VERSITA

Can I Just Check... ? Effects of Edit Check Questions on Measurement Error and Survey Estimates

Peter Lugtig[1] and Annette Jäckle[2]

Household income is difficult to measure, since it requires the collection of information about all potential income sources for each member of a household. We assess the effects of two types of edit check questions on measurement error and survey estimates: within-wave edit checks use responses to questions earlier in the same interview to query apparent inconsistencies in responses; dependent interviewing uses responses from prior interviews to query apparent inconsistencies over time. We use data from three waves of the British Household Panel Survey (BHPS) to assess the effects of edit checks on estimates, and data from an experimental study carried out in the context of the BHPS, where survey responses were linked to individual administrative records, to assess the effects on measurement error. The findings suggest that interviewing methods without edit checks underestimate non-labour household income in the lower tail of the income distribution. The effects on estimates derived from total household income, such as poverty rates or transition rates into and out of poverty, are small.

Key words: Dependent interviewing; validation study; record linkage; British household panel survey; income; poverty.

1. Introduction

Household income is a key measure of social welfare and as such important for policy analyses. Some surveys, such as the European Social Survey, ask one household member a single question about their income: *"Using this card, please tell me which letter describes your household's total income, after tax and compulsory deductions, from all sources? If you don't know the exact figure, please give an estimate."* Surveys for which income is a key outcome measure more commonly ask a host of questions about each potential source of income, including questions about receipt status, timing of receipt and amounts received. Total income has to be computed from these questions and aggregated over all income sources and all household members. In both cases, reporting on household income is a difficult task for respondents. As a result, household income is likely to be measured with error and estimates derived from it, such as poverty rates or income dynamics over time, may be biased.

In this article we assess the effects of edit check questions, which are incorporated into the questionnaire to detect and correct potential reporting errors, on estimates derived

from detailed questions about household income. We examine the effects of both within-wave and cross-wave edit checks in the measurement of non-labour household income. *Within-wave edit checks* use information collected earlier in the same interview to check the consistency of answers. For example, respondents can be queried about sources they have not reported, but for which they are likely to be eligible, judging from responses given earlier in the interview (Pennell 1993). *Cross-wave edit checks* are specific to longitudinal surveys. They use information provided in previous interviews to check the longitudinal consistency of responses. For example, respondents can be queried about sources they have reported in the past, but not in the current interview (see Jäckle 2009; Mathiowetz and McGonagle 2000). Cross-wave edit checks are typically referred to as 'dependent interviewing' (DI) and we follow this convention.

The key question examined here is to what extent edit checks affect estimates of household income and poverty. Previous studies evaluating the effects of DI have mainly focused on measurement error in receipt status for individual income sources. These studies have shown that some non-labour income sources are considerably underreported and that DI improves reporting for non-labour income (Lynn et al. 2012). Other studies have examined measurement error in the timing of receipt and shown that DI reduces errors in monthly transition rates (Moore et al. 2009) and spell durations (Jäckle 2008). The effects on monetary amounts have not been examined to our knowledge. Neither have the effects on estimates related to total (household) income. Although the reduction of error in individual survey questions can be substantial, it is not a priori clear what effect this methodological improvement has on estimates that are derived from a series of detailed questions about all components of household income.

We contribute to this literature by examining to what extent edit checks affect estimates of household income, poverty rates and transitions into and out of poverty. For this purpose we use three waves of the British Household Panel Survey (BHPS), in which both within-wave edit checks and DI are used in a quasi-experimental way for the collection of non-labour income data. The analyses of the BHPS data illustrate to what extent edit checks affect estimates derived from income data. The BHPS data however do not allow any conclusions about the effects of edit checks on measurement error and resulting biases in estimates. We therefore complement these analyses using data from an experimental study carried out in the context of the BHPS, which linked survey responses to individual administrative records.

The results suggest that traditional methods of interviewing that do not use edit checks for non-labour income sources underestimate household income in the lower tail of the income distribution. Estimated poverty status and poverty transitions however hardly change. The changes in estimates appear to reflect a reduction in measurement error in the reporting and duration of receipt, thus reflecting an improvement of data accuracy.

