



Australian Government

**Department of Infrastructure, Transport,
Regional Development, Communications and the Arts**

Bureau of Communications,
Arts and Regional Research

Australian households and the affordability of telecommunications

Evidence from Household Income and Labour
Dynamics in Australia (HILDA) data

Working paper

November 2023



The Department of Infrastructure, Transport, Regional Development, Communications and the Arts acknowledges the Traditional Custodians of Country of the land on which we work and live. We recognise and respect the continuing connections to land, waters and communities. We pay our respects to them and their cultures and to their Elders both past and present and to all Aboriginal and Torres Strait Islander people.

© Commonwealth of Australia 2023

ISBN 978-1-922879-14-1

November 2023

Ownership of intellectual property rights in this publication

Unless otherwise noted, copyright (and any other intellectual property rights, if any) in this publication is owned by the Commonwealth of Australia (referred to below as the Commonwealth).

Disclaimer

The material contained in this publication is made available on the understanding that the Commonwealth is not providing professional advice, and that users exercise their own skill and care with respect to its use, and seek independent advice if necessary.

The Commonwealth makes no representations or warranties as to the contents or accuracy of the information contained in this publication. To the extent permitted by law, the Commonwealth disclaims liability to any person or organisation in respect of anything done, or omitted to be done, in reliance upon information contained in this publication.

Creative Commons licence

With the exception of (a) the Coat of Arms; (b) the Department of Infrastructure, Transport, Regional Development and Communications photos and graphics; and (c) [OTHER], copyright in this publication is licensed under a Creative Commons Attribution 4.0 Australia Licence.

Creative Commons Attribution 4.0 Australia Licence is a standard form licence agreement that allows you to copy, communicate and adapt this publication provided that you attribute the work to the Commonwealth and abide by the other licence terms.

Further information on the licence terms is available from <https://creativecommons.org/licenses/by/4.0/>

This publication should be attributed in the following way: © Commonwealth of Australia 2023

Use of the Coat of Arms

The Department of the Prime Minister and Cabinet sets the terms under which the Coat of Arms is used. Please refer to the Commonwealth Coat of Arms - Information and Guidelines publication available at <http://www.pmc.gov.au>.

Contact us

This publication is available in hard copy or PDF format. All other rights are reserved, including in relation to any departmental logos or trade marks which may exist. For enquiries regarding the licence and any use of this publication, please contact:

Director – Creative Services

Communication Branch

Department of Infrastructure, Transport, Regional Development and Communications and the Arts

GPO Box 594

Canberra ACT 2601

Australia

Email: publishing@infrastructure.gov.au; Website: www.infrastructure.gov.au

Table of contents

1. Key findings	5
2. Introduction	6
3. Related literature	7
4. Data	8
Telecommunications expenditure in HILDA	8
Trends in affordability	9
Telecommunications — a necessity	11
5. Household spending on telecommunications – summary statistics	12
6. Household spending on telecommunications – OLS estimates	13
Households which spend less on telecommunications	15
Households which spend more on telecommunications	16
7. ‘Low-income, low spending’ and ‘low-income, high spending’ groups – descriptive statistics	17
8. ‘Low-income, low spending’ and ‘low-income, high spending’ groups – logistic regression estimates	19
Low-income, high spending households	19
Low-income, low spending households.....	22
9. Conclusion and discussion	24
10. Acknowledgements	25
Attachment A — Trends in the key variables	26
Attachment B — Significance of differences in means for demographic and socioeconomic groups	27
Attachment C — OLS regression estimates	29
Attachment D — Predicted average communication expenditures using OLS estimates	36
Attachment E — Log transformations	38
Attachment F — Logistic regression estimates	39
Attachment G — Logistic regression diagnostics	48
11. References	49

List of figures and tables

Figure 1: Share of household disposable income spent on telecommunications, mean and median.	9
Figure 2: Average household telecommunications spending and disposable income.	10
Figure 3: Average share of household disposable income spent on telecommunications by equivalised disposable income decile, 2006 and 2021.....	10
Figure 4: Average share of disposable income spent on necessity goods and services, all households and equivalised disposable income decile 1, 2021.....	11
Figure 5: Average share of disposable household income spent on telecommunications by characteristic (2006 – 2021).	13
Figure 6: Average marginal effects calculated using OLS estimates, ‘Some’ and ‘All’ models.....	14
Figure 7: Share of low-income households with relatively high or low telecommunications expenditures.	18
Figure 8: Average share of income spent on telecommunications by low-income households.....	18
Figure 9: Average marginal effect calculated using logistic regression estimates, ‘Low-income, high spending’ (LIHS) group, ‘Some’ and ‘All’ models.	21
Figure 10: Average marginal effect calculated using logistic regression estimates, ‘Low-income, low spending’ (LILS) group, ‘Some’ and ‘All’ models.....	23
Table 1: Household disposable income and telecommunications expenditure, by wave.	26
Table 2: Difference in mean telecommunications expenditure share by characteristic, all waves.	27
Table 3: OLS regression results, ‘Some’ and ‘All’ models without interaction terms.	29
Table 4: OLS regression results, ‘Some’ and ‘All’ models with interaction terms.....	31
Figure 11: Predicted share of telecommunications spending in household disposable income, ‘Some’ model.....	36
Figure 12: Predicted share of telecommunications spending in household disposable income, ‘All’ model.....	37
Figure 13: Histogram visuals of natural log transformations.	38
Table 5: Logistic regression estimates, LIHS group, ‘All’ and ‘Some’ models without interaction terms.....	40
Table 6: Logistic regression estimates, LIHS group, ‘All’ and ‘Some’ models with interaction terms.....	42
Table 7: Logistic regression estimates, LILS group, ‘All’ and ‘Some’ models without interaction terms. ..	44
Table 8: Logistic regression estimates, LILS group, ‘All’ and ‘Some’ models with interaction terms.....	45
Table 9: Key logistic regression diagnostics, LIHS and LILS groups, ‘Some’ and ‘All’ models.....	48

1. Key findings

This paper uses Household Income and Labour Dynamics in Australia (HILDA) data to examine telecommunications spending by Australian households. It does so with an aim to better identify vulnerable households who are most at risk of potential digital exclusion.

Our analysis builds on previous research by the Bureau of Communications, Arts, and Regional Research (BCARR 2017; 2020) and Breunig and McCarthy (2020). To assess the affordability of telecommunications for households, this study augments the econometric approach developed by Breunig and McCarthy (2020) with additional variables found to be important to telecommunications affordability.

Continuing from Breunig and McCarthy's (2020) analysis, this paper also applies regression modelling to identify two low-income household groups that are at-risk of potential digital exclusion. The first group, 'low-income, high spending', spends much more on telecommunications relative to their income. Their behaviour may not be financially sustainable. The second group, 'low-income, low spending', spends much less on telecommunications relative to their income, and this lack of spending may not be enough for these households to realise the benefits of digital connectedness.

Key findings from this analysis include:

- **Affordability of telecommunications services is improving.** Australian households now spend a lower share of their disposable income on telecommunications. Household spending on telecommunications services dropped from a peak of 4.1 per cent of disposable income in 2008, to 3 per cent in 2021.
- **Affordability improvements in 2021 were driven by an increase in household disposable income and a marked drop in nominal telecommunications spending.**
- **Household spending on telecommunications, which includes telephone bills, calls and internet charges, behaves like other necessities such as food.** The share of telecommunications spending in household disposable income declines as household income grows. Low-income households spend a much larger share of their household budget on telecommunications than do higher income households.
- When controlling for income and other factors, we identified the following **households that spend less on telecommunications relative to their disposable income: single person households, households where English is spoken poorly, older person households, Aboriginal or Torres Strait Islander households, immigrant households, households with low education, or where members have a disability/long term health condition or are not employed.** Low telecommunications spending as a share of disposable income could reflect an inability for these households to afford a reliable, quality internet connection, and puts these groups at heightened risk of potential digital exclusion.
- When controlling for income and other factors, we found a **higher spend on telecommunications relative to their disposable income in households where members: work from home, are financially stressed, are younger or middle-aged, where children are present, or if the household is located either in a rural area, or in a more prosperous area.** Higher spending as a share of disposable income on telecommunications may reflect their financial difficulties, inappropriate internet bundle, higher demand and use of the internet amongst these groups, or higher costs of their telecommunications service delivery.

- **The share of low-income households with high, potentially unsustainable, spending on telecommunications increased from 5.6 per cent in 2010 to 5.8 per cent in 2021 to be at the highest level in 12 years.** Households were more likely to belong to this group if members worked from home, were younger or middle aged, in financial stress, had children, were Aboriginal or Torres Strait Islander, were immigrants from a non-English speaking background, or if the household size was larger or in a rural area.
- **Households identified as ‘low-income, low spending’ represent a small but stable share of Australian households (approximately 0.6 per cent of all households).** These households spend less than one per cent of their disposable income on telecommunications, which puts them at high risk of potential digital exclusion.

2. Introduction

The importance of access to affordable telecommunications services and the devices which support them has increased. From attending lectures and working from home to facilitating entertainment and keeping in touch with family and friends, access to telecommunications is integral to our everyday life. The affordability of telecommunications is an important enabler for Australians to participate effectively in society.

The affordability of telecommunications affects how digitally inclusive we are as a nation. Telecommunications affordability issues can create a ‘digital divide’ leaving some Australians unable to fully participate in digital life. Understanding the characteristics of households facing affordability barriers is essential to addressing digital exclusion.

This paper uses data from the HILDA survey to examine household spending on telecommunications and the socioeconomic factors affecting the affordability of telecommunications. This analysis extends earlier research by BCARR (2017, 2020) and Breunig and McCarthy (2020). It identifies the socioeconomic factors linked to the affordability of telecommunications – measured by the share of a household’s disposable income that is spent on telecommunications.

This study also seeks to understand the composition of low-income households that have potentially unsustainable spending on telecommunications. The two groups identified are at heightened risk of potential digital exclusion either because they overspend on telecommunications (and risk running into financial difficulties), or underspend (and therefore may not be fully connected). By identifying these vulnerable household groups, the research aims to inform policies designed to improve telecommunications affordability and digital inclusion.

The paper is set out as follows. Section 3 reviews the small number of studies which examine the drivers of telecommunications spending in Australia. Sections 4 and 5 describe the data we use and key trends in household spending on telecommunications. Section 6 discusses the characteristics of households spending more or less on telecommunications. Section 7 examines trends in low-income households that have either very low or very high spending on telecommunications. Section 8 identifies the socioeconomic characteristics of these vulnerable households. Section 9 concludes and discusses some of the limitations of this research.

3. Related literature

There is a small body of literature on the patterns and socioeconomic drivers of telecommunications expenditures in Australia. Previous BCARR studies drew on HILDA data to identify groups of households which spend a higher than average proportion of their disposable income on telecommunications. These were: households with at least one member who was: Aboriginal or Torres Strait Islander, not employed, aged 65 or older, a student, and households in rural areas (BCARR 2017, 2020).

Breunig and McCarthy (2020) took this analysis further and modelled the relationship between a household's characteristics and its share of disposable income spent on telecommunications using HILDA data over a ten-year period (2006 to 2015). They found that households with younger people spent more on telecommunications (as a share of their disposable income), while households with older people spent less. In general, higher income households spent a lower proportion of income on telecommunications. Once income and household size were controlled for, poor health, Indigenous status, speaking English poorly, being an immigrant from a non-English speaking country, and not being employed (unemployed or not in the labour force) were all found to be associated with lower than average telecommunications spending (as a proportion of disposable household income). In contrast, living in remote or rural areas and being in financial stress were found to be associated with higher than average telecommunications spending.

Breunig and McCarthy also found that telecommunications spending behaves like a necessity – households on low incomes spend a high share of their disposable income on telecommunications, but this share of spending becomes smaller as household incomes grow. Further, they identified two groups of low-income households with unsustainable patterns of spending on telecommunications: one group overspending, and the other group underspending. The first group spends a very high share of their household budget on telecommunications; while the second group spend 'too little' in some sense and may miss out on the benefits associated with using telecommunications. Both groups, Breunig and McCarthy argued, were at heightened risk of digital exclusion because of their spending patterns.

Using the same dataset, Ali, Alam, Taylor and Rafiq (2019) examined the relationship between the affordability of telecommunications and household income distribution and socioeconomic inequality. The authors tested two measures of telecommunications affordability: a measure of household annual expenditure on telecommunications, and a composite index measuring aspects of digital inclusion, such as the share of household income spent on telecommunications and the total internet data allowance per dollar of expenditure. Using both measures, they found that socioeconomic advantage translates into digital advantage by impacting affordability of telecommunications. In particular, they found that the greater household socioeconomic advantage, the greater its affordability of telecommunications. Further, Ali *et al.* found that while affordability was negatively correlated with age, being located in a major city or an urban area had a positive effect on affordability.

Thomas, McCosker, Parkinson, Hegarty, Featherstone, Kennedy, Holcombe-James, Ormond-Parker, Ganley (2023) used the Internet Usage Survey to construct a measure of telecommunications affordability that served as an input into the Australian Digital Inclusion Index. They found that in 2022, on average, 4 per cent of Australians would need to pay more than 10 per cent of their household income to gain quality, reliable connectivity. This proportion increases to 27 per cent in the lowest income quintile (households with a total income of less than \$33,800 per annum). In this group the affordability of telecommunications services is particularly low. Lower affordability was also found for those unemployed, aged above 75 years, those renting from a public housing authority, those that did not complete secondary school, people with disability, those receiving income support and Aboriginal and Torres Strait Islanders.

4. Data

This analysis uses data from the HILDA survey, a longitudinal survey that collects information on respondents' economic and demographic characteristics. The survey started in 2001 and has been collected annually since. The original sample collected in 2001 was 13,969 individuals and 7,682 households across all Australian states and territories. The sample was topped up in 2011 with an additional 4,009 individuals and 2,153 households added. In 2021, the most recent year of HILDA data available, the sample comprised 16,549 individuals in 9,358 households.¹ See Watson and Wooden (2010) for an in-depth discussion of the HILDA data.

Telecommunications expenditure in HILDA

HILDA started to collect information on household spending on telecommunications from 2006, allowing for 15 years of data to be analysed. The key variable of interest is a derived variable within HILDA, *_hxytlii* which captures household annual expenditure on telecommunications. This variable is based on responses to the HILDA Self-Completion Questionnaire on monthly expenses on telephone rent and calls, and internet charges. This variable has been imputed if a respondent did not answer this question.² While the information collected through this variable allows us to estimate household spending on telecommunications, it is not possible to examine the types of products purchased, or the quality, quantity and prices of services consumed. The information collected also does not allow us to account for the prices of devices other than phones (such as laptops, smart tvs, tablets) used to connect to internet.

We construct a measure of telecommunications affordability by dividing a household annual spending on telecommunications by its total disposable income.³ High values of this measure indicate households with the lowest levels of telecommunications affordability. Low values indicate that a household can more easily afford telecommunications. We drop observations where household disposable income is negative or when the values of our constructed measure of affordability are below or equal to zero, or where observations are greater than or equal to one.⁴ Values above one are implausible. Households with negative income, which generate the values below zero, have business or investment losses and this negative reported income may not be a good measure of their ability to consume goods and services.

¹ Individuals refer to the number of adult persons that completed an individual interview. There may be other adult members in a responding household who were not interviewed, along with children under 15 years of age. Responding households refer to the number of households for which interviews were completed with at least one eligible member of the household.

² In the regression modelling we control for whether or not this variable was imputed. Around 16 per cent of all observations have been imputed.

