DOI: 10.1111/1475-5890.12353

Fiscal Studies
The Journal of Applied Public Economics
IFS Institute for Fiscal Studies

SYMPOSIUM PAPER

Understanding Society: the income data

Paul Fisher | **Omar Hussein**

Institute for Social and Economic Research, University of Essex; Understanding Society

Correspondence
Paul Fisher, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, UK.
Email: pfishe@essex.ac.uk

Submitted: November 2023

Abstract

We introduce the income data of Understanding Society, the UK Household Longitudinal Study. First, we show that the data are widely used in academic and policy research. We then discuss the pros and cons of different types of data on household incomes. We go on to describe the income content of Understanding Society, emphasising key details of data collection and data processing – specifically the derivation of net household income totals. We perform a quality assessment that compares Understanding Society estimates of net household incomes to those from a reliable cross-sectional source – the Households Below Average Income series. We conclude that the Understanding Society income data are of high quality, and so are an excellent source for research on the income distribution or incomes more generally. We finish with a discussion of future directions for income data collection in the study.

KEYWORDS
data quality, measurement error, validation

JEL CLASSIFICATION
C81, C83, D31

1 | INTRODUCTION

Social scientists have long been interested in population incomes and matters pertaining to the income distribution. Research questions that require accurate measurement of household incomes include the monitoring of poverty, inequality, and their dynamics, or understanding social gradients in socio-economic outcomes – ranging from health to attitudes. Household surveys are the most widely used and comprehensive method for measuring population incomes. Administrative income records, if available, offer several key advantages over surveys, but they suffer from limitations too, including non-coverage of population subgroups (such as tax non-filers) or important income sources (such as

family transfers), and an absence of demographic controls or key outcome variables. In the context of the UK, one of the few longitudinal data sources on incomes is Understanding Society, the UK Household Longitudinal Study.

In this paper, we introduce the income data of Understanding Society. First, we motivate the collection of income data in a household survey by discussing the pros and cons of different types of data on household incomes. We then describe the income questionnaire content and outline the derivations and processing that are undertaken to arrive at the (publicly available) household income variables. As the main UK panel data on household incomes, Understanding Society underpins much research, and so it is important to document its quality. This paper presents a quality assessment by comparing the Understanding Society income distribution to a gold standard cross-sectional source – the Households Below Average Income series (HBAI). While the HBAI is also derived from a survey – the Family Resources Survey (FRS) – and is not a register of population incomes, it is considered to be of high quality as it undergoes careful editing and checking by the Department for Work and Pensions, based on their access to administrative records and knowledge of the tax and benefit system.

Earlier validations have assessed the income data of both Understanding Society and the former British Household Panel Survey (BHPS) – the predecessor to Understanding Society. For the BHPS, Jenkins (2011) presents the net income series and demonstrates that it tracks very closely its counterpart from the FRS, although the BHPS gives lower estimates of inequality which do not track the FRS exactly. In this spirit, Francesconi, Sutherland and Zantomio (2010) compare earnings distributions of the BHPS and FRS but at three points in time and, while estimates from the initial comparison line up closely, a gap is present for the latter comparisons that could be the result of survey attrition. Differently, Postel-Vinay and Sepahsalari (2023) consider average earnings around the Understanding Society–BHPS seam and show that an average earnings series closely matches an equivalent derived from Office for National Statistics (ONS) sources. For the first eight waves of Understanding Society, it has been established that distributions of household net income and earnings match closely FRS counterparts, as do measures of inequality.[1]

Since the latest validations, there has been a trend to online, rather than in-person, interviewing, which has quickened since the COVID-19 pandemic. This could potentially affect both response quality and response rates to the survey. More generally, survey response rates have declined across the world in recent decades, and so it is important to continuously monitor the impact of this trend on data quality. Finally, like all panel surveys, Understanding Society has suffered from attrition that may threaten its representativeness. However, boost samples have been added since wave 1 (a further boost sample is planned for wave 14), which may counteract the effects of attrition. As we approach the wave 13 release of the survey, the present paper provides an updated quality assessment in the face of the above trends.

Our main finding is that the household net income distributions derived from Understanding Society and the FRS track each other closely over the last decade. This pattern holds, with only minor exceptions, for all of the measures we study; selected income percentiles and percentile ratios, poverty and inequality measures. Where the surveys deviate is in their assessment of subcomponents of income: earnings track each other quite closely, but Understanding Society estimates relatively more state benefit income with the gap increasing over time – this latter fact may reflect observed differences in sample compositions. While we are not able to perform longitudinal comparisons directly – as there is no longitudinal version of the FRS – we interpret our findings as indicative that Understanding Society has a strong basis for longitudinal, as well as cross-sectional, research on the income distribution.

The results of this paper support a substantial body of both academic and policy research that uses Understanding Society income data. These data have been used to study income volatility

[1] Fisher et al., 2019; Fisher, 2019.

following the Great Recession[2] and the effects of state benefit cuts on income dynamics,[3] including the introduction of the bedroom tax[4] and the effects on long-run incomes of Universal Credit.[5] It was a critical resource in understanding the short-term economic impacts of the COVID-19 pandemic and, in particular, with respect to pre-pandemic economic positions.[6] Distinct research agendas have explored trends in the dynamics of pensioner[7] and ethnic minority[8] poverty, to estimate intergenerational income mobility,[9] and to explore the role of higher education expansion in income inequalities.[10] Many published papers feature income as a key control variable, for example, to document income gradients in health,[11] life satisfaction,[12] child subjective well-being,[13] the effect of job loss on fertility,[14] and energy use.[15] The income data have also been used to identify benefit recipients[16] and to understand the role of benefit cuts in the support for Brexit.[17]

Numerous policy reports have also been underpinned by Understanding Society income data, which form the basis for official estimates on income dynamics.[18] The data have been further used in the analysis of poverty and poverty measurement,[19] distributional analysis,[20] low pay,[21] gender wage gaps,[22] pension saving[23] and health.[24]

The paper proceeds as follows. In Section 2, we discuss the relative strengths of different types of data on household incomes. In Section 3, we set out the Understanding Society income data collection and post-data-collection processing. In Section 4, we set out our quality assessments comparing Understanding Society to a cross-sectional gold standard. Finally, we conclude in Section 5, setting out some future directions for income data collection in the study.

2 | DATA SOURCES ON HOUSEHOLD INCOMES

In this section, we discuss the strengths and weaknesses of the different data sources on household incomes. We finish with a discussion of the advantages of panel data on incomes over cross-sectional income data.

[2] Avram et al., 2022.

[3] Mari and Keizer, 2023.

[4] Gibbons, Sanchez-Vidal and Silva, 2020.

[5] Brewer et al., 2020.

[6] Crossley, Fisher and Low, 2021; Crossley et al., 2022; Brewer and Tasseva, 2021.

[7] Kanabar, 2017; Barnes, 2022.

[8] Perez Hernandez, Kanabar and Nandi, 2018.

[9] Rohenkohl, 2019.

[10] Carrieri, Davillas and Jones, 2022.

[11] Davillas, Jones and Benzeval, 2019; Johnston, Kung and Shields, 2021; Davillas and Jones, 2021; Parra-Mujica et al., 2023.

[12] Brown, Gray and Roberts, 2015; FitzRoy and Nolan, 2020, 2022; Kaiser, 2020; Davillas, Burlinson and Liu, 2022.

[13] Main, 2019.

[14] Di Nallo and Lipps, 2023.

[15] Baltruszewicz et al., 2023.

[16] Katikireddi et al., 2017, 2018; Borbely, 2022.

[17] Fetzer, 2019.

[18] Department for Work and Pensions, 2023b.

[19] Social Metrics Commission, 2020; Cebula et al., 2023.

[20] HM Treasury, 2020; Ray-Chaudhuri et al., 2023.

[21] Cominetti et al., 2023.

[22] Andrew et al., 2021a, 2021b.

[23] Cribb, Karjalainen and O'Brien, 2023.

[24] Thomas et al., 2023.

2.1 | Survey versus administrative data

Administrative records represent a high-quality and appealing data source for estimating population incomes. The records are constructed by linking individual data held by different agencies. For example, information on state benefit receipt held by a benefits agency would be combined with income tax returns held by tax authorities, creating a complete income record for each individual. As administrative records contain the full population, the income distribution can be estimated with high precision. Analysis of small subgroups is also possible, provided such subgroups can be identified in the records. For example, tax records have been the central data source in studying top income shares.[25]

Unlike surveys, administrative records do not suffer from non-response that threatens their representativeness. Nor do they suffer from reporting errors, such as the under-reporting of state benefits that is widespread in surveys.[26] Although a number of recent studies have documented important errors in administrative income data too.[27,28]

The availability of administrative income data differs by country. The Nordic countries in particular have long-standing and high-quality administrative registers that have been developed and maintained over several decades. Researchers have good access to the register micro-data and they have been widely used in economic research. Several other European countries are in the planning stages of using register data in official income statistics.[29] But, in other countries, usable administrative data do not exist, or different datasets may not be linked where they are held by different agencies. Researcher access may also be limited where data owners perceive the planned research falls outside the interests of their organisations, where legal barriers prevent data sharing, or where they are focused (and trained) on administration, rather than the production of research data.