2.   Data

2.1.   *The British Household Panel Survey (BHPS)*

The BHPS is a panel survey of the UK population that started in 1991 with a clustered and stratified address-based sample of 5,500 households. All household members aged 16+

are interviewed annually and followed as long as they remain in the UK. The individual response rates, conditional on response in the prior wave, are around 94% (RR1 – AAPOR 2011) in the waves we used for our analyses (Waves 15-17). All in all, 49.6% of the original sample members with an interview at Wave 1 completed an interview in Waves 15-17 (Taylor et al. 2009). Proxy interviews were held with about 1% of sample members in each wave and treated as missing data in our analyses.

Income data are collected in two sections of the questionnaire: one on labour earnings and another on non-labour income (including state cash transfers, private pensions, private transfers and investment income). Edit checks are only used for non-labour income and we therefore focus on those questions. We do however include data about labour income to assess the effects of edit checks on derived measures of total household income.

In the original version, respondents are shown a series of four showcards, listing 34 potential income sources, and asked which of these they have received during the reference period. Fieldwork takes place between September and January each year. Respondents are asked to report about the period since the start of fieldwork in the previous year. This means that, depending on the month in which a respondent is interviewed, the recall period covers between 12 and 16 months: *"Please look at this card and tell me if, since September 1st <previous calendar year>, you have received any of the types of income or payments shown, either just yourself or jointly?"* For each income source reported, respondents are then asked a series of follow-up questions about the timing and amounts of receipt: *"And for which months since September 1st <previous calendar year> have you received <source>?"*, *"How much was the last payment of <source> you received?"*, and *"What period did that cover?"*

From 2005 the BHPS added within-wave edit checks for those cash transfers, for which questions earlier in the same interview predict eligibility: Pension Credit, Disability Benefits, Income Support, Jobseeker's Allowance, Child Benefit and Housing Benefit. For example, respondents above the state retirement age who have not reported a state pension are asked *"Can I just check, do you currently receive the State Retirement Pension?"*

From 2006 onwards, reactive dependent interviewing (RDI) was added for all non-labour income sources (listed in Subsection 2.4). Respondents are first asked the original question. For any income sources reported in the previous but not the current interview, they are asked a follow-up question: *"Can I just check, according to our records you have in the past received <source>. Have you received <source> at any time since <date of interview>?"* (see Jäckle et al. 2007).

Although the BHPS data are not experimental, the public release file identifies which income sources were reported in response to the initial question, which in response to the within-wave edit check, and which in response to the RDI follow-up question. This enables a quasi-experimental comparison of the effects of the interviewing method on responses and estimates.

2.2. The Experimental Validation Study

The experimental study was carried out using the former European Community Household Panel (ECHP) low-income subsample for Great Britain. This sample was surveyed as part of the BHPS (using the BHPS survey procedures and questionnaires) from 1997 until

funding expired in 2001. In 2003 the sample was interviewed once more for methodological purposes. The experimental survey included a split-ballot experiment comparing independent and dependent interviewing for various sections of the questionnaire. In addition, respondents were asked for permission to link to their records on receipt of 17 different state cash transfer programmes held by the Department for Work and Pensions (the department in charge of administering cash transfers). The transfer programmes included Child Benefit, Housing Benefit, Working Families' Tax Credit, different types of Disability Allowances, Income Support, Jobseeker's Allowance and State Pensions.

The response rate for the experimental survey was 89% (N=1,033, RR1–AAPOR 2011), of which 77% gave consent for the record linkage (Jäckle et al. 2004), of which 74% were successfully linked. A related study by Jenkins et al. (2006) found that households that had reported receipt of means-tested state cash transfers in a previous wave of the survey were more likely to consent to the data linkage, but that household income was not related to consent. The linkage was performed independently five times using deterministic (exact) matching on National Insurance Number (the UK social security number) or sex with two or three out of date of birth, postcode, first line of address, first name, and family name (see Jenkins et al. 2008 for details on the linkage methodology). Results for each respondent were pooled to identify a single match. For 12 of the 14 respondents who were matched to more than one person in the administrative records, the modal match (which matched on at least three of the five criteria) was used as the correct match. The other two cases were inspected visually to determine the correct match. Although some problems with the linkage variables cannot be excluded, Jenkins et al. (2008) suggested that those not linked were probably respondents who had not received state cash transfers during the time frame of interest. The authors estimated that the true non-match rate was about one quarter, since 29% of respondents never reported receiving any of the relevant state cash transfers in any of the annual interviews between 1999 and 2003.

