

Fisher, Paul

**Working Paper**

## Does repeated measurement improve income data quality?

ISER Working Paper Series, No. 2016-11

**Provided in Cooperation with:**

Institute for Social and Economic Research (ISER), University of Essex

*Suggested Citation:* Fisher, Paul (2016) : Does repeated measurement improve income data quality?, ISER Working Paper Series, No. 2016-11, University of Essex, Institute for Social and Economic Research (ISER), Colchester

This Version is available at:

<https://hdl.handle.net/10419/163537>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Does repeated measurement improve income data quality?



**Paul Fisher**

Institute for Social and Economic Research  
University of Essex

No. 2016-11  
October 2016

INSTITUTE FOR SOCIAL  
& ECONOMIC RESEARCH

## **Non-technical summary**

Social scientists studying material living standards commonly analyse data from surveys that measure the income of the same individual at multiple points in time (panel data) e.g. large-scale household panel surveys or panel data collected as part of field experiments. It is well known that respondents to a survey tend to under-report their income, but researchers commonly assume that the amount of under-reporting is stable across waves of a panel. This assumption is critical, as if reporting behaviour changes, then estimates of how material living standards evolve over time would be confounded with the changes in reporting behaviour and give a misleading picture of changes in real living standards. In this paper we provide evidence on how the underreporting of income in a large scale household panel survey changes across waves.

Our main finding is that repeated interviewing results in better quality income data. The improvements in quality are large and are strongest for unearned income sources (largely pensions and state benefits) and take effect in the initial waves of the panel. Our estimates indicate that the effect of being interviewed for a second time is to increase the mean of reported monthly income by a substantial £142 (8 percent).

As to why these effects occur, dependent interviewing (DI) – a common recall device used in household panel surveys that reminds respondents of their reports at an earlier interview – takes effect only after a first survey interview. It can explain approximately one third of the observed increase in reported income. On the remaining two thirds, given the use of the same interviewers, infrastructure and questionnaires at both waves, this points to changes in the reporting behaviour of survey respondents. Indeed, we present evidence suggestive of a reduction in respondent confidentiality concerns following the first interview. This is backed up with an examination of refusal rates on the income variables which fall off most sharply between waves one and two. We confirm our results by presenting evidence of similar effects in another leading panel survey (British Household Panel Survey). More generally they could be expected to extend to other sensitive areas of data collection.

If our evidence is interpreted as reducing under-reporting of income, it suggests that income data provided as part of panel surveys offers some quality advantages over that collected from cross-sectional surveys deriving both from the use of DI and also through being able to improve respondent cooperation through repeated measurement.

# Does repeated measurement improve income data quality?\*

Paul Fisher<sup>†</sup>

Institute for Social and Economic Research, University of Essex

October 2016

## Abstract

This paper presents evidence that the quality of survey data on household incomes systematically improves across waves of a panel. Our estimates indicate that the effect of being interviewed for a second time is to increase the mean of reported monthly income by £142 (8 percent). Dependent interviewing - a recall device commonly used in panel surveys - takes effect only after a first interview. It explains approximately one third of the observed increase. The remaining share is attributed to changes in respondent reporting behaviour (panel conditioning). Our analysis suggests that falls in respondent confidentiality concerns are important in explaining the result.

Keywords: income, measurement error, household survey, panel conditioning

---

\*I would like to thank Jonathan Burton; Mike Brewer; Nick Buck; Brian Bucks; Emanuele Ciani; Tom Crossley; Annette Jackle; Peter Lynn and Steve Pudney for their useful comments. I also thank seminar participants at the 'Improving the measurement of household finances' workshop, University of Essex; and the International Association for Research in Income and Wealth conference 2016, Dresden, Germany. Financial support received from the ESRC is also gratefully acknowledged, award numbers ES/K005146/1, ES/N00812X/1 and ES/N006534/1. Data from the Understanding Society Survey has been made available through the UK Data Archive and has been used with permission.

<sup>†</sup>Address: Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. e-mail: [pfishe@essex.ac.uk](mailto:pfishe@essex.ac.uk).

# 1 Introduction

Social scientists interested in material living standards commonly analyse data from surveys that measure the income of the same individual at multiple points in time e.g. large-scale household panel surveys or panel data collected as part of field experiments. It is known that state transfers (Meyer and Sullivan (2003, 2011); Lynn et al. (2012a); Brewer, Etheridge, and O’Dea (forthcoming)) and self-employment income (Hurst, Li, and Pugsley (2014)) are under-reported in household surveys. Less is known about whether measurement error is on average stable across waves of a panel. In this paper we provide evidence that the quality of measured income systematically improves across the early waves of a leading panel survey. Such improvements in income data quality represent a benefit of repeated interviewing. However, estimates of distributional change based on the early waves of the panel will confound true changes with data quality changes and be biased.

There are two reasons to think that reported income may not be comparable across waves of a panel. First, panel conditioning (PC) effects may operate where panel participants change their behaviour as a result of being part of the panel. PC improves data quality if it reflects a respondents improved understanding of the questionnaire content or a growing trust in the interviewer or data holders. PC reduces data quality if respondents learn to strategically answer questions with the aim of reducing the interview length. Related, the data will become unrepresentative if survey participation leads to changes in real behaviour (see Crossley et al. (forthcoming)). Second, dependent interviewing (DI) - a tool that reminds survey respondents of their reports at the previous interview - will lead to differences in data quality between the baseline and subsequent interviews where it takes effect.<sup>1</sup>

Only a few studies have examined the stability of measurement error in income across waves of a panel. David and Bollinger (2005) find that false negative reporting of US food

---

<sup>1</sup>Jenkins (2011) notes that when DI was introduced in the British Household Panel Survey there were “no obvious discontinuities in income series”. A third reason is non-random attrition but we address this directly in the identification section.

stamps is highly correlated across wave one and two of the Survey of Income and Program Participation. Das, Toepoel, and van Soest (2011) propose a methodology to quantify PC effects, by comparing responses from first-time responders in refreshment samples to more experienced panel members and making assumptions on the attrition process.<sup>2</sup> In this spirit, Halpern-Manners, Warren, and Torche (2014) note that experienced panel members in the US General Social Survey are less likely to refuse to answer questions about their income. Similarly, Frick et al. (2006) find that experienced panel members report higher income in the German Socio-Economic Panel. Nevertheless, despite the large interest of social scientists in living standards, we know of no study that has performed a systematic analysis of how measurement error in income and its components evolves across waves of a panel.

In this study we provide evidence on the comparability of reported income across the waves of a large general purpose panel survey: the UK Household Longitudinal Study. The novelty of our approach is that it exploits two features of the survey design to separate changes in reporting behaviour from true income changes (and does not require data linkage or refreshment samples). First, the fieldwork period for adjacent waves overlaps by one year, giving us random samples of individuals being interviewed at different waves of the panel but in the same calendar year, removing time effects. Second, UKHLS uses reactive DI meaning that we can observe exactly which individuals would have failed to report an income source in its absence. The latter allows us to decompose the observed data quality changes into shares due to PC and DI.

Our approach offers several advantages over other studies exploring measurement error in income data that have used small scale administrative data linkage or refreshment samples (eg. Lynn et al. (2012a); David and Bollinger (2005)). First, the large sample sizes available mean we can precisely estimate the effects of interest. Second, our data are representative of the Great Britain population (and not subsamples such as the poor or

---

<sup>2</sup>Taking this approach, Van Landeghem (2014) finds a drop in a stated utility measure across the first-rounds of interviews in two panel surveys.

individuals covered by tax records) meaning that we can study how effects vary by representative subgroups of interest such as pensioners, working age groups and families. Third, our analysis covers a comprehensive set of income sources including earnings, investment income, and a total of 39 unearned sources. This enables us to identify precisely which income sources are most sensitive to prior panel participation.

Our main finding is that repeated interviewing improves the quality of income data (in particular unearned income sources and state transfers) and this occurs across the initial waves of the panel. Being interviewed for a second time, relative to the first, increases reported gross monthly income by £142 or 8 percent and about one third of the difference can be explained by DI, with the other two thirds due to PC. As to why these effects occur, we are able to rule out interviewer and fieldwork agency learning as explanations. Given the use of the same infrastructure and questionnaires at both waves, this points to changes in the reporting behaviour of survey respondents. Indeed, we present evidence showing reductions in respondent confidentiality concerns and falls in income refusal rates following a first interview.

On the broader implications of our results, we present evidence suggestive of similar effects in another leading panel survey (British Household Panel Survey) and more generally they might be expected to extend to other sensitive areas of data collection. This suggests that income data collected as part of panel surveys offers some quality advantages over similar cross-sectional data - deriving both from the use of DI and also through being able to improve respondent cooperation through repeated measurement.

The paper proceeds as follows: the next section describes the data and compares our estimates of the UK income distribution to those from official cross-sectional sources. Section 3 discuss the identification strategy and the empirical results are presented in section 4. Section 5 explores the mechanisms behind our results and section 6 concludes. A supplementary appendix includes additional materials.

## 2 Data

### 2.1 UKHLS data

This paper makes use of data from the UK Household Longitudinal Study (UKHLS) that began in 2009. UKHLS is a large general purpose social survey that, in addition to income, collects information on a wide range of topics including: education, health, happiness, household organisation and objectively measured bio-markers. It will replace the former BHPS as the data source for official UK Government statistics on poverty dynamics. We work with the main ‘General Population Sample Great Britain’ sub-sample (GPSGB), which is representative of the Great Britain household population at wave one.

Due to the large samples, the fieldwork for each wave is spread out equally over a two year period. At wave 1, selected sampling units were randomly allocated across the 24 month interview period. All adult members (aged 16 or over) in households selected to be part of the panel at wave one are interviewed annually.<sup>3</sup> Together, these features imply an overlap of one year in the fieldwork period of consecutive waves. This overlap is an issue we return to and exploit in identification.

At wave one, interviews were conducted in 24,797 households with 41,586 individuals receiving an interview. As with all household panel surveys, there is an initial drop-off in individual response rates and 75.4 percent of wave one respondents completed an interview at wave two with a further 1.9 percent completing a proxy interview.<sup>4</sup>

### 2.2 Income variables and dependent interviewing

UKHLS includes a detailed set of questions on income receipt. These are collected for each sample member in a face-to-face interview that is conducted by computer-assisted personal interviewing (CAPI). An exact list of the income questions used in analysis is included in

---

<sup>3</sup>It is the issue date which is fixed at 12 monthly intervals and not the actual date of interview. The two may differ by some months in order to maximise the chances of an individual response.

<sup>4</sup>Response rates were lower for: those under 30, in urban locations, being a renter and expecting to move home at wave 1 (Lynn et al. (2012b)).



appendix B.

Data collection of the income components occurs in different modules. An employees module asks for gross (and then net) pay at last payment and then the usual pay if last and usual differed. A self-employment module asks self-employees for their share of the profit or loss figure on their most recent accounts or, where this is not available, for an estimate of their usual monthly or weekly self-employment income. All respondents receive the household finances module that records the amount received in interest and dividends in the last 12 months.

An ‘unearned income and state benefits’ module uses DI. It first asks respondents to examine a list of 9 broad types of payment and indicate which they are currently receiving. Respondents are then filtered to lists of specific sources (up to 39) where they indicate those that they receive.<sup>5</sup> DI is used to check whether any sources not reported but reported at wave t-1 are currently received. A final stage asks for the amounts for each reported source, the period it covered and whether the income was received solely or jointly.