Administrative records are not without their weaknesses and household survey data on incomes retain several advantages, which we detail below. First, even where the full population is covered, item non-coverage can occur where key sources of income are excluded from administrative records. One such example is non-taxable transfers to family and friends. Crossley, Fisher and Low (2021) highlight the widespread use of such transfers in the early stages of the COVID-19 pandemic in the UK.[30] Distinctly, important types of income may not be reported to authorities to evade tax.

Second, household surveys permit the construction of income totals at the household level, which is the level at which much income distribution analysis is performed, including UK official poverty statistics. In contrast, administrative tax records are collected at the tax unit level and it is not possible to identify members of different tax units living in the same household.[31]

Third, surveys can provide a consistent set of questions not influenced by changes to administrative systems or even political interference. However, survey content is sometimes revised, but this tends to be rarer for established surveys and only when there is a research need to do so – rather than an administrative need.

[25] Piketty and Saez, 2003; Atkinson, 2005.

[26] See, for example, Brewer, Etheridge and O'Dea (2017) and Meyer and Mittag (2019).

[27] Bingley and Martinello, 2017; Wilhelm, 2018; Garin, Jackson and Koustas, 2022.

[28] Garin et al. (2022) study a supposed increase in self-employment which is seen in US tax records in the 2000–14 period, but not major labour force surveys. The authors demonstrate that exogenous variation in tax incentives changes reporting behaviour and conclude that it is strategic reporting to tax authorities, rather than real changes in work, that can explain most of the observed patterns in the tax data. Bingley and Martinello (2017) finds a positive error variance in Danish administrative earnings data, although much smaller than the survey-based equivalents. Wilhelm (2018) finds evidence of errors in the US Social Security Administration's measure of earnings originating from young individuals with high earnings growth.

[29] Jantti, Törmälehto and Marlier, 2013.

[30] Crossley et al. (2021) report that 12 per cent of individuals in households experiencing an earnings loss at the start of the pandemic received a transfer from friends and family.

[31] See Larrimore, Mortenson and Splinter (2021) for a comparison of tax-based and household-based income distribution estimates.

Fourth, where a researcher wants to analyse income alongside other variables (including basic demographics), administrative income records fall short when they are not linked to wider databases on other topics. Further, many variables that economists wish to analyse alongside income, such as beliefs and attitudes, are not collected for administrative purposes and so cannot appear in any administrative dataset.

A distinct type of data is produced where a household survey is linked to administrative income records in order to benefit from some of the strengths of both data sources. We comment on this possibility in the Understanding Society context in Section 5.

2.2 | Specialist versus non-specialist income surveys

The specialist income survey produces the highest-quality income data, amongst the different types of survey. Specialisation comprises of thorough questioning on income sources, an emphasis on document checking, the use of specialist interviewers and time-intensive post-data-collection processing (data cleaning, imputation, etc.). Yet surveys that do not specialise in income can also produce reliable income data where the data quality depends on the degree of income specialisation. At one end, we have surveys with a focus on other topic domains that ask only summary or single questions on income. Summary questions have been shown to perform well but suffer from under-reporting.[32] At the other end of the spectrum are surveys such as Understanding Society that have a high degree of income specialisation but to a lesser extent than the specialist income survey.

In addition to income variables, both types of survey nearly always contain basic socio-economic characteristics. General purpose surveys go further still and collect information on a wide range of topics in addition to income. For example, Understanding Society collects data on topics including household finances, the labour market, health, expectations, attitudes, education, family, fertility, transport and neighbourhoods, and civic engagement.

2.3 | Researcher designed surveys

A popular alternative is for a researcher to collect their own income data using a convenience sample. These are readily available samples of unknown composition but are often made to match a few known population totals (such as age and gender) either through the use of quota sampling or through re-weighting ex-post to the known totals. A strong advantage of this approach is that the survey content can be tailored to the specific research question of the researcher and not be restricted by the questions of established surveys. Such surveys can also be quicker to field and then analyse in real time. In contrast, administrative reporting deadlines (such as tax self-assessments) can occur many months after the reference period, delaying the availability of administrative data. Household surveys also face release lags where they are complicated, as extensive post-fieldwork processing must be completed.

There are also notable downsides of the analysis of convenience samples. First, it is not possible to generalise results from the sample to the wider population without strong assumptions on sample inclusion and response.[33] Second, resource constraints can affect survey quality whether it be through, for example, an absence of testing of survey questions, low power, or the capacity to field and process detailed questions on income. Third, issues of replication are raised as the wider research community may find it difficult to access the data, or the detailed documentation that accompanies large-scale surveys can be missing.

[32] Micklewright and Schnepf, 2010; Crossley, Fisher and Hussein, 2023.

[33] See Crossley et al. (2021) for a discussion of this topic.

2.4 | Panel data on incomes

Whereas cross-sectional data are defined by individuals being observed only once, in panel data individuals are followed over time and observed periodically. A key advantage is that the researcher can control for certain types of omitted variable bias (time-invariant fixed effects). An additional advantage when the panel data come from surveys is that more contextual variables become available over time as new questions are asked and information accumulates across the waves of the panel. Long-standing household panel studies also facilitate the study of intergenerational income (and wealth) mobility as both parent and children are observed as adults.[34]

Cohort studies, which are panels defined by those born in a particular birth cohort (e.g., a particular week in a given year), have distinct advantages but they are not the tool for monitoring population incomes as they do not represent the full population in any given year. Rather, their main strength is that they are rich in information on child development and the home environment, so that adult outcomes can, over time, be related to childhood. While most cohort studies collect income data, a compromise is often made by using summary income questions, rather than the highest-quality detailed question sets. A larger spacing between survey interviews than is seen in the typical household panel also makes them less suited to measuring shorter-term changes in income.

Household panel data on incomes, such as Understanding Society, through their following rules, start out and remain representative of the population in any given year. They allow for the study of income mobility and within-individual changes, in addition to the repeated snapshots of the income distribution which is possible with cross-section data. This also means that the impact of life events can be studied by income prior to an event; for example, the home environment is observed prior to birth (unlike in a cohort study) and household income is observed prior to an economic shock.

The main disadvantage of panel surveys is attrition, which occurs when participants in the study drop out over time. Selective attrition can bias estimates of population parameters which survey weighting can aim to correct. Administrative income data, however, do not suffer from survey attrition, although individuals may fall out of administrative records where they fall out of their coverage; for example, a benefit claimant might fall out of benefit records once they stop receiving a benefit or a taxpayer might fall out of tax records if their income fell below the income tax thresholds.

In sum, each income data source has its strengths and weaknesses depending on the requirements of the researcher. Where a stand-alone analysis of the income distribution is to be performed, administrative income records likely perform well, but such records are nearly always incomplete and can be difficult to access. Where such records are unavailable or incomplete, a specialist income survey should be preferred. In the large number of cases where other variables are to be analysed in conjunction with income, household survey data will be strongest and household panel data bring their own unique advantages. In the next section, we discuss the specific details of the income data collection of Understanding Society.

3 | WHAT IS COLLECTED AND HOW?

Whilst Understanding Society is a general purpose survey, and not a specialist income survey, substantial effort and questionnaire time is allocated to the collection of income data. In collecting income data, international best practice is followed.[35] Being a longitudinal and not a cross-sectional survey, it can take advantage of longitudinal data collection tools (dependent interviewing) that have been shown to improve income data quality,[36] in addition to the use of in-interview soft checks and the encouragement of document checking by respondents where appropriate.

[34] See, e.g., Levell and Sturrock (2023).

[35] See United Nations (2011).

[36] Lynn et al., 2012.

Understanding Society aims to collect data on household incomes net of taxes and National Insurance contributions. To do this, it asks each individual in the household about each of their income sources. Household total income can be calculated by summing over sources and then individuals in a household. Benefits of the source-by-source approach are that it enables the analysis of separate income sources (e.g., individual earnings) and also gives the researcher a degree of flexibility over what sources are counted as income (e.g., market incomes excluding state transfers). A comprehensive set of income sources is collected, up to 46 in total, which can be summarised as: earnings from main and second jobs, social security benefits, state and private pensions, private transfers and investment income.