In the experimental survey, three versions of questions on non-labour income components were randomly assigned: independent interviewing (INDI, N=348 respondents of which N=262 consented to the record linkage), reactive dependent interviewing (RDI, N=344 respondents of which N=274 consented) and proactive dependent interviewing (PDI, N=341 respondents of which N=263 consented). With PDI respondents were reminded upfront of each source they had reported in the previous interview, and asked whether they had received the source since. Since the BHPS uses RDI for the income questions, the analyses presented here focus on the comparison of INDI and RDI. The INDI version used the original BHPS question, as described in Subsection 2.1. The RDI version had an added edit check question, again as described in Subsection 2.1.

Respondents in both experimental conditions were asked the same series of follow-up questions, described in Subsection 2.1, about the timing and amounts of each income source. The administrative records contain information about which of the 17 potential state cash transfers listed above each respondent has received, including the exact start and end dates of receipt and weekly amounts received. The data stem from the database system used by the state to administer transfer payments and are generated in the process of payments being made. This means that the administrative data reflect the actual dates and amounts of payments and can therefore be considered high quality (except for Housing Benefit data which are from decentralized databases and less reliable). A few transfer

types included in the survey are not included in the records (Widowed Mother's Allowance, War Disability Pension, Council Tax Benefit). Some cash transfer types (Disability Living Allowance, Child Benefit) are recorded as a single source, while the survey collects separate information about different components (e.g., care component vs. mobility component). For comparability, we derived variables from the experimental survey data that reflect the data structure and definitions of the record data.

2.3. Comparability of the BHPS and Experimental Survey Data

Although the survey data from the experimental study and the BHPS are based on the same design, there are several differences between the surveys which are relevant to our analyses:

(1) Time frames: The BHPS data are from 2005, 2006 and 2007, while the experimental survey data are from 2001 and 2003.
(2) Sample composition: the BHPS is a general population sample, while the experimental survey data overrepresent low-income households and may be affected by selection bias due to non-consent to linkage.
(3) Dependent interviewing method: the BHPS used RDI for all sample members in 2006 and 2007, while the experimental survey used INDI in 2001 and experimentally allocated respondents to a DI treatment in 2003.
(4) Within-wave edit checks: the BHPS used within-wave edit checks for questions on cash transfer receipt in 2005, 2006 and 2007 surveys, while the experimental survey did not use any within-wave edit checks.

These differences between the survey data from the BHPS and the survey data in the experimental validation study mean that it is not clear a priori whether the results from the validation study are likely to apply to the BHPS survey. We report on additional analyses we have carried out to verify the comparability of data from the two surveys in the discussion in Section 4.

2.4. Data Description

For analysis purposes, we group the income sources into four components of non-labour income: State cash transfers, private pensions, other transfers, and investments. This grouping corresponds to the derived income components provided with the BHPS public release file and consists of the following income components:

(1) State cash transfers: four types of national insurance pensions and tax credits, ten types of disability-related cash transfers and tax credits, two types of income support, Housing Benefit, Council Tax Benefit, Jobseeker's Allowance, Child Benefit, Maternity Allowance, Working Families' Tax Credit, Child Tax Credit.
(2) Private pensions: three types of private pensions.
(3) Other transfers: education grants, sickness insurance, maintenance/foster allowance, payments from trade unions/friendly societies, payments from absent family members, other payments.
(4) Investment income: rent from boarders/lodgers, rent from other properties.

Labour income also contributes to household income. We do, however, not examine this component separately, because edit check questions were not used for the collection of labour income data.

3.   Results

3.1.   *Effects of Edit Checks on Survey Estimates*

The number of income sources reported in the BHPS is documented in Table 1, for the 2005 survey (Wave 15), 2006 (Wave 16), and 2007 (Wave 17). In Wave 16, for example, respondents reported receipt of a total of 8,170 state cash transfers when asked the original BHPS question. Respondents for whom information collected earlier in the interview suggested that they might be eligible for additional cash transfers were then asked the within-wave edit check, whereupon they reported a further 165 sources. Finally, all respondents were queried about sources they had reported in the previous interview using the RDI edit check question, whereupon they reported a further 615 income sources. As described in Subsection 2.1, the within-wave edit check questions were only used for state cash transfers. For the other types of non-labour income only the RDI edit check was used.