³ Household disposable income is derived by subtracting the negative financial year disposable income (*_hifditn*) from positive financial year disposable income (*_hifditn*). As this variable is primarily used as a proportion, it has not been adjusted for inflation. In our regression analysis, we include time dummies which control for changes in the overall price of telecommunications and inflation in household income.

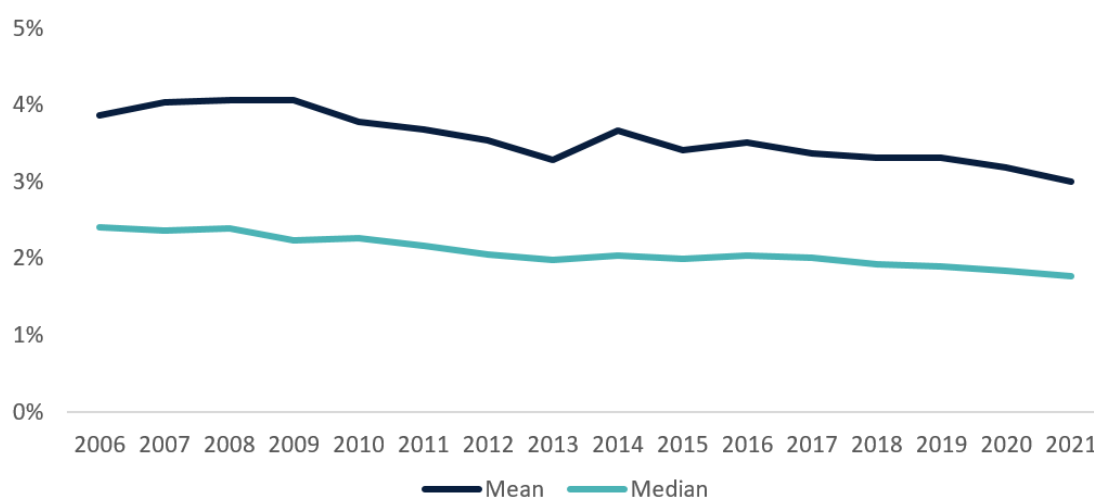
⁴ 392 observations (0.3 per cent) were removed as they had negative disposable income. 4,597 observations (3.4 per cent) were dropped as their spending on telecommunications as a share of disposable income was equal to zero or was greater than or equal to one.

Trends in affordability

Household spending on telecommunications as a share of disposable income has fallen in recent years (Figure 1). Average household expenditure on telecommunications declined from a high point of 4.1 per cent of disposable income in 2008, down to 3 per cent in 2021. The median household share of income spent on telecommunications is lower than the mean (average) share spent on telecommunications.⁵ We describe household characteristics associated with higher than average telecommunications spending in section 5 (Figure 5).

Over the time period analysed, reductions in the share of household income spent on telecommunications have largely been driven by increases in disposable income rather than reduced spending on telecommunications. In 2014, there was an uptick in household spending on telecommunications. This increase was likely due to Australians engaging more intensively online, downloading more data and making greater use of mobile handsets, including smartphones.⁶ The consumption of telecommunications around this time was also substantially impacted by the merging of telecommunications technology and media content, and the rise of online platforms and over-the-top services (BCAR 2016).

Figure 1: Share of household disposable income spent on telecommunications, mean and median.



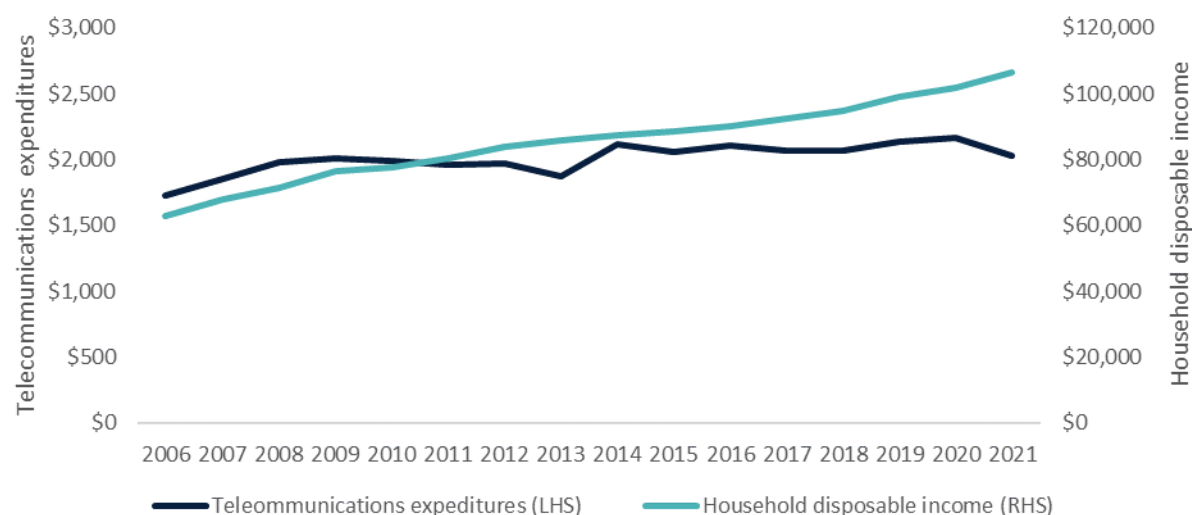
Source: The HILDA Survey, Release 21; BCARR calculations.

For the majority of the time period analysed, household spending on telecommunications was constant, at approximately \$2,000 a year (Figure 2). In 2021, the most recent year of HILDA data available, there was a marked decline in telecommunications spending. This lower-than-average spending could be partially driven by a range of assistance packages provided by telecommunications companies during COVID-19 (Wiwatowska 2020). It may also be a result of a drop in the prices of telecommunications observed during this period.⁷ Unfortunately, HILDA does not collect information that would allow us to distinguish between the prices, quantity and quality of telecommunications consumed.

⁵ There is a high proportion of households spending very little of their disposable income on telecommunications therefore our variable is strongly positively skewed. For this reason, we use in our OLS model a log of the share of telecommunications expenditures in disposable income. For more on this see: *Attachment E — Log transformations*

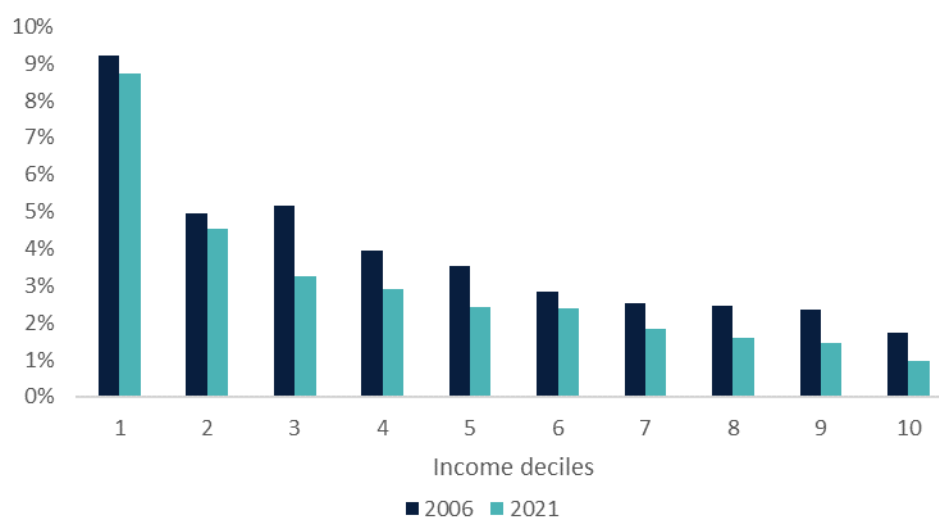
⁶ The total volume of data downloaded in Australia during the June quarter of 2014 was 53 per cent higher than the volume downloaded for the same period in 2013. Video and audio content contributed to the continued growth in the volume of data downloaded: 44 per cent of adult Australians (6.4 million) streamed music, movies, TV programs, video clips or radio — a 21 percentage point increase over the past five years (ACMA 2014, ACCC 2014).

⁷ ACCC (2022) found that in 2021-22, telecommunications prices dropped across all types of services. Feature-adjusted price indexes dropped by 6 per cent, 11.6 per cent and 15.3 per cent for fixed broadband, mobile phone services and mobile broadband, respectively.

Figure 2: Average household telecommunications spending and disposable income.

Source: The HILDA Survey, Release 21; BCARR calculations.

We further analyse telecommunications spending by income deciles – ten equal-sized categories of households ranked from lowest (decile 1) to highest (decile 10) by their disposable income.⁸ In the 15 years to 2021, telecommunications spending shares declined across all income deciles (Figure 3). This decline was particularly pronounced for households in higher income deciles where the share of disposable income spent on telecommunications dropped by nearly half.

Figure 3: Average share of household disposable income spent on telecommunications by equivalised disposable income decile⁹, 2006 and 2021.

Source: The HILDA Survey, Release 21; BCARR calculations.

⁸ Household income deciles are a way to divide a population into ten equal groups, or 'deciles', based on their income levels. Each decile represents 10 per cent of the population, sorted in ascending order of income. In 2021, the lowest decile of household equivalised disposable income was \$26,250 per annum or less, and the highest decile was \$112,240 per annum or more.

⁹ Household equivalised disposable income was constructed by dividing household disposable income by the square root of the number of individuals in the household.

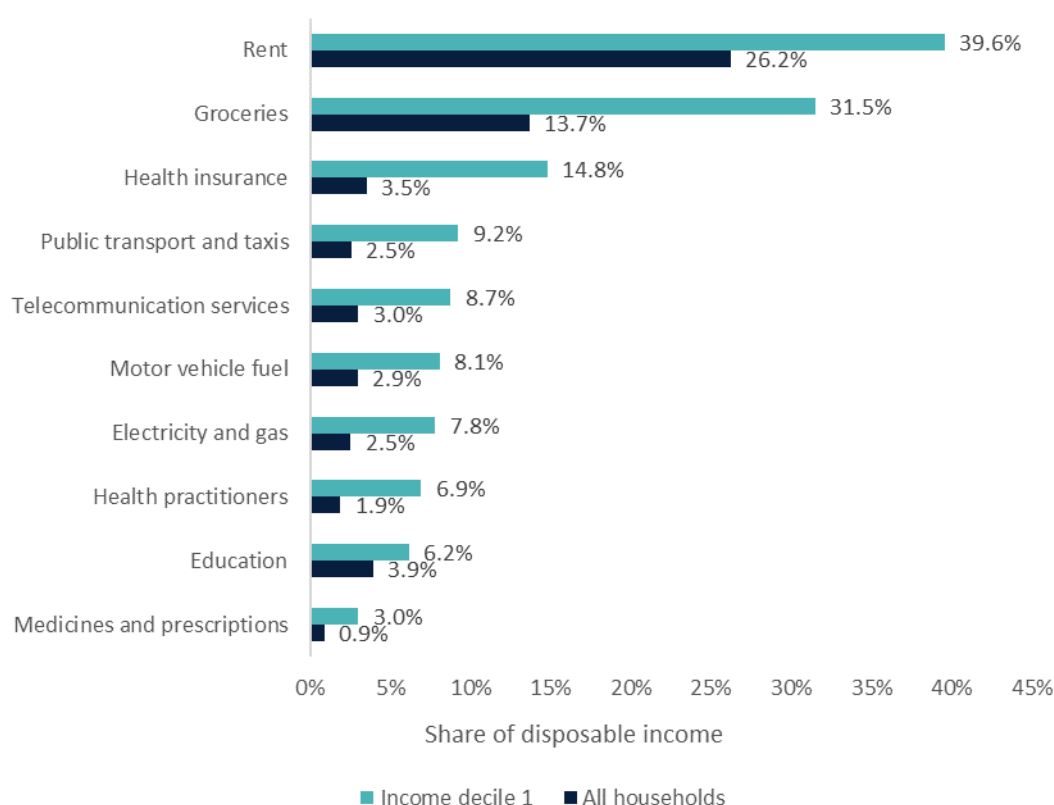
As discussed in Breunig and McCarthy (2020), household spending on telecommunications tends to behave like spending on other necessities, such as food. That is, as incomes grow, the share of disposable income spent on them tends to decline. Despite the recently observed drop in nominal spending on telecommunications, households in the lowest income decile continue to spend a significant portion of their disposable income on telecommunications compared to households in the higher income deciles.

Telecommunications — a necessity

In 2021, telecommunications services were the fifth largest spending category as a share of disposable income, behind rent, groceries, health insurance and public transport for households in the lowest income decile (Figure 4). During this time, households in the lowest income decile spent on average 5.7 percentage points more of their disposable income share on telecommunications compared to the average expenditure across all households (8.7 per cent compared to 3.0 per cent).

If we reproduce Figure 4 for each year of the data, we observe that over the 15 years to 2021, spending on telecommunications in the poorest households ranked consistently high, with an average of 9 per cent of their disposable income spent on telecommunications. We do not present these figures in the paper. Over the period discussed, a household in the lowest income decile spent between 5 and 6 percentage points more of their disposable income on telecommunications than an average household.

Figure 4: Average share of disposable income spent on necessity goods and services, all households and equivalised disposable income decile 1, 2021.



Source: The HILDA Survey, Release 21; BCARR calculations.

5. Household spending on telecommunications – summary statistics

In this section, we seek to identify characteristics of households spending higher and lower shares of their income on telecommunications. This can inform policy designed to improve digital inclusion of Australian households. Using HILDA data, we classify households into two groups to reflect the possible characteristics of the household: ‘some’ where at least one household member exhibits a given characteristic and ‘all’ in which all household members exhibit the given characteristic. For example, a household could have one member from a non-English speaking background and one from an English-speaking background. This household would be classified as having ‘some’ members from a non-English speaking background but not as having ‘all’ members from a non-English speaking background.

Figure 5 compares the average share of disposable income spent on telecommunications across a variety of household characteristics that the previous literature has identified as being associated with low/high levels of telecommunications affordability. Averages in Figure 5 are calculated across the entire HILDA sample, from 2006 to 2021. Figure 5 uses dotted bars to indicate where the differences from the average were not statistically significant (that is, if their p-values were above 10 per cent).¹⁰ Further details of the level of significance and standard errors are in Table 2 of *Attachment B – Significance of differences in means demographic and socioeconomic groups*.