A further key definitional issue is the reference period for reporting. Understanding Society aligns with official UK (cross-sectional) income statistics by asking about 'current income' or income around the time of a survey interview. Under current income measurement, the reference period for reporting is chosen by the survey respondent and the reported amounts can then be standardised by the researcher post-data-collection.[37] In the publicly released Understanding Society files, income amounts are converted to monthly equivalents.[38] Understanding Society makes an exception for income from investments and savings, which is received more normally on an annual basis, and is asked about for the last 12 months. The approach of current income measurement is considered best to maximise response.[39]

The diverse topic coverage of Understanding Society leads to a questionnaire design with multiple modules defined by topic. We review the most relevant modules here. The 'current employment module' establishes whether a respondent is currently employed and, if so, whether an employee or self-employed. The 'employee's' module asks for net (take-home) and gross pay at last payment.[40] To aid accurate reporting, respondents are encouraged to consult a payslip and an in-interview check alerts a respondent where reports of gross pay are less than or equal to net pay. The 'self-employment' module asks self-employees for their share of the net profit or loss on their most recent accounts or, where not available, an estimate of their net usual monthly or weekly self-employment income. In web-mode, where respondents fail to provide amounts, additional motivational statements appear, which have been shown to increase respondent willingness to answer the question.[41] The 'second jobs module' asks about gross income from all second or odd jobs in the last month, and from wave 13, the module will be expanded to ask for details of each job separately. The 'unearned income and state benefits module' collects information on whether a respondent is currently receiving any of 42 income sources (covering social security benefits, state and private pensions, private transfers and other income streams), the amount received and whether it is received jointly with a partner. The latter module makes use of dependent interviewing to reduce spurious change in reports of unearned sources between waves.[42]

Further modules relevant for the construction of household incomes are the pensions module (pension membership, contribution rates, and pension drawdown) and, from wave 12, a student loans module asking about repayment behaviour of any student loan. Separately, variables relating to housing costs (not the focus of this paper) and local tax measures are collected in the household questionnaire.

To derive a final household net income variable from the raw survey reports is a time-intensive process that demands extensive computer programming. The programs are updated on an annual

[37] While there is a research interest in annual measures of income, competing demands on interview space mean that annual measures of income are not collected.

[38] In Understanding Society, most respondents choose to report on a monthly period code (e.g., for earnings, 71.5 per cent chose monthly at the latest wave).

[39] United Nations, 2011.

[40] Deductions referred to by the questionnaire are: tax, National Insurance, pensions, student loan repayments, union dues, etc.

[41] Al Baghal and Lynn, 2015.

[42] Lynn et al., 2012.

basis as part of the Understanding Society data release. The final net income variables, contained in the annual public release, consist of household net income and its subcomponents, individual-level equivalents and equivalents that are gross of taxes. A full list of the Understanding Society derived income variables can be found in the Understanding Society user guide.[43]

Specifically, there are several aspects to the household net income derivations. Monthly amounts are calculated for each income source and some basic data cleaning is performed to identify reporting errors where they are clear. Where the same income source is reported twice in a household (e.g., the same state benefit is reported by both members of a couple), this must be identified to avoid double-counting of income. Next, missing data – both item and unit missing – are filled by imputation (complete household non-response is addressed by weighting[44]). Understanding Society makes use of several imputation methods including the longitudinal method of Little and Su (1989), which aims to reduce the introduction of spurious income change between waves. To preserve the link between net and gross pay, earnings are imputed gross and then tax and national insurance are calculated by simulation (i.e., applying the current year's tax schedule), to arrive at a net figure. Occupational pension contributions are deducted where they have been reported in the pensions module. Likewise, where only one of net or gross earnings is item-missing, the other is calculated by simulation, rather than by imputation. Note that estimation of tax and National Insurance contributions is based on labour income only. Further matters of detail of the net income derivations are provided in Fisher (2019).

A final step makes deductions for household-level taxes by linking external information on council tax to the survey.[45] In particular, information on demographics and discounts received (council tax reduction) are taken from the survey, while a property's council tax band and a local authority's council tax rate are linked to external and publicly available records.[46] Gross council tax liability can be calculated from the externally linked bands and rates, whilst net council tax liability can be obtained by applying the discounts as calculated from the survey reports.[47]

The Understanding Society derived income variables – in particular, household net income – are the focus of our quality assessment in the next section.

4 | CROSS-SECTIONAL COMPARISONS TO THE FAMILY RESOURCES SURVEY, 2010–19

In this section, we compare the Understanding Society income data to an alternative UK data source on household incomes. Our comparator data source is the HBAI,[48] which is derived from the FRS. We choose the HBAI as it is the data source for official UK (cross-sectional) statistics on the income distribution and it is considered to be of high quality. The comparisons extend those in Fisher et al. (2019), which themselves build on the substantial body of work done by Stephen Jenkins for the former BHPS.[49]

There are two main differences between our data sources that we need to account for in our comparisons. First, the timing of fieldwork differs. The HBAI fieldwork period corresponds to a financial year (April–March), whereas fieldwork for each Understanding Society wave is spread out

[43] Institute for Social and Economic Research, 2022.

[44] See Lynn, Cabrera-Álvarez and Clarke (2023).

[45] Council tax is a UK household-level tax where each residential property is assigned to one of eight bands based on the property value. The tax is set as a fixed amount for each band, which is set at the local authority level.

[46] Council tax bands are linked, at the address level, to administrative records from the UK Valuations Office. Council tax band is also collected in the household questionnaire of the survey. Council tax rates are merged to local-authority-level data.

[47] Council tax reduction is collected in the state benefit section of the questionnaire. A 'single person discount' is applied where only one person resides in a property, according to the survey reports.

[48] Department for Work and Pensions, 2023a.

[49] Jenkins, 2011.

over two calendar years. Second, as HBAI newly samples the UK population each year, new immigrant households are included in the sample frame, but this is not the case for Understanding Society with the exception of wave 1 and wave 6 where an immigrant boost was implemented.

We address the differences as follows. To make the reference periods comparable, we construct Understanding Society financial year samples by combining data from multiple waves, that is, months 4–15 from wave n are combined with months 16–24 from wave $n-1$ and months 1–3 from wave $n+1$. This means that while our analysis covers all currently available Understanding Society waves, it excludes data collected before April 2010 in wave 1 and after March 2020 in wave 12, as it is not possible to construct full financial year windows for this period. Our comparisons therefore correspond to the financial years 2010–19. To make the coverage of the two surveys comparable, we remove new immigrant households from the HBAI sample, where Understanding Society excludes these populations.[50]

We perform other key data adjustments as follows. All amounts are expressed in 2015–16 prices using a bespoke monthly consumer prices index (CPI) produced by the ONS that is used in official UK income statistics. We equivalise household net income using the OECD-modified scale and express all amounts in monthly equivalents. Cross-sectional weights are applied to each sample so that the data represent the UK population excluding new immigrants as above. As we work with financial year samples, we are required to perform a rescaling of the Understanding Society weights so that observations from latter waves in the same financial year are not under-represented. We drop households larger than nine individuals from Understanding Society to mirror their treatment in the public HBAI files.

We first compare the surveys on a set of demographic variables and at two points in time almost ten years apart (2010 and 2019). Our aim here is not to say which survey produces estimates closer to the true population totals, but rather to emphasise the demographic similarities and differences between the samples and how they have evolved since the start of Understanding Society. Where we observe demographic differences, the unconditional income distributions of the different surveys should not be expected to match exactly. Our demographic comparisons are limited by the set of variables that are comparable in both surveys (benefit unit type, housing tenure, and region). The 2010 comparisons are reported in columns 1–3 of Table 1 and the 2019 ones in columns 4–6. Additionally, column 7 of the table contains difference-in-difference (DID) estimates to test formally for divergence in the (weighted) sample compositions over time.

Table 1 shows that the samples differ in their demographic profiles in 2010. Several of the differences are both statistically significant and economically meaningful, and so deserve comment. Understanding Society estimates a lower share of individuals without children (32.3 per cent versus 35.6 per cent) but a larger share of pensioners (22.2 per cent versus 21.1 per cent) and working age with children (45.5 per cent versus 43.3 per cent). It also estimates a lower share of individuals living in the private rented sector (13.7 per cent versus 16.9 per cent) but a large share living in social housing (17.9 per cent versus 16.4 per cent) or owning their home with a mortgage (41.2 per cent versus 38.9 per cent). By the time of the second comparison, many of the differences have changed in magnitude. Most notably, there have been relative increases in Understanding Society in the share of pensioners (DID estimate of 1.9 per cent) and working age without children (DID estimate of 3.1 per cent), but relative decreases for working age with children (DID estimate of −5 per cent). We might therefore expect Understanding Society to estimate relatively more pension income that increases over time, but less earnings.

We now turn to our income comparisons. Figure 1 shows selected percentiles of the household net income distributions for our two surveys. There is a striking similarity between the two surveys, both in the level of estimates and in their time trend. For example, the estimates of the median in 2019 are very close indeed, being £1,526 (Understanding Society) and £1,472 (HBAI). Larger differences are

[50] Specifically, we remove new immigrant households in the years 2010–13 and from 2015 onwards from the HBAI data. New immigrants are defined as immigrants since 2009/10 for the 2010–13 period, and immigrants since 2014 for the period 2015 onwards.