   The results suggest that RDI edit checks were more effective at increasing reporting of income sources than within-wave edit checks. Depending on the type of income and the survey year, about 1-2% of total income sources were reported in response to the within-wave edit checks, whereas 5-12% were reported in response to RDI edit check questions.

   The sample sizes in the experimental survey and validation data are documented in Table 2. Of the sample allocated to INDI, 262 respondents consented to the linkage. These respondents reported a total of 338 state cash transfers in the 2003 survey, while the administrative records for these respondents, corresponding to the same time period, list 374 cash transfers. The fact that respondents in aggregate reported fewer income sources in the survey than they received according to the records is a first indication of underreporting. In contrast, respondents allocated to RDI reported 401 state cash transfers with 407 recorded in the administrative data, suggesting that RDI improved the aggregate reporting of income sources.

Table 1.   Number of income sources reported in the BHPS

		Cash transfers	Pensions	Other transfers	Investment
Wave 15	INDI	8,088	1,717	426	274
	WWEC	117 (1.4%)	–	–	–
Wave 16	INDI	8,170	1,776	515	323
	WWEC	165 (1.8%)	–	–	–
	RDI	615 (6.9%)	121 (6.4%)	55 (9.6%)	39 (10.8%)
Wave 17	INDI	7,895	1,846	501	302
	WWEC	157 (1.8%)	–	–	–
	RDI	506 (5.9%)	94 (4.8%)	49 (8.9%)	42 (12.2%)

Notes: Number of respondents in Wave 15: 8,538; Wave 16: 8,484; Wave 17: 8,322.
INDI: independent interviewing, WWEC: within-wave edit check, RDI: reactive dependent interviewing. Percentages represent the percent of total income sources of a given type reported in response to the WWEC or RDI.

Table 2. *Number of income sources in the experimental validation data*

Experimental treatment group	Cash transfers	
	in records	in survey
INDI	374	338
INDI+RDI	407	401

Notes: INDI: independent interviewing, RDI: reactive dependent interviewing.
Based on 2003 survey.

3.1.1. Effects on the Distribution of Household Income

To examine whether edit checks affect estimates of household income, we use Waves 15 to 17 of the BHPS. Table 3 shows estimates of the equivalised annual household income distribution for the population of Great Britain. The estimates are based on all members of surveyed households, adjusted for differences in household size using the McClements equivalence scale (Taylor et al. 2009) and weighted for nonresponse. The first column indicates the estimated cut-off points between percentiles of the income distribution, including only amounts associated with income sources reported in response to the INDI questions. The income measures based on INDI include imputed values if the receipt status, amount received or dates of receipt are missing for any of the income sources (Taylor et al. 2009). The second column indicates by how much the income percentile changes when income sources reported in response to the within-wave edit checks are included. For Waves 16 and 17, the third column indicates by how much the INDI estimate changes if sources reported both in response to the within-wave edit checks and the RDI follow-up questions are included.

Within-wave edit checks have a considerable effect, increasing estimated income percentiles below median income, for example increasing household income for the fifth percentile by 6% at Wave 16. RDI has an additional effect, increasing the fifth percentile by a further four percentage points to 10%. The effects of RDI and edit checks are largest for people in the lowest percentile, fall monotonically across percentiles, and are zero or close to zero for all percentiles above the median. The effect on median income is small: when sources reported in response to the within-wave edit checks are included, the estimated median increases by less than 0.3% in each of the three waves, and by a further 1% at Waves 16 and 17 when responses to RDI are included.

Edit checks on non-labour income sources therefore increase estimates of household income at the lower end of the income distribution, where non-labour income from cash transfers, pensions, and other transfers represents a major component of total income. For households with higher levels of income, these sources are less important, while non-labour income from investments may contribute a large part of total income. Nonetheless, the edit checks do not have any effect at the upper tail of the income distribution.

