0.191	0.748***	0.189	0.476***	0.154	0.741***	0.201
Major city	0.794***	0.168	0.773***	0.173	0.282**	0.139	0.437***	0.146
Inner regional	0.385***	0.133	0.541***	0.141	0.370***	0.112	0.343***	0.117
Outer regional	0.300**	0.146	0.473***	0.157	0.014	0.124	-0.027	0.140
Remote	0.476*	0.256	0.287	0.252	-0.085	0.238	-0.370*	0.217
SEIFA quintile (base=quintile 1 (most disadvantaged))								
Quintile 2	0.035	0.119	-0.154	0.126	0.011	0.099	0.121	0.108
Quintile 3	0.007	0.123	0.129	0.129	0.175	0.107	0.119	0.113
Quintile 4	-0.038	0.122	0.041	0.127	0.077	0.106	-0.006	0.112
Quintile 5 (least disadvantaged)	-0.043	0.132	-0.006	0.136	-0.010	0.113	-0.098	0.113
Moved	-0.185*	0.101	0.051	0.105	-0.169**	0.083	-0.172*	0.089
Has internet	0.663***	0.133	0.650***	0.160	0.952***	0.112	0.916***	0.149
Constant	1.090***	0.366	1.500***	0.384	0.508	0.342	1.345***	0.357

Note: * $p \geq 0.05$ & $p < 0.1$, ** $p \geq 0.01$ & $p < 0.05$, *** $p < 0.01$. Standard errors allow for clustering of people in households.

Table A3: Logistic regression coefficients for model predicting completion of individual interview by telephone (vs face-to-face)

Variable	Waves 18		Wave 19		Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Female	-0.032	0.048	-0.072	0.047	-0.022	0.067	0.106***	0.033
Age (/10)	0.340***	0.103	0.289***	0.103	0.749***	0.131	0.139*	0.075
Age squared (/10 ²)	-0.055***	0.011	-0.053***	0.011	-0.076***	0.013	-0.023***	0.007
Dwelling (base=separate house)								
Semi-detached	-0.133	0.142	-0.212	0.145	-0.291*	0.177	-0.029	0.113
Flat	0.259**	0.112	0.224**	0.111	0.228	0.169	0.277***	0.106
Other private dwelling	1.112***	0.316	1.033***	0.291	-0.266	0.579	0.325	0.310
Non-private dwelling	-0.024	0.443	-1.072*	0.621	-0.064	0.381	0.308	0.349
Renter	0.045	0.093	-0.016	0.091	-0.523***	0.120	-0.225***	0.075
Country of birth (base=Australia)								
Main English-speaking country	-0.179	0.118	0.248**	0.106	-0.220	0.159	0.026	0.086
Not main English-speaking country	-0.441***	0.120	-0.409***	0.119	-0.816***	0.130	-0.278***	0.089
Part responding HH	1.980***	0.110	2.044***	0.110	-0.138	0.187	0.444***	0.113
Marital status (base=married)								
De-facto	-0.156	0.115	-0.253**	0.113	-0.158	0.161	0.003	0.088
Separated	-0.028	0.179	-0.082	0.177	-0.219	0.228	-0.117	0.139
Divorced	0.072	0.139	0.029	0.139	-0.106	0.185	-0.040	0.101
Widowed	-0.087	0.200	-0.292	0.213	-0.222	0.211	-0.107	0.122