TABLE 1 Demographic profiles, 2010 and 2019

	Financial year = 2010			Financial year = 2019			DID
	UKHLS	HBAI	Diff.	UKHLS	HBAI	Diff.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Benefit unit type							
Pensioner	22.2	21.1	1.0***	23.2	20.3	2.9***	1.9***
Working age with children	45.5	43.3	2.2***	39.0	41.9	−2.9***	−5.0***
Working age without children	32.3	35.6	−3.2***	37.7	37.8	−0.1	3.1***
Tenure type							
Owned outright	27.0	27.8	−0.8*	31.8	30.6	1.2*	2.0***
Owned with mortgage	41.2	38.9	2.3***	35.5	36.5	−1.0	−3.3***
Social housing	17.9	16.4	1.5***	19.2	16.5	2.8***	1.3*
Private rented	13.7	16.9	−3.2***	12.5	16.5	−4.0***	−0.8
Region							
North East	4.2	4.1	0.1	4.3	4.0	0.2	0.1
North West	11.1	11.2	−0.1	11.4	11.2	0.2	0.3
Yorkshire and the Humber	8.4	8.4	0.0	9.1	8.3	0.8**	0.8*
East Midlands	7.5	7.2	0.3	7.7	7.2	0.5	0.2
West Midlands	8.5	8.8	−0.3	8.8	8.9	−0.1	0.2
East of England	9.3	9.3	0.0	9.8	9.4	0.4	0.3
London	11.9	12.9	−0.9**	11.0	12.6	−1.6***	−0.7
South East	13.5	13.7	−0.2	13.6	13.7	−0.1	0.0
South West	8.5	8.4	0.2	8.4	8.5	−0.1	−0.3
Wales	5.3	4.9	0.4*	4.8	4.9	−0.1	−0.5
Scotland	9.1	8.4	0.7***	8.1	8.3	−0.2	−0.9**
Northern Ireland	2.6	2.9	−0.2**	2.9	2.9	0.0	0.3
Number of individuals	80,708	57,811		45,312	41,491		456,155

Note: The table reports percentages. Understanding Society financial years and weights are constructed by pooling the relevant waves. Standard errors are clustered at the household level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.
Source: Understanding Society and HBAI.

observed as we move below the median, which are statistically significant, although the gap between the estimates of the two surveys remains fairly stable across time. In particular, Understanding Society gives slightly higher estimates of percentiles 1–25, relative to HBAI. In contrast, at the very top of the distribution, HBAI gives higher estimates, where in 2019 p99 is £5,899 (Understanding Society) and £8,066 (HBAI). The latter difference likely reflects the known difficulties in measuring the incomes of the very richest and the statistical adjustments that are made to top incomes in the HBAI.[51]

To further examine the differences over time, selected percentile ratios are presented in Figure 2 (90–50, 75–25 and 50–10 ratios). The estimated ratios are always similar for both surveys and the levels have changed little over the observation period.

We now turn to commonly estimated measures of poverty and inequality. In what follows, we trim the top and bottom 1 per cent of each sample as measures of inequality may be sensitive to

[51] The very richest households have had their incomes adjusted in the HBAI in an imputation method that makes use of the Survey of Personal Incomes (SPI), which is drawn from tax records. The SPI suggests that the unadjusted incomes at the top of the HBAI distribution are understated.

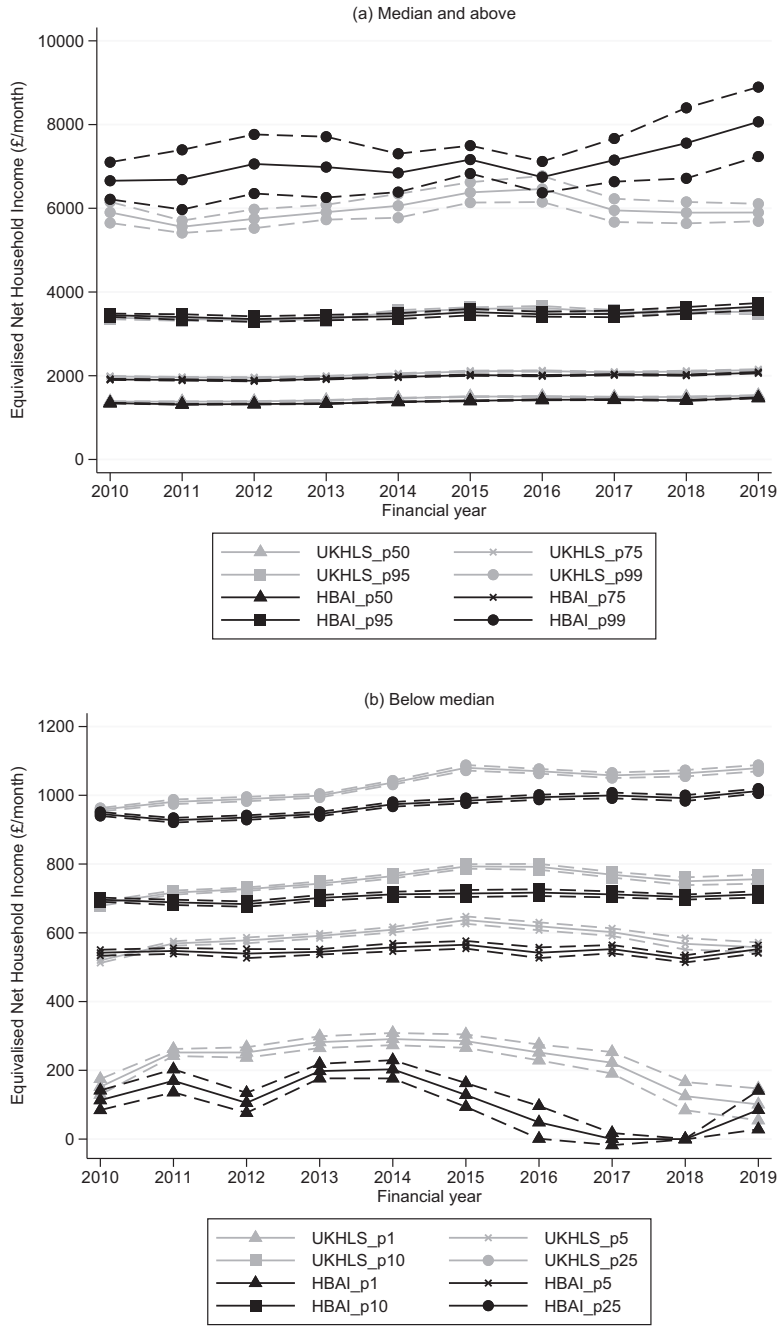

FIGURE 1 Selected income percentiles by survey

Note: Income is measured at the household level, equivalised using the OECD-modified scale, and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves. The dashed lines show 95 per cent confidence intervals (bootstrapped standard errors).
Source: Understanding Society and HBAI.

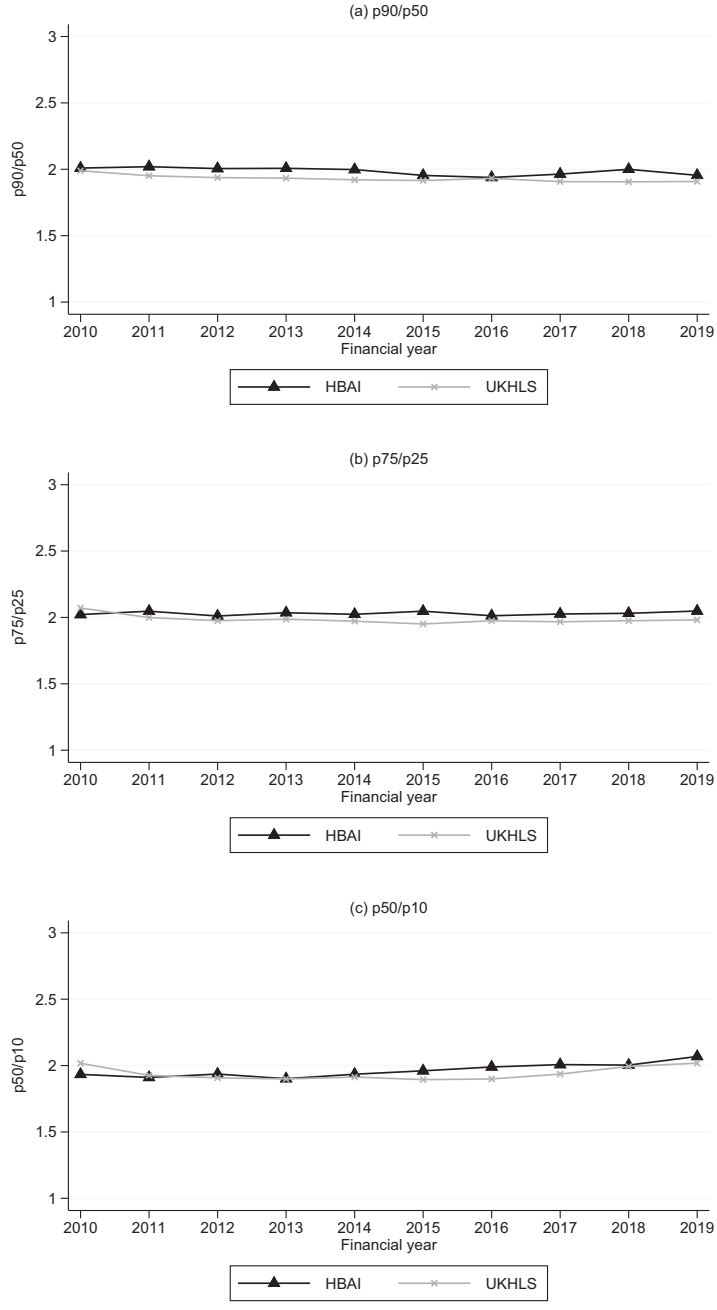

FIGURE 2 Selected percentile ratios by survey

Note: Income is measured at the household level, equivalised using the OECD-modified scale, and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves. Estimation samples drop the richest and poorest 1 per cent of observations.
Source: Understanding Society and HBAI.