One dimension of data quality is the extent to which respondents refuse to answer a question. Figure 1 plots trends in refusal rates (refusing to provide an amount), for a sample of respondents who completed a full-interview at each of waves one (2009-10) to five (2014-15).<sup>6</sup> The refusal rates fall across all sources as the panel ages indicating improvements in data quality. The biggest drop occurs for self-employment profit which starts at 42.0 per cent in wave 1 and reaches a minimum over the five waves of 34.2 per cent at wave 4. The drop off in refusal rates is notably sharper for all sources between waves 1 and waves 2. For example, non-response for self-employment profit falls from 42.0 per cent to 37.1 per cent and the earnings refusal rate from 12.0 per cent to 10.3 per cent.<sup>7</sup>

In our main analysis, we replace missing values with the standard longitudinal imputes

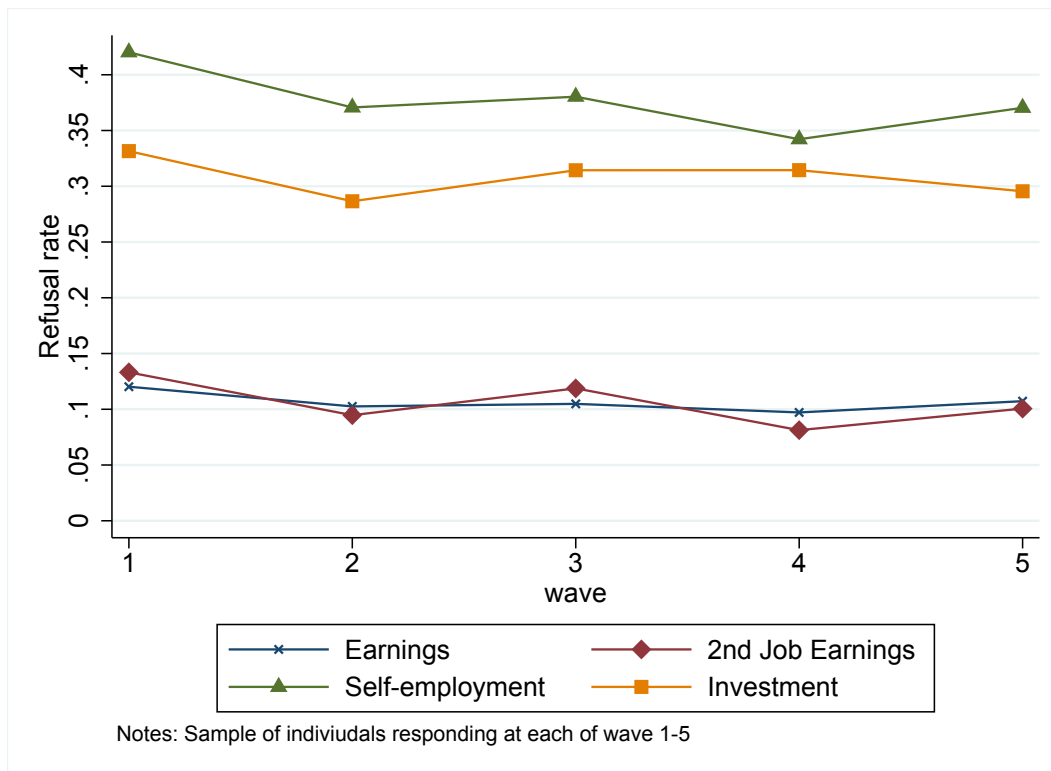
---

<sup>5</sup>Respondents meeting certain criteria are automatically prompted with the specific benefit showcards eg. those of retirement age are shown the pensions showcard; the long-term sick are shown the ‘disability benefits’ showcard.

<sup>6</sup>Refusals are counted as ‘refusals’ + ‘don’t knows’ where a ‘don’t know’ could be a polite refusal.

<sup>7</sup>The refusal rates rates for ‘unearned income and state benefits’ are low. For example, refusal rates on the pension showcard for our balanced sample are: 0.007, 0.003, 0.004, 0.003, and 0.004.

Figure 1: Income refusal rates by wave



produced by the data providers. There can be two other types of missing income data: missing an individual interview (unit non-response) and missing an individual interview but agrees that a proxy answers a shorter interview on their behalf. We address these issues directly in the methodology section.

### 2.3 Comparison to a cross-sectional income survey

In this section we look for evidence of reporting changes by comparing the UKHLS income distribution to that from a cross-sectional income survey and examining how any differences evolve over time. The cross-sectional survey we use is the Family Resources Survey – the source for UK official statistics on the income distribution.<sup>8</sup> As the FRS is

<sup>8</sup>The FRS is a purpose built income survey that collects information from a random sample of approximately 20,000 households each year. Our analysis is based on the ‘Households Below Average Income’ data sets which are produced by the Department for Work and Pensions as the basis for official UK statistics on the income distribution. The HBAI data-set includes income variables that have been adjusted to better measure top incomes. In order that our data sources are comparable, we use the unadjusted HBAI variables.

a specialist income survey, and UKHLS a general purpose survey, the FRS is expected to provide better coverage of income.<sup>9</sup> Crucially though, if repeated measurement leads to reporting improvements, the difference is expected to reduce as UKHLS ages.

Estimates of selected quantiles of the income distribution from the two surveys are shown in figure 2. Focusing on the lower half of the distribution, whilst there is a clear similarity between the estimates the difference between them is changing over-time. The strongest differences occur between waves 1 and 2, followed by waves 2 and 3. For example, at wave 1 the FRS gives higher estimates of incomes for the 1st, 5th, 10th and 25th percentiles but by wave 3 the pattern has reversed. The top half of the panel shows the median, 75th, 95th and 99th percentiles. The estimates line-up remarkably closely.<sup>10</sup>

Given the strongest convergence occurs between wave one and two, in our main analysis we concentrate on reporting behaviour at these waves.

### 3 Methodology

We construct a sample of individuals allocated for interview in 2010 but who are responding at different waves of the panel. Using a treatment/control terminology, the wave one 2010 sub-sample forms the control group (interviewed for the first time) and a different group of individuals responding to wave two in 2010 (interviewed for the second time) forms the treated group. With non-random attrition, there will be compositional differences between the groups that would bias our comparison. We remove the differences by restricting the analysis to a balanced sample of respondents who completed a full-interview at both waves one and two, under the assumption that the wave 2 interview outcome is statistically independent of the wave one survey year allocation, an assumption we return to below.<sup>11</sup>

We perform ordinary least squares regressions of our income components of interest

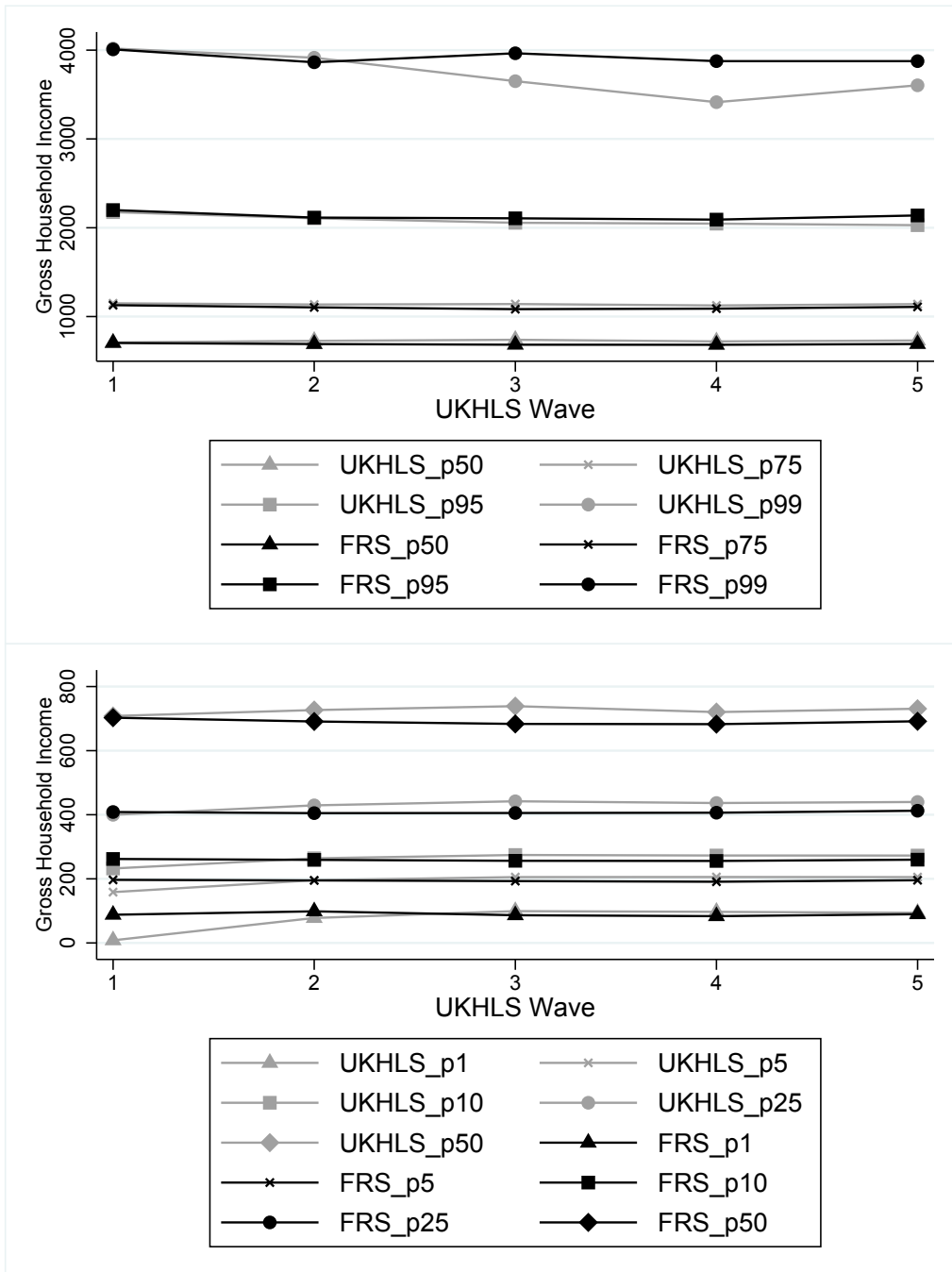
---

<sup>9</sup>The FRS tries to reduce under-reporting, for example, through detailed interviewer training and extensive checking of respondent documentation during an interview.

<sup>10</sup>Estimates of the 99th percentiles are diverging from wave 3 onwards but measurement error is known to be large for the very highest incomes in household surveys (see for example Bricker et al. (2015)).

<sup>11</sup>This implies that individuals who had a proxy interview at either wave one or two are excluded.

Figure 2: Selected quantiles of UKHLS/FRS (2009/10-2013/14)



Notes: The FRS corresponds to a financial year (April to March) and a UKHLS wave to two calendar years. To account for differences in the fieldwork period of the two surveys, we pool two consecutive FRS data sets when comparing to a single UKHLS wave. All figures are expressed in 2014-15 prices using the bespoke monthly CPI price index used in the official UK income statistics and produced by the Office for National Statistics.

(earnings, benefits and unearned income, and investment income) on an indicator for being a wave 2 respondent, with the coefficient on this indicator giving our estimate of the wave 2 reporting difference on the mean. To estimate the reporting effect net of DI, we set to zero any income source for which an individual received a DI reminder and then re-estimating our main coefficient of interest.

There are three threats to validity of this design that would lead to compositional differences between the treatment and control groups and so bias our results. First, an inequivalence between the 2009 and 2010 wave 1 responding samples. Second, a time effect in attrition making the decision to attrite statistically dependent on the initial survey year allocation. Third, the ageing effect of the panel resulting in individuals responding for the second time in 2010 being one year older than those responding for the first time. We directly tested violations one and two and found little evidence to suggest they are a concern.<sup>12</sup> Moreover, we include a wide range of control variables in our regression models - including a full set of age dummies. The fact that UKHLS covers a large number of topics, including determinants of income, makes this strategy persuasive. Given that the control variables are potentially also subject to panel conditioning, we focus on controls with low item non-response rates and that we judge are unlikely to be sensitive areas of questioning. The full list controls is given in the footnote to figure 3. Formally, we estimate:

$$Y_i = \alpha + \beta_1 wave\_2_i + \beta_3 X_i + \epsilon_i \quad (1)$$

where  $Y_i$  is the income component of interest,  $wave\_2_i$  an indicator variable taking the value 1 for wave 2 respondents,  $X_i$  a vector of controls and  $\epsilon_i$  an error term with  $E[\epsilon_i | wave\_2, X] = 0$ .  $\beta_1$  is our coefficient of interest, estimates of which are presented in the results. We report standard errors robust to heteroskedasticity.<sup>13</sup>

---

<sup>12</sup>We estimated an attrition model including interaction terms with sample year. None of the interactions were statistically significant. We also compared sample means of the 2009 and 2010 wave 1 samples. Full results are included in table A1, appendix A.

<sup>13</sup>We also tried clustering standard errors at the level of the Primary Sampling Unit but it made little difference to the estimated standard errors.

The validity of our identification strategy can be checked by comparing sample means of our treatment and control groups in their baseline (wave one) interview to assess differences in composition. The large sample makes it possible to detect even small differences where they occur. On demographics, the two subsamples are balanced in terms of sex, age, ethnicity (half a percentage point differences in the share of Indian and Chinese), qualifications and marital status (although nearly 1ppt difference in the share that are single and never married). For household composition, we see no differences in the mean number employed, of working age, or number of couples or single parents in a household. There are small differences in the mean number of people, children, and the age of the youngest child. Larger statistically significant differences are seen for living in social housing (1.6ppt) and having children (1.9 ppt) but not for the shares that own, rent or mortgage their home or number of bedrooms. Overall we interpret the comparison as indicating that the samples are well balanced and where there are small differences, as stated above, they are accounted for in our regression models. Full results can be found in table A2, appendix A.