3.1.2. Effects on Estimated Poverty Rates

To examine whether edit checks affect estimated poverty rates, we again use Waves 15 to 17 of the BHPS. Replicating the official UK poverty definition, we define the poverty threshold as 60% of median household income: any individual living in a household with

Table 3. Distribution of equivalised annual household income

Percentile	Wave 15		Wave 16			Wave 17		
	INDI (£)	INDI+ WWEC (% Δ)	INDI(£)	INDI+ WWEC (% Δ)	INDI+ WWEC+ RDI (% Δ)	INDI (£)	INDI+ WWEC (% Δ)	INDI+ WWEC+ RDI (% Δ)
1	1609	49	436	183	210	842	95	137
2	3740	6	2780	31	35	3205	26	32
5	6047	5	6073	6	10	6385	4	7
10	8353	1	8549	2	6	8630	2	6
25	13594	0	13881	0	3	13847	1	4
50	25192	0	25267	0	2	25594	0	1
75	40921	0	41106	0	1	42472	0	1
90	57812	0	58602	0	0	61828	0	0
95	71091	0	73977	0	0	75365	0	0
98	89872	0	93273	0	0	96199	0	0
99	107793	0	109815	1	1	115145	0	0

Notes: Based on all BHPS enumerated household members, Wave 15: 11,700; Wave 16: 11,611; Wave 17: 11,374. Data are weighted for nonresponse and initial sample selection probabilities. INDI: Income derived from independent interviewing questions only, INDI+WWEC: INDI plus within-wave edit checks, INDI+WWEC+RDI: INDI+WWEC+ reactive dependent interviewing. (£): annual equivalised household income. (% Δ): percentage change compared to INDI only.

Table 4. Poverty rates (%)

Wave	Interviewing method	'Poor'	INDI: 'poor' Edit check: 'not poor'	INDI: 'not poor' Edit check: 'poor'
15	INDI	18.6	–	–
	INDI+WWEC	18.5	0.8	0.0
16	INDI	18.9	–	–
	INDI+WWEC	18.8	0.9	0.1
	INDI+WWEC+RDI	18.4	4.2	0.4
17	INDI	18.4	–	–
	INDI+WWEC	18.2	1.2	0.0
	INDI+WWEC+RDI	17.9	3.9	0.3

Notes: Based on all weighted BHPS enumerated household members, Wave 15: 11,700; Wave 16: 11,611; Wave 17: 11,374. INDI: independent interviewing. WWEC: within-wave edit check. RDI: reactive dependent interviewing.

less income is classified as poor. Official poverty statistics use 60% of current income to define the threshold for poverty (Brewer et al. 2009). We use annual income instead, in order to examine the net effects of edit checks on all questions related to household income, including questions about the timing of receipt during the year. In addition, Böheim and Jenkins (2006) show that there are few differences between poverty indicators based on current and annual income.

The results in Table 4 suggest that the edit checks somewhat reduce estimated poverty rates, but the effects are small: both in Waves 16 and 17, adding responses to edit checks and RDI reduced the poverty rate by 0.5 percentage points. Nonetheless, some individuals are classified differently depending on the interviewing method. For example, in Wave 16, 4.2% of individuals are classified as 'poor' based on the INDI questions, and as 'not poor' when the income sources reported in response to the edit checks are added. Similarly, the third column shows that 0.4% of individuals classified as 'not poor' with INDI are classified as 'poor' when information from the edit checks is added. These are probably households whose income is only just above the poverty threshold based on the INDI data, and who did not report any additional income sources in response to the checks or RDI. Since median income, and therefore also the poverty threshold, increases slightly when the edit check responses are included, these respondents slip just below the poverty threshold.