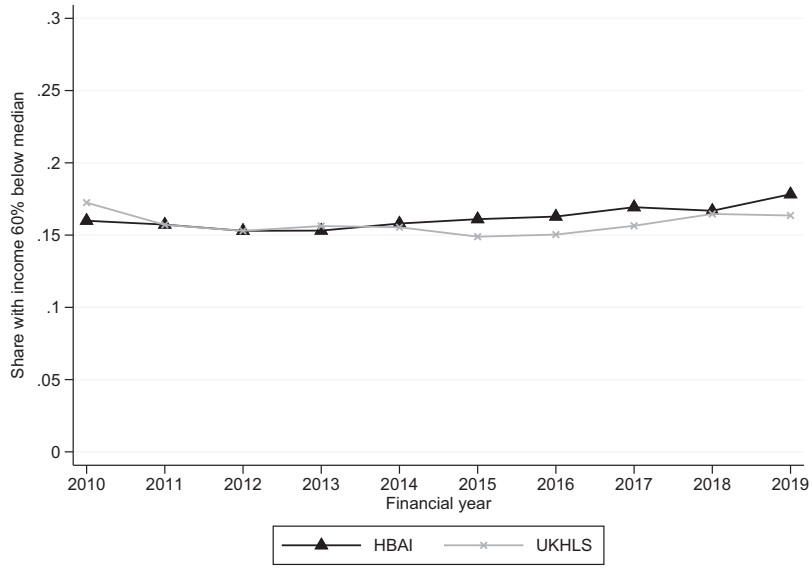

FIGURE 3 Relative poverty by survey (income <60 per cent median)

Note: Income is measured at the household level, equivalised using the OECD-modified scale, and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves. Estimation samples drop the richest and poorest 1 per cent of observations.
Source: Understanding Society and HBAI.

outliers. Figure 3 plots the standard UK poverty rate for each survey.[52] The two surveys line up very closely according to this measure. Poverty rates decrease until 2012 and then increase slowly thereafter. For the first half of the period, the difference across the surveys is negligible, while in the second half Understanding Society tends to give slightly lower rates. For example, at the end of the observation period, the estimated poverty rates are 16.4 per cent (Understanding Society) and 17.8 per cent (HBAI).

Figures 4 and 5 plot trends in a selection of inequality measures (90:10 ratio, Gini, Theil, Atkinson and GE2) for each data source. Looking across the measures, as with the poverty rate, both surveys give similar estimates and trends, with some exceptions that we comment on below. Measured inequality is again typically lower according to Understanding Society by a small amount. The trends in the 90:10 ratio also differ until 2015, where we see a much stronger decline in Understanding Society than what is observed in HBAI. These two features of the data may relate to the compositional differences of the two surveys as above. We note that the pattern of lower inequality for Understanding Society mirrors what is seen in the former BHPS at its later waves.[53] A final point of difference is the spike in inequality in 2017–18 in HBAI that is not seen in Understanding Society.

Figure 6 provides some evidence from each survey on who is where in the income distribution by showing the distribution of benefit unit types for the top, bottom and middle quintiles of the net household income distribution. The same broad patterns are seen in each survey, for example, with the working age without children concentrated in the top quintile (50.7 per cent (UKHLS) and 51.5 per cent (HBAI) in 2019), and the working age with children at the bottom (45.8 per cent (UKHLS), 47.6 per cent (HBAI) in 2019). The compositions of each quintile remain largely stable over time in

[52] That is, the share of individuals living in households with income 60 per cent below the median.

[53] Jenkins, 2011.

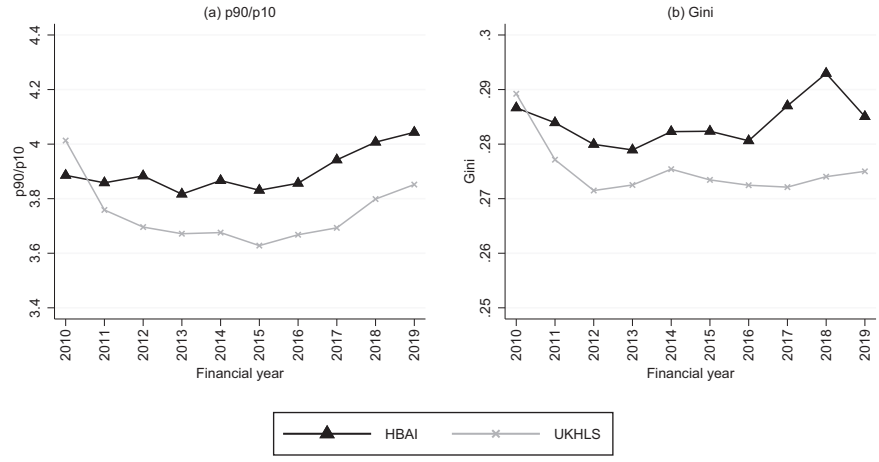

FIGURE 4 Inequality by survey (Gini, 90:10 ratio)

Note: Income is measured at the household level, equivalised using the OECD-modified scale, and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves. Estimation samples drop the richest and poorest 1 per cent of observations.
Source: Understanding Society and HBAI.

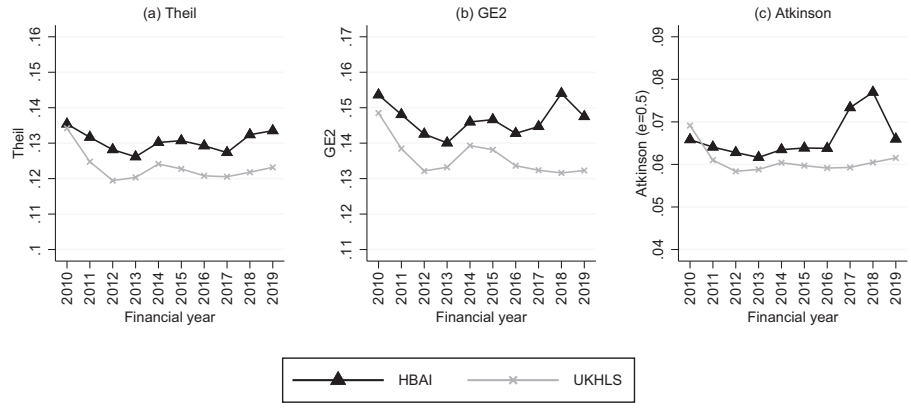

FIGURE 5 Alternative measures of inequality

N*ote*: Income is measured at the household level, equivalised using the OECD-modified scale, and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves. Estimation samples drop the richest and poorest 1 per cent of observations.
Source: Understanding Society and HBAI.

both surveys. The differences in benefit unit type seen in Table 1 are generally mirrored across the quintiles, but less so for the third quintile where the surveys are slightly closer in their composition.

Next we perform analysis of the major and most widely analysed subcomponents of income, which we do separately for working age and pensioner benefit units. Here, we work with (unequivalised) gross household income, as its subcomponents are directly comparable across the datasets. Panel A of Figure 7 presents means and medians of gross household income, earnings and state benefit income for the working age, while Panel B reports the same for pensioners but where earnings are replaced with income from occupational pensions – the more widely received income source for that group.