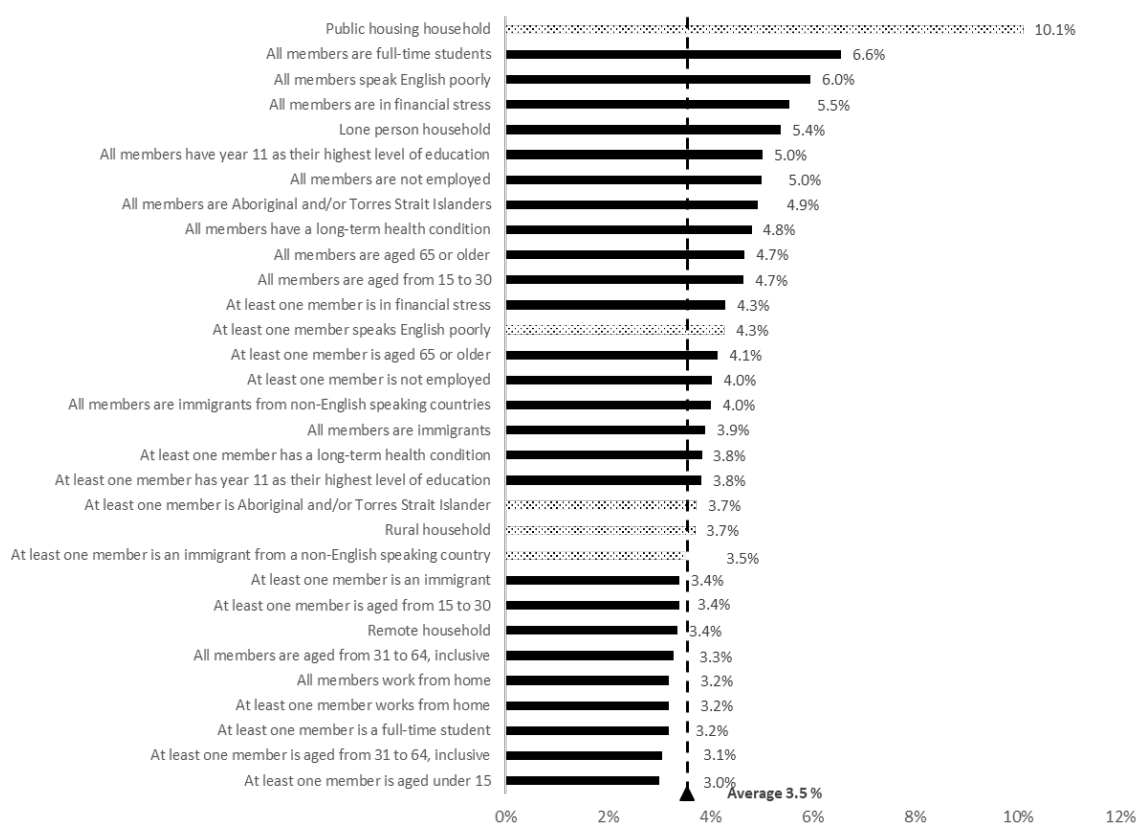
On average, the share of disposable income spent on telecommunications is higher in households where:

- all members are full-time students;
- all members speak English ‘not well’ or ‘not very well’;
- at least one member was in financial stress;
- the household is a lone person household;
- at least one member has year 11 as their highest level of education;
- at least one member is not employed;
- at least one member has a long-term health condition;
- at least one member is aged 65 or older;
- all members are aged from 15 to 30;
- all members are Aboriginal or Torres Strait Islander;
- all members are immigrants to Australia; and
- all members are immigrants from non-English-speaking countries.¹¹

Spending shares on telecommunications are typically higher among lower income households and in groups considered vulnerable. While this information is helpful in identifying households with barriers to affordability, it does not account for the confounding effect of other variables on them. For example, Aboriginal or Torres Strait Islander households might have higher telecommunications spending not because they are spending more on telecommunications, but because their average disposable income is lower and their household size is larger than average.

¹⁰ Statistical significance was determined using t-tests. Here, we compare whether households with the identified characteristic have an expenditure share that is statistically different to the average across all households — 3.5 per cent. A p-value that is greater than 10 per cent is considered not statistically significant, while one that is less than one per cent is viewed as highly statistically significant. When the difference is not statistically significant, we conclude that the observed differences could simply be due to the particular sample of individuals and should not be interpreted as a difference in the population.

¹¹ Non-English-speaking countries refer to countries other than Australia, UK, New Zealand, Canada, US, Ireland and South Africa.

Figure 5: Average share of disposable household income spent on telecommunications by characteristic (2006 – 2021).

Note: Dotted bars indicate differences that were not statistically significant (p-values were above 10 per cent).

Source: The HILDA Survey, Release 21; BCARR calculations.

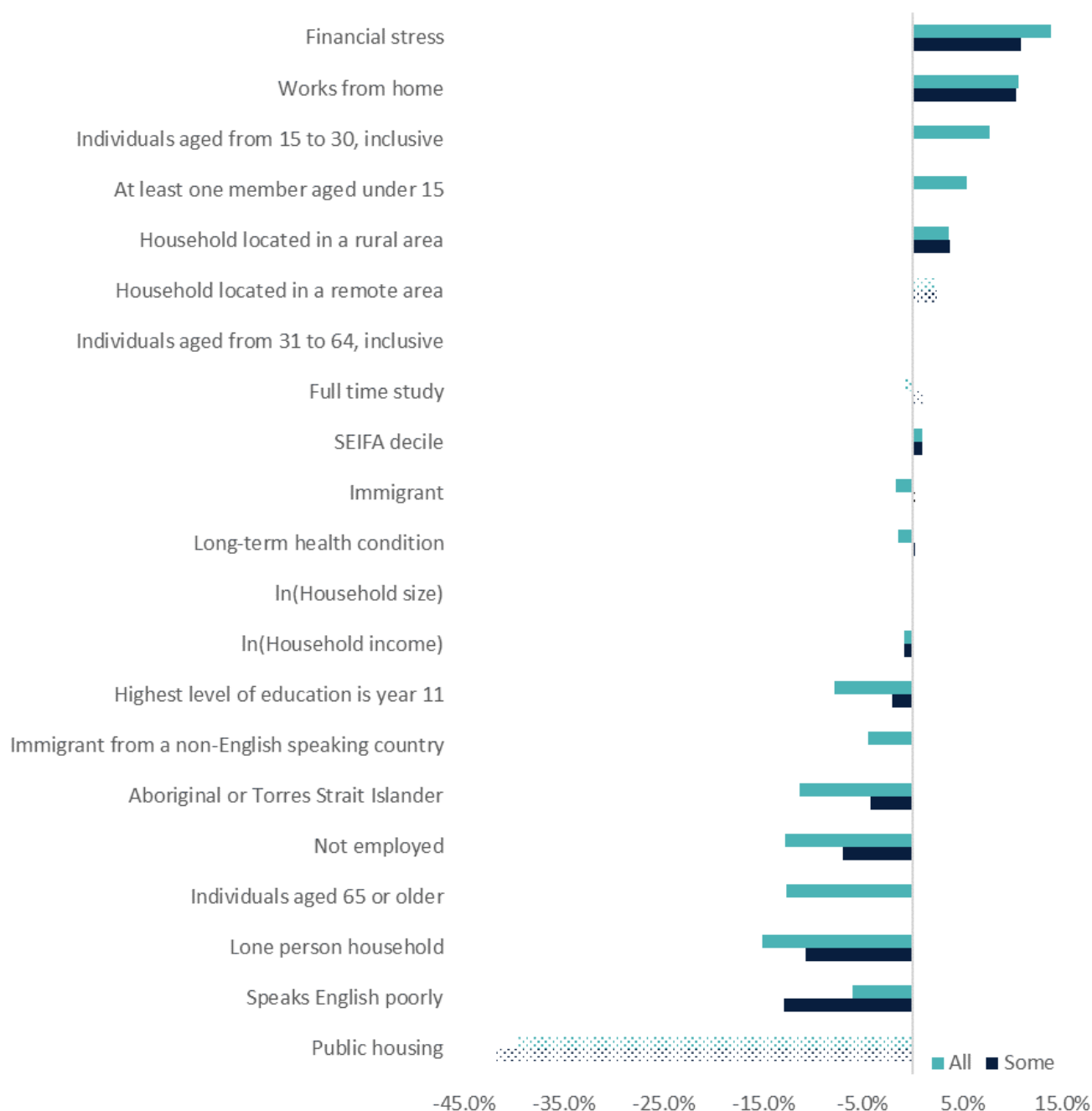
6. Household spending on telecommunications – OLS estimates

In this section, we use an ordinary least squares (OLS) regression to estimate the effect of a particular household characteristic on telecommunications spending, while holding all other household characteristics constant. For example, we use OLS regression to estimate whether the presence of a household member who is not employed impacts telecommunications spending for the household, holding all the other characteristics constant (such as income and household size, etc). Detailed regression results are presented in *Attachment C — OLS regression estimates*.

We measure the strength of the association between the response and independent variables by calculating their average marginal effects. Figure 6 illustrates the marginal effects of the variables used in the OLS regressions. The marginal effects for a binary, independent variable¹² refer to the percentage difference in the share of telecommunications expenditure between a household that exhibits a given characteristic and one that does not.¹³ Aside from this characteristic, the two households are identical. For instance, Figure 6 shows that households with some members not employed spent on average 7 per cent less of their income on telecommunications compared to households where all members are employed.

¹² When a household exhibits a given characteristic, the binary variable is set to 1, and 0 otherwise.

¹³ For the variables that are expressed in natural logs, such as log of household size and log of household income, marginal effects refer to their elasticity with respect to the response variable, i.e., percentage change in the response variable due to a one per cent change in a given variable while holding all other variables constant.

Figure 6: Average marginal effects calculated using OLS estimates, 'Some' and 'All' models.

Notes: 'Some' refers to OLS regression estimates using variables set to 1 if some members of the household displayed a given characteristic. 'All' refers to the regression estimates using variables set to 1 if all members of the household displayed this particular characteristic.

Dotted bars indicate marginal effects that were not statistically significant (p-values were above 10 per cent).

Marginal effects of log of household size and log of household income refer to the percentage change in the response variable from a 1 per cent change in their levels while holding all the other control variables constant. For other variables which are binary in nature, marginal effects refer to the percentage change in the response variable if a given variable changes from 0 to 1 while holding all the other variables constant. For SEIFA, the marginal effect refers to the percentage change in the response variable if the SEIFA index increases by 1 decile.

Source: The HILDA Survey, Release 21; BCARR calculations.

As expected, household income has a negative effect on the share of the household budget that is spent on telecommunications. On average, a 1 per cent increase in household disposable income results in a 0.9 per cent decline in its telecommunications spending share. This finding is consistent with the inverse relationship between telecommunications expenditure shares and the higher income deciles, shown earlier in section 4 of the paper (Figure 3).

Households which spend less on telecommunications

Lower spending on telecommunications (as a proportion of disposable income) and negative marginal effects are found for households where:

- at least one member who speaks English poorly;
- household member lives alone;
- every member is aged 65 or older;
- at least one member is not employed;
- at least one member is Aboriginal or Torres Strait Islander;
- every member is an immigrant from a non-English speaking country;
- at least one member has year 11 as their highest level of education;
- every member is an immigrant; and
- every member has a long-term health condition.

Importantly, most of the characteristics of households that spend less on telecommunications (as a proportion of disposable income), shown above, have the same characteristics as households which had higher than average shares of telecommunications spending (Figure 5). The results of the regression analysis confirmed that these variables are also strongly negatively correlated with income. Expenditure shares on telecommunications were high not because these households were spending more on telecommunications but because their disposable incomes were low. Once we control for income (and household size), households with these characteristics are associated with a lower propensity to spend on telecommunications and, therefore, are at potential risk of digital exclusion. For example, Figure 5 shows that households with all or some members who are Aboriginal or Torres Islanders spent a higher than average share of their disposable income on telecommunications. The OLS regression showed that once controlled for income and other characteristics, these households actually spent less on telecommunications as demonstrated by their negative marginal effect illustrated in Figure 6.

While the information collected in HILDA does not allow us to draw conclusions about the volume of telecommunication services consumed, some of the groups identified in our research as spending less on telecommunications have also been identified in other research as consuming less broadband (Kenny, Kenny and Gehan, (2023)). For example, an older person living alone is likely to use broadband far less and spend considerably less on telecommunications services than a family of five.

There are instances in which a household will have an even lower propensity to spend on telecommunications when they exhibit more than one of these characteristics at the same time. For instance, the above list identifies that an Aboriginal and Torres Strait Islander household spends less on telecommunications than a non-Aboriginal and Torres Strait Islander household. However, the Aboriginal and Torres Strait Islander household will spend an even lower share of income on telecommunications if it is also larger in size. Similarly, households with member(s) aged 65 or older spend an even lower share on telecommunications if their members live alone. Likewise, lone person households spent a lower share on telecommunications during the COVID-19 pandemic in 2020 and 2021.¹⁴

¹⁴ For the estimates of the interaction effects in the OLS regression see Table 4 in *Attachment C — OLS regression estimates*.

Households which spend more on telecommunications

At the other end of the spectrum, there are households that spend more of their disposable income on telecommunications (Figure 6). The predicted share of disposable income spent on telecommunications was higher and marginal effects positive for households:

- where at least one member reported being under financial stress;
- with at least one member working from home;
- with every member aged from 15 to 30;
- with children present;
- located in a rural area;
- with every member aged from 31 to 64; and
- located in higher Socioeconomic Indexes for Areas (SEIFA) deciles¹⁵.

Household size is also positively associated with expenditure on telecommunications. An increase of 1 per cent in household size, increases the telecommunications expenditure share by between 0.08 and 0.05 per cent, on average. This finding is in line with other research pointing to household size being one of the main predictors of data usage and broadband traffic (Kenny, Kenny and Gehan (2023)).

Higher spending on telecommunications (as a proportion of income) for households in financial stress could represent their financial difficulties, or it could be a factor contributing to it. Higher telecommunications spending for households with children might be closely correlated with their generally larger household size and therefore higher internet consumption. Higher expenditures on telecommunications for younger households (with at least one member aged from 15 to 30) may reflect the greater number of activities performed online by younger Australians and the importance of connectivity to education. ACMA (2022) found that over 90 per cent of those aged from 18 to 34 used the internet to email, access news, shop and stream audio content. More than half of this age group also used the internet to work from home and to attend telehealth consults, while 45 per cent used the internet to study from home.

Higher telecommunications spending by households in rural and remote areas may reflect the higher costs of telecommunications in these areas – for example, internet services using satellite cost more than equivalent plans using fixed-line broadband in Australian cities. They may also be a sign of a lack of alternative telecommunication providers and limited competition in these areas.

Lastly, higher telecommunications spending for households in areas of greater socioeconomic advantage is in line with the other reported research findings (Ali *et al.* 2019). Higher telecommunications spending in these areas could be because wealthier and less disadvantaged areas generally have better access to the latest telecommunications infrastructure including high-speed internet, advanced mobile networks and a wider range of telecommunications providers.

There are instances in which a household will have an even higher spending on telecommunications when they exhibit more than one of these characteristics at the same time.¹⁶ For example, households in rural areas and households with at least one full-time student spend even more on telecommunications as their household size increases. Similarly, households with full-time students residing alone tend to spend more on telecommunications.

While the above analysis finds that households where: members work from home, were financially-stressed, or located in a rural area, generally spend more on telecommunications, our analysis has revealed that during the COVID-19 pandemic in 2020 and 2021, these households actually spent less than

¹⁵ Socio-Economic Indexes for Areas (SEIFA) are area-based deciles, calculated by dividing the areas, ordered by disadvantage, into 10 equally sized groups. Decile 1 contains the most disadvantaged areas.

¹⁶ For estimates of the interaction effects in the regression see Table 4 in *Attachment C — OLS regression estimates*.

they did before.¹⁷ This is consistent with a general drop in telecommunications spending for all households during this time, as discussed in section 4 of this paper (Figure 2). The only households that spent more on telecommunications during the COVID-19 pandemic were households with children, and those with some or all members having a long-term health condition.

7. 'Low-income, low spending' and 'low-income, high spending' groups – descriptive statistics

This section examines trends in low-income households that have either very low or very high spending on telecommunications. Thomas *et al.* (2023) found that low income households are at heightened risk of digital exclusion. This research identifies low income households at potential risk of digital exclusion to help inform policies aimed at improving digital connectivity outcomes for low-income Australians.

Continuing on from earlier BCARR research (2017, 2020) and the work of Breunig and McCarthy (2020), this analysis uses HILDA data to identify two groups of low-income households, each with very different patterns of telecommunications spending that puts them at potential risk of digital exclusion. The first group, 'low-income, low spending' (LILS) households, have:

- household disposable income below half the median level; and
- share of household disposable income spent on telecommunications below half of the median share.