Variable	Waves 18		Wave 19		Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE	<i>b</i>	SE
Never married	0.021	0.108	-0.052	0.108	-0.034	0.163	0.144	0.089
Number of children	-0.094**	0.044	-0.138***	0.042	-0.192***	0.052	-0.072**	0.033
Number of adults	-0.263***	0.068	-0.359***	0.062	0.037	0.063	-0.116***	0.041
Education (base=Year 11 or below)								
Year 12	0.112	0.095	0.208**	0.096	0.527***	0.138	0.195***	0.073
Certificate	0.067	0.086	0.109	0.089	0.025	0.113	0.225***	0.062
Diploma	0.039	0.115	0.118	0.116	0.484***	0.174	0.392***	0.085
Graduate degree	0.258**	0.104	0.339***	0.103	0.792***	0.166	0.498***	0.082
Post-graduate	0.277**	0.118	0.256**	0.119	0.681***	0.197	0.532***	0.091
Employment and hours (base=not in labour force)								
Unemployed	0.220	0.143	-0.059	0.154	-0.183	0.171	-0.301**	0.119
Employed part time	-0.200**	0.097	-0.235**	0.097	0.130	0.138	-0.035	0.072
Employed full time (35-54 hours)	-0.024	0.093	-0.094	0.091	-0.070	0.129	0.077	0.070
Employed full time (55 plus hours)	0.400***	0.131	0.224*	0.129	-0.242	0.215	0.225**	0.115
Serious long term health condition	0.119	0.098	-0.043	0.100	-0.374***	0.117	-0.225***	0.065
Location (base=Sydney)								
Melbourne	0.004	0.137	-0.002	0.135	0.640***	0.193	0.419**	0.180
Brisbane	0.272*	0.159	0.091	0.160	0.294	0.202	-1.923***	0.145
Adelaide	-0.252	0.201	-0.373*	0.219	0.363	0.237	-2.703***	0.152
Perth	0.004	0.186	0.000	0.179	-0.552***	0.196	-3.948***	0.153
Major city	0.199	0.162	0.170	0.159	0.685***	0.251	-0.546***	0.175
Inner regional	0.438***	0.138	0.442***	0.137	0.617***	0.183	-1.550***	0.138
Outer regional	1.187***	0.144	1.251***	0.146	0.514**	0.213	-2.318***	0.146
Remote	2.466***	0.211	2.485***	0.202	0.955	0.640	-1.900***	0.232
SEIFA quintile (base=quintile 1 (most disadvantaged))								
Quintile 2	-0.024	0.121	-0.062	0.119	0.254*	0.149	-0.086	0.088
Quintile 3	0.143	0.120	0.088	0.116	0.203	0.158	0.062	0.093
Quintile 4	0.274**	0.123	0.257**	0.121	0.390**	0.169	0.187*	0.100
Quintile 5 (least disadvantaged)	0.133	0.132	0.221*	0.131	0.479***	0.183	0.242**	0.108
Moved	0.390***	0.096	0.388***	0.090	-0.180	0.143	0.111	0.085
Constant	-2.749***	0.350	-2.192***	0.349	1.312***	0.460	2.637***	0.277

Note: * $p \geq 0.05$ & $p < 0.1$, ** $p \geq 0.01$ & $p < 0.05$, *** $p < 0.01$. Standard errors allow for clustering of people in households.

Table A4: Logistic regression coefficients for model predicting use of showcards in individual interview

Variable	Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE
Female	0.099***	0.030	0.091***	0.030
Age (/10)	0.029	0.064	-0.030	0.064
Age squared (/10 ²)	0.004	0.007	0.003	0.007
Dwelling (base=separate house)				
Semi-detached	-0.019	0.088	0.170*	0.090
Flat	0.041	0.075	0.013	0.075
Other private dwelling	-0.588**	0.261	-0.854***	0.255
Non-private dwelling	-0.976***	0.244	-0.638**	0.257
Renter	-0.173***	0.056	0.016	0.058

Country of birth (base=Australia)				
Main English-speaking country	0.064	0.073	0.040	0.075
Not main English-speaking country	-0.309***	0.067	-0.257***	0.068
Part responding HH	-0.655***	0.081	-0.341***	0.079
Marital status (base=married)				
De-facto	-0.231***	0.069	-0.141*	0.073
Separated	-0.530***	0.109	-0.253**	0.118
Divorced	-0.472***	0.084	-0.375***	0.085
Widowed	-0.427***	0.108	-0.524***	0.105
Never married	-0.359***	0.068	-0.339***	0.072
Number of children	-0.101***	0.027	-0.057**	0.029
Number of adults	-0.062**	0.031	0.029	0.033
Education (base=Year 11 or below)				
Year 12	0.002	0.060	0.052	0.062
Certificate	0.066	0.053	-0.074	0.055
Diploma	0.177**	0.072	0.138*	0.073
Graduate degree	0.325***	0.067	0.355***	0.068
Post-graduate	0.348***	0.075	0.402***	0.077
Employment and hours (base=not in labour force)				
Unemployed	-0.188**	0.089	0.086	0.105
Employed part time	0.044	0.057	0.196***	0.061
Employed full time (35-54 hours)	-0.094*	0.056	0.114**	0.057
Employed full time (55 plus hours)	-0.644***	0.092	-0.248***	0.093
Serious long term health condition	-0.064	0.056	-0.012	0.055
Location (base=Sydney)				
Melbourne	-0.015	0.084	-0.388***	0.082
Brisbane	-0.486***	0.093	-0.700***	0.094
Adelaide	-0.122	0.110	0.342***	0.122
Perth	0.693***	0.119	1.644***	0.157
Major city	-0.369***	0.099	-0.054	0.099
Inner regional	-0.317***	0.083	0.075	0.084
Outer regional	-0.484***	0.095	0.275***	0.100
Remote	-0.375**	0.189	-0.203	0.188
SEIFA quintile (base=quintile 1 (most disadvantaged))				
Quintile 2	0.132*	0.073	0.118	0.074
Quintile 3	0.068	0.075	0.184**	0.077
Quintile 4	0.052	0.075	0.102	0.078
Quintile 5 (least disadvantaged)	0.173**	0.083	0.300***	0.083
Moved	-0.655***	0.062	-0.417***	0.064
Has internet	0.501***	0.085	0.441***	0.098
Constant	0.811***	0.222	0.470**	0.233