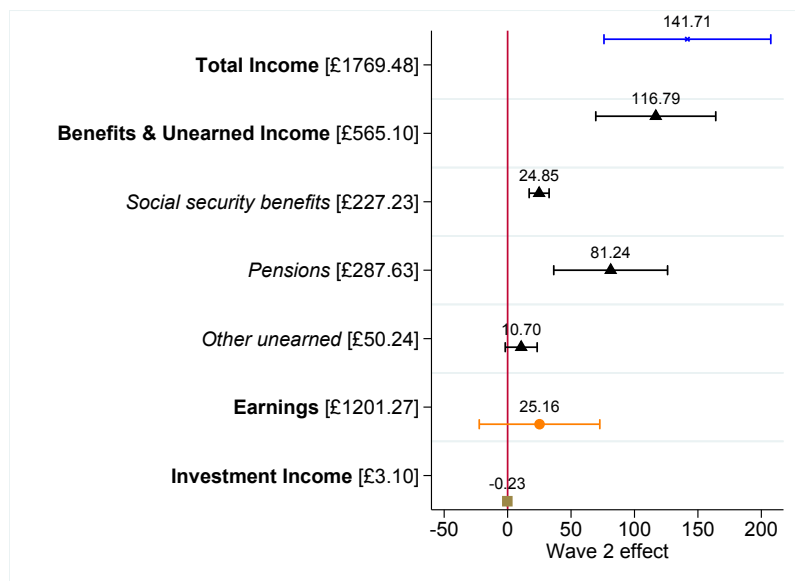
## 4 Results

### 4.1 Differences in reporting between waves 1 and 2

Figure 3 presents estimates of  $\beta_1$  from equation (1). The figure shows results from models estimated separately for total: income; benefits and unearned income (social security benefits; pensions; and other unearned income); earnings; and investment income. For details of the main pension types and social security benefits see table 1.

Second-time responders reported a total monthly income that was £141.71 or 8 percent higher than first-time reporters and this represents a causal effect of being interviewed for a second time. Decomposing the effect shows it is driven by the benefits and unearned income sub-component. In particular, the strongest effects occur for social security benefits and pensions, where we see statistically significant increases in reported income of £24.85

Figure 3: Effect of second interview on reported income



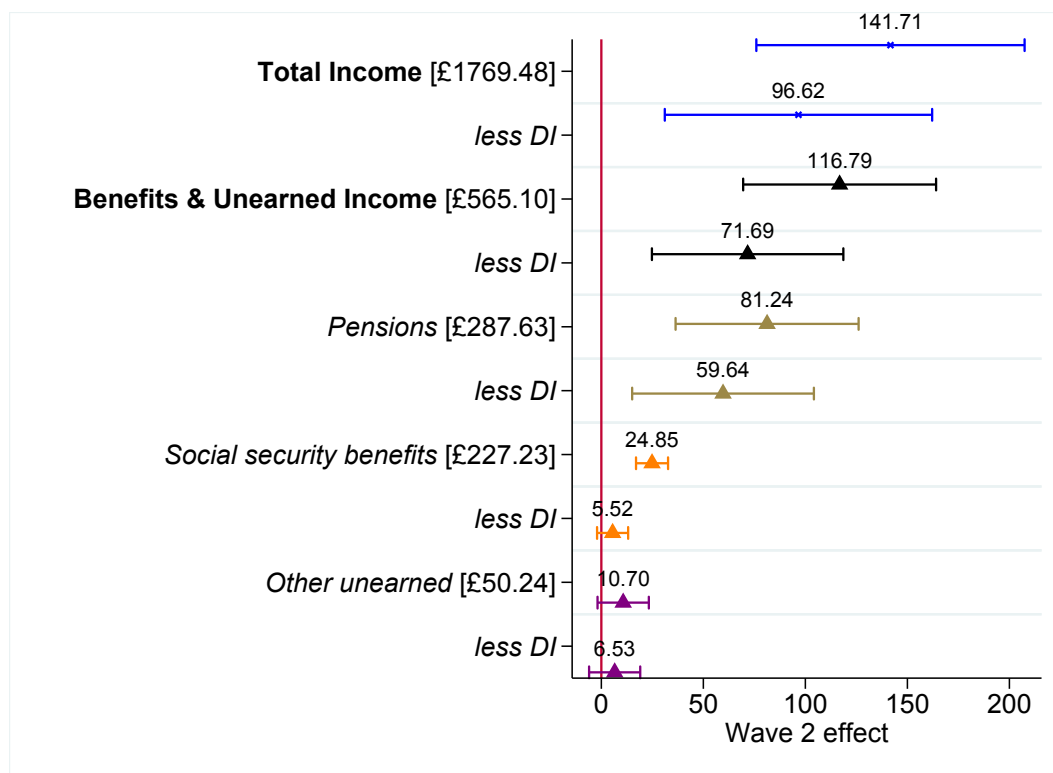
Notes: The point estimates presented correspond to  $\beta_1$  from equation 1 with the full-set of controls. Means from the baseline (wave 1) interview are reported in square brackets. Confidence intervals are calculated using robust standard errors. The controls are dummy variables for (number of categories in parenthesis): sex, age in one year bands, ethnicity (6), highest qualification (6), retired, student, relationship status (5), housing type (4), long-standing illness, household size (16), number of children (11), region (12) and interview month (12).

(10.9 percent) and £81.24 (28.2 percent), respectively. These numbers imply substantial differences in the quality of reported data across the first two waves of the panel. Failure to account for this reporting difference would give a highly misleading picture of real changes in the income distribution over the period.

We would like to examine the extent to which the change in reporting behaviour of wave two respondents is due to panel conditioning. Figure 4 presents results from re-estimating equation (1) but by setting a reported wave 2 amount to zero where a dependent interviewing prompt was triggered. As expected, the estimates fall in size but surprisingly they remain large and statistically significant. For example, the wave 2 effect on total income falls by around a third from £141.71 to £96.62. That is, wave 2 respondents report a total income which is on average 5.5 percent higher per month and this effect represents a change in reporting behaviour not due to dependent interviewing.

We also explored the possibility of heterogenous treatment effects by estimating mod-

Figure 4: Panel Conditioning effect of second interview on reported income



Notes: see figure 3 notes.

els separately for subsamples of: pensioners, working age with children and working age without children.<sup>14</sup> In the interests of space, we only briefly review the results here. The effects are strongest for the pensioner subsample and are concentrated in the ‘benefits and unearned income’ component of income. The wave 2 effect is to increase reporting of this category by a large 24.1 percent. Moreover, 73.8 percent of this reported increase is due to PC and not DI. For the ‘working age without children’ subsample, the effects are weaker in absolute value but are proportionally large. Benefits and unearned income increase by 35.9 percent of the wave 1 mean and 57.5 percent is due to PC and not DI. Finally, for the ‘working age with children’ subsample, the effects are smaller and statistically significant only for the total effect in ‘benefits and unearned income’ (7.4 percent of the wave 1 mean). The interested reader can find the full set of results in figures A1-A6, appendix A.

We also explored the extent to which imputation can reduce the data quality differences

<sup>14</sup>We define ‘pensioners’ as those of UK state pension age (60 for women and 65 for men) and ‘Working age’ those below it.



across waves. Details are provided in section A2 of appendix A.

An important question is whether the differences in reporting of benefits and unearned income across waves 1 and 2 of the panel are due to changes in reporting of receipt or changes in the amounts reported. Table 1 explores this matter by presenting estimates of equation (1) separately for each of 12 of the most widely received pensions and social security benefits in the data. Columns 1 and 2 refer to receipt, and 3 and 4 to the (conditional) reported amounts. The odd columns show the total effect (PC + DI), and the even, the panel conditioning effects only.

Table 1: Effect of wave 2 interview on reporting of selected unearned sources

	<b>Receipt</b>		<b>Amount</b>	
	DI + PC	PC	DI + PC	PC
<b>Pensions</b>				
State pension <sup>†</sup> [23.62, £528.24]	1.327*** (0.204)	0.477* (0.214)	6.289 (4.091)	6.696 (4.127)
Private pension [5.00, £455.50]	1.932*** (0.264)	0.883*** (0.255)	90.497 (61.799)	109.403 (70.711)
Employer pension [14.84, £822.87]	1.497*** (0.337)	0.381 (0.335)	279.624* (121.585)	274.081* (126.274)
Spouse employer pension [2.76, £649.47]	0.666*** (0.175)	0.316 (0.171)	-2.274 (221.224)	36.725 (238.808)
<b>Social security benefits:</b>				
<b>Family benefits</b>				
Working Tax Credit [6.33, £254.43]	1.154*** (0.278)	0.598* (0.274)	-0.105 (8.499)	-3.016 (8.616)
Child benefit [24.04, £145.75]	0.535 (0.312)	0.078 (0.312)	-0.563 (1.020)	-0.664 (1.014)
Child tax credit [17.68, £297.48]	1.108*** (0.337)	0.295 (0.336)	-3.796 (5.870)	-2.424 (5.921)
<b>Disability benefits</b>				
Incapacity benefit [2.71, £409.44]	0.805*** (0.193)	0.453* (0.188)	5.199 (10.554)	3.205 (10.786)
Disability living allowance [4.96, £305.65]	1.497*** (0.256)	0.853*** (0.249)	-12.521 (8.455)	-16.577 (8.539)
<b>Low income benefits</b>				
Income support [4.81, £355.62]	0.334 (0.231)	0.078 (0.228)	-1.840 (12.003)	-1.548 (12.167)
Housing benefit [10.57, £372.02]	1.058*** (0.288)	-0.119 (0.286)	-4.354 (6.238)	-3.370 (6.455)
Council tax benefit [12.10, £101.06]	1.969*** (0.332)	0.356 (0.327)	3.173 (4.542)	2.424 (4.720)

N = 29528

Notes: see figure 3 notes.

<sup>†</sup> The state pension is a part contribution based benefit. It is paid from age 65 for men. For women, it is in the process of being increased from 60 to 65.

We observe increases in reporting of all of the 12 unearned income sources as a result of being interviewed for a second time. The effects are statistically significant with the exception of income support and the magnitudes of the effects are non-trivial. For example, the effect for the state pension<sup>15</sup> is to increase reporting by 1.31 percentage points or 5.6 percent of the wave 1 mean. Column (2) shows that a sizeable share of the observed pattern is attributable to panel conditioning and the panel conditioning effects are concentrated in disability benefits and pensions, although the effect for Working Tax Credit is also statistically significant. For example, the panel conditioning effect for the state pension is nearly 0.4 percentage points or 2.0 percent of the wave 1 mean.<sup>16</sup>

Interestingly, these results are in contrast with David and Bollinger (2005) who found that false negative reporting of food stamp receipt in the US Survey of Income and Program Participation was stable across waves. One possibility is that a lack of statistical power in their study made it difficult to detect small reporting changes where they occurred.

Moving to the reported amounts in columns 3, we see that 11/12 of the estimated coefficients are statistically insignificant, with the exception of employer pensions. Column 4 confirms this finding when estimating the PC effect only. Put together, the results of this section tell us that it is the receipt of unearned sources that changes with panel experience but with no indication of increases in the reported amounts. Combining the facts that state transfers are known to be under-reported in survey data and item non-response rates drop-off at wave two, suggests that the observed changes can be interpreted as improvements in data quality as the panel ages.

## 4.2 Is reporting behaviour stable after the second interview?

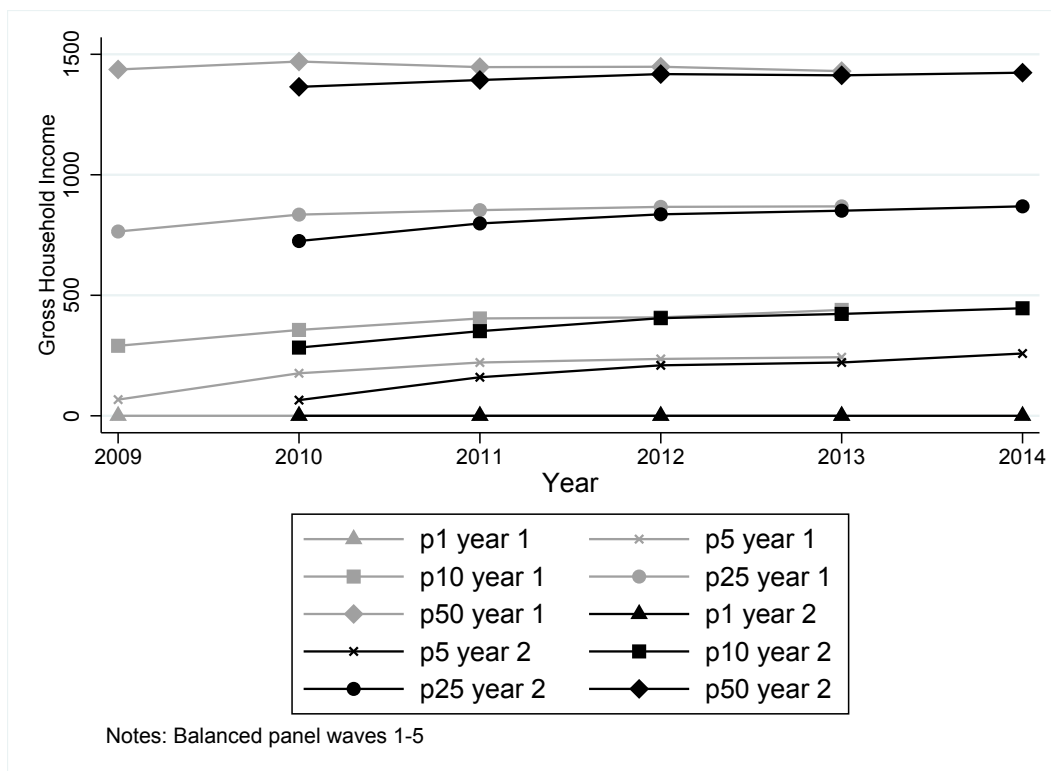
To explore this matter, figures 5 and 6 plot estimates of selected quantiles of gross income separately for the survey year one and two subsamples. The difference between the series

---

<sup>15</sup>See table 1 notes.