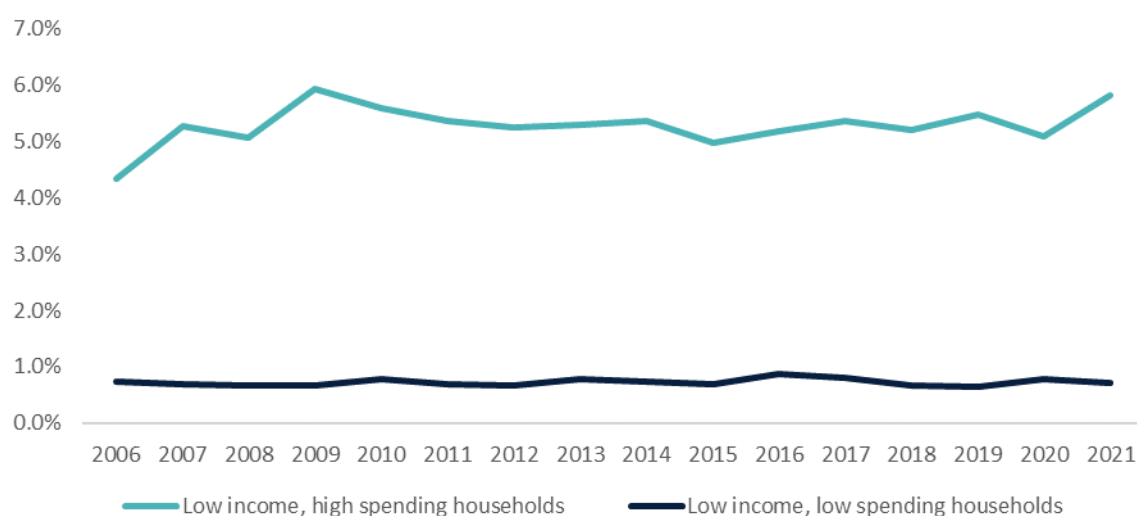
The second group, 'low-income, high spending' (LIHS) households, have:

- household disposable income below half the median level, and
- share of household disposable income spent on telecommunications over three times the median share.

The first group, LILS, is at potential risk of digital exclusion from underspending on telecommunications, and may miss out on benefits of digital connectedness or improvements in digital ability. The second group, LIHS, is at potential risk of digital exclusion from overspending on telecommunications, which may be financially unsustainable or cause periods of disconnection e.g. high cost, low data pre-paid mobile plans.

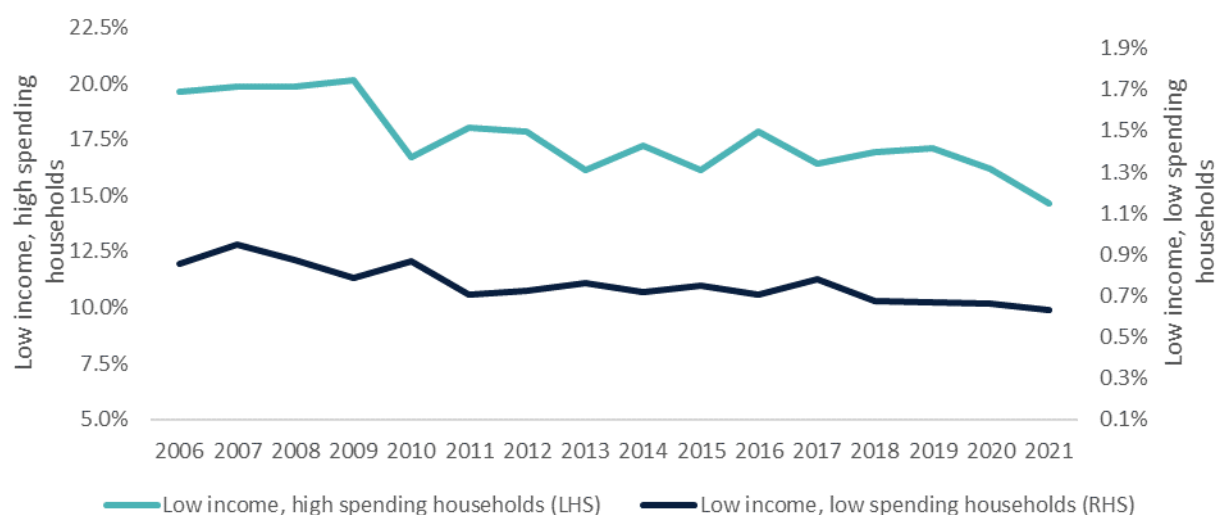
Figure 7 shows that the share of these two groups in the population remained stable over the 2006-2021 period. The proportion of LIHS households remained relatively unchanged, representing approximately 5 to 6 per cent of Australian households. There was a slight increase in the share of LIHS households in 2009, following the Global Financial Crisis, and most recently in 2021, during the COVID-19 pandemic. In contrast, the proportion of LILS households was much lower, representing 0.6 to 0.7 per cent of Australian households over the 15-year period.

¹⁷ For estimates of the interaction effects in the OLS regression see Table 4 in *Attachment C — OLS regression estimates*.

Figure 7: Share of low-income households with relatively high or low telecommunications expenditures.

Source: The HILDA Survey, Release 21; BCARR calculations.

While the sizes of these low-income groups were stable over the 2006-2021 period, the average share of disposable income spent on telecommunications declined for both groups over this timeframe (Figure 8). The overall decline was more pronounced for LIHS households, where the average share of disposable income spent on telecommunications declined from its peak of 20.1 per cent in 2009 to a low of 14.6 per cent in 2021. For the LILS households, the drop in the average share of telecommunications expenditures decreased from its peak of 0.9 per cent in 2007 to a low of 0.6 per cent in 2021.

Figure 8: Average share of income spent on telecommunications by low-income households.

Source: The HILDA Survey, Release 21; BCARR calculations.

8. 'Low-income, low spending' and 'low-income, high spending' groups – logistic regression estimates

In this section, we investigate which characteristics make households more likely to fall into LIHS and LILS groups, and may indicate a heightened risk of potential digital exclusion. Understanding these characteristics is important for designing policies aimed at improving digital inclusion of vulnerable groups of the Australian population.

We determine the change in the probability of a household falling into either the LIHS or LILS group if a specific household characteristic is present. We estimate a series of logistic regressions using HILDA data to predict the change in a household's likelihood of belonging to the LIHS or LILS group when controlling for their income, household size and other socioeconomic characteristics. The response variables used in our regressions are coded 1 if a household belongs either to the LIHS or LILS group, and 0 otherwise. With the exception of household disposable income, household size and SEIFA deciles, all other variables used were binary in nature, i.e. they took the value of 1 if some or all of household members exhibited the specific characteristic and 0 otherwise.

We estimate four separate logistic regression models for each of the LIHS and LILS groups:

- model 1 – where control variables included household size, household income, time dummies, SEIFA indexes, a set of variables capturing socioeconomic characteristics displayed by 'some' members of the households;
- model 2 – as in model 1, but replacing a set of variables capturing socioeconomic characteristics displayed by 'some' members of the households with variables capturing socioeconomic characteristics displayed by 'all' members of the household;
- model 3 – as model 1, but adding statistically significant interactions with household income, household size and a COVID-19 time dummy; and
- model 4 - as model 2, but adding statistically significant interactions with household income, household size and a COVID-19 time dummy.

The detailed model estimates are included in *Attachment F — Logistic regression estimates*. In what follows we use the average marginal effects to discuss the factors impacting the likelihood of a household being in the LIHS or LILS group. Figure 9 and Figure 10 illustrate the percentage change in the probability of belonging either to the LIHS or LILS group if a given variable is changed by one unit of measure, holding all the other variables constant. For example, Figure 9 shows that a household where all members work from home is, on average, 1.9 per cent more likely to belong to the LIHS group compared to a household where nobody works from home. In contrast, a household where some members are not employed is on average 1.5 per cent less likely to be in the LIHS group compared to a household where no members are not employed.

Low-income, high spending households

More generally, as depicted in Figure 9, a household has a higher predicted probability of being in the LIHS group, and by extension a higher probability of being at potential risk of digital exclusion from unsustainable expenditure on telecommunications if:

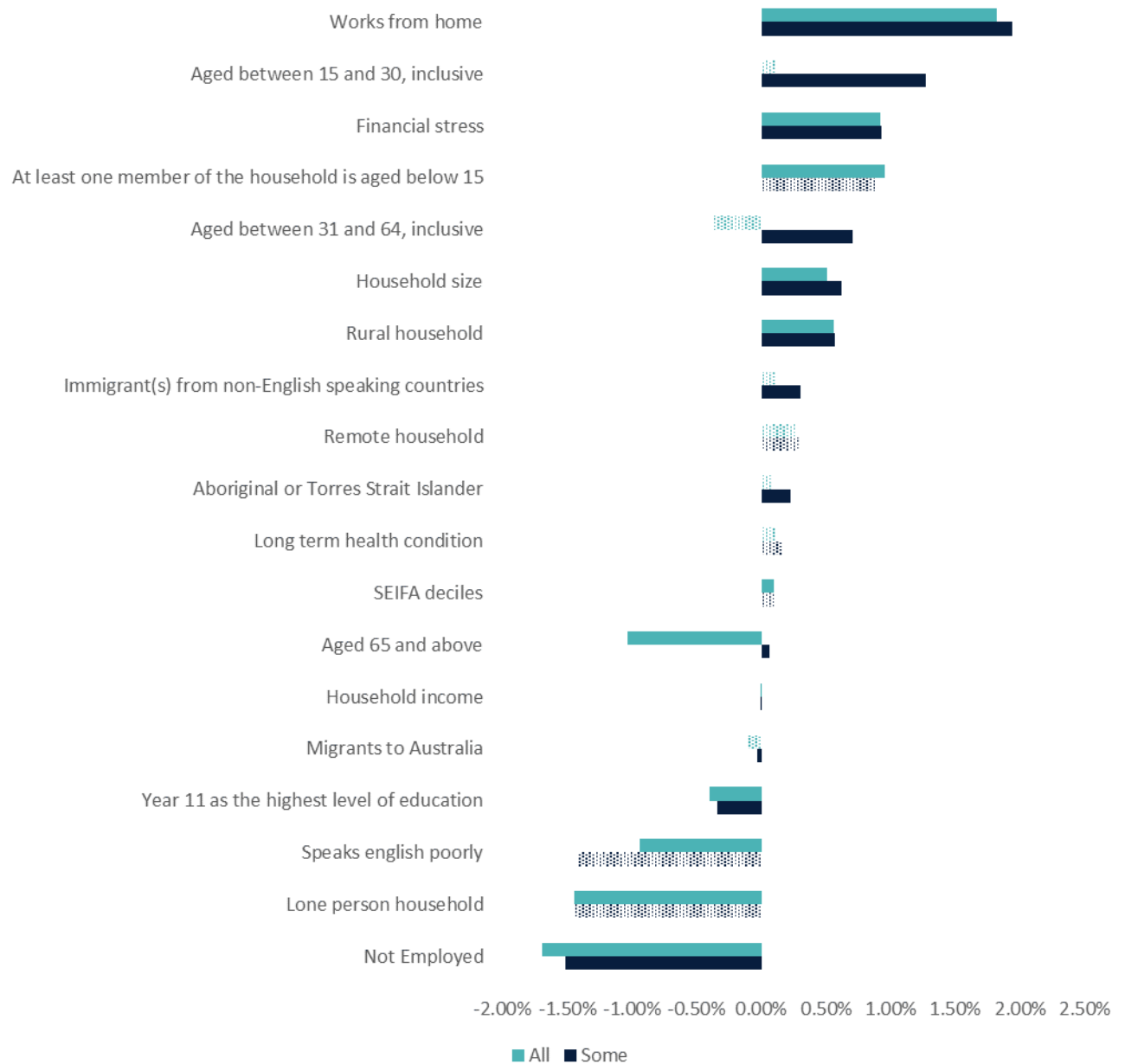
- at least one member works from home;
- some, but not all members are of working age (aged from 15 to 64);
- at least one member is in financial stress;
- there are children present;
- household size is larger;
- the household is in a rural area;
- some, but not all members are immigrants from a non-English speaking background; and
- some, but not all members are Aboriginal or Torres Strait Islander.

Some of these household characteristics suggest there might be high spending on telecommunications, irrespective of income. For instance, households where members work from home, or larger households may require broadband plans with faster speeds or higher data allowances, which are generally more expensive. Alternatively, households in rural locations may face higher costs of telecommunications services, potentially from a less-competitive market (Thomas et al. 2023). At the other end of the spectrum, there are households that are less likely to belong to the LIHS group. These are households where:

- at least one member is not employed;
- the household is a lone person household;
- all members speak English poorly;
- at least one member has year 11 as their highest level of education; and
- all members are aged 65 or older.

These households are less likely to be at risk of potential digital exclusion from unsustainable expenditure on telecommunications services. These results largely consolidate findings from the earlier OLS regression analysis, described in section 6 of this paper. Households where some or all members experienced financial stress, work from home, have children or have younger members spend more on telecommunications on average, and have higher odds of being in the LIHS group compared to households that do not exhibit these characteristics. Similarly, households with members that are not employed, aged 65 or older, or lone person households spent less on average on telecommunications, and they have lower odds of being in the low-income, high spending group than households that do not display these characteristics.

Figure 9: Average marginal effect calculated using logistic regression estimates, 'Low-income, high spending' (LIHS) group, 'Some' and 'All' models.



Notes: 'Some' refers to logistic regression estimates using control variables set to 1 if some members of the household displayed a given characteristic. 'All', refers to the regression estimates using control variables set to 1 if all members of the household displayed the particular characteristic.

Dotted bars indicate marginal effects that were not statistically significant (p-values were above the 10 per cent).

Marginal effect refers to the percentage change in the probability of belonging to the LIHS if a given variable is changed by one unit of measure, holding all the other control variables constant.

Source: The HILDA Survey, Release 21; BCARR calculations.

Low-income, low spending households

Unlike households in the LIHS group which spend heavily on telecommunications, households in the LILS group spend very little on telecommunications, which puts them at potential risk of digital exclusion due to a lack of digital connectivity. The sample of these households is very small, representing just 0.9 per cent of the entire sample. Therefore, many of the variables used in our model were not statistically significant. Figure 10 illustrates that a household was more likely to fall into the LILS group if it displayed the following characteristics:

- at least one member has year 11 as their highest level of education;
- at least one member is not employed;
- all members are aged from 15 to 30; or
- all members are aged 65 or older.

A low level of spending on telecommunications may be the result of lower needs for online connection, or potentially a low level of digital ability. However, it could also reflect households' need to allocate their limited incomes elsewhere, to products that are seen as more essential.

The likelihood of a household falling into the LILS group is statistically lower when a household:

- is larger in size;
- is a lone person household;
- includes some, but not all members aged from 31 to 64;
- includes some, but not all members aged 65 or older;
- includes some, but not all members aged from 15 to 30;
- includes at least one member working from home; or
- includes at least one member who is a migrant to Australia.

The results of the LILS model largely consolidate findings from the earlier OLS regression results, described in section 6. For example, households where at least one member is not employed, or has a lower education and households with older residents spend less on telecommunications, and are more likely to be in the LILS group.

Figure 10: Average marginal effect calculated using logistic regression estimates, 'Low-income, low spending' (LILS) group, 'Some' and 'All' models.



Notes: 'Some' refers to logistic regression estimates using control variables set to 1 if some members of the household displayed a given characteristic. 'All', refers to the regression estimates using control variables set to 1 if all members of the household displayed the particular characteristic.

Dotted bars indicate marginal effects that were not statistically significant (p-values were above 10 per cent).