Note: * p>=0.05 & p<0.1, ** p>=0.01 & p<0.05, *** p<0.01. Standard errors allow for clustering of people in households.

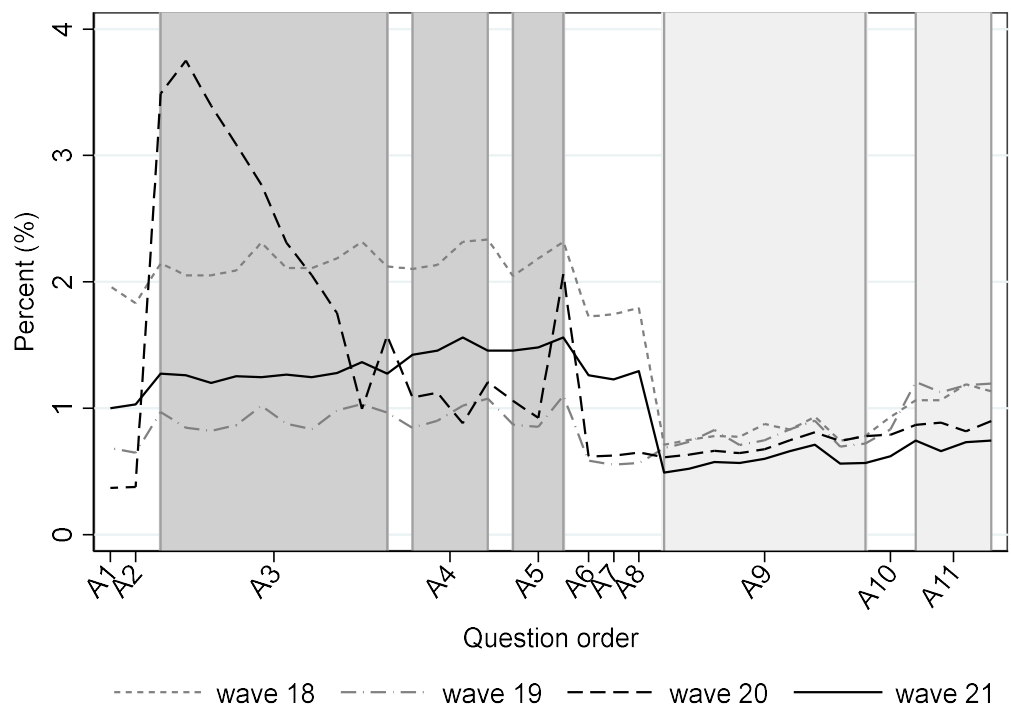
Table A5: Logistic regression coefficients for model predicting completion of SCQ online vs hardcopy