<sup>16</sup>It is worth noting that the conditional pension amounts have the largest means in the table. This explains why even small differences in the reporting of pension receipt can have sizable effects on the income distribution.

Figure 5: Income trends by survey year

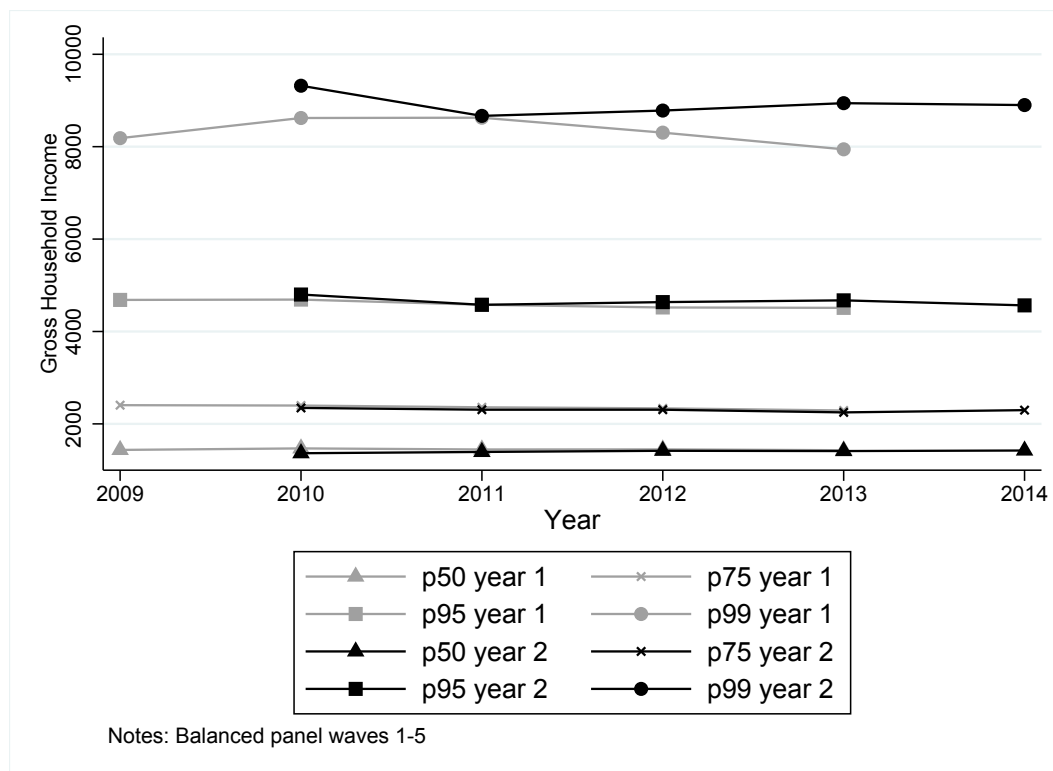


can be attributed to the fact that the survey year one participants have one more year of survey experience relative to the survey year two participants. The samples are restricted to respondents who completed an interview at each of the first five waves, to remove compositional changes due to attrition.

Figure 5 shows the bottom half of the distribution showing the median, 1st, 5th, 10th, 25th and 50th percentiles. The two series line up closely but in the early years of the panel there are notable differences between the two. For each quantile, the survey year 1 sample gives higher estimates of income in both 2010 (waves 1 and 2) and to a lesser extent in 2011 (waves 2 and 3) - but not from 2012 (waves 3 and 4) onwards. We interpret this as evidence that the data quality improvements from repeated measurement occur early in the lifetime of the panel.

Figure 6 repeats the exercise for the upper half of the distribution showing the 50th, 75th, 95th and 99th percentiles. The estimates from the different subsamples line-up remarkably closely and so it would seem that any survey effects are confined to the bottom

Figure 6: Income trends by survey year

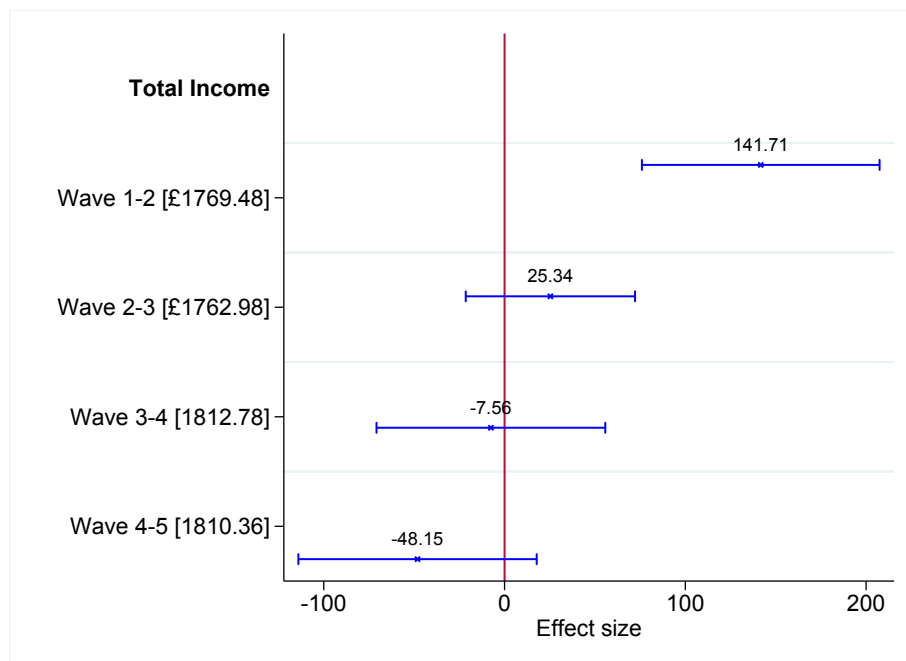


half of the distribution.

To provide estimates of the stability of reporting behaviour following wave 2, we apply our methodology to later consecutive wave pairs. Figure 7 presents the results for the waves 1 and 2 (2010), waves 2 and 3 (2011), waves 3 and 4 (2012) and waves 4 and 5 (2013) pairs. Following the wave 1 and 2 pair, the estimated differences in income are always small and statistically insignificant. This supports our preliminary conclusion that the biggest changes in response behaviour occurs between the first and second interviews.

To comment on the implications of our findings for distributional analysis, analysis based on a short-panel of the early waves of data will suffer from bias. In contrast, estimates of distributional change based on the full panel will benefit from the data quality improvements that derive from repeated measurement. Distributional estimates based on the wave one cross-section will also suffer from under-reporting problems (relative to later cross-sections), but insofar as wave one of a panel is a cross-section, analysis from a similar cross-sectional survey could be expected to suffer from comparable amounts of

Figure 7: Reporting differences at later waves



Notes: For a list of control variables, see the notes to figure 3. Analysis samples correspond: to 2010 (wave 1-2), 2011 (wave 2-3), 2012 (wave 3-4) and 2013 (wave 4-5) and are restricted to respondents who had a full interview at both corresponding waves.

under-reporting.

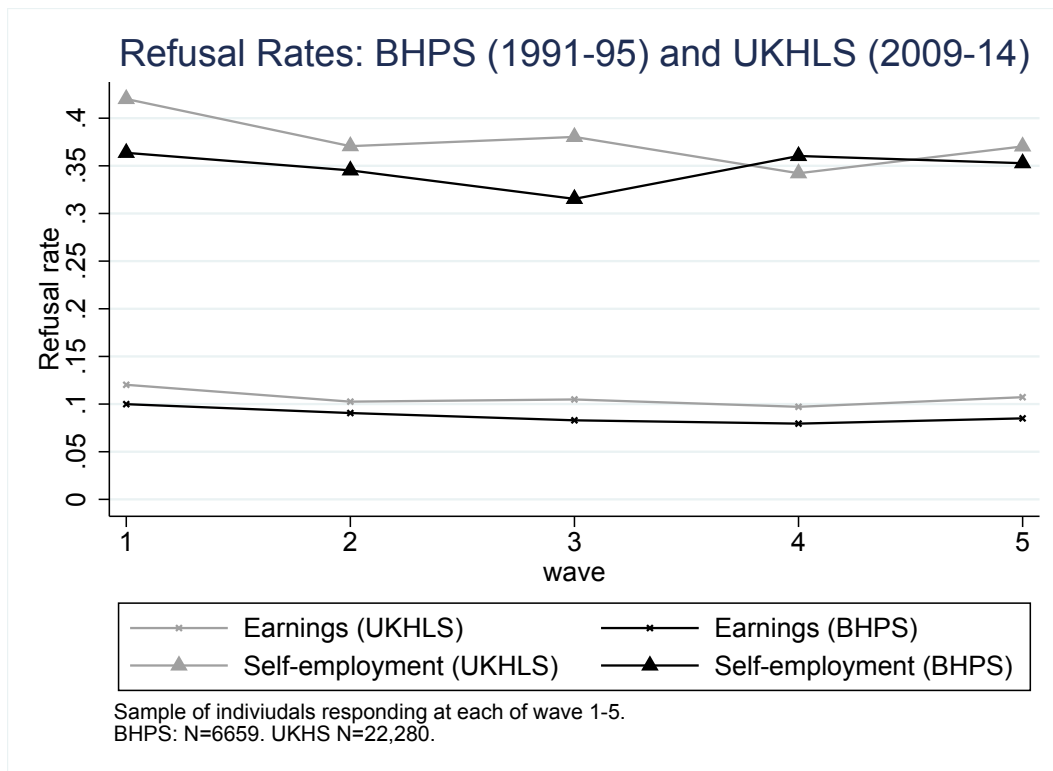
### 4.3 Evidence from the British Household Panel Survey

It is important to know whether the reporting pattern established is a peculiarity of UKHLS or a more general feature of income data collection. We turn to the predecessor of UKHLS - the British Household Panel Survey (BHPS). Following the start of the BHPS in 1991, refreshment samples were added in 1999 and again in 2000 giving us new samples of first-time responders with which we can examine changes in reporting behaviour as their panel experience grows. Dependent interviewing was not introduced in BHPS until 2006 and so our results relate to the effects of panel conditioning only.

We compare trends in BHPS and UKHLS item non-response rates at the start of each panel.<sup>17</sup> We focus on earnings from main job and self-employment profit as the questions

<sup>17</sup>We cannot compare early estimates from BHPS to the FRS as the later began only in 1994.