Marginal effect refers to the percentage change in the probability of belonging to the LILS if a given variable is changed by one unit of measure, holding all the other control variables constant.

Source: The HILDA Survey, Release 21; BCARR calculations.

9. Conclusion and discussion

In this paper we have analysed household spending on telecommunications in Australia using Household Income and Labour Dynamics in Australia (HILDA) data. We demonstrated that telecommunications expenditure behaves like spending on other necessities, such as food. As household incomes grow, the share of disposable income spent on telecommunications tends to decline.

Our analysis revealed that between 2006 and 2021, household spending on telecommunications dropped in both absolute terms and as a share of disposable income, signalling an improvement in telecommunications affordability.

We identified groups that appear to spend less on telecommunications relative to their income: those that speak English poorly, lone person households, those aged 65 years or older, those who are not employed, Aboriginal or Torres Strait Islanders, immigrants, people with lower levels of education and those with long-term health conditions.

We also identified groups that spend more on telecommunications relative to their income: those under financial stress, people working from home, households located in a rural area, households with young and middle-aged members, households with children, and households located in areas of higher relative advantage.

Further, we used HILDA data to monitor trends in low-income households identified by BCARR (2017, 2020) and Breunig and McCarthy (2020) to be at heightened risk of digital exclusion due to their very high or very low spending on telecommunications. We examined the socioeconomic characteristics of households belonging to these groups.

Overall, we found that the proportion of Australian households belonging to low-income, high spending (LIHS) and low-income, low spending (LILS) to be around 5-6 per cent and 0.6-0.7 per cent, respectively. We found that the proportion of LIHS households increased in 2021 while the proportion of LILS households remained relatively steady over the 2006-2021 period. The steady shares of low-income households over time suggests that the observed improvements in telecommunications affordability are not necessarily being realised by low-income groups. However, it is important to note that digital inclusion is a function of more than just affordability, and so these low-income households at potential risk of digital exclusion may be experiencing barriers to inclusion from limited internet ability or internet access.

We discovered that the following groups are more likely to belong to the LIHS group: those working from home, immigrants from non-English speaking backgrounds, those of working age, 3 in financial stress, households with children, larger households or households located in rural Australia.

Households are more likely to belong to the LILS group if their member(s) are not employed, they are older Australians, have lower levels of education, and are younger households. Due to small sample sizes, the statistical significance and magnitude of the above identified socioeconomic factors were very low. Importantly, this is a finding in itself as this group does not represent a large enough proportion of the population to allow for definitive statistical inference.

Our analysis has its limitations. The scope of our research was limited by our key dependent variable – household expenditure on telecommunications services, which did not allow us to distinguish between different types of telecommunication services, or their quantity and quality. More comprehensive data would allow us to calculate measures of affordability of telecommunications with greater scope and precision.

Further, small sample sizes prevented us in some cases from drawing stronger conclusions about the characteristics of households being at potential risk of digital exclusion due to their unsustainable spending patterns on telecommunications.

Lastly, our study fails to control for the structure and regulations in the telecommunications market due to a lack of available data. These factors can potentially impact the cost of accessing and using telecommunications, and hence impact their affordability.

Finally, the limited scope of HILDA prevented us from looking at other aspects of digital exclusion such as digital literacy. We will seek to complement this research with an analysis of individuals that do not have access to the internet, a measure which HILDA data does capture. The intent of proposed further research is to better understand the level of digital exclusion that exists in Australia, and the socioeconomic characteristics, if any, of individuals and households that are digitally excluded.

10. Acknowledgements

We would like to acknowledge and thank Professor Robert Breunig for his STATA code, and the advice and expertise he shared with us in the peer review of our analysis.

Attachment A — Trends in the key variables

Table 1: Household disposable income and telecommunications expenditure, by wave.

Wave	Year	Household	Telecommu ications Expenditure Share Mean	Telecommu ications Expenditure Share Standard Deviation	Telecommu ications Expenditure Share Median	Disposable Income Mean	Disposable Income Standard Deviation	Disposable Income Median
6	2006	6,850	0.039	0.062	0.024	62,684	45,975	54,081
7	2007	6,768	0.040	0.066	0.024	67,887	54,077	58,374
8	2008	6,770	0.041	0.068	0.024	71,574	53,788	61,677
9	2009	6,948	0.041	0.075	0.022	76,600	55,998	67,538
10	2010	7,024	0.038	0.060	0.023	77,518	57,588	66,621
11	2011	9,214	0.037	0.063	0.022	80,443	60,214	67,640
12	2012	9,229	0.035	0.061	0.021	83,767	61,784	72,365
13	2013	9,214	0.033	0.057	0.020	85,744	65,594	73,307
14	2014	9,263	0.037	0.071	0.020	87,239	68,344	74,035
15	2015	9,279	0.034	0.061	0.020	88,480	66,213	75,418
16	2016	9,383	0.035	0.064	0.020	90,075	67,920	77,549
17	2017	9,389	0.034	0.058	0.020	92,586	74,539	79,071
18	2018	9,288	0.033	0.061	0.019	94,546	71,007	80,884
19	2019	9,327	0.033	0.061	0.019	98,873	77,340	83,287
20	2020	9,165	0.032	0.059	0.018	101,801	74,815	87,499
21	2021	8,982	0.030	0.056	0.018	106,357	80,473	90,620

Attachment B — Significance of differences in means for demographic and socioeconomic groups

Table 2: Difference in mean telecommunications expenditure share by characteristic, all waves.

Category	True Mean	True Std. error	False Mean	False Std. error	Difference Mean	Difference Std. error
At least one member is Aboriginal or Torres Strait Islander	0.037	0.0012	0.0354	0.0006	0.002	0.0015
At least one member is a migrant to Australia	0.034	0.0010	0.0363	0.0005	-0.002**	0.0009
At least one member is unemployed or not in the labour force	0.040	0.0006	0.0305	0.0007	0.010***	0.0006
At least one member had difficulty paying electricity, gas or telephone bills on time	0.043	0.0008	0.0343	0.0006	0.009***	0.0008
At least one member speaks English 'not well' or 'not at all'	0.043	0.0054	0.0352	0.0004	0.008	0.0053
At least one member has a long-term health condition	0.038	0.0005	0.0335	0.0008	0.005***	0.0009
At least one member is an immigrant from a non-English speaking country	0.035	0.0014	0.0356	0.0004	0.000	0.0013
At least one member has year 11 as their highest level of education	0.038	0.0006	0.0336	0.0007	0.005***	0.0007
At least one member is a full-time student	0.032	0.0007	0.0362	0.0006	-0.004***	0.0009
At least one member works from home	0.032	0.0010	0.0358	0.0005	-0.012***	0.0007
Household has at least one member under the age of 15	0.030	0.0009	0.0376	0.0005	-0.008***	0.0008
Household has at least one individual aged from 15 to 30	0.034	0.0007	0.0364	0.0006	-0.002***	0.0008
Household has at least one individual aged from 31 to 64	0.031	0.0006	0.0462	0.0008	-0.016***	0.0007
Household has at least one individual aged 65 or older	0.041	0.0007	0.0336	0.0006	0.008***	0.0008
All members of the household are Aboriginal or Torres Strait Islander	0.049	0.0021	0.0353	0.0005	0.014***	0.0022
All members are migrants to Australia	0.039	0.0014	0.0346	0.0004	0.004***	0.0013
All members are unemployed or not in the labour force	0.050	0.0009	0.0300	0.0006	0.020***	0.0008
All members had difficulty paying electricity, gas or telephone bills on time	0.055	0.0014	0.0343	0.0005	0.021***	0.0014

Category	True Mean	True Std. error	False Mean	False Std. error	Difference Mean	Difference Std. error
All members speak English 'not well' or 'not at all'	0.060	0.0127	0.0353	0.0005	0.024*	0.0125
All members have long-term health condition	0.048	0.0009	0.0324	0.0007	0.016***	0.0012
All members are immigrants from non-English speaking countries	0.040	0.0022	0.0348	0.0004	0.005**	0.0020
All members have year 11 as their highest level of education	0.050	0.0012	0.0326	0.0006	0.018***	0.0012
All members are full-time students	0.066	0.0039	0.0349	0.0005	0.031***	0.0040
All members work from home	0.0319	0.0010	0.0358	0.0005	-0.004***	0.0008
All members are aged from 15 to 30	0.047	0.0012	0.0338	0.0005	0.013***	0.0012
All members are aged from 31 to 64	0.033	0.0008	0.0373	0.0005	-0.004***	0.0007
All household members are aged 65 or older	0.047	0.0010	0.0331	0.0006	0.013***	0.0009
Lone person household	0.054	0.0012	0.0297	0.0005	0.024***	0.0011
Public housing household	0.101	0.0573	0.0355	0.0005	0.066	0.0572
Rural household	0.037	0.0013	0.0353	0.0006	0.002	0.0013
Remote household	0.034	0.0039	0.0355	0.0005	-0.002	0.0037

Notes: *Difference significant at 10 per cent level; **Difference significant at 5 per cent level; ***Difference significant at 1 per cent level.

The 'true' column is the expenditure share if the household is described by the sentence in column one. The 'false' column is the expenditure share for households for which the statement in column one does not describe them. Standard errors have been adjusted for HILDA survey design using the jack knife method.

Source: The HILDA Survey, Release 21; BCARR calculations.

Attachment C — OLS regression estimates

Table 3 presents OLS regression results for the models without interaction terms. The dependent variable is the telecommunications expenditure share. Explanatory variables include household income, household size, time (wave) dummies, and a range of household characteristics.¹⁸

Table 3: OLS regression results, ‘Some’ and ‘All’ models without interaction terms.

Dependent variable: ln (telecommunications expenditure share)	Model 1 Coef.	Model 1 Std. error	Model 2: some Coef.	Model 2: some Std. error	Model 3: all Coef.	Model 3: all Std. error
ln(household income)	-0.776***	0.0033	-0.840***	0.0040	-0.856***	0.0040
ln(household size)	0.207***	0.0045	0.094***	0.0111	0.126***	0.0106
wave 6 is omitted category						
wave 7	0.061***	0.0129	0.066***	0.0125	0.067***	0.0125
wave 8	0.107***	0.0129	0.116***	0.0126	0.117***	0.0126
wave 9	0.087***	0.0128	0.099***	0.0125	0.100***	0.0125
wave 10	0.114***	0.0128	0.143***	0.0125	0.137***	0.0125
wave 11	0.091***	0.0121	0.108***	0.0117	0.109***	0.0117
wave 12	0.087***	0.0121	0.109***	0.0117	0.110***	0.0117
wave 13	0.043***	0.0121	0.066***	0.0119	0.067***	0.0119
wave 14	0.078***	0.0121	0.102***	0.0120	0.104***	0.0120
wave 15	0.067***	0.0121	0.092***	0.0120	0.093***	0.0120
wave 16	0.106***	0.0120	0.132***	0.0120	0.133***	0.0120
wave 17	0.113***	0.0120	0.140***	0.0120	0.142***	0.0120
wave 18	0.089***	0.0121	0.119***	0.0119	0.121***	0.0119
wave 19	0.088***	0.0121	0.122***	0.0121	0.122***	0.0121
wave 20	0.12***	0.0121	0.153***	0.0119	0.158***	0.0119
wave 21	0.091***	0.0122	0.124***	0.0119	0.129***	0.0119
1 if expenditure is imputed	0.011*	0.0058	0.005	0.0063	0.008	0.0063
1 if income is imputed	0	0.0050	0.009*	0.0050	0.016***	0.0050
Aboriginal or Torres Strait Islander			-0.106***	0.0123	-0.111***	0.0196
Immigrant			0.004	0.0060	-0.006	0.0084

¹⁸ Log-transformations were used for the variables representing telecommunication expenditure share, disposable household income and household size. These (natural) log transformations changed the probability distributions of these variables so that they were closer to a normal distribution (see: *Attachment E — Log transformations*)

Dependent variable: In (telecommunications expenditure share)	Model 1 Coef.	Model 1 Std. error	Model 2: some Coef.	Model 2: some Std. error	Model 3: all Coef.	Model 3: all Std. error
Speaks English poorly			-0.086***	0.0164	-0.022	0.0312
Long-term health condition			0.000	0.0045	-0.011*	0.0060
Immigrant from a non-English speaking country			-0.029***	0.0077	-0.045***	0.0111
Not employed			-0.070***	0.0053	-0.150***	0.0070
Financial stress			0.100***	0.0065	0.125***	0.0094
Highest level of education is year 11			-0.021***	0.0047	-0.076***	0.0064
Full time study			0.044***	0.0073	-0.037**	0.0158
Works from home			0.071***	0.0049	0.069***	0.0075
At least one member aged under 15			0.044***	0.0077	0.041***	0.0076
Individuals aged from 15 to 30			0.094***	0.0066	0.029***	0.0079
Individuals aged from 31 to 64			0.018***	0.0068	-0.052***	0.0059
Individuals aged 65 or older			-0.078***	0.0080	-0.092***	0.0084
Rural household			0.046***	0.0061	0.043***	0.0061
Remote household			0.025	0.0160	0.021	0.0160
Public housing			-0.486*	0.2939	-0.446	0.2897
Lone person household			-0.101***	0.0095	-0.061***	0.0093
SEIFA decile			0.010***	0.0008	0.009***	0.0008
Constant	4.529***	0.0350	5.245***	0.0427	5.487***	0.0443
Sample size	136,093	136,093	136,063	136,063	136,063	136,063
R-squared	34.2%	34.2%	35.6%	35.6%	35.6%	35.6%

Notes: *Significant at 10 per cent level; **Significant at 5 per cent level; ***Significant at 1 per cent level.

Source: The HILDA Survey, Release 21; BCARR calculations.

The first column in Table 3 includes the names of the independent variables used. The table includes:

- model 1 (columns 2 and 3) which includes independent variables on income, household size, a range of time dummies for particular HILDA waves and a dummy variable set to 1 if telecommunication expenditure or income data has been imputed;
- model 2 (columns 4 and 5) model 1 variables plus a vector of controls that apply if 'some' members of the household displayed a given characteristic; and
- model 3 (columns 6 and 7) model 1 variables *plus* a vector of controls that apply if 'all' members of the household displayed a given characteristic.

Models 2 and 3 in addition to the explanatory variables used in Breunig and McCarthy (2020) include variables capturing:

- the socioeconomic disadvantage of households;
- residing alone status;
- educational attainment;
- working from home status; and
- the presence of full-time students.

As shown in Table 3, these added variables were statistically significant.¹⁹ The final models presented in Table 4 add interaction terms for household income, size and a dummy variable capturing the COVID-19 period (2020 and 2021). Inclusion of these interaction terms improves model fit and allows for a more nuanced interpretation of the effects of the main explanatory variables since they are allowed to vary by other characteristics.

In all models summarised in Table 3, household income is strongly negatively correlated with the telecommunications expenditure share, confirming the inverse relationship we observed when looking at the distribution of telecommunications expenditure across different income deciles. On average, a one per cent increase in household income decreases the share of telecommunications expenditure in disposable income by between 0.8 and 0.9 per cent. Household size has a strong positive impact on telecommunication expenditures. A one per cent increase in household size increases telecommunications spending by between 0.1 and 0.2 per cent.

The time dummies for each HILDA wave control for changes in the overall price level (inflation) in telecommunications expenditure and household disposable income. They are significant for each year. The imputation dummies for expenditure shares and household income control for any systematic information or error that could be contained in the imputed values.

Many of the characteristics associated with different telecommunications expenditure shares highlighted in section 5 of this report remain statistically significant when controlling for household income and size, however, many of them change sign. As shown in Table 3, with the exception of the variable indicating the presence of full-time students, the estimated coefficients and their statistical significance do not differ much across the 'all' and 'some' models.²⁰

Table 4 shows that many of the control variables have statistically significant interactions with household income, size and the COVID-19 (2020, 2021) dummy.