Variable	Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE
Female	0.065	0.046	0.056	0.038
Age (/10)	-0.311***	0.101	0.220**	0.086
Age squared (/10 ²)	-0.026***	0.010	-0.060***	0.009
Dwelling (base=separate house)				
Semi-detached	0.196	0.130	0.299**	0.120
Flat	0.239**	0.116	0.119	0.107
Other private dwelling	-0.102	0.301	-0.606**	0.291
Non-private dwelling	-0.245	0.367	-0.383	0.439
Renter	-0.274***	0.090	-0.125	0.078
Country of birth (base=Australia)				
Main English-speaking country	0.407***	0.100	0.337***	0.096
Not main English-speaking country	-0.336***	0.102	-0.292***	0.094
Part responding HH	-0.002	0.131	-0.150	0.107
Telephone interview	3.905***	0.160	2.461***	0.078
Marital status (base=married)				
De-facto	-0.157	0.116	-0.315***	0.093
Separated	-0.520***	0.156	-0.434***	0.137
Divorced	-0.515***	0.112	-0.468***	0.106
Widowed	-0.427***	0.131	-0.541***	0.140
Never married	-0.878***	0.109	-0.612***	0.095
Number of children	0.062	0.048	-0.102***	0.037
Number of adults	-0.019	0.046	-0.063	0.042
Education (base=Year 11 or below)				
Year 12	0.388***	0.092	0.200**	0.079
Certificate	0.308***	0.073	0.192***	0.069
Diploma	0.805***	0.104	0.563***	0.095
Graduate degree	1.018***	0.104	0.673***	0.089
Post-graduate	1.080***	0.116	0.731***	0.097
Employment and hours (base=not in labour force)				
Unemployed	-0.145	0.144	0.098	0.133
Employed part time	0.441***	0.090	0.248***	0.076
Employed full time (35-54 hours)	0.427***	0.090	0.157**	0.076
Employed full time (55 plus hours)	0.340**	0.172	-0.104	0.126
Serious long term health condition	-0.338***	0.072	-0.261***	0.069
Location (base=Sydney)				
Melbourne	-0.198	0.133	-0.155	0.125
Brisbane	0.020	0.150	-0.317**	0.134
Adelaide	0.021	0.165	-1.074***	0.173
Perth	0.216	0.162	-0.702***	0.152
Major city	-0.075	0.158	-0.515***	0.136
Inner regional	-0.107	0.129	-0.500***	0.114
Outer regional	-0.361**	0.144	-0.401***	0.133
Remote	-0.659***	0.248	-0.368	0.239
SEIFA quintile (base=quintile 1 (most disadvantaged))				
Quintile 2	0.233**	0.101	0.057	0.097

Variable	Wave 20		Wave 21	
	<i>b</i>	SE	<i>b</i>	SE
Quintile 3	0.394***	0.107	0.345***	0.101
Quintile 4	0.508***	0.111	0.457***	0.106
Quintile 5 (least disadvantaged)	0.700***	0.122	0.565***	0.114
Moved	0.168	0.104	0.032	0.087
Has internet	1.914***	0.108	2.213***	0.139
Constant	-2.011***	0.374	-2.074***	0.314

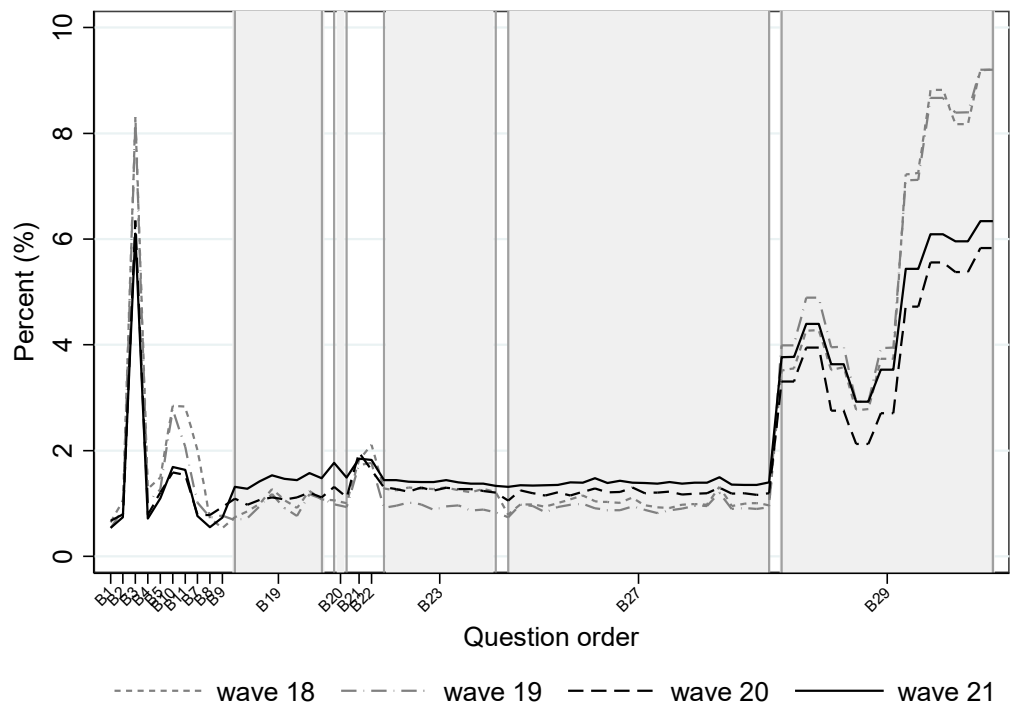
Note: * $p \geq 0.05$ & $p < 0.1$, ** $p \geq 0.01$ & $p < 0.05$, *** $p < 0.01$. Standard errors allow for clustering of people in households.