Figure 8: Comparison of UKHLS and BHPS item non-response rates by wave

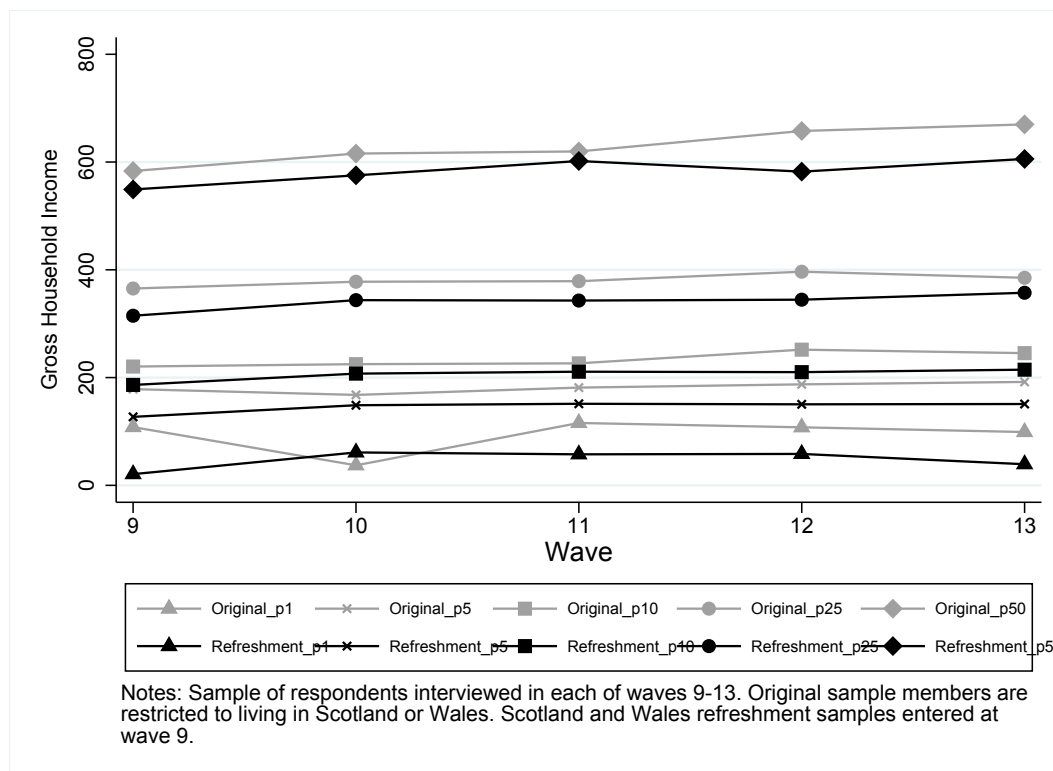


are identical in both surveys. Figure 8 shows that following wave 1, we see a fall in item non-response rates in both surveys, although the fall appears to be stronger in the UKHLS sample.<sup>18</sup> Between waves one and two the earnings refusal rate fell from 12 to 10 percent and self-employment from 42 to 37 percent in UKHLS and from 10 to 9 percent and 36 to 35 percent in BHPS. Fitting a linear regression line through the data points confirms the negative trend in both surveys: for earnings the coefficient on the wave trend is -0.41 for BHPS and -0.31 for UKHLS; and for self-employment for BHPS -0.07 and -1.28 for UKHLS.

Figure 9 plots selected income quantiles for a sample of respondents interviewed in each of waves 9-13 of the BHPS and living in Scotland or Wales, separately by whether they form part of the refreshment sample or were an original sample member. At wave 9, we

<sup>18</sup>We observe that the level of non-response is similar in both surveys. The rates are slightly higher in UKHLS. For example, at wave 1 earnings refusal rates are 10 and 12 percent; and self-employment refusal rates are 36 and 42 percent, for BHPS and UKHLS, respectively. Given that the BHPS started in the early 1990's and UKHLS some 18 years later, the differences may reflect the decline in data quality over time that has been observed in numerous surveys and across countries (see Meyer and Sullivan (2015)).

Figure 9: BHPS Scotland and Wales refreshment samples (wave 9)



observe that the refreshment sample gives lower values of percentiles 1, 5, 10 and 25 but that the differences notably decrease at wave 10. Thereafter, the gaps remain relatively stable.<sup>19</sup>

We find similar results for the Northern Ireland refreshment sample (figure A8, appendix A). For percentiles 1 and 5, we again observe that, relative to the original sample, the refreshment sample provides lower estimates in wave 11 but by an amount that noticeably decreases at wave 12. Thereafter, the gap between the two estimates is relatively stable. For higher percentiles, there was no obvious reporting difference and this fact could reflect underlying differences in the shape of the NI and UK income distributions.

<sup>19</sup>The refreshment sample tends to give lower estimates compared to the original sample and this reflects compositional differences between the groups.

## 5 Explaining the panel conditioning effect

In this section we present evidence on possible explanations for the improvements in reported income seen in UKHLS. We group the explanations according to: implementer learning; respondent-interviewer rapport; and respondent learning.<sup>20</sup>

### 5.1 Implementer learning

At wave 1 of a panel, implementers (fieldwork agency, interviewers) have no previous waves from which to draw on experience. Implementers accrue experience over the first wave of the panel, raising the possibility that they are better able to elicit responses at the wave 2 interviews.

In order for implementer learning to explain our results, two conditions need to be met, and we consider both to be implausible. First, implementers must benefit from their experience at wave 2 2010 but not at wave 1 2010, even though the two were being collected at the same point in time.<sup>21</sup> Second, substantial implementer learning would have to occur beyond the first full year of data collection (when the biggest learning might be expected) as the fieldwork agency (and interviewers) already had a full year of field experience (wave 1 2009) before the period of our analysis sample.

We compare estimates of the income distribution from UKHLS wave 1 (2009 respondents) with the FRS 2009; and UKHLS wave 1 (2010 respondents) with the FRS 2010. If the problem is with implementers, rather than respondents, then the 2010 comparison should be more favourable. Table 2 shows the results from this comparison. Columns 2 and 3 shows estimated 2009 percentiles from the FRS and UKHLS, respectively, and column 4 shows their ratio, where a ratio greater than 1 indicates that UKHLS underestimates a percentile relative to FRS. Columns 5-7 repeats the analysis but for the 2010 (wave 1)

---

<sup>20</sup>A separate possibility is that survey participation leads to behavioural change as in Crossley et al. (forthcoming). This looks implausible given that in the present paper the main effects are concentrated in pensions, which respondents cannot manipulate in the short-term.

<sup>21</sup>We show that there are few differences in the 2010 wave 1 and 2 interviewer experience distributions (see table A3, appendix A).



Table 2: Comparison of UKHLS wave 1 to FRS by calendar year

Percentile	FRS 2009	UKHLS 2009	Ratio (col 2/col 3)	FRS 2010	UKHLS 2010	Ratio (col 5/col 6)	Ratio of ratios (col 7/col 4)
5	199.14	161.00	1.24	194.54	152.16	1.28	1.03
10	264.05	237.21	1.11	259.56	230.77	1.12	1.01
15	317.08	297.02	1.07	309.95	286.26	1.08	1.01
20	365.61	355.30	1.03	358.02	339.49	1.05	1.02
25	412.32	411.89	1.00	404.49	395.35	1.02	1.02
30	463.11	466.50	0.99	457.18	455.04	1.00	1.01
35	518.96	530.16	0.98	513.14	509.83	1.01	1.03
40	578.80	593.04	0.98	568.75	572.68	0.99	1.02
45	644.98	654.25	0.99	627.74	634.09	0.99	1.00
50	710.95	721.08	0.99	694.89	699.68	0.99	1.01
55	779.18	794.58	0.98	761.77	779.61	0.98	1.00
60	854.60	876.58	0.97	838.50	846.41	0.99	1.02
65	942.85	961.07	0.98	920.23	933.10	0.99	1.01
70	1034.04	1057.79	0.98	1008.65	1023.25	0.99	1.01
75	1142.53	1172.90	0.97	1112.50	1144.67	0.97	1.00
80	1273.40	1313.70	0.97	1236.48	1274.61	0.97	1.00
85	1451.13	1487.80	0.98	1409.49	1453.66	0.97	0.99
90	1702.50	1745.69	0.98	1666.63	1726.21	0.97	0.99
95	2259.56	2193.66	1.03	2138.24	2170.68	0.99	0.96

Notes: Analysis is based on the ‘Households Below Average Income’ data sets and for household gross income before deductions. The FRS corresponds to a financial year (April to April) and a UKHLS a full calendar year.

calendar year.

In 2009, the UKHLS estimates typically match closely with the FRS ones but UKHLS misses income at the bottom of the distribution and most notably for percentiles 5, 10 and 15 where the ratios are 1.24, 1.11 and 1.07, respectively. A remarkably similar pattern is seen for the 2010 calendar year. Column 8 presents the ratio of ratios and it is always close to one indicating little change in the relative difference between the surveys over time. It reaches an absolute maximum of 1.03 for the fifth percentile, which if anything, suggests that the coverage of UKHLS got worse relative to FRS in 2010 relative to 2009. We conclude that there is no evidence to suggest fieldwork agency learning lead to improvements in data quality across the first two years of the panel.

## 5.2 Respondent-interviewer rapport

If respondent-interviewer rapport is established at the first interview, then it could potentially lead to greater respondent cooperation at the second. A simple comparison of

reported income between respondents that did and did not change interviewer at wave 2 would be uninformative as interviewer assignment is not random. However, we are able to perform difference-in-differences analysis that enables us to remove time invariant differences between those that changed interviewer and not.

Specifically, we augment equation (1) with a dummy variable for a respondent having different interviewers at wave 1 and 2, and its interaction with our main wave 2 indicator. If familiarity with the interviewer drives the increases in reported income, then we would expect the interaction to be negative and significant. In that the interaction would be biased by any time varying factors associated with changing interviewer (eg. a location move associated with changing income and interviewer), the controls included in the regression model would work to reduce the bias. To further test the interviewer experience mechanism, we add to the model a continuous variable for the number of interviews completed since the start of the survey, alongside controls for interviewer age and sex.

Column 1 of table 3 replicates our main finding for total income, alongside the difference-in-differences results in column 2. Both the changed interviewer dummy and its interaction with the wave 2 dummy are highly insignificant indicating that the interviewer played a limited role in increasing the reporting of income. Moreover, our estimated main effects (wave 2 dummy) remain positive, statistically significant and are relatively stable. The other interviewer controls, are statistically insignificant including the continuous interviewer experience variable. We conclude that the respondent-interviewer rapport cannot explain the changes in reporting behaviour across waves.

### **5.3 Respondent learning**

Respondents may have an improved comprehension of the complex interview or have updated their beliefs about the trustworthiness of the data holders following a first interview. In the presence of even small doubts over confidentiality, respondents may be willing to misreport their income in a survey, if the cost of lying is low (Hurst, Li, and Pugsley

Table 3: Effect of changing interviewer on reported total income

	(1)	(2)
Wave 2	141.71*** (33.542)	126.07** (41.346)
Changed interviewer [36.70 ]		5.56 (39.568)
Changed interviewer X Wave 2		45.77 (74.202)
<b>Interviewer:</b>		
No. Interviews completed [100.00 ]		-0.29 (0.262)
Female [54.55 ]		4.21 (32.577)
Age [57.68 ]		2.59 (1.859)
Observations	29528	29528

Notes: Robust standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Estimates of equation 1 with full set of controls (see figure 3 notes).

(2014)). We use information collected by the interviewers on respondent behaviours during the interview to provide evidence on these two possibilities.

First, to examine differences in respondent comprehension across waves, we exploit interviewer reports of how well the respondent understood the questions during the interview (on a 5 point scale). Second, to examine how confidentiality concerns may have lessened, we have available interviewer reports of whether a respondent was ‘suspicious’ about the study after the interview (3 point scale) and whether prior to the interview, the household respondent had questions about ‘confidentiality’ (binary variable). As the later was recorded before the individual interviews took place, it should better reflect changes in the latent confidentiality concerns of the household, in contrast to the other two measures, which may reflect an interviewers interpretation of response behaviour during the interview. We estimate equation (1) for the 3 outcomes where we recoded the interviewer observations into

binary indicators.<sup>22</sup> Full details of the questions and the construction of the interviewer observation variables are provided in appendix B.

Table 4: Effect of wave 2 interview on respondent behaviours

	(1)	(2)	(3)
	Misunderstood questions	Suspicious	Queries confidentiality
Wave 2	-0.01 (0.005)	-0.09*** (0.003)	-0.16*** (0.003)
<i>N</i>	29502	29502	29365

Notes: Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Estimates of equation 1 with full set of controls (see figure 3 notes).

Means of the dependent variables are: 0.31, 0.12, 0.18, respectively.

Table 4 shows the results. Sample means are reported in the footnote to the table. We find no evidence that being interviewed for a second time improved respondent understanding of the interview with the effects being small and statistically insignificant. In contrast, we observe that interviewers rated respondents as being less suspicious after the second interview and were also less likely to have confidentiality queries.

The final step in the causal chain is to show that confidentiality queries are related to respondent willingness to reveal income information. We do this using item non-response as a measure of willingness to reveal income and show that in a cross-section it is related to confidentiality concerns.<sup>23</sup> Table 5 presents the results and confirms that confidentiality concerns are predictors of item non-response in the wave 1 cross-section.

An interpretation consistent with these findings is that the first interview reveals information to respondents about the trustworthiness of the data holders. At the start of the panel, respondents have doubts about the survey organisation, a stranger to them, who may share their sensitive data with third party organisations. But following the first interview respondents learn that the data holders are reliable and that their data have not been shared. By the time the second interview comes, respondents have updated their beliefs

<sup>22</sup>Our results are insensitive to changes in the chosen thresholds.

<sup>23</sup>We focus on the ‘queries confidentiality’ measure as it is measured before the individual interviews and so there is no concern that it reflects item non-response during the interview.