Table 4: OLS regression results, 'Some' and 'All' models with interaction terms.

Dependent variable: ln (telecommunications expenditure share)	Model 2: some Coef.	Model 2: some Std. error	Model 3: all Coef.	Model 3: all Std. error
ln(household income)	-0.871***	0.0086	-0.846***	0.0072
ln(household size)	0.123***	0.0125	0.205***	0.0140
wave 6 is omitted category				

¹⁹ To control for unobserved heterogeneity of households we also estimated the model summarized in the 1st column of Table 3 in first differences. We obtain similar coefficients on income (- 0.91) and household size (0.33). We decided to use the model in levels as it allowed us to include the socio-economic characteristics of households, many of which are time-constant.

²⁰ The variable capturing the presence of full-time students in the household indicates very different households. The households with a least one student are mostly families with children (55.6 per cent of the sample) and the households with 'all' full-time students refer mostly to lone person households (64.8 per cent of the sample). These households have very different telecommunications expenditure patterns as captured by our models.

Dependent variable: ln (telecommunications expenditure share)	Model 2: some Coef.	Model 2: some Std. error	Model 3: all Coef.	Model 3: all Std. error
wave 7	0.067***	0.0124	0.068***	0.0125
wave 8	0.117***	0.0125	0.118***	0.0126
wave 9	0.102***	0.0125	0.103***	0.0125
wave 10	0.144***	0.0125	0.138***	0.0125
wave 11	0.111***	0.0117	0.111***	0.0117
wave 12	0.112***	0.0117	0.112***	0.0117
wave 13	0.070***	0.0119	0.070***	0.0119
wave 14	0.107***	0.0120	0.107***	0.0120
wave 15	0.096***	0.0120	0.097***	0.0120
wave 16	0.137***	0.0120	0.136***	0.0120
wave 17	0.147***	0.0120	0.146***	0.0120
wave 18	0.126***	0.0119	0.124***	0.0119
wave 19	0.129***	0.0121	0.127***	0.0121
wave 20	0.202***	0.0177	0.139***	0.0166
wave 21	0.176***	0.0178	0.112***	0.0167
1 if expenditure is imputed	0.008	0.0063	0.010	0.0064
1 if income is imputed	0.008	0.0050	0.015***	0.0051
Aboriginal or Torres Strait Islander	-1.047***	0.1993	0.047*	0.0270
Immigrant	0.044***	0.0096	0.035***	0.0107
Speaks English poorly	0.629**	0.2675	-0.061**	0.0316
Long-term health condition	-0.173***	0.0633	-0.020***	0.0065
Immigrant from a non-English speaking country	0.412***	0.0899	0.499***	0.1084
Not employed	-0.739***	0.0774	-0.610***	0.0852
Financial stress	-0.487***	0.1177	-0.468***	0.1351
Highest level of education is year 11	-0.388***	0.0639	-0.052***	0.0088
Full time study	-0.278**	0.1205	-0.688***	0.2391
Works from home	1.051***	0.0930	0.916***	0.1229
At least one member aged under 15	0.478***	0.1038	0.557***	0.1032
Individuals aged from 15 to 30	0.252***	0.0917	0.414***	0.1248
Individuals aged from 31 to 64	0.020***	0.0075	0.357***	0.0879
Individual aged 65 or older	-0.038***	0.0119	0.169**	0.0682
Rural household	0.019	0.0115	0.018	0.0115

Dependent variable: ln (telecommunications expenditure share)	Model 2: some Coef.	Model 2: some Std. error	Model 3: all Coef.	Model 3: all Std. error
Remote household	0.024	0.0160	-0.513**	0.2348
Public housing	-0.421	0.2958	-0.396	0.2976
Lone person household	0.028	0.0836	-0.125***	0.0132
SEIFA decile	0.010***	0.0008	0.010***	0.0008
Aboriginal or Torres Strait Islander * household income	0.108***	0.0190	n.a.	n.a.
Long-term health condition*household income	0.015***	0.0057	n.a.	n.a.
Speaks English poorly*household income	-0.076***	0.0268	n.a.	n.a.
Immigrant from a non-English speaking country * household income	-0.039***	0.0080	-0.049***	0.0100
Not employed*household income	0.061***	0.0069	0.047***	0.0081
Financial stress*household income	0.059***	0.0114	0.056***	0.0126
Highest level of education is year 11*household income	0.033***	0.0058	n.a.	n.a.
Full time study*household income	0.023**	0.0114	0.045**	0.0217
Works from home*household income	-0.085***	0.0081	-0.072***	0.0109
At least one member aged under 15 *household income	-0.039***	0.0091	-0.046***	0.0091
Individuals aged from 15 to 30*household income	-0.015*	0.0081	-0.024**	0.0115
Individuals aged from 31 to 64*household income	n.a.	n.a.	-0.026***	0.0079
Remote household*household income	n.a.	n.a.	0.048**	0.0210
Lone person household*household income	-0.012	0.0078	n.a.	n.a.
Long-term health condition*2020/21	0.036***	0.0122	0.043***	0.0161
Immigrant from a non-English speaking country * 2020/21	n.a.	n.a.	-0.037	0.0227
Financial stress*2020/21	-0.056***	0.0186	-0.067**	0.0282
Works from home*2020/21	-0.050***	0.0129	-0.078***	0.0179
At least one member aged under 15*2020/21	0.029*	0.0148	0.024	0.0150
Individuals aged from 15 to 30*2020/21	n.a.	n.a.	0.055***	0.0213
Individuals aged from 31 to 64*2020/21	-0.035***	0.0136	0.046***	0.0159
Individual aged 65 or older*2020/21	n.a.	n.a.	0.092***	0.0192
Rural household*2020/21	-0.055***	0.0165	-0.058***	0.0166
Lone person household*2020/21	-0.042***	0.0149	-0.051***	0.0157
Aboriginal or Torres Strait Islander*household size	-0.258***	0.0260	-0.218***	0.0320

Dependent variable: ln (telecommunications expenditure share)	Model 2: some Coef.	Model 2: some Std. error	Model 3: all Coef.	Model 3: all Std. error
Immigrant*household size	-0.053***	0.0094	-0.070***	0.0125
Speaks English poorly*household size	0.110***	0.0407	n.a.	n.a.
Not employed*household size	n.a.	n.a.	-0.054***	0.0145
Financial stress*household size	-0.064***	0.0145	n.a.	n.a.
Highest level of education is year 11*household size	n.a.	n.a.	-0.036***	0.0122
Full time study*household size	0.057***	0.0161	0.170**	0.0725
Individuals aged from 15 to 30*household size	n.a.	n.a.	-0.103***	0.0199
Individuals aged from 31 to 64*household size	n.a.	n.a.	-0.097***	0.0153
Individual aged 65 or older*household size	-0.064***	0.0130	-0.367***	0.0949
Rural household*household size	0.035***	0.0113	0.036***	0.0113
Full time study* Lone person household	n.a.	n.a.	0.205***	0.0761
Individual aged 65 or older* Lone person household	n.a.	n.a.	-0.132**	0.0679
Constant	5.571***	0.0958	5.283***	0.0787
Sample size	136,063	136,063	136,063	136,063
R-squared	36 %	36 %	35.9%	35.9%

Notes: *Significant at 10 per cent level; **Significant at 5 per cent level; ***Significant at 1 per cent level. n.a. – not included in the model as not statistically significant.

Source: The HILDA Survey, Release 21; BCARR calculations.

To interpret the coefficients in Table 4, refer to estimates of marginal effects shown in section 6. Due to the presence of interaction terms in Table 4, only the direction the interaction terms can be described here.

Table 4 shows that household income continues to be negatively correlated with household expenditures on telecommunications, even after interacting household income with other characteristics. This relationship is weaker for households where some members are Aboriginal or Torres Strait Islander, have long-term health condition, are not employed, are experiencing financial stress, have low levels of education and are full-time students. This is also true for households where all members are not employed, are under financial stress, are full-time students, and households in remote areas. On the other hand, negative correlation of household income with household spending on telecommunications is further enhanced by some or all members of the households: working from home, being immigrants from a non-English speaking country, being aged from 15 to 30 and there being a child in a household.

Household size continues to be positively correlated with the share of telecommunications spending in disposable income, but this correlation is weaker in households where some or all member are: Aboriginal or Torres Strait Islander, immigrants, or aged 65 or older. In contrast, households in rural areas or where some or all members are full-time students tend to spend even more on telecommunications as they increase in their size.

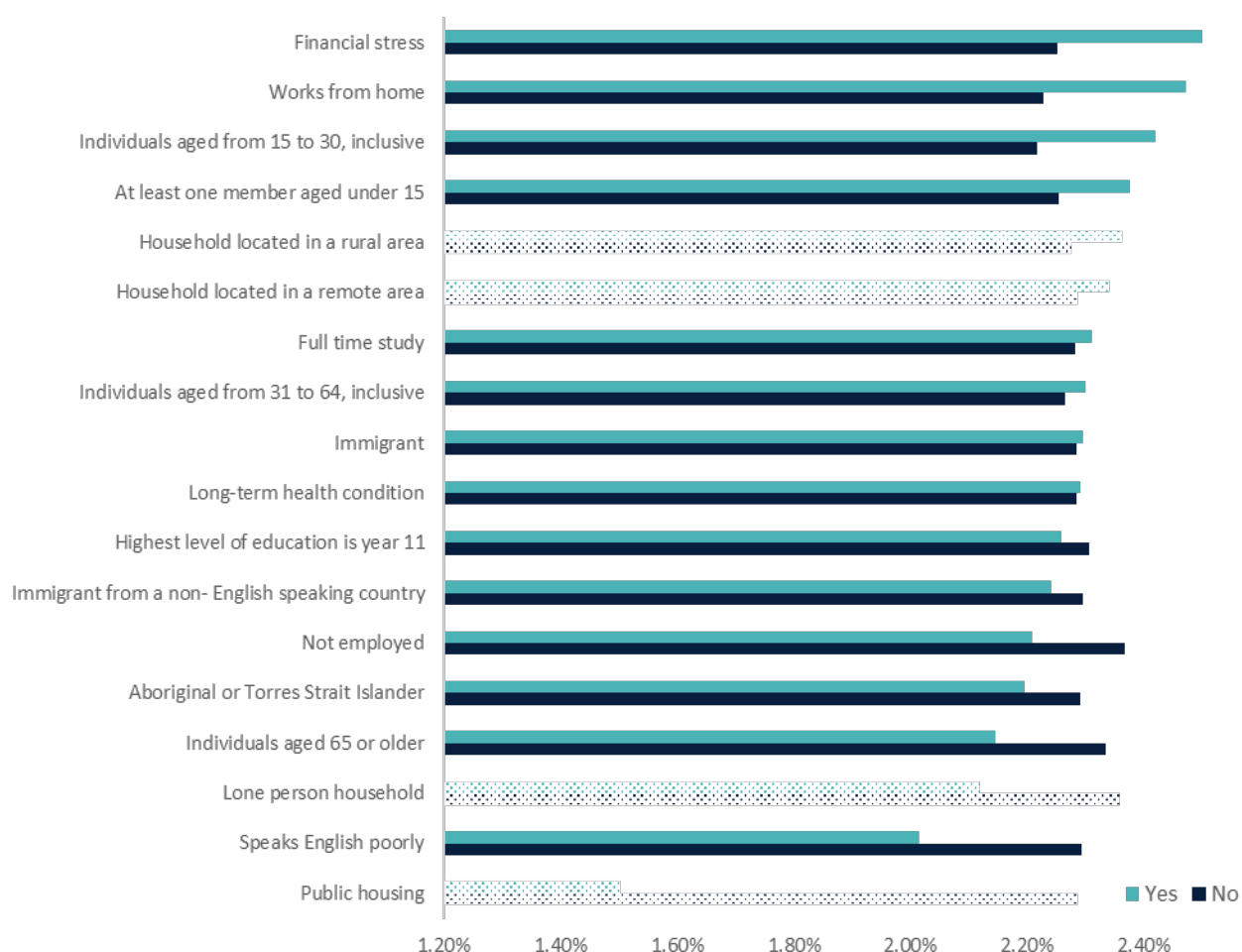
COVID-19 appeared to have a strong significant effect on telecommunications spending of a number of households. Households in rural areas, lone person households and households where all or some members are in financial stress or working from home tended to spend less on telecommunications during that time. Conversely, households with children or having all or some of their members suffering longer-term health conditions tended to spend more.

Testing of the interaction terms revealed that lone person households who studied full time tended to spend more on telecommunications. Lone person households spent less on telecommunications if they were comprised of persons aged 65 or older.

Attachment D — Predicted average communication expenditures using OLS estimates

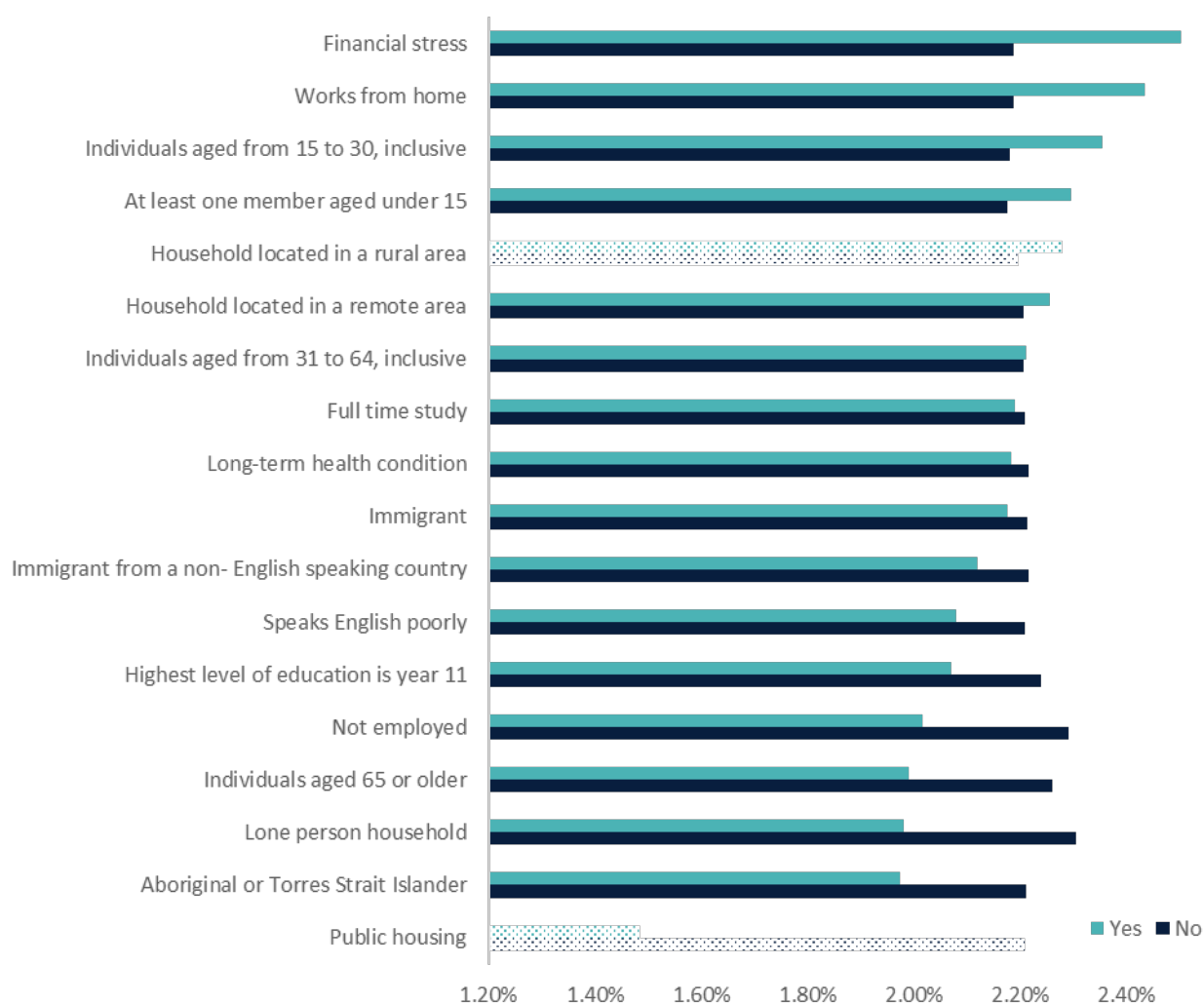
In addition to quantifying the relationship between household expenditure shares on telecommunications and the characteristics of households, regression estimates allow us to predict the share of income spend on telecommunications for a given household while holding the values of all the other control variables constant. Figure 11 and Figure 12 illustrate the predicted values of the expenditure share of telecommunications for particular socioeconomic groups of interest, with other variables set to their average in the data.