Figure A1: Proportion missing in SCQ Section A



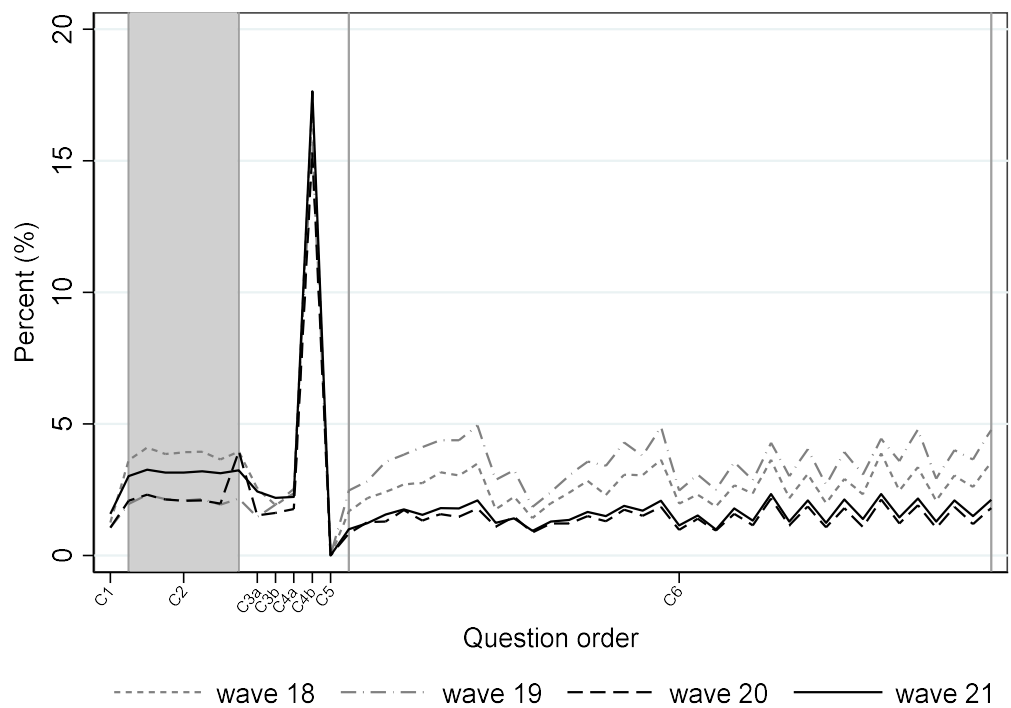
Note: Dark shading corresponds to multi-item questions using the carousel style questions and light shading corresponds to accordion style questions using the accordion style questions in wave 20. In wave 21, the carousel style questions were converted to accordion style questions. Question numbering refers to the wave 21 SCQ.