3.1.3. Effects on Estimated Poverty Transitions

To examine whether edit checks affect the longitudinal consistency of poverty classifications across waves, we again use BHPS data. Table 5 shows the transitions in poverty status between Waves 15 and 16, and Waves 16 and 17, based on the INDI data only, adding the within-wave edit check data, and further adding the RDI data. The edit checks have little effect on transition rates in both wave pairs: in the INDI data about 76% of individuals were living in non-poor households in both waves, 13% were poor in both waves, around 5% entered poverty and a further 5% exited poverty from one wave to the next. These estimates are similar when data from the within-wave edit checks and RDI are added. The lack of effects is surprising, since we would have expected RDI to increase the

Table 5.   Transition rates into and out of poverty (%)

Wave	Transition type	INDI	INDI+WWEC	INDI+WWEC+ RDI
15-16	Persistent non-poor	76.3	76.4	–
	Persistent poor	13.1	13.0	–
	Transition into poverty	5.6	5.7	–
	Transition out of poverty	4.9	5.0	–
16-17	Persistent non-poor	76.3	76.5	77.0
	Persistent poor	13.2	13.2	12.9
	Transition into poverty	5.0	4.8	4.8
	Transition out of poverty	5.5	5.5	5.3

Notes: Based on all weighted BHPS enumerated household members, Wave 15-16:10,278; Wave 16-17: 9,692. INDI: independent interviewing. WWEC: within-wave edit check. RDI: reactive dependent interviewing.

consistency of responses across waves, and by implication to reduce changes in household income and resulting changes in poverty status across waves.

In sum, both within-wave edit checks and RDI increase estimates of household income at the lower end of the distribution, but neither method has much effect on poverty classifications or transitions. The next section examines whether the changes in household income reflect an improvement in data accuracy.

3.2.   Effects of RDI and Edit Checks on Measurement Errors

We use the validation data to examine various aspects of measurement error related to the estimates presented in Subsection 3.1. We examine measurement error in receipt status, amounts of income, duration of receipt, and transitions in receipt status between waves. For each of these aspects we examine the extent of measurement error with independent interviewing, and how this changes with RDI edit checks. Note that as described in Subsection 2.1, RDI was only applied to questions about receipt, not to questions about dates or amounts of receipt. The RDI edits on the receipt questions nonetheless affect responses to the amounts and duration questions, since sources that are not reported by default have zero amounts and durations associated with them. The analyses of the experimental validation data are unweighted.

We expect the changes in responses with RDI to reflect a reduction in the various aspects of measurement error, and therefore expect the changes in estimates in Subsection 3.1 to reflect improvements in data accuracy.

3.2.1.   Effects of RDI on Measurement Error in Receipt of Income Sources

We first examine the effect of RDI on measurement error in individual reports of non-labour income *receipt*. We compare responses to the experimental survey with individual register data. For each potential income source, we derive indicators of whether or not the source was received at any point during the reference period. Separate indicators are derived for the survey and the record data and used to classify all potential income sources for each respondent: *true negatives* are income sources which were neither received according to the survey, nor according to the records; *true positives* are income sources which were received both according to the survey and the records; *false negatives* are

Table 6.   Effect of RDI on measurement error in income receipt reported by individuals

	Sample sizes (N) and row percentages (in brackets)				Error rates (%)	
	True negative	False negative	False positive	True Positive	False negative rate	False positive rate
INDI	3257 (88.8%)	73 (2.0%)	26 (0.7%)	312 (8.5%)	19.0	0.8
RDI	3377 (88.0%)	58 (1.5%)	30 (0.8%)	371 (9.7%)	13.5	0.9

Notes: The sample includes all respondents (INDI=262, RDI=274), multiplied by 14 potential income sources. Columns are defined in the text. INDI: Independent Interviewing, RDI: Reactive Dependent Interviewing.

income sources which were received according to the records, but not reported in the survey; *false positives* are income sources which were not received according to the records, but reported in the survey.

To account for the possibility that respondents may report income sources which are recorded in the name of a different household member in the record data, income sources are counted as 'true positives' if there is a record for the source in the name of another household member. This was the case for 3% of income sources reported by the INDI sample, and 5% of sources reported by the RDI sample. Table 6 indicates the number of potential income sources which are classified as true/false positives/negatives. Assuming that the record data represent the true values, we interpret 'false negatives' as indicators of underreporting, and 'false positives' as overreporting. The last two columns indicate the corresponding error rates: the false negative rate is the number of false negatives as a proportion of sources received according to the records; the false positive rate is the number of false positives, as a proportion of the sources not received according to the records.

The results indicate that the main type of error is underreporting: with INDI 19.0% of sources recorded in the records are not reported in the survey, while overreporting hardly occurs (less than 1%). RDI reduces the false negative rate to 13.5% and does not have any effect on overreporting. The increase in the reporting of income sources with RDI therefore represents a reduction in net measurement error in receipt of non-labour income sources.