Table 5: Effect of confidentiality concerns on item non-response

	(1)	(2)	(3)	(4)	(5)
	Earnings	2nd Job	Self-employment	Investment	Pensions
<b>Respondent queries:</b>					
purpose [0.43 ]	0.04*** (0.006)	0.04* (0.020)	-0.01 (0.024)	0.05*** (0.009)	0.00 (0.002)
interview length [0.29 ]	-0.00 (0.006)	0.01 (0.022)	0.01 (0.025)	0.02* (0.010)	0.00 (0.002)
panel design [0.03 ]	-0.01 (0.016)	-0.03 (0.056)	-0.12* (0.054)	-0.09*** (0.023)	0.00 (0.007)
confidentiality [0.19 ]	0.10*** (0.008)	0.09** (0.028)	0.10*** (0.028)	0.11*** (0.011)	0.01** (0.003)
incentive/payment [0.05 ]	0.00 (0.013)	0.01 (0.046)	-0.01 (0.052)	0.02 (0.022)	-0.00 (0.005)
other query [0.03 ]	0.04* (0.018)	0.07 (0.065)	0.01 (0.074)	0.08*** (0.025)	0.00 (0.006)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	19015	1795	2197	12862	10937

Notes: Robust standard errors in parentheses. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

Sample of wave 1 respondents.

For a list of control variables, see the notes to figure 3.

about the trustworthiness of the survey organisation, and are so more open in revealing details of their personal finances.

One of our findings was of PC effects for pensions, including the state pension. This raises the question of why state pension reporting is affected by PC - when on the face of it, it is not a confidential area of questioning. One explanation is that details of state pension receipt are collected alongside more sensitive pension types (eg. private pensions) as part of a single pensions question. It may be that respondents unwilling to disclose their sensitive pension types simply choose not to engage with the pensions question at all.

## 6 Conclusions

We have shown that the quality of income data collected as part of a large-scale household panel survey improves over the life-time of the panel due to changes in respondent reporting behaviour. The largest changes in reported income are concentrated across the first waves

of the panel and in unearned income sources, particularly pensions and disability benefits. The effect sizes are large and have until this point gone unnoticed, potentially as it is difficult to distinguish changes in reporting behaviour from real changes in living standards, without linked administrative records. The novelty of our approach is that it does not require data linkage, but makes use of unique features of the survey design of the Understanding Society survey as a quasi-experiment. We also show that similar income data quality improvements are found in another leading panel survey, suggesting that our results may be generalisable to other surveys.

The use of income data from repeated survey measures is commonplace in social science research, including the use of large scale household surveys and purpose built surveys implemented as part of field experiments. Our results suggest that researchers analysing data from the early waves of a panel or with short panels should proceed with caution. One possibility is that researchers may want to consider adjusting data from the first waves of data collection. Our findings are also relevant for studies based on cross-sectional data, which essentially forms wave one of a panel and so are indicative of the types of income sources that may be under-reported.

Our work is suggestive that respondent confidentiality concerns play a role in the observed patterns. Addressing these during data collection may bring data quality improvements. Separately, other sensitive variables collected as part of survey data eg. voting intentions, illicit behaviours may also show similar effects. Both of the later points are left for future work.

## References

- Brewer, Mike, Ben Etheridge, and Cormac O’Dea. forthcoming. “Why are households that report the lowest incomes so well-off?” *Economic Journal* .
- Bricker, Jesse, Alice M. Henriques, Jacob Krimmel, and John Sabelhaus. 2015. “Measuring Income and Wealth at the Top Using Administrative and Survey Data.” *Finance and*

- Economics Discussion Series 2015-30, Board of Governors of the Federal Reserve System (U.S.).
- Crossley, Thomas, Jochem de Bresser, Liam Delaney, and Joachim Winter. forthcoming. “Can survey participation alter household saving behavior?” *Economic Journal* .
- Das, Marcel, Vera Toepoel, and Arthur van Soest. 2011. “Nonparametric Tests of Panel Conditioning and Attrition Bias in Panel Surveys.” *Sociological Methods & Research* 40 (1):32–56.
- David, Martin H. and Christopher R. Bollinger. 2005. “I didn’t tell, and I won’t tell: dynamic response error in the SIPP.” *Journal of Applied Econometrics* 20 (4):563–569.
- Frick, Joachim R., Jan Goebel, Edna Schechtman, Gert G. Wagner, and Shlomo Yitzhaki. 2006. “Using Analysis of Gini (ANoGi) for Detecting Whether Two Sub-Samples Represent the Same Universe: The SOEP Experience.” *Sociological Methods Research* 34 (4):427–468.
- Halpern-Manners, Andrew, John Robert Warren, and Florencia Torche. 2014. “Panel Conditioning in the General Social Survey.” *Sociological Methods & Research* .
- Hurst, Erik, Geng Li, and Benjamin Pugsley. 2014. “Are household surveys like tax forms? Evidence from income underreporting of the self-employed.” *Review of economics and statistics* 96 (1):19–33.
- Jenkins, Stephen P. 2011. *Changing Fortunes: Income Mobility and Poverty Dynamics in Britain*. Oxford University Press.
- Lynn, Peter, Annette Jäckle, Stephen P. Jenkins, and Emanuela Sala. 2012a. “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* 175 (1):289–308.

- Lynn, Peter J, Jonathan Burton, Olena Kaminska, Gundi Knies, and Alita Nandi. 2012b. “An initial look at non-response and attrition in Understanding Society.” Tech. rep., Understanding Society Working Paper Series: 2012-02.
- Meyer, Bruce D. and James X. Sullivan. 2003. “Measuring the Well-Being of the Poor Using Income and Consumption.” NBER Working Papers 9760, National Bureau of Economic Research, Inc.
- . 2011. “Viewpoint: Further results on measuring the well-being of the poor using income and consumption.” *Canadian Journal of Economics* 44 (1):52–87.
- Meyer, Wallace KC Mok, Bruce D. and James X. Sullivan. 2015. “Household Surveys in Crisis.” *Journal of Economic Perspectives* 29 (4):199–226.
- Van Landeghem, Bert. 2014. “A test based on panel refreshments for panel conditioning in stated utility measures.” *Economics Letters* 124 (2):236–238.



# Does repeated measurement improve income data quality?

Paul Fisher

Institute for Social and Economic Research, University of Essex

October 2016

**Supplementary material**

# Appendix A

## A1. Additional tables and figures

Table A1: Comparison of UKHLS wave one 2009 and 2010 responding samples

	2009	2010	Mean Diff	SE	N_Year One	N_Year Two
sex	0.4524	0.4555	-0.0031	0.0049	21447	20139
age	46.8445	47.0688	-0.2243	0.1800	21447	20139
age (bands):						
10-19 years old	0.0627	0.0603	0.0023	0.0024	21447	20139
20-29 years old	0.1470	0.1467	0.0003	0.0035	21447	20139
30-39 years old	0.1718	0.1654	0.0065	0.0037	21447	20139
40-49 years old	0.1881	0.1915	-0.0034	0.0038	21447	20139
50-59 years old	0.1571	0.1575	-0.0004	0.0036	21447	20139
60-69 years old	0.1432	0.1458	-0.0026	0.0034	21447	20139
70+	0.1301	0.1329	-0.0027	0.0033	21447	20139
ethnicity:						
white	0.9136	0.9046	0.0090**	0.0029	20646	19339
mixed	0.0104	0.0108	-0.0004	0.0010	20646	19339
indian and chinese	0.0217	0.0270	-0.0052***	0.0015	20646	19339
other asian	0.0239	0.0260	-0.0021	0.0016	20646	19339
african or black caribbean	0.0231	0.0226	0.0005	0.0015	20646	19339
other	0.0074	0.0090	-0.0016	0.0009	20646	19339
highest qualification:						
Degree	0.2025	0.2081	-0.0055	0.0040	21419	20104
Other higher degree	0.1128	0.1125	0.0003	0.0031	21419	20104
A-level	0.1894	0.1911	-0.0017	0.0039	21419	20104
GCSE	0.2121	0.2115	0.0006	0.0040	21419	20104
other	0.1111	0.1045	0.0066*	0.0030	21419	20104
no qualification	0.1720	0.1723	-0.0003	0.0037	21419	20104
marital status:						
married or civil partnership	0.5087	0.5084	0.0003	0.0049	21443	20132
cohabiting	0.1258	0.1244	0.0013	0.0032	21443	20132
single and never married	0.2139	0.2189	-0.0050	0.0040	21443	20132
divorced or separated	0.0886	0.0846	0.0040	0.0028	21443	20132
widowed	0.0630	0.0637	-0.0007	0.0024	21443	20132
long-standing illness or impairment						
SF-12 Physical Component Summary	49.2970	49.4700	-0.1731	0.1174	20164	18642
SF-12 Mental Component Summary	50.6627	50.5925	0.0702	0.1020	20164	18642
housing tenure:						
owned	0.3048	0.2982	0.0066	0.0045	21394	20098
mortgage	0.3928	0.3810	0.0118*	0.0048	21394	20098
rent	0.1360	0.1425	-0.0065	0.0034	21394	20098
social housing	0.1628	0.1751	-0.0123***	0.0037	21394	20098
# bedrooms	2.9238	2.9173	0.0064	0.0099	21430	20117
# people	2.7510	2.7841	-0.0330*	0.0134	21447	20139
# children	1.7237	1.7310	-0.0072	0.0151	6733	6590

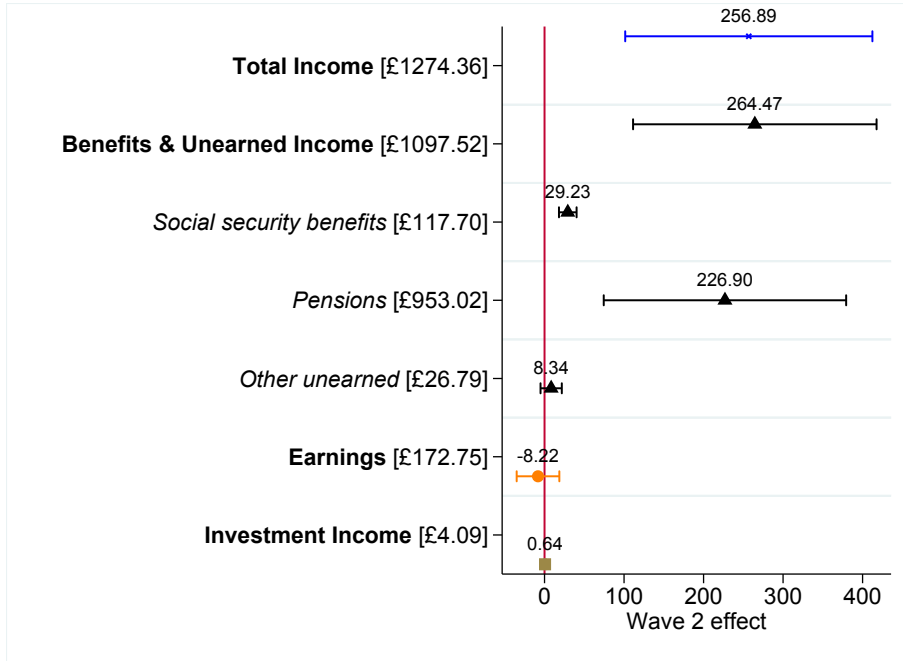
Notes: \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001.

Table A2: Estimation sample: comparison of baseline (wave one) characteristics by year of interview

	2010	2009	Mean Diff	SE	N_2010	N_2009
sex	0.4292	0.4281	-0.0011	0.0057	14731	14912
age	48.4870	48.4485	-0.0385	0.2068	14731	14912
<i>ethnicity:</i>						
white	0.9233	0.9283	0.0050	0.0030	14720	14899
mixed	0.0092	0.0095	0.0003	0.0011	14720	14899
indian and chinese	0.0221	0.0171	-0.0050**	0.0016	14720	14899
other asian	0.0195	0.0203	0.0008	0.0016	14720	14899
african or black caribbean	0.0186	0.0193	0.0007	0.0016	14720	14899
other	0.0073	0.0054	-0.0018*	0.0009	14720	14899
<i>highest qualification:</i>						
Degree	0.2119	0.2102	-0.0018	0.0047	14726	14906
Other higher degree	0.1203	0.1200	-0.0003	0.0038	14726	14906
A-level	0.1861	0.1829	-0.0032	0.0045	14726	14906
GCSE	0.2124	0.2076	-0.0048	0.0047	14726	14906
other	0.1051	0.1121	0.0070	0.0036	14726	14906
no qualification	0.1641	0.1672	0.0030	0.0043	14726	14906
<i>marital status:</i>						
married or civil partnership	0.5294	0.5315	0.0020	0.0058	14725	14911
cohabiting	0.1229	0.1241	0.0013	0.0038	14725	14911
single and never married	0.1883	0.1784	-0.0099*	0.0045	14725	14911
divorced or separated	0.0932	0.0976	0.0043	0.0034	14725	14911
widowed	0.0662	0.0684	0.0022	0.0029	14725	14911
<i>economic status:</i>						
self-employed	0.0744	0.0665	-0.0079**	0.0030	14729	14911
employed	0.4627	0.4787	0.0160**	0.0058	14729	14911
unemployed	0.0562	0.0496	-0.0067*	0.0026	14729	14911
retired	0.2433	0.2417	-0.0016	0.0050	14729	14911
student	0.0473	0.0493	0.0020	0.0025	14729	14911
long-term sick or disabled	0.0356	0.0388	0.0033	0.0022	14729	14911
other	0.0806	0.0754	-0.0052	0.0031	14729	14911
usual weekly hours worked	35.6694	35.8480	0.1786	0.2312	7067	7430
long-standing illness or impairment	0.3884	0.3910	0.0026	0.0057	14712	14901
SF-12 Physical Component Summary	49.4445	49.1819	-0.2627*	0.1336	14652	14849
SF-12 Mental Component Summary	50.7113	50.8040	0.0928	0.1155	14652	14849
<b>household</b>						
<i>tenure:</i>						
owned	0.3174	0.3270	0.0096	0.0054	14712	14879
mortgage	0.3881	0.3961	0.0079	0.0057	14712	14879
rent	0.1192	0.1167	-0.0024	0.0038	14712	14879
social housing	0.1722	0.1565	-0.0157***	0.0043	14712	14879
# bedrooms	2.9264	2.9256	-0.0008	0.0115	14722	14902
# people	2.7206	2.6762	-0.0444**	0.0155	14731	14912
any children	0.3269	0.3076	-0.0193***	0.0054	14731	14903
# children	1.7317	1.7315	-0.0002	0.0178	4815	4584
age of youngest child	6.6287	6.3709	-0.2578*	0.1017	4815	4584
# employed	1.2321	1.2543	0.0222	0.0121	14731	14912
# working age	1.7032	1.7004	-0.0028	0.0143	14731	14912
# couples	0.7184	0.7209	0.0025	0.0055	14731	14912
# single parents	0.0650	0.0612	-0.0039	0.0029	14731	14912