Figure A2: Proportion missing in SCQ Section B



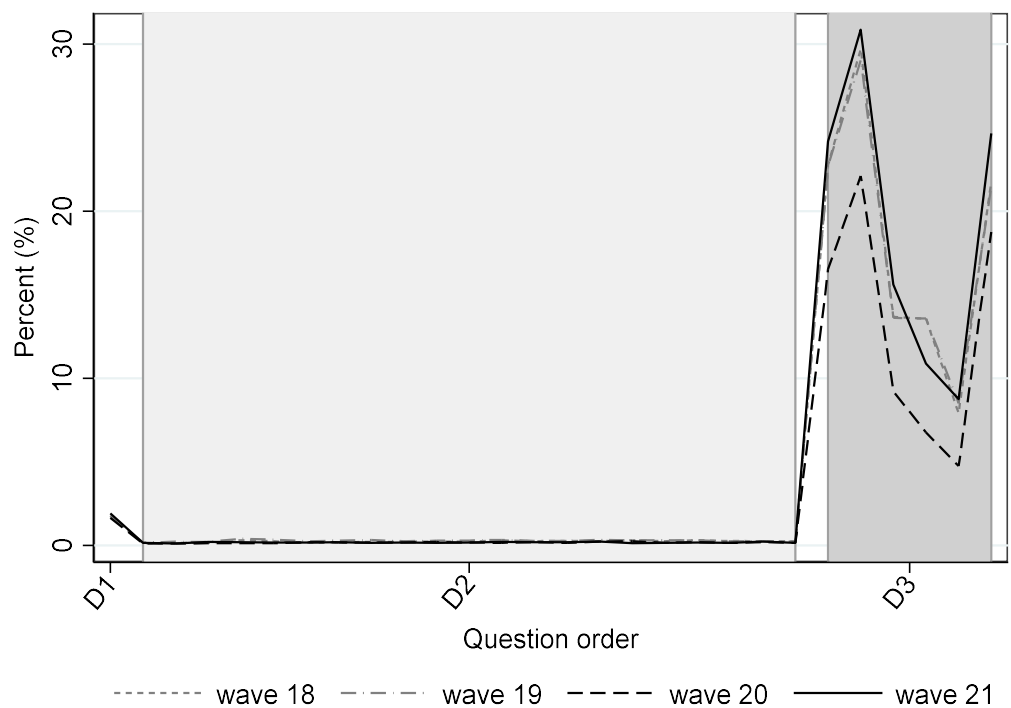
Note: Light shading corresponds to multi-item questions using the accordion style questions. Question numbering refers to the wave 21 SCQ.

Figure A3: Proportion missing in SCQ Section C



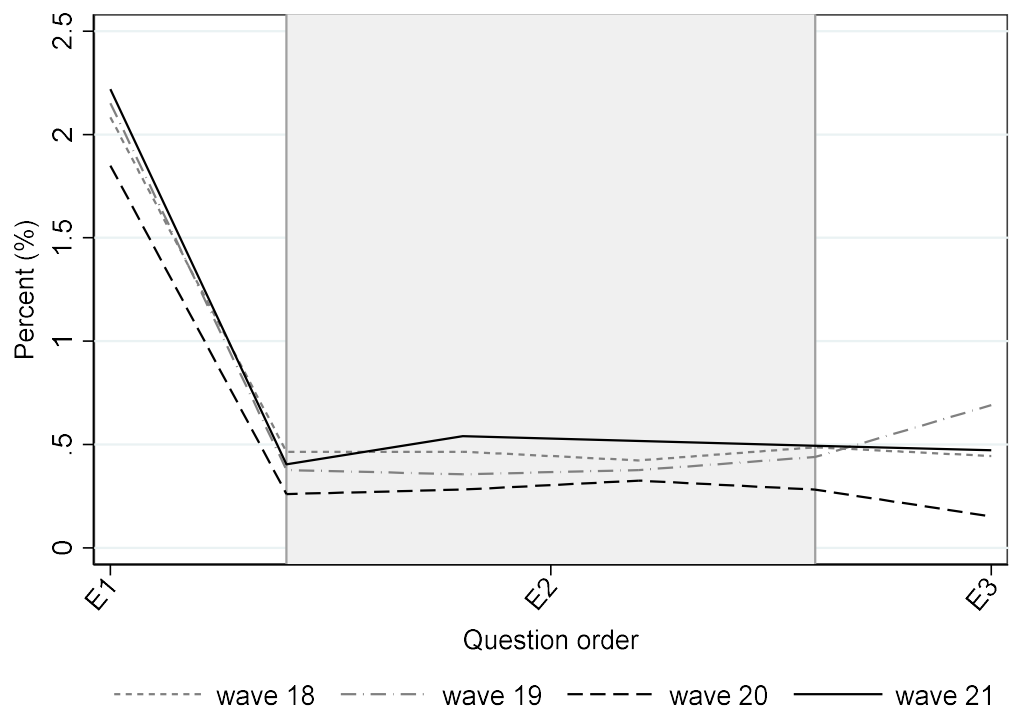
Note: Dark shading corresponds to multi-item questions using the carousel style questions in wave 20. In wave 21, the carousel style questions were converted to accordion style questions. Question numbering refers to the wave 21 SCQ.

Figure A4: Proportion missing in SCQ Section D



Note: Dark shading corresponds to multi-item questions using the carousel style questions and light shading corresponds to multi-item questions using the accordion style questions in wave 20. In wave 21, the carousel style questions were converted to accordion style questions. Question numbering refers to the wave 21 SCQ.

Figure A5: Proportion missing in SCQ Section E



Note: Light shading corresponds to multi-item questions using the accordion style questions. Question numbering refers to the wave 21 SCQ.