3.2.2.   Effects of RDI on Measurement Error in the Amounts of Non-Labour Income

Second we test the effects of RDI on measurement error in the *amount* of income, again comparing the survey reports to the individual records. For each source we derive the amount of the last payment during the reference period according to the survey and according to the records. The amounts are standardised to weekly amounts, for comparability with the format in which they are recorded in the administrative data. We then calculate the error in amounts of receipt as the difference between the survey and the record. In the final step, we calculate the mean error over all cash transfers and respondents. The analysis includes sources reported either in the survey, or the records, or both. Housing benefits were excluded in this step as we found large irresolvable consistencies between the records and survey data.

With INDI, weekly non-labour income is underreported by £4.60 on average (95% confidence interval (CI) from −9.1 to −0.2). With RDI the error increases to £5.90

(95% CI from − 9.9 to − 1.9). This suggests that although RDI reduces underreporting of receipt, it does not help respondents report the amounts received.

3.2.3.   Effects of RDI on Measurement Error in Duration of Receipt

Third, to assess the effects of RDI on measurement error in reported *duration* of receipt, we again compare the survey and administrative data. For each income source we calculate the error as the difference between the number of months of receipt according to the survey and the records. The analysis is restricted to receipt between September 1st 2001 and September 1st 2002, for comparability with the BHPS data. The base includes all sources either reported in the survey, or recorded in the administrative data, or both, but excludes true negatives. In the case of overreporting where a record exists in the name of a different household member, the survey duration is compared to the record duration for the other household member. We then calculate the mean error over all income sources and respondents.

   With INDI receipt is underreported by 1 month on average (95% CI from − 1.4 to − 0.4). With RDI the mean error is no longer significantly different from zero (95% C.I. from − 0.4 to 0.4). This suggests that RDI reduces measurement errors in reported duration of receipt of cash transfers.

3.2.4.   Effects of RDI on Measurement Error in Transitions of Cash Transfer Receipt Across Waves

Fourth, we evaluate whether RDI reduces measurement error in reported *transitions* of receipt across waves. We classify each potential income source for each respondent according to the type of transition between the 2001 survey and the 2003 survey as continued non-receipt, continued receipt, transition off receipt, and transition onto receipt. Each potential income source is classified separately based on the survey data and the record data. We then pool the results for all income sources and compare the transition types derived from the survey and records to identify errors in transition classifications.

   Overall, the transition type is misclassified for 4% of potential income sources with both INDI and RDI. Since RDI was only used in the 2003 interview, the interviewing method cannot have affected the wave 2001 status. Therefore Table 7 focuses on errors in the classification of transition types, conditional on the 2001 status being reported correctly in the survey. The rows indicate the respondents' transition statuses (pooled over all potential income sources) according to the records. The columns indicate the percentage of income sources for which the 2003 status was misclassified in the survey, resulting in an error in the transition type.

   'Continued non-receipt' is reported well with INDI and not improved with RDI: the error rates are 0.5% with both methods. With RDI more respondents are correctly classified as having 'continued receipt': the error rates are reduced from 11.3% to 3.2%. However, respondents who 'transitioned onto' cash transfer receipt are more likely to be misclassified with RDI: the error rate unexpectedly increases from 20.4% to 38.6%. Further investigation (not shown in the table) suggests that RDI respondents who transitioned onto receipt are more likely to be misclassified as 'continued non-receipt'. This is a surprising finding, since non-receipt in the previous interview does not trigger any RDI questions. 'Transitions off' cash transfer receipt tends to be reported correctly with

Table 7. *Effect of RDI on measurement error in transitions onto and off cash transfer receipt, conditional on correct classification in the 2001 survey*

Transition in Records	INDI			RDI		
		N mis-classified	% mis-classified	N	N mis-classified	% mis-classified
Continued non-receipt	3230	16	0.5	3348	15	0.5
Transition on	49	10	20.4	57	22	38.6
Continued receipt	284	32	11.3	317	10	3.2
Transition off	16	0	0.0	15	2	13.3

Notes: The sample includes all respondents (INDI=262, RDI=274), multiplied by 14 potential income sources, and excluding sources reported incorrectly in the first interview (INDI=89, RDI=99). INDI: independent interviewing, RDI: reactive dependent interviewing.