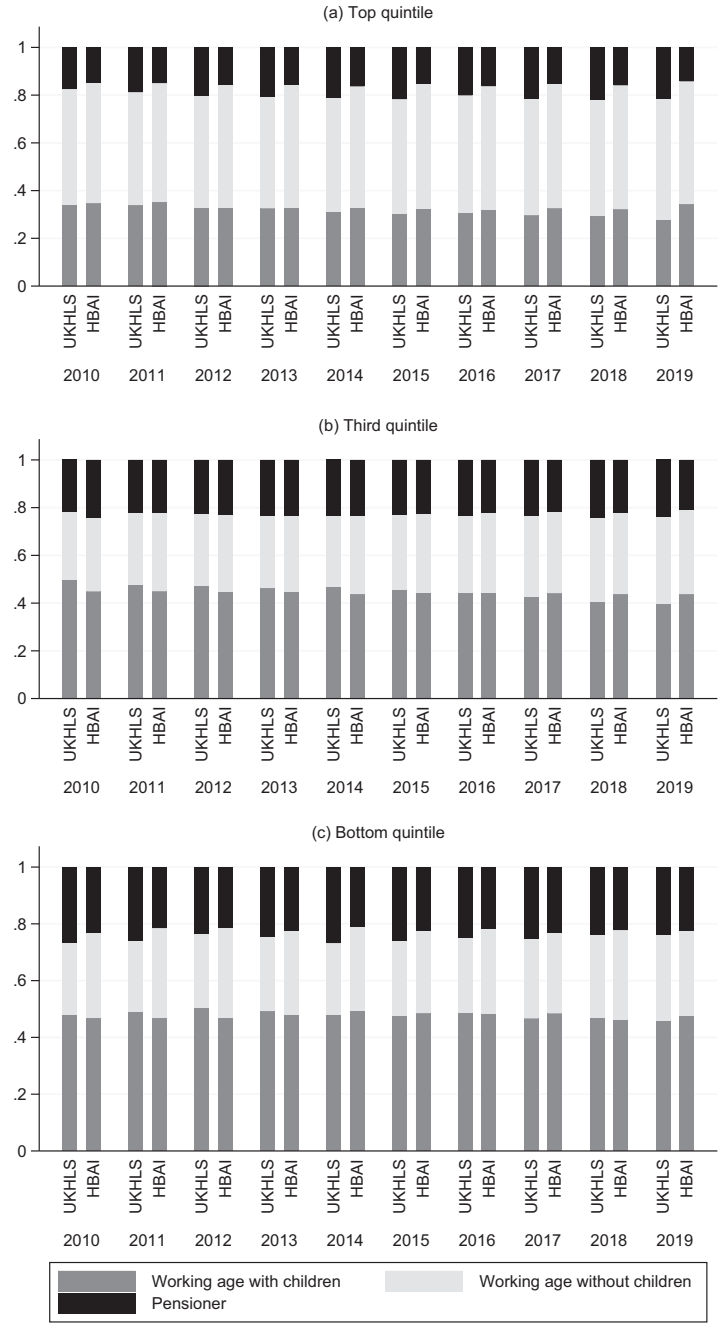

FIGURE 6 Quintile compositions

Note: Income is measured at the household level, equivalised using the OECD-modified scale (rescaled so that the first adult has a weight of 0.67, subsequent adults and children above the age of 14 a weight of 0.33, and children under the age of 14 a weight of 0.2), and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves.
Source: Understanding Society and HBAI.

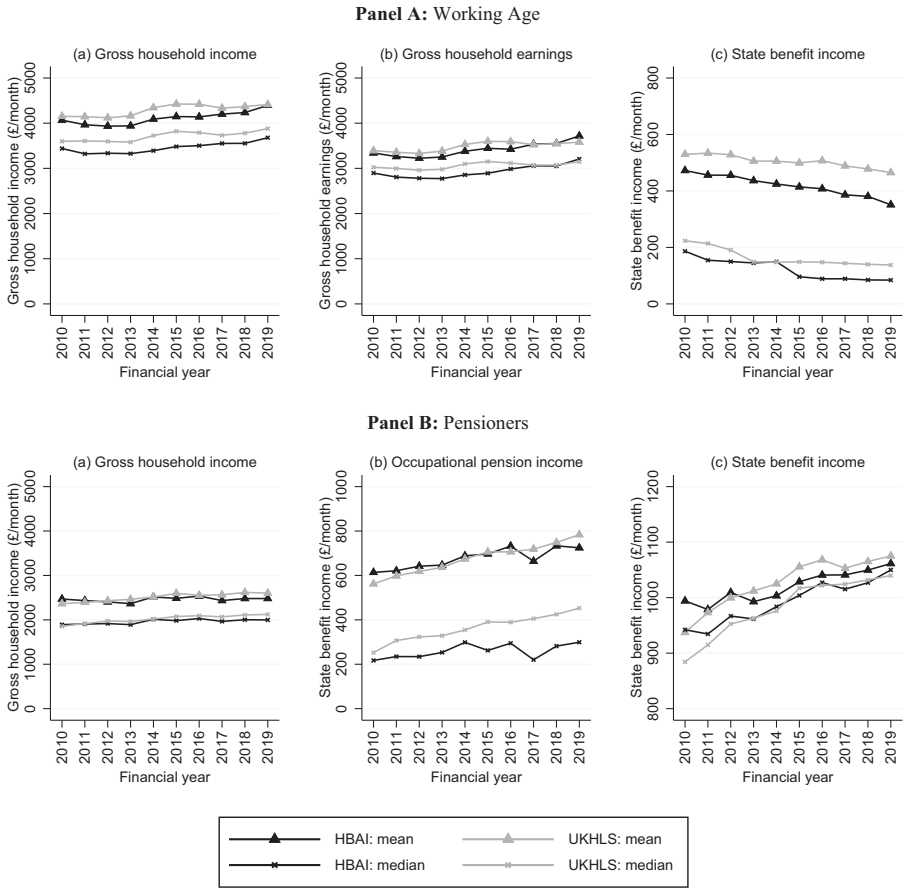

FIGURE 7 Gross household income components (mean and medians)

Note: Income is measured at the household level, equivalised using the OECD-modified scale, and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves. Estimation samples drop the richest and poorest 1 per cent of observations.
Source: Understanding Society and HBAI.

Household gross income as measured in each survey accords well, although Understanding Society tends to give slightly higher estimates of total income and more so for the working age. Estimates of earnings of the working age are similar in both datasets but those of benefit income are less similar (Panel A, Figure 7). In 2010, mean state benefit income in Understanding Society is £531 while it is only £472 in HBAI; the trends also diverge subsequently as HBAI estimates decrease more strongly.[54] Both surveys also give a similar understanding of income trends of pensioners (Panel B, Figure 7), which steadily increase across the period, but slightly more so for Understanding Society. One explanation for the divergence between the surveys is that they reflect the compositional differences of the two samples, as seen in Table 1.

To focus more closely on the (proportional) differences in the estimates and their evolution, we present the ratio of means (Understanding Society/HBAI) for each of the income components. Figure 8 shows that the largest ratio occurs for state benefit income of the working age. Understanding Society estimates start out proportionally larger than the HBAI ones (ratio of 1.12) and the ratio increases

[54] While it might be assumed that HBAI state benefit data are more reliable as they undergo extensive editing on a case-by-case basis, Understanding Society benefits from panel conditioning and dependent interviewing in the collection of its benefits data, both of which have been shown to increase reporting (Fisher, 2019).

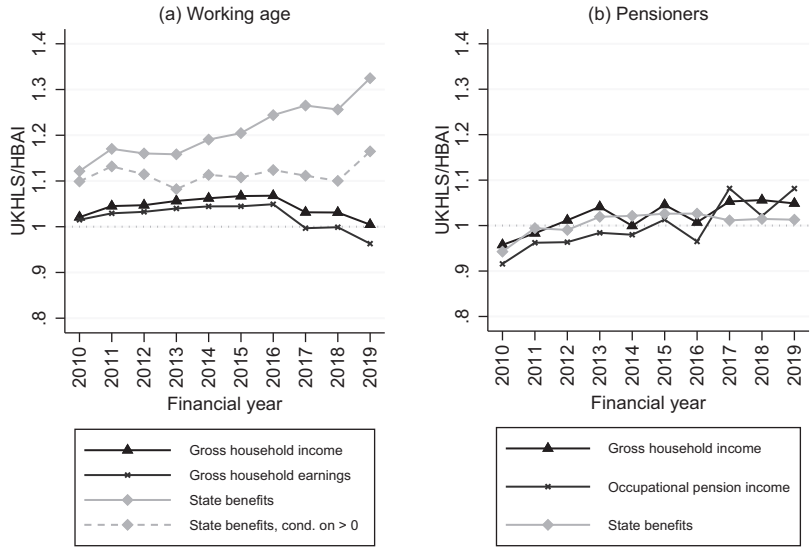

FIGURE 8 Gross household income components (ratio of means)

Note: Income is measured at the household level, equivalised using the OECD-modified scale, and net of taxes, National Insurance contributions and other deductions. All figures are expressed in 2015–16 prices using the bespoke monthly CPI produced by the ONS and used in the official UK income distribution statistics. Understanding Society financial years are constructed by pooling the relevant waves. Estimation samples drop the richest and poorest 1 per cent of observations.
Source: Understanding Society and HBAI.

quite strongly thereafter. However, when restricting to individuals in households with positive benefit income, the ratios are smaller and fairly stable over time. The share with zero benefit income is similar in both surveys at the start of the period (27 per cent (UKHLS) and 28.5 per cent (HBAI) in 2010) and increases over time, but more strongly for HBAI (35.7 per cent (UKHLS) and 43.4 per cent (HBAI) in 2019). For pensioners, a similar pattern is seen across their income components where Understanding Society estimates start out proportionally lower and steadily increase across the period to end up proportionally higher. Of the pensioner components, the ratios increase more modestly for household income (0.96 (2010); 1.05 (2019)) and state benefits (0.94 (2010); 1.01 (2019)) but more strongly for occupational pensions (0.92 (2010); 1.08 (2019)).

A caveat to the results of this section is that we have treated the data as a time-series of cross-sections, but of course many uses of panel data are longitudinal, requiring balanced samples. This raises the question of how our comparisons would fare for different (appropriately weighted) balanced samples or, put differently, how effective are survey weights in correcting for selection into such balanced samples. Here, we refer the reader to Lynn, Cabrera-Álvarez and Clarke (2023), who provide new evidence on the performance of the Understanding Society longitudinal weights in correcting for such selection.