Figure 11: Predicted share of telecommunications spending in household disposable income, ‘Some’ model.



Note: ‘Yes’ and ‘No’ indicate whether a household belongs to a particular group of interest. Dotted bars indicate predictions that were not statistically significant, (i.e., p-values were above 10 per cent).

Source: The HILDA Survey, Release 21; BCARR calculations.

Figure 12: Predicted share of telecommunications spending in household disposable income, 'All' model.

Note: 'Yes' and 'No' indicate whether a household belongs to a particular group of interest. Dotted bars indicate predictions that were not statistically significant (i.e., p-values were above 10 per cent).

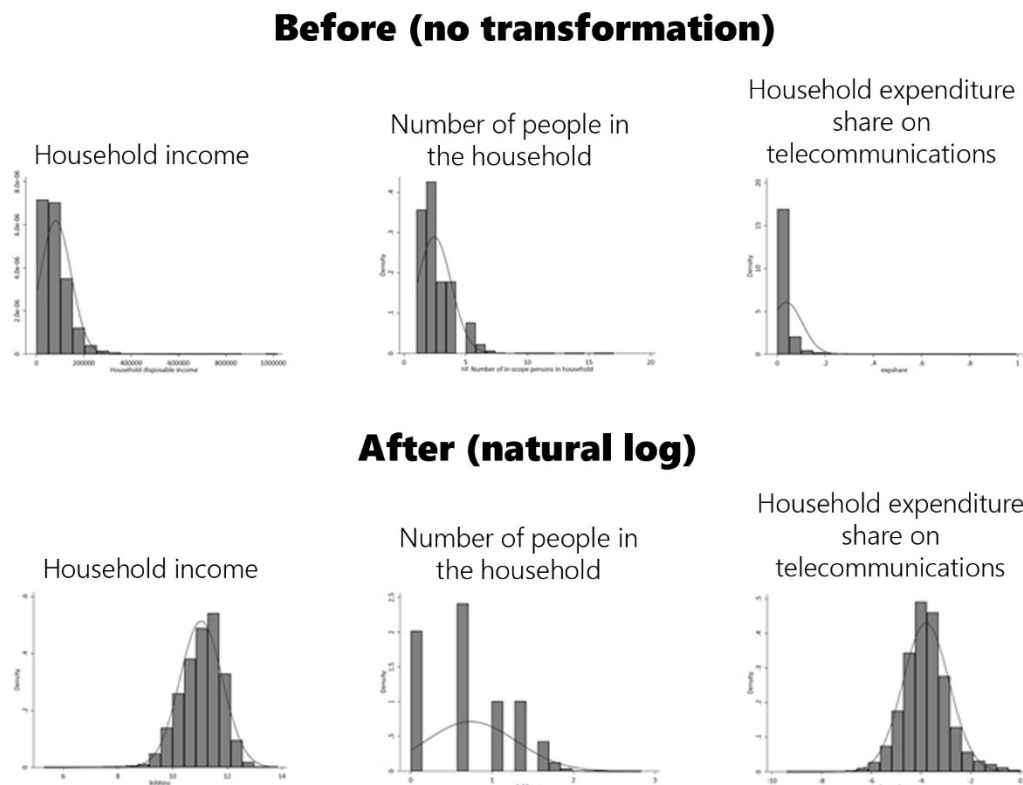
Source: The HILDA Survey, Release 21; BCARR calculations.

Attachment E — Log transformations

Log transformations were applied to some of the variables used in the regression modelling to improve their distributions for the analysis. As shown in the histograms below, prior to the log transformations, 'the share of household income spent on telecommunications', 'household income', and 'the number of people in the household' variables were positively skewed. These variables were much more likely to take on a lower value than a higher value. Log transformations of these variables resulted in a more normally distributed form.

Log transformations changed the interpretation of marginal effects calculated from the OLS regression coefficients. The marginal effects of log of household size and log of household income referred to the percentage change in the response variable from one per cent change in their levels while holding all the other control variables constant. For other variables, binary in nature, marginal effects referred to the percentage change in the response variable if a given variable changes from 0 to 1 while holding all the other variables constant.

Figure 13: Histogram visuals of natural log transformations.



Source: The HILDA Survey, Release 21; BCARR calculations.

Attachment F — Logistic regression estimates

To analyse the likelihood of a household being a part of the low-income low spending (LILS), or low-income high spending (LIHS) groups we apply a binomial logistic regression to HILDA data. The main difference between this regression and the OLS regression described in *Attachment C — OLS regression estimates* is that its response variable is not continuous but binary, coded as 0 and 1. Because the response variable is binary, the logistic regression makes a different set of assumptions about the underlying data.²¹

The response variables used in our regressions are coded 1 if a household belongs either to the LIHS or the LILS group and 0 otherwise. With the exception of household disposable income, household size and SEIFA deciles, all of the control variables in our models are binary in nature, coded 1 if some or all household members exhibited a specific socioeconomic characteristic and 0 otherwise.

The sample sizes used in this analysis are much smaller compared to those used in the OLS regression. The logistic regression on the LIHS group uses a sample of 8,581 households which is 6.3 per cent of the total sample used in the OLS analysis. The logistic regression using LILS group as a response variable is even smaller as it uses 1,169 households which is only 0.9 per cent of the original sample used in the OLS regression. In doing so, both logistic models are predicting outcomes which are rare in the population and this affects the statistical significance and accuracy of predictions. As the model estimates show, this is particularly true for the logistic model estimating the probability of belonging to the LILS group where many control variables turn out to be statistically insignificant, or have a very small magnitude (Table 7 and Table 8).

Table 5 and Table 7 present the estimates of logistic regressions for LIHS and LILS respectively not including any interactions terms. Table 6 and Table 8 present estimates of the same models including statistically significant interactions of with income, household size and a COVID-19 dummy. We interpret the coefficients in Table 6 and Table 8 when discussing marginal effects in section 8 of the paper. Due to the presence of interactions, this attachment discusses only the direction of the effect of the interaction terms.

In all four tables, in the column labelled ‘Some’, the odds ratios and robust standard errors refer to variables set to 1 if some members of the household displayed a given characteristic. In the column labelled ‘All’, the odds ratios and robust standard errors refer to variables set to 1 if all members of the household displayed this particular characteristic. As model estimates show, this distinction between ‘Some’ and ‘All’ does not matter in practice as in most cases the differences in the odd ratios and their statistical significance does not differ much between the two.

We report the estimates of the logistic regression in the form of odds ratios, which is the ratio of two probabilities. For example, for the LIHS group, the odds ratio refers to the probability of being in the LIHS group when exhibiting a characteristic, over the probability of being LIHS group when not exhibiting the same characteristic. An odd ratio greater than one means the household exhibiting the given characteristic is more likely to be in the LIHS group. An odd ratio less than one means the household exhibiting the given characteristic is less likely to be in the LIHS group. For example, a household where some members work from home has 52.8 per cent higher odds of being in the LIHS group than household where no members work from home. Contrastingly, a lone person household in the same sample has 40 per cent lower odds of being in the LIHS group than a non-lone person household.

Notably the coefficients for the wave dummies are quite large, especially in the later years of data. When we controlled for income across all waves, i.e. interacted household income with each wave dummy

²¹ For a detailed analysis of the differences between the OLS and logistic regressions and assumptions them underlying see Hosmer and Lemeshow (2000).

separately, the dummies were no longer as large. The income interactions with time were not kept in the final model. Their exclusion does not affect any of the other reported results.

Table 5: Logistic regression estimates, LIHS group, 'All' and 'Some' models without interaction terms.

LIHS	Some Odds ratio ²²	Some Std error	All Odds ratio	All Std error
wave 7	1.505***	0.1248	1.506***	0.1248
wave 8	1.675***	0.1409	1.678***	0.1411
wave 9	2.332***	0.1958	2.33***	0.1956
wave 10	2.685***	0.2256	2.678***	0.2249
wave 11	2.633***	0.2141	2.639***	0.2149
wave 12	3.257***	0.2676	3.263***	0.2687
wave 13	3.718***	0.3068	3.737***	0.3083
wave 14	4.011***	0.333	4.036***	0.3355
wave 15	4.384***	0.366	4.43***	0.3701
wave 16	4.89***	0.4136	4.914***	0.416
wave 17	5.439***	0.4511	5.454***	0.4527
wave 18	5.414***	0.4699	5.429***	0.472
wave 19	6.891***	0.5964	6.917***	0.5986
wave 20	8.163***	0.7278	8.198***	0.7308
wave 21	10.81***	0.9636	10.861***	0.9681
Household income	1***	0	1***	0
Household size	1.112**	0.0511	1.119**	0.0503
Aboriginal or Torres Strait Islander	0.973	0.0765	1.031	0.0898
Migrants to Australia	0.988	0.0483	0.97	0.052
Not employed	0.622***	0.0288	0.616***	0.0268
Financial stress	1.272***	0.0544	1.301***	0.0606
Speaks English poorly	0.671***	0.0712	0.75**	0.1043
Long term health condition	1.044	0.0346	1.04	0.0348
Immigrant(s) from non-English speaking countries	1.102	0.0677	1.04	0.0701
Year 11 as the highest level of education	0.905***	0.0294	0.891***	0.0294
Aged 65 or older	0.916	0.0667	0.811***	0.0549
Aged from 31 to 64	1.095	0.0718	0.948	0.0635

²² Odds ratio calculates the relationship between a variable (the household characteristic) and the likelihood of an event occurring (being in the specific low-income group).

LIHS	Some Odds ratio ²²	Some Std error	All Odds ratio	All Std error
Aged from 15 to 30	1.26***	0.0862	1.078	0.0768
At least one member is aged below 15	1.112	0.088	1.147*	0.0914
Rural household	1.178***	0.0527	1.173***	0.0524
Remote household	1.068	0.1205	1.055	0.1194
Lone person household	0.594***	0.0357	0.63***	0.0379
Works from home	1.528***	0.0945	1.638***	0.1194
SEIFA deciles	1.028***	0.0057	1.027***	0.0057
_cons	6.363***	0.8584	6.998***	0.9979
N Obs	136,063	136,063	136,063	136,063
Log Pseudolikelihood²³	-15499.2	-15499.2	-15499.8	-15499.8
Pseudo R2²⁴	51.6%	51.6%	51.6%	51.6%

Notes: *Significant at 10 per cent level; **Significant at 5 per cent level; ***Significant at 1 per cent level. n.a. – not included in the model as not statistically significant.

Source: The HILDA Survey, Release 21; BCARR calculations.

Table 7 presents the LIHS regression output after controlling for interaction terms. After adding the interaction terms the pseudo r-squared and log pseudolikelihood values improve, implying the model has greater explanatory power and fits the data better²⁵.

In interpreting the model with interactions, a larger household size is associated with higher odds of being in the LIHS group. However, these odds are lower where a household has some Aboriginal or Torres Strait Islander members, or some or all members in financial stress. This supports the findings from the OLS regression, that larger Aboriginal and Torres Strait Islander households spend less on telecommunications than smaller ones. Households where all members have a relatively low level of education are less likely to be in the LIHS group, and this effect was even larger during COVID-19. Interestingly, households with low education are also less likely to be in the LILS group, and therefore are at lower risk of digital exclusion from unsustainable or insufficient expenditure patterns. Rural households are more likely to be in the LIHS group, but during the pandemic years their likelihood of being in this group dropped, as confirmed by the OLS regression showing that their expenditure on telecommunications services decreased within this period. We also tested for interactions with household disposable income, and while a number were significant, the magnitude of the coefficients indicated that there is no association between the response variable and the interaction term.

²³ The log pseudolikelihood measures the goodness of fit for a model. A value closer to 0 indicates a better model fit.

²⁴ The pseudo r-squared measures goodness-of fit and indicates the share of the variation in the dependent variable that can be explained by the independent variables. A value closer to 1 (or 100 per cent) indicates an improved model-fit.

²⁵ More variation in the dependent variable can be explained by the control variables.

Table 6: Logistic regression estimates, LIHS group, 'All' and 'Some' models with interaction terms.

LIHS	Some Odds ratio	Some Std error	All Odds ratio	All Std error
wave 7	1.5091***	0.1263	1.518***	0.1271
wave 8	1.683***	0.1432	1.695***	0.1439
wave 9	2.341***	0.1978	2.35***	0.1987
wave 10	2.659***	0.2258	2.675***	0.2268
wave 11	2.628***	0.2156	2.657***	0.2182
wave 12	3.234***	0.267	3.269***	0.2707
wave 13	3.66***	0.303	3.714***	0.308
wave 14	3.939***	0.3287	4.003***	0.3346
wave 15	4.322***	0.3623	4.418***	0.3704
wave 16	4.803***	0.4083	4.89***	0.4157
wave 17	5.318***	0.4441	5.381***	0.4496
wave 18	5.282***	0.4616	5.361***	0.469
wave 19	6.699***	0.5826	6.809***	0.5931
wave 20	13.561***	1.9773	13.957***	2.1557
wave 21	17.875***	2.6556	18.251***	2.8631
Household size	1.148*	0.0824	1.192***	0.06
Aboriginal or Torres Strait Islander	1.542***	0.2453	1.025	0.0875
Aboriginal or Torres Strait Islander * Household size	0.76***	0.0616	n.a.	n.a.
Immigrant to Australia	1.163	0.112	0.969	0.0518
Household income	1***	0	1***	0
Immigrant to Australia * Household income	1*	0	n.a.	n.a.
Not employed	0.4***	0.0464	0.346***	0.0327
Not employed * Household income	1***	0	1***	0
Not employed * Household size	1.106*	0.0652	n.a.	n.a.
Financial stress	1.609***	0.1403	1.783***	0.1737
Financial stress * Household size	0.851***	0.0436	0.787***	0.0516
Speaks English poorly	0.642***	0.0711	0.752**	0.1047
Long term health condition	0.883	0.07	0.89	0.0716
Long term health condition * Household income	1**	0	1**	0

LIHS	Some Odds ratio	Some Std error	All Odds ratio	All Std error
Immigrant(s) from non-English speaking background	1.089	0.0686	1.029	0.0694
Year 11 as the highest level of education	0.907***	0.0296	0.916**	0.0318
Year 11 as the highest level of education * 2020/21	n.a.	n.a.	0.79**	0.0817
Aged 65 or older	2.223***	0.464	0.725***	0.0548
Aged 65 or older * 2020/21	n.a.	n.a.	1.246**	0.1237
Aged 65 or older * Household income	1***	0	n.a.	n.a.
Aged from 31 to 64	2.432***	0.4673	0.713***	0.0883
Aged from 31 to 64 * Household income	1***	0	1***	0
Aged from 15 to 30	1.957***	0.3829	0.515***	0.0684
Aged from 15 to 30 * Household income	1**	0	1***	0
At least one member is aged below 15	1.959***	0.3743	1.975***	0.3773
At least one member is aged below 15 * Household income	1***	0	1***	0
Rural household	1.207***	0.0582	1.204***	0.0579
Rural household * 2020/21	0.78*	0.1113	0.776*	0.1103
Remote household	3.266***	1.2129	3.304***	1.2318
Remote household * Household income	1***	0	1***	0
Lone person household	0.483***	0.0592	0.493***	0.0556
Lone person household * Household income	1***	0	1***	0
Lone person household * 2020/21	0.649***	0.0741	0.653***	0.0753
Works from home	2.633***	0.4821	1.618***	0.1238
Works from home * household income	1***	0	n.a.	n.a.
SEIFA deciles	1.032***	0.006	1.033***	0.006
SEIFA deciles * 2020/21	0.958**	0.0166	0.95***	0.0168
_cons	3.633***	1.0454	14.459***	2.7129
N Obs	136063	136063	136063	136063
Log pseudolikelihood	-15410.9	-15410.9	-15414.3	-15414.3
Pseudo R2	51.9%	51.9%	51.8%	51.8%

Notes: *Significant at 10 per cent level; **Significant at 5 per cent level; ***Significant at 1 per cent level. n.a. – not included in the model as not statistically significant.