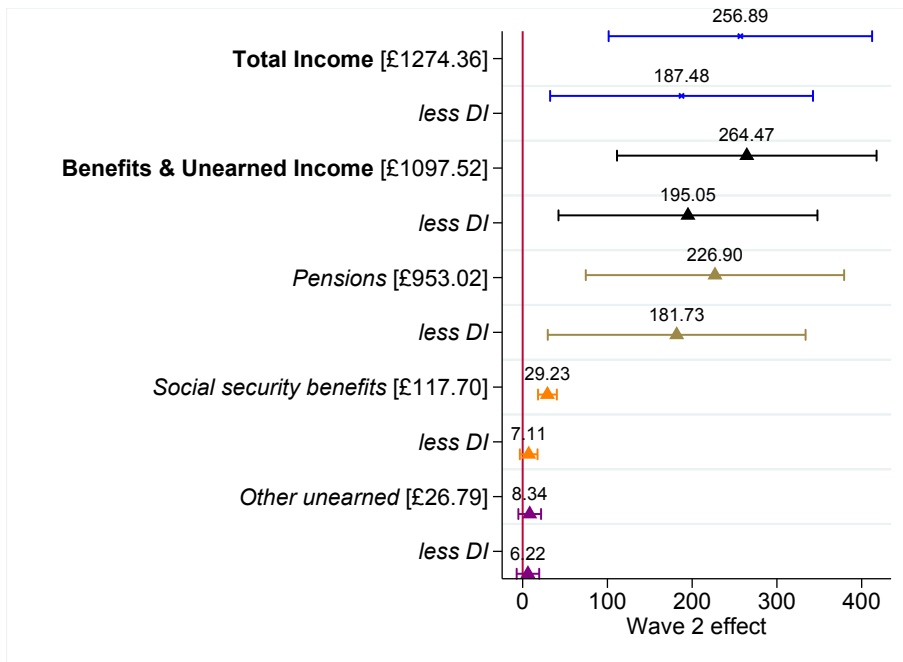
Notes: \*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001. Balanced sample of respondents who complete a full-interview at wave 1 and wave 2.

Figure A1: Effect of second interview on reported income (pensioner sample)



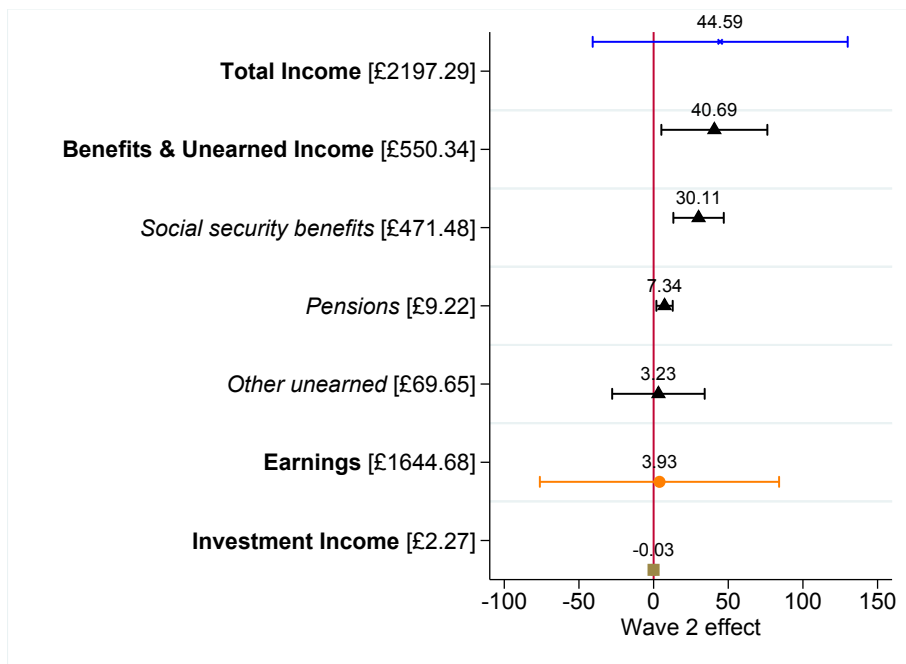
Notes: see figure 3 notes. Sample restricted to respondents of state pension age.

Figure A2: Panel Conditioning effect of second interview on reported income (pensioner sample)



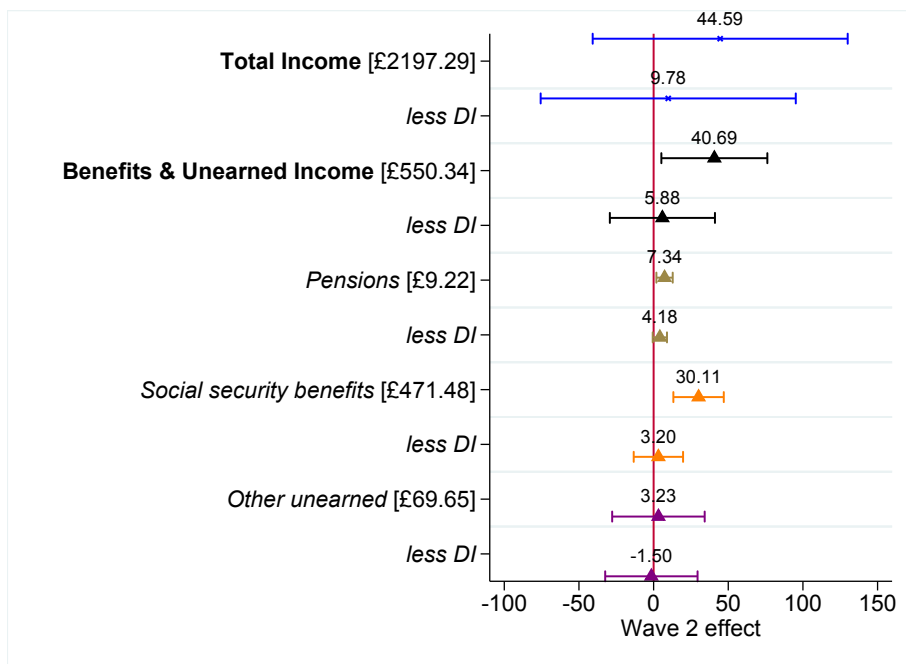
Notes: see figure 3 notes. Sample restricted to respondents of state pension age.

Figure A3: Effect of second interview on reported income (working age with children sample)



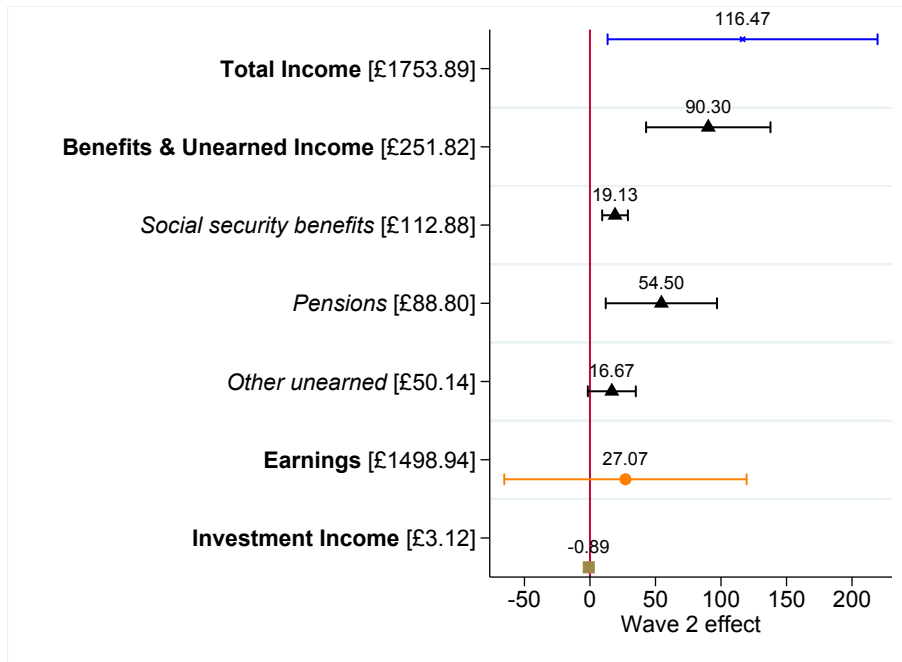
Notes: see figure 3 notes. Sample restricted to respondents less than state pension age and with children.

Figure A4: Panel Conditioning effect of second interview on reported income (working age with children sample)



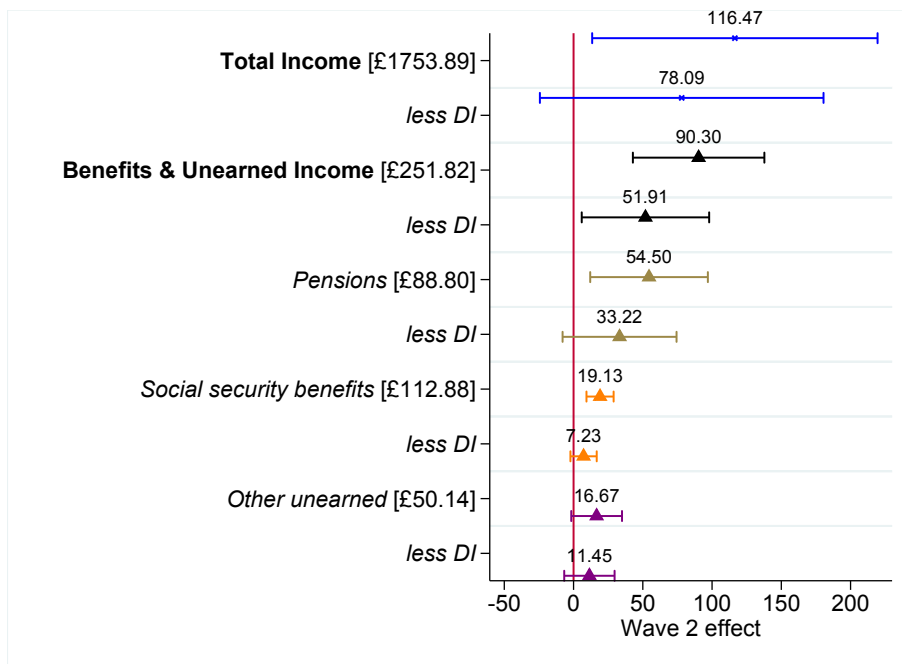
Notes: see figure 3 notes. Sample restricted to respondents less than state pension age and with children.

Figure A5: Effect of second interview on reported income (working age no children sample)



Notes: see figure 3 notes. Sample restricted to respondents less than state pension age and without children.

Figure A6: Panel Conditioning effect of second interview on reported income (working age no children sample)



Notes: see figure 3 notes. Sample restricted to respondents less than state pension age and without children.

Table A3: Interviewer experience: number of interviews completed

Percentile	Wave 1 (2010)	Wave 2 (2010)
1	2	4
5	11	16
10	22	27
25	47	52
50	85	94
75	139	148
90	199	207
95	237	246
99	324	330
Mean	100.00	108.09
sd	71.2	73.4

Notes: Sample is defined as in the methodology section.