This section has presented two independent attempts at measuring the UK income distribution: one derived from Understanding Society, and the other from the HBAI. Estimates from each data source have given a similar picture of household incomes in the UK and how they have evolved over the last decade. Where differences occur, we see that Understanding Society, relative to HBAI, tends to slightly overestimate incomes at the bottom of the distribution and underestimate them at the top. This means that Understanding Society tends to give slightly lower inequality levels than does HBAI, although the basic inequality trends align well. We have also shown that Understanding Society estimates marginally more pensioners; and more state benefit income with the gap, relative to HBAI, increasing over time.

5 | FUTURE DIRECTIONS

Understanding Society income data are widely used by the scientific and policy communities and so it is important to monitor its quality. We have detailed its approach to income data collection and processing, and shown that its income distributions compare well with a cross-sectional gold standard. In this section, we set out some future directions for income data collection of the study.

Our results lead us to emphasise the value of keeping the income data collection consistent as the panel ages. As the income series grows, it will enable the study of longer-term trends, in addition to short-term trends, which could span decades or be across generations. Questionnaire and fieldwork revisions can, where they change measurement properties of the data, undermine the long-run stability of the series.

That being said, where data quality improvements can be made with a high degree of confidence, they are desirable. We therefore emphasise the importance of testing and validations in the development of the income content. We see two main ways in which the scientific scope of the data could be increased: (1) developing a better understanding of measurement errors and also how they might be reduced in fieldwork; and (2) expanding the scope of the data through the collection of new variables or data linkages.

Burton et al. (2020) is one ongoing attempt to better understand reporting errors of Understanding Society respondents through the use of income summary screens during data collection. An income summary screen reflects back to survey respondents information they previously supplied by category and a total, and then allows them an opportunity to change erroneous answers. In a similar experiment, information is collected on the full household budgeting identity and respondents are invited to adjust their answers where they are inconsistent with each other.[55] While these new data collection tools show promise, further proof of concept is required before they can be rolled out as part of the main Understanding Society data collection.

Historically, it has proved problematic to link UK household surveys to administrative records, but the Longitudinal Linkage Collaboration of which Understanding Society is a part, has been set up to pursue exactly this goal. There are two main ways in which individual record linkage could be exploited to improve the estimates of household income. The first is simple replacement of the survey reports with equivalents from administrative data in cases where the administrative data are known to be more accurately measured and use the survey reports where they are higher quality. This approach likely works well when the administrative data cover the full population and do not suffer from non-coverage, so that reports for the full survey sample could be replaced. Types of government record that could be exploited for such individual-level linkages include HM Revenue and Customs records on tax self-assessments and Department for Work and Pensions records on state benefit receipt.

A second approach would use the administrative data to better process the survey data. The linked data would be used to understand the error properties of both data sources, which would then inform the editing and imputation of the survey reports. This approach could be effective when administrative data suffer from non-coverage and so replacing the survey reports with the administrative ones is not possible for the full survey sample or where restrictions limit access to the linked dataset (such as requiring special licences) or where definitional differences make direct substitution of the survey reports with the administrative ones problematic. One such small-scale linkage that Understanding Society has recently completed is to the National Employment Savings Trust (NEST) records.[56] This linkage covers a small subgroup of respondents who belong to a NEST pension scheme, and provides an opportunity to learn about errors in both survey and NEST reported earnings (as well as providing variables on pension membership for substantive research on pensions).

Such data combinations are not without their complications and future work should seek to understand how to perform them while addressing issues such as non-consent to linkage, definitional

[55] Brewer et al., 2018; Burton et al., 2017.

[56] Institute for Social and Economic Research, 2023.

mismatch (such as differing reference periods), and the presence of errors in both data sources. Work that seeks to better understand non-response patterns in the survey and the effectiveness of different imputation methods when using linked datasets should also be a priority for future work.

The above considered the case of individual-level linkages, but linkage to public datasets that occur at the address level should also prove fruitful. Understanding Society already links to public data on a properties council tax band, which when combined with Local Authority data on council tax rates, and demographics, allows for the accurate calculation of property council tax liability. Another publicly available dataset that could be linked to is house price (price paid) data held by the Valuations Office. Such information could supplement the housing data already collected by the survey to improve estimates of housing costs of homeowners.

We finish by noting that Understanding Society collects a broader range of financial information beyond income, including wealth, savings, debt, and subjective finances. Related types of economic data are also being collected in new data collection trials such as through the use of mobile apps[57] and 'event triggered data collection'. It remains for future work to document the quality of those data.

ACKNOWLEDGEMENTS

We thank James Banks, Michaela Benzeval, Pablo Cabrera-Álvarez, Monica Costa Dias and Tom Crossley for useful comments on this paper. Responsibility for any mistakes is the authors' alone. Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The Households Below Average Income and Understanding Society research data are distributed by the UK Data Service.

ORCID

Paul Fisher https://orcid.org/0000-0002-4066-1810
Omar Hussein https://orcid.org/0009-0002-3112-9029

REFERENCES

Al Baghal, T., & Lynn, P. (2015), Using motivational statements in web-instrument design to reduce item-missing rates in a mixed-mode context, *Public Opinion Quarterly*, 79, 568–79.

Andrew, A., Bandiera, O., Costa-Dias, M., & Landais, C. (2021a), Women and men at work, IFS Deaton Review of Inequalities, https://ifs.org.uk/inequality/women-and-men-at-work/.

Andrew, A., Bandiera, O., Dias, M. C., & Landais, C. (2021b), The careers and time use of mothers and fathers, IFS report, https://ifs.org.uk/publications/careers-and-time-use-mothers-and-fathers.

Atkinson, A. B. (2005), Top incomes in the UK over the 20th century, *Journal of the Royal Statistical Society Series A: Statistics in Society*, 168, 325–43.

Avram, S., Brewer, M., Fisher, P., & Fumagalli, L. (2022), Household earnings and income volatility in the UK, 2009–2017, *Journal of Economic Inequality*, 20, 345–69.

Baltruszewicz, M., Steinberger, J. K., Paavola, J., Ivanova, D., Brand-Correa, L. I., & Owen, A. (2023), Social outcomes of energy use in the United Kingdom: household energy footprints and their links to well-being, *Ecological Economics*, 205, 107686.

Barnes, M. (2022), Longer-term pensioner poverty and poverty transitions: a quantitative analysis of the Understanding Society survey (USoc) – final report, https://www.understandingsociety.ac.uk/research/publications/547124.

Bingley, P., & Martinello, A. (2017), Measurement error in income and schooling and the bias of linear estimators, *Journal of Labor Economics*, 35, 1117–48.

Borbely, D. (2022), The impact of housing subsidy cuts on the labour market outcomes of claimants: evidence from England, *Journal of Housing Economics*, 57, 101859.

Brewer, M., Burton, J., Crossley, T. F., Fisher, P., Gaia, A., Jäckle, A., & Winter, J. (2018), Another look at whether the rich save more: evidence from a survey experiment, prepared for the 2018 IARIW conference, Copenhagen, Denmark.

Brewer, M., Etheridge, B., & O'Dea, C. (2017), Why are households that report the lowest incomes so well-off? *The Economic Journal*, 127, F24–F49.

[57] Jäckle, Burton and Couper, 2023.