Source: The HILDA Survey, Release 21; BCARR calculations.

Table 8 and Table 9 show the logistic regression output for the LILS models, excluding and including interaction terms respectively. Following the addition of interaction terms, the Pseudo R2 and Log Pseudolikelihood values for the LILS model improve.

Table 7: Logistic regression estimates, LILS group, 'All' and 'Some' models without interaction terms.

LILS	Some Odds ratio	Some Std error	All Odds ratio	All Std error
wave 7	0.897	0.1723	0.896	0.1724
wave 8	0.909	0.1767	0.907	0.1761
wave 9	1.022	0.1954	1.02	0.1949
wave 10	1.223	0.2266	1.221	0.226
wave 11	1.212	0.2135	1.206	0.2125
wave 12	1.419**	0.2447	1.414**	0.2439
wave 13	1.595***	0.2722	1.586***	0.2712
wave 14	1.594***	0.2738	1.585***	0.2722
wave 15	1.549**	0.269	1.536**	0.2671
wave 16	1.599***	0.2775	1.585***	0.2751
wave 17	1.734***	0.2955	1.726***	0.2943
wave 18	1.485**	0.2655	1.471**	0.2631
wave 19	1.729***	0.3028	1.715***	0.3
wave 20	2.056***	0.3551	2.027***	0.3504
wave 21	2.042***	0.3584	2.014***	0.3537
Household income	1***	0	1***	0
Household size	0.633***	0.0766	0.67***	0.0786
Aboriginal or Torres Strait Islander	1.249	0.1765	1.124	0.1898
Migrants to Australia	0.835*	0.0856	0.821*	0.094
Not Employed	1.293***	0.1236	1.538***	0.1445
Financial stress	0.992	0.0888	0.998	0.1005
Speaks English poorly	0.912	0.1966	0.691	0.2058
Long term health condition	1.073	0.0724	1.038	0.0701
Immigrant(s) from non-English speaking countries	1.18	0.1533	1.294*	0.1835
Year 11 as the highest level of education	1.27***	0.0843	1.346***	0.0924
Aged 65 or older	0.704**	0.106	1.324**	0.1872
Aged from 31 to 64	0.659***	0.09	1.315**	0.1828
Aged from 15 to 30	0.735**	0.1061	1.506***	0.2231
At least one member is aged below 15	0.862	0.1463	0.772	0.1299

LILS	Some Odds ratio	Some Std error	All Odds ratio	All Std error
Rural household	0.949	0.0906	0.946	0.0901
Remote household	0.996	0.2443	1.008	0.2477
Lone person household	0.558***	0.0781	0.537***	0.075
Works from home	0.678**	0.1096	0.762	0.1536
SEIFA deciles	0.955***	0.0113	0.956***	0.0113
_cons	0.236***	0.0746	0.099***	0.0314
N Obs	136,063	136,063	136,063	136,063
Log pseudolikelihood	-5504.2	-5504.2	-5494.2	-5494.2
Pseudo R2	18.2%	18.2%	18.3%	18.3%

Notes: *Significant at 10 per cent level; **Significant at 5 per cent level; ***Significant at 1 per cent level. n.a. – not included in the model as not statistically significant.

Source: The HILDA Survey, Release 21; BCARR calculations.

Larger households have lower odds of being in the LILS group. Interestingly though, a larger household that is composed of some or all Aboriginal or Torres Strait Islander members has significantly *higher* odds of being in this group. The same can be said for households where some members are not employed, work from home, or where some or all members have a relatively low level of education, all of which have lower odds of being in the LILS group, but have higher odds of being in the LILS group as their household size increases. This is particularly notable, as these households have been identified as having low telecommunications spending, an issue that is only amplified when that low level of expenditure is being split over a larger number of people. This may indicate a higher risk of potential digital exclusion. COVID-19 resulted in households with children present being much more likely to be present in the LILS group than at other times, indicating that their potential risk of being digitally excluded increased during this time. We also tested for interactions with household disposable income, and while a number were significant, the magnitude of the coefficients indicated that there is no association between the response variable and the interaction term.

Table 8: Logistic regression estimates, LILS group, 'All' and 'Some' models with interaction terms.

LILS	Some Odds ratio	Some Std error	All Odds ratio	All Std error
wave 7	0.909	0.1751	0.893	0.1722
wave 8	0.913	0.1781	0.899	0.1751
wave 9	1.033	0.1983	1.025	0.1962
wave 10	1.21	0.225	1.205	0.2237
wave 11	1.203	0.2126	1.19	0.2103
wave 12	1.39*	0.2404	1.374*	0.2377
wave 13	1.52**	0.2602	1.512**	0.2591
wave 14	1.53**	0.2637	1.511**	0.2606
wave 15	1.503**	0.2619	1.486**	0.259
wave 16	1.536**	0.2672	1.523**	0.2653

LILS	Some Odds ratio	Some Std error	All Odds ratio	All Std error
wave 17	1.67***	0.286	1.657***	0.2835
wave 18	1.4*	0.2513	1.383*	0.2481
wave 19	1.63***	0.2869	1.609***	0.283
wave 20	1.654	0.5564	2.808***	0.8723
wave 21	1.672	0.5476	2.782***	0.8346
Household size	0.222***	0.078	0.715**	0.1179
Aboriginal or Torres Strait Islander	0.339***	0.1086	0.437***	0.1386
Aboriginal or Torres Strait Islander * Household size	1.959***	0.2775	1.626***	0.2266
Immigrant to Australia	0.72***	0.0923	0.83	0.0945
Household income	1***	0	1***	0
immigrant to Australia * Household income	1**	0	n.a.	n.a.
Not employed	0.282***	0.1	0.621***	0.0713
Not employed * Household size	1.953**	0.5608	n.a.	n.a.
Not employed * Household income	1***	0	1***	0
Not employed * 2020/21	0.552***	0.1123	0.349***	0.0694
Financial stress	0.999	0.09	1.02	0.1026
Speaks English poorly	0.859	0.1858	0.706	0.2098
Long term health condition	1.381**	0.2024	1.048	0.0695
Long term health condition * Household size	0.831**	0.0752	n.a.	n.a.
Immigrant(s) from non-English speaking background	1.196	0.1553	1.273*	0.1799
Year 11 as the highest level of education	0.824	0.151	0.892	0.1452
Year 11 as the highest level of education * Household size	1.402***	0.1822	1.344***	0.1437
Lone person household	0.112***	0.0202	0.158***	0.027
Lone person household * Household income	1***	0	1***	0
Lone person household * 2020/21	1.633**	0.3944	1.787**	0.4241
Aged 65 or older	0.719**	0.1127	1.267*	0.1791
Aged from 31 to 64	0.847	0.1341	1.26	0.1794
Aged from 31 to 64 * Household income	1***	0	n.a.	n.a.
Aged from 31 to 64 * 2020/21	1.659***	0.2988	n.a.	n.a.

LILS	Some Odds ratio	Some Std error	All Odds ratio	All Std error
Aged from 15 to 30	0.753*	0.1145	1.506***	0.2322
Aged from 15 to 30 * 2020/21	n.a.	n.a.	0.502***	0.1286
At least one member is aged below 15	0.859	0.1599	0.617***	0.1157
At least one member is aged below 15 * 2020/21	3.291***	1.3365	3.965***	1.6008
Rural household	1.809*	0.5526	1.86*	0.6055
Rural household * Household size	0.628**	0.1336	0.622**	0.1432
Remote household	0.402	0.2433	0.954	0.2345
Remote household * Lone person household	3.102*	2.0563	n.a.	n.a.
Works from home	0.411**	0.1484	0.633**	0.1287
Works from home * Household size	1.472**	0.2753	n.a.	n.a.
SEIFA deciles	0.951***	0.0112	1.048	0.0323
SEIFA deciles * Household size	n.a.	n.a.	0.93***	0.0202
_cons	3.752**	2.1379	0.396**	0.1447
N Obs	136063	136063	136063	136063
Log Pseudolikelihood	-5381.9	-5381.9	-5369.7	-5369.7
Pseudo R2	19.9%	19.9%	20.2%	20.2%

Notes: *Significant at 10 per cent level; **Significant at 5 per cent level; ***Significant at 1 per cent level. n.a. – not included in the model as not statistically significant.

Source: The HILDA Survey, Release 21; BCARR calculations.

Attachment G — Logistic regression diagnostics

As discussed earlier in *Attachment F — Logistic regression estimates*, the logistic regressions are estimated on a much smaller sample than the OLS regressions. The logistic regressions, and in particular the ones analysing the likelihood of belonging to the low-income, low spending group (LILS), estimate an outcome that occurs very infrequently. This makes it difficult to predict the outcome from the model. We have taken this into account when running diagnostic tests on our models.

We conducted three diagnostic tests to assess the validity of our logistic regression models. The first and second tests were run to identify the proportion of observations correctly predicted and the sensitivity and specificity of the models. Sensitivity is the fraction of $y_j = 1$ observations that are correctly classified. Specificity is the percentage of $y_j = 0$ observations that are correctly classified. The third test was the Hosmer-Lemeshow goodness of fit test.

The results of the first two tests that we conducted on the final models (including statistically significant interactions) are summarised in Table 9. The proportion of correctly predicted outcomes is above 81% in the LILS models, and 88% in the LIHS models. Using a cut-off point of 0.5 in the specificity and sensitivity analysis, we identified that our models are good at correctly predicting individuals that do not fall into the low-income groups, or specificity, but they are not as good at correctly predicting individuals that do fall into the low-income groups, or sensitivity. To balance the sensitivity and specificity of our predictions we adjusted the cut-off point for our models. The cut-off point was determined graphically and was set at 0.11 for the LIHS models, and 0.02 for the LILS models.

Table 9: Key logistic regression diagnostics, LIHS and LILS groups, 'Some' and 'All' models.

Diagnostic indicator	LIHS Some	LIHS All	LILS Some	LILS All
Specificity	89%	88.9%	82.3%	81.9%
Sensitivity	89.2%	89.4%	77.9%	77.3%
Correctly predicted	89%	89%	82.2%	81.9%
Correctly predicted cases	5.6%	5.6%	0.7%	0.7%
Correctly predicted non-cases	83.4%	83.3%	81.6%	81.2%
Defined cut-off point	0.11	0.11	0.02	0.02

Source: The HILDA Survey, Release 21; BCARR calculations.

The logistic regression was conducted on non-log transformed variables as the use of transformed variables worsened the models' fit (pseudo r^2 and log pseudolikelihood values were lower in these models). The household characteristics specified in each model were determined through a series of tests which we used to identify if a variable improved the model's ability to explain the variation in the dependent variable. Households with one or more members that are a migrant to Australia from a non-English speaking background, or households where one or more members are Aboriginal or Torres Strait Islanders were often not statistically significant, but we included them in the analysis regardless to be able to control for these socioeconomic characteristics of households. Full-time study did not improve the model's ability to predict and was not included.

The Hosmer-Lemeshow test assesses the model's goodness-of-fit. The test should produce a larger p-value (closer to 1) as it is indicative of a good model fit, where a small p-value (< 0.05) indicates poor model fit. All four models get a p-value of 1 in the chi-squared test for goodness-of-fit. These are not reported here.

11. References

- ACCC. 2014. *ACCC telecommunications reports 2013-14*.
https://www.accc.gov.au/system/files/906_ACCC%20Telecommunications%20reports%202013%E2%80%9314_web_2-June-2015.pdf.
- . 2022. “Communications market report, 2021-22.” https://www.accc.gov.au/system/files/22-71RPT_Communications%20Market%20Report_FA.pdf.
- ACMA. 2022. *Communications and media in Australia: How we use internet*.
<https://www.acma.gov.au/publications/2022-12/report/communications-and-media-australia-how-we-use-internet>.
- . 2014. *Communications Report 2013-14*. <https://www.acma.gov.au/sites/default/files/2019-08/research-communications-trends-2014.pdf>.
- BCAR. 2016. “The communications sector: recent trends and developments. October 2016.”
<https://www.infrastructure.gov.au/sites/default/files/the-communications-sector-recent-trends-and-developments.pdf>.
- BCARR. 2020. “Affordability of communications services for low income households.”
<https://www.infrastructure.gov.au/departments/media/publications/affordability-communications-services-low-income-households>.
- . 2017. “Trends and drivers in the affordability of communications services for Australian households.”
<https://www.infrastructure.gov.au/departments/media/publications/trends-and-drivers-affordability-communications-services-australian-households>.
- Breunig Robert, McCarthy Owen. 2020. “Household telecommunications expenditure in Australia.” *Telecommunications Policy* Volume 44, Issue 1.
- David W. Hosmer, Stanley Lemeshow. 2000. *Applied Logistic Regression. Second edition*. John Wiley & Sons, Inc.
- Julian Thomas, Anthony McCosker, Sharon Parkinson, Kieran Hegarty, Daniel Featherstone, Jenny Kennedy, Indigo Holcombe-James, Lyndon Ormond-Parker, Lauren Ganley. 2023. *Measuring Australia’s Digital Divide: the Australian Digital Inclusion Index 2023*.
https://www.digitalinclusionindex.org.au/wp-content/uploads/2023/07/ADII-2023-Summary_FINAL-Remediated.pdf.
- Mohammad Afshar Ali, Khorshed Alam, Brad Taylor, Shuddhasattwa Rafiq. 2019. “Do income distribution and socio-economic inequality affect ICT affordability? Evidence from Australian household panel data.” *Economic Analysis and Policy* 317-328.
- Nicole Watson, Mark Woden. 2010. “The HILDA Survey: Progress and Future Developments.” *Australian Economic Review* 326-336.
- Robert Kenny, Charles Kenny, Zack Gehan. 2023. “What Drives Broadband Traffic.” *Telecommunications Policy* Volume 47, Issue 9.
- Wiwatowska, Anula. 2020. *Every telco’s COVID-19 assistance package [Updated]*.
<https://www.whistleout.com.au/Broadband/Guides/every-telco-COVID-19-coronavirus-assistance-package>.