## A2. Effects of imputation

If the observed patterns in reporting behaviour are driven by differences across waves in the extent to which respondents are willing to reveal details of their income, then this raises the question of whether imputation of missing data masks the extent of reporting differences across waves. To explore this issue, we estimate equation 1 for income and its subcomponents, where we set to zero any source which has missing income information. In this way, changes in the extent of missingness are not corrected by imputation and would be reflected in our estimates.

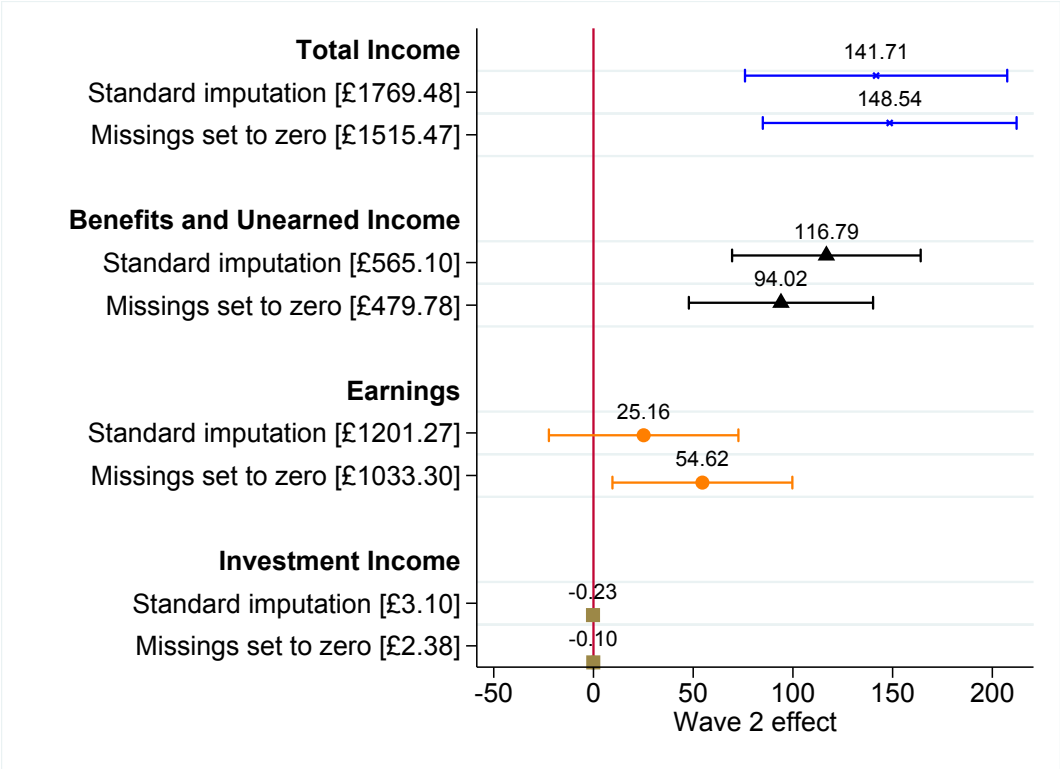
Figure A7 presents the estimates (labelled imputes to zero) alongside our main results (labelled standard imputes). Estimates of the effect on total income are stable but this hides offsetting changes within its components. For earnings, the estimated effects with zero imputes are larger relative to those using standard imputes and they have become statistically significant. This implies that imputes produced as standard in most household surveys help in reducing some of the differences due to differential reporting behaviour across the early waves of the panel.

In contrast, for benefits and unearned income, we observe that the effects have slightly fallen in magnitude when using zero values for missing amounts. This results as our main effects for benefits and unearned income are driven by changes in reported receipt, but some of the reporting increases are now not binding (where an amount is missing and now set to zero).

Overall, the results suggest imputation can work to correct data quality differences across the initial waves but only for amounts and not in case of false negative reporting, which is not corrected by imputation.



Figure A7: Effect sizes when setting missing amounts to zero

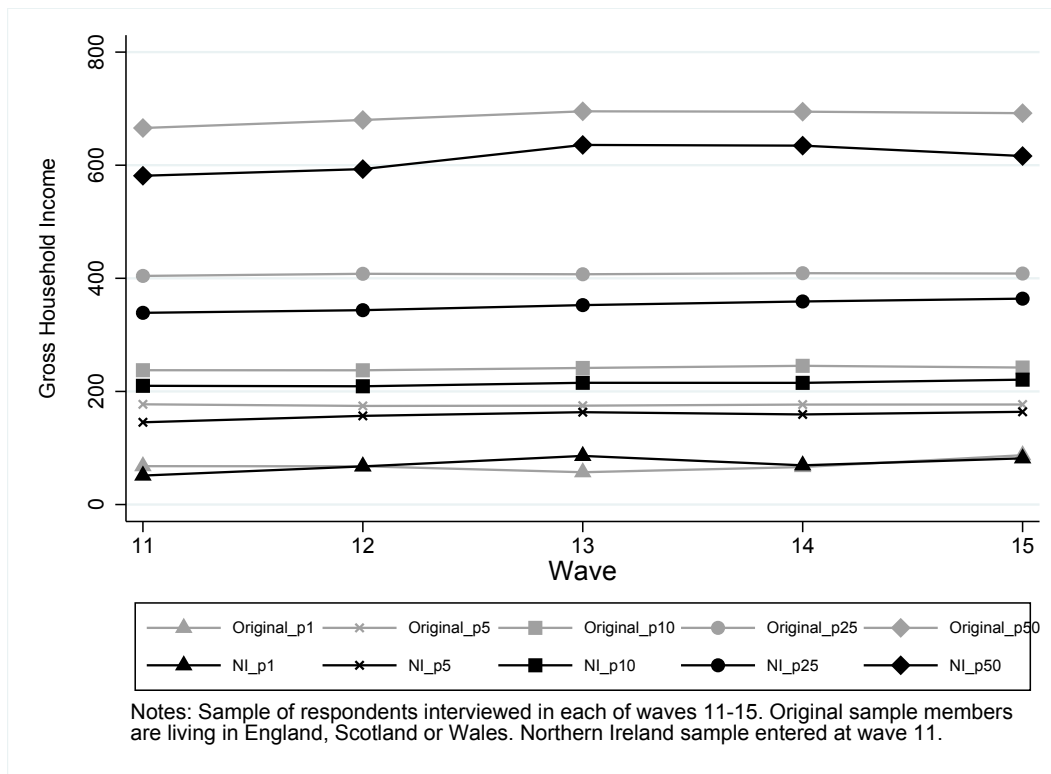


Notes: see figure 3 notes. ‘Imputes to zero’ sets missing amounts to zero. ‘Standard imputation’ replaces missing amounts with the imputes of the data providers.

### A3. BHPS Northern Ireland refreshment sample

Figure A8 shows results from the analysis of the BHPS Northern Ireland refreshment sample. Northern Ireland did not form part of the original BHPS sample and so our ‘original’ comparison sample consists of respondents living in England, Scotland or Wales.<sup>24</sup> For percentiles 1 and 5, we again observe that, relative to the original sample, the refreshment sample provides lower estimates in wave 11 but by an amount that noticeably decreases at wave 12. Thereafter, the gap between the two estimates is relatively stable. For the higher percentiles, presented, there is no obvious reporting difference and this fact could reflect underlying differences in the shape of the NI and UK income distributions.

Figure A8: BHPS Northern Ireland refreshment sample (wave 11)



<sup>24</sup>NI was not included in the HBAI data until 2002/03.

## Appendix B (Data appendix)

This appendix provides details of the income questions asked in Understanding Society, including details on the use of dependent interviewing.

### B1. Benefits and unearned income module

The showcard of 9 broad payment types lists: Unemployment-related benefits or National Insurance Credits; Income Support; Sickness, disability or incapacity benefits; Any sort of pension including a private pension or the State pension; Child Benefit; Tax credits; Any other family related benefits or payment; Housing or Council Tax Benefit; Income from any other state benefit.

The 39 income sources are allocated as:

#### **Pensions:**

- 1) ni retirement/state retirement (old age) pension
- 2) pension, previous employer
- 3) pension from a spouse's previous employer
- 4) private pension/annuity
- 5) widow's or war widow's pension

#### **Social security benefits:**

- 6) widowed mother's allowance / widowed parent's allowance / bereavment allowance
- 7) pension credit (incl. guarantee credit saving credit)
- 8) severe disablement allowance
- 9) industrial injury disablement allowance
- 10) disability living allowance
- 11) attendance allowance
- 12) carer's allowance (was invalid care allowance)
- 13) war disablement pension
- 14) incapacity benefit

- 15) income support
- 16) job seeker's allowance
- 17) national insurance credits
- 18) child benefit (incl. lone-parent child benefit payments)
- 19) child tax credit
- 20) working tax credit (incl. disabled person's tax credit)
- 21) maternity allowance
- 22) housing benefit
- 23) council tax benefit
- 24) foster allowance / guardian allowance
- 25) rent rebate
- 26) rate rebate
- 27) employment and support allowance
- 28) return to work credit
- 29) in-work credit for lone parents

**Other unearned income:**

- 30) educational grant (not student loan or tuition fee loan)
- 31) trade union / friendly society payment
- 32) maintenance or alimony
- 33) payments from a family member not living here
- 34) rent from boarders or lodgers (not family members) living here
- 35) rent from any other property
- 36) sickness and accident insurance

**Other:**

- 37) other disability related benefit or payment
- 38) any other regular payment
- 39) income from any other state benefit

A scripting error at wave 1 meant that amounts were not collected for respondents who reported receipt of sources 37-39. The coverage of these sources is small so we deduct them from our income totals following wave 1 to ensure consistency of our totals across waves.

At the end of stage one above, respondents who fail to report a source at wave t but reported it at wave t-1 are asked:

Can I just check, according to our records, you have in the past received [x]. Are you currently receiving [x], either just yourself or jointly?

## **B2. Employee earnings**

Q1) Can I just check, did you do any paid work last week - that is in the seven days ending last Sunday - either as an employee or self-employed?

Q2) Even though you weren't working did you have a job that you were away from last week?

Q3) Are you an employee or self-employed?

Q4) If an employee on Q3): The last time you were paid, what was your gross pay - that is including any overtime, bonuses, commission, tips or tax refund but before any deductions for tax, National Insurance or pension contributions, union dues and so on? How long a period did that cover?

Q5) Is this the amount you usually receive (before any statutory sick pay or statutory maternity, paternity or adoption pay)?

Q6) If 'no' to Q5): How much are you usually paid? How long a period did that cover?

Q7) If 'no' to Q5): And is that before or after any deductions for tax, National Insurance, union dues and so on or are there usually no deductions at all made from your salary?

## **B3. Self-employee earnings**

Q6) If a self-employee on Q3): In this job/business are annual business accounts prepared for the Inland Revenue for tax purposes?

Q7) If yes to Q6): What was the amount of (your share of) the profit or loss figure shown on these accounts for this period? (And month/year accounts began and ended)

Q8) Does this figure relate to profit or loss?

Q9) Can i just check, is that figure before deduction of income tax?

Q10) Can i just check, is that figure before deduction of National Insurance?

Q11) If no to Q6): After paying for any materials, equipment or goods that you use(d) in your work, what was your weekly or monthly income, on average, from this job/business over the last 12 months?

Q12) Was that weekly or monthly income?

Q13) Can i just check, is that figure before deduction of income tax?

Q14) Can i just check, is that figure before deduction of National Insurance?

#### **B4. Second job earnings**

Q15) Do you currently earn any money from a second job, odd jobs, or from work that you might do from time to time, apart from any main job you have?

Q16) If yes to 15): Before tax and other deductions, how much do you earn from your second and all other occasional jobs in a usual month?

#### **B5. Investment income**

Q17) In the past 12 months how much have you personally received in the way of dividends or interest from any saving and investments you may have?

Where respondents cannot give an exact amount in 17) they are presented with a series of unfolding brackets where they can bound their annual investment income. For individuals reporting bounds, the data providers impute an amount.

## **B6. Interviewer observations**

**Misunderstood questions:** In general, how would you describe the respondents understanding of the question?

- 1 Excellent
- 2 Good
- 3 Fair
- 4 Poor
- 5 Very poor

Responses 2-4 are coded as one and category 1 as zero.

**Suspicious:** Was the respondent suspicious about the study after the interview was completed?

- 1 No, not at all suspicious
- 2 Yes, somewhat suspicious
- 3 Yes, very suspicious

Responses 2-3 are coded as one and category 1 as zero.

**Queries confidentiality:** Did the household respondent query any of the following topics?

- 1 purpose (e.g. 'Whats the purpose? Whats all this about?')
- 2 interview length (e.g. 'How long will this take?')
- 3 panel design (e.g. 'Youll be coming back next year?')
- 4 confidentiality (e.g. 'Whos going to see the answers?')
- 5 incentive/payment (e.g. 'Whats in it for us/me?')
- 6 other query

A 0/1 indicator is constructed from the responses to item 4.