Brewer, M., Joyce, R., Waters, T., & Woods, J. (2020), A method for decomposing the impact of reforms on the long-run income distribution, with an application to universal credit, *Economics Letters*, 192, 109230.
Brewer, M., & Tasseva, I. V. (2021), Did the UK policy response to COVID-19 protect household incomes? *Journal of Economic Inequality*, 19, 433–58.
Brown, S., Gray, D., & Roberts, J. (2015), The relative income hypothesis: a comparison of methods, *Economics Letters*, 130, 47–50.
Burton, J., Chisholm, T., Coulter, A., Crossley, T. F., Fisher, P., Jäckle, A., Hasbrouck, H., & Winter, J. (2017), Understanding how people conceptualise household finances, Understanding Society Working Paper 2017-12, https://www.understandingsociety.ac.uk/research/publications/524691.
Burton, J., Crossley, T. F., Fisher, P., Gaia, A., & Jäckle, A. (2020), Understanding and reducing income reporting error in household surveys, Understanding Society Working Paper 2020-01, https://www.understandingsociety.ac.uk/research/publications/525973.
Carrieri, V., Davillas, A., & Jones, A. M. (2022), Equality of opportunity and the expansion of higher education in the UK, *Review of Income and Wealth*, forthcoming, https://doi.org/10.1111/roiw.12613.
Cebula, C., Collingwood, A., Earwaker, R., Elliott, J., Matejic, P., Taylor, I., & Wenham, A. (2023) UK poverty 2023: the essential guide to understanding poverty in the UK, Joseph Rowntree Foundation report, https://www.jrf.org.uk/report/uk-poverty-2023.
Cominetti, N., McCurdy, C., Thwaites, G., & Vieira-Marques, R. (2023), Low pay Britain 2023: improving low-paid work through higher minimum standards, Resolution Foundation report, https://economy2030.resolutionfoundation.org/reports/low-pay-britain-2023/.
Cribb, J., Karjalainen, H., & O'Brien, L. (2023), The gender gap in pension saving, IFS Report R250, Institute for Fiscal Studies, https://ifs.org.uk/publications/gender-gap-pension-saving.
Crossley, T. F., Fisher, P., & Hussein, O. (2023), Assessing data from summary questions about earnings and income, *Labour Economics*, 81, 102331.
Crossley, T. F., Fisher, P., & Low, H. (2021), The heterogeneous and regressive consequences of COVID-19: evidence from high quality panel data, *Journal of Public Economics*, 193, 104334.
Crossley, T. F., Fisher, P., Low, H., & Levell, P. (2022), A year of COVID: the evolution of labour market and financial inequalities through the crisis, *Oxford Economic Papers*, 75, 589–612.
Davillas, A., Burlinson, A., & Liu, H-H. (2022), Getting warmer: fuel poverty, objective and subjective health and well-being, *Energy Economics*, 106, 105794.
Davillas, A., & Jones, A. M. (2021), Unmet health care need and income-related horizontal equity in use of health care during the COVID-19 pandemic, *Health Economics*, 30, 1711–16.
Davillas, A., Jones, A. M., & Benzeval, M. (2019), Chapter 22 – The income-health gradient: evidence from self-reported health and biomarkers in Understanding Society, in M. Tsionas (ed.), *Panel Data Econometrics: Empirical Applications*, Elsevier, 709–41.
Department for Work and Pensions (2023a), Households Below Average Income, 1994/95–2021/22, [data collection], 17th Edition. UK Data Service. SN: 5828, DOI: http://doi.org/10.5255/UKDA-SN-5828-15.
Department for Work and Pensions (2023b), Income dynamics: income movements and the persistence of low income, 2010 to 2021, Technical report, Department for Work and Pensions, https://www.gov.uk/government/statistics/income-dynamics-2010-to-2021/income-dynamics-income-movements-and-the-persistence-of-low-income-2010-to-2021.
Di Nallo, A., & Lipps, O. (2023), How much his or her job loss influences fertility: a couple approach, *Journal of Marriage and Family*, 85, 873–97.
Fetzer, T. (2019), Did austerity cause Brexit? *American Economic Review*, 109 (11), 3849–86.
Fisher, P. (2019), Does repeated measurement improve income data quality? *Oxford Bulletin of Economics and Statistics*, 81, 989–1011.
Fisher, P., Fumagalli, L., Buck, N., & Avram, S. (2019), Understanding Society and its income data, Understanding Society Working Paper 2019-08, https://www.understandingsociety.ac.uk/research/publications/525786.
FitzRoy, F. R., & Nolan, M. A. (2020), Education, income and happiness: panel evidence for the UK, *Empirical Economics*, 58, 2573–92.
FitzRoy, F. R., & Nolan, M. A. (2022), Income status and life satisfaction, *Journal of Happiness Studies*, 23, 233–56.
Francesconi, M., Sutherland, H., & Zantomio, F. (2010), A comparison of earnings measures from longitudinal and cross-sectional surveys: evidence from the UK, *Journal of the Royal Statistical Society Series A: Statistics in Society*, 174, 297–326.
Garin, A., Jackson, E., & Koustas, D. (2022), New gig work or changes in reporting? Understanding self-employment trends in tax data, University of Chicago, Becker Friedman Institute for Economics Working Paper 2022–67.
Gibbons, S., Sanchez-Vidal, M., & Silva, O. (2020), The bedroom tax, *Regional Science and Urban Economics*, 82, 103418.
HM Treasury (2020), Impact of COVID-19 on working household incomes: distributional analysis as of May 2020, London: HM Treasury, https://www.gov.uk/government/publications/impact-of-covid-19-on-working-household-incomes-distributional-analysis-as-of-may-2020.
Institute for Social and Economic Research (2022), Understanding Society: Waves 1–12, 2009–2021 and harmonised BHPS: Waves 1–18, 1991–2009, User guide, Colchester: University of Essex.
Institute for Social and Economic Research (2023), Linked nest auto-enrolment pensions dataset, 2014–2022, [data collection]. UK Data Service. SN: 9127, DOI: http://doi.org/10.5255/UKDA-SN-9127-1.

Jäckle, A., Burton, J., & Couper, M. P. (2023), Understanding Society: minimising selection biases in mobile data collection, *Fiscal Studies*, 44.
Jantti, M., Törmälehto, V.-M., & Marlier, E. (2013), The use of registers in the context of EU–SILC: challenges and opportunities, Publications Office of the European Union, https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/KS-TC-13-004.
Jenkins, S. P. (2011), *Changing Fortunes: Income Mobility and Poverty Dynamics in Britain*, Oxford: Oxford University Press.
Johnston, D., Kung, C., & Shields, M. (2021), Who is resilient in a time of crisis? The importance of financial and non-financial resources, *Health Economics*, 30, 3051–73.
Kaiser, C. (2020), People do not adapt. new analyses of the dynamic effects of own and reference income on life satisfaction, *Journal of Economic Behavior & Organization*, 177, 494–513.
Kanabar, R. (2017), In or out? Poverty dynamics among older individuals in the UK, *Journal of Pension Economics & Finance*, 16, 509–53.
Katikireddi, S., Molaodi, O., Gibson, M., Dundas, R., & Craig, P. (2017), OP76: evaluating the health impacts of restrictions to income support for lone parents: a natural experiment study using Understanding Society, *Journal of Epidemiology and Community Health*, 71, A38.
Katikireddi, S. V., Molaodi, O. R., Gibson, M., Dundas, R., & Craig, P. (2018), Effects of restrictions to income support on health of lone mothers in the UK: a natural experiment study, *The Lancet Public Health*, 3, e333–40.
Larrimore, J., Mortenson, J., & Splinter, D. (2021), Household incomes in tax data: using addresses to move from tax-unit to household income distributions, *Journal of Human Resources*, 56, 600–631.
Levell, P., & Sturrock, D. (2023), Using Understanding Society to study intergenerational mobility in the UK, *Fiscal Studies*, 44, (this issue), https://doi.org/10.1111/1475-5890.12352.
Little, R., & Su, H. (1989), Item nonresponse in panel surveys, in D. Kasprzyk et al. (eds), *Panel Surveys*, New York: Wiley.
Lynn, P., Cabrera-Álvarez, P., & Clarke, P. (2023), Sample composition and representativeness on Understanding Society, *Fiscal Studies*, 44, (this issue), https://doi.org/10.1111/1475-5890.12357.
Lynn, P., Jäckle, A., Jenkins, S. P., & Sala, E. (2012), The impact of questioning method on measurement error in panel survey measures of benefit receipt: evidence from a validation study, *Journal of the Royal Statistical Society Series A: Statistics in Society*, 175, 289–308.
Main, G. (2019), Money matters: a nuanced approach to understanding the relationship between household income and child subjective well-being, *Child Indicators Research*, 12, 1125–45.
Mari, G., & Keizer, R. (2023), Families of austerity: benefit cutbacks and family stress in the UK, Mimeo.
Meyer, B. D., & Mittag, N. (2019), Using linked survey and administrative data to better measure income: implications for poverty, program effectiveness, and holes in the safety net, *American Economic Journal: Applied Economics*, 11, 176–204.
Micklewright, J., & Schnepf, S. V. (2010), How reliable are income data collected with a single question? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173, 409–29.
Parra-Mujica, F., Johnson, E., Reed, H., Cookson, R., & Johnson, M. (2023), Understanding the relationship between income and mental health among 16- to 24-year-olds: analysis of 10 waves (2009–2020) of Understanding Society to enable modelling of income interventions, *Plos One*, 18, e0279845.
Perez Hernandez, V., Kanabar, R., & Nandi, A. (2018), Low income dynamics among ethnic minorities in Great Britain, ISER Working Paper Series 2018-06, https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2018-06.
Piketty, T., & Saez, E. (2003), Income inequality in the United States, 1913–1998, *Quarterly Journal of Economics*, 118, 1–41.
Postel-Vinay, F., & Sepahsalari, A. (2023), Labour mobility and earnings in the UK, 1992–2017, *Economic Journal*, 133 (656), 3071–98.
Ray-Chaudhuri, S., Waters, T., Wernham, T., & Xu, X. (2023), Living standards, poverty and inequality in the UK: 2023, IFS Report R265, London: Institute for Fiscal Studies, https://ifs.org.uk/publications/living-standards-poverty-and-inequality-uk-2023.
Rohenkohl, B. (2019), Intergenerational income mobility in the UK: new evidence using the BHPS and Understanding Society, Sheffield Economic Research Paper Series (SERPS), No. 2019017.
Social Metrics Commission (2020), Measuring poverty 2020, Report, London: Social Metrics Commission.
Thomas, C., Jung, C., Statham, R., & Quilter-Pinner, H. (2023), Healthy people, prosperous lives, Interim report of the IPPR, Commission on Health and Prosperity, IPPR.
United Nations (2011), *Canberra Group Handbook on Household Income Statistics*, 2nd edn, New York and Geneva: United Nations.
Wilhelm, D. (2018), Testing for the presence of measurement error, Technical report, cemmap working paper.

How to cite this article: Fisher, P., & Hussein, O. (2023), Understanding Society: the income data. *Fiscal Studies*, 44, 377–97. https://doi.org/10.1111/1475-5890.12353