INDI, but the number of transitions is very small. With RDI however 13.3% of transitions are misclassified, mostly as 'continued receipt'. This could be due to respondents falsely confirming a receipt status presented to them from the previous interview. A potential cause of the findings for transitions onto and off receipt might be found with the interviewers. With DI designs, they might be more focused on reducing errors in continued receipt than on picking up transitions onto and off receipt (Sala et al. 2009). Since the number of transitions onto and off receipt is small, we would however interpret these results with caution.

In sum, RDI reduces various aspects of measurement error in the reporting of state cash transfers: RDI reduces underreporting of any receipt, of the duration of receipt within one wave, and of continued receipt across waves. RDI does not reduce overall misclassification rates in transitions, although the nature of misclassifications changes.

4. Discussion and Conclusion

4.1. Summary of Results

The motivation for this study was to examine what effect methodological innovations that are expected to reduce measurement error have on substantive estimates. In this case, substantive conclusions are affected by whether or not edit checks are used to collect income data. Methodological studies designed to evaluate the effects of alternative data collection methods on data quality often only examine answers to individual survey questions. Evaluations of the impact on data quality however further need to relate to the actual uses of the survey data. In this spirit, we examine the effects of within-wave edit checks and RDI on derived estimates, and subsequently whether these effects reflect a decrease in measurement error. For this purpose we exploit a unique combination of data sets: we use data from the BHPS, a large-scale panel survey which has implemented within-wave edit checks and RDI for questions on non-labour income components in a quasi-experimental way, and from an experimental validation study based on the BHPS survey design.

We use the experimental study to assess the effects of RDI on different aspects of measurement error, and the BHPS data to assess the effects of RDI and within-wave edit

checks on estimates of household income and poverty. The results suggest that both the within-wave edit checks and RDI increase estimates of total household income in the lower tail of the income distribution. Neither method has much effect on estimated poverty rates or estimated rates of transitions into and out of poverty. The increase in household income reflects an increase in data accuracy: RDI reduces underreporting without affecting overreporting; RDI reduces underreporting of months of receipt and reduces erroneous transitions off income receipt and underreporting of continued receipt across waves.

In our view, the effects of RDI on measurement error are considerable; for example, the underreporting rate is reduced by about 29% compared to independent interviewing. The effects on estimates of household income and poverty are arguably small. This suggests that while within-wave edit checks and RDI may have large effects on measurement error in responses to individual survey questions, the combined effects, in this case over different survey items and different household members, may be small. This conclusion may however be open to interpretation, since a reduction in the estimated poverty rate by a mere 0.5 percentage points affects around 300,000 individuals in the population of Great Britain.

4.2. Limitations

Our study has several limitations. First, we only examined the effects of edit checks on non-labour components of household income. Non-labour income makes up almost 90% of total gross household income for the lowest income quintile. The higher the quintile however, the less important non-labour income is; non-labour income comprises 65% of total income for the second income quintile, 39% for the third, 18% for the fourth and 9% for the highest quintile (Brewer et al. 2009). We cannot draw conclusions about the likely effects of the introduction of edit checks on labour income, because the nature of measurement error in labour and non-labour income is quite different. With non-labour income, measurement error is mainly in the form of underreporting: respondents fail to report receipt of a source, and by default the amount received is set to zero. Overreporting is rare (see Table 6 and Bound et al. 2001). With labour income, respondents are less likely to underreport receipt, as they are unlikely to underreport being in work. Instead, measurement error occurs in the amount of earnings. Errors tend to be negatively correlated with true values and to cancel out across respondents (Bound et al. 2001). The nature of measurement error has implications for the potential effects edit checks can have. For non-labour income sources, we expect edit checks to reduce underreporting and therefore to increase estimates of household income for households for which non-labour income sources represent a major proportion of total income. This corresponds to our findings. For labour income, we would expect edit checks on average to decrease the earnings reported by those with low earnings, and to increase the earnings reported by those with high earnings. As a result, we would expect estimates of household income to change at both ends of the distribution, but because labour income comprises a larger part of total income, we expect the largest changes for the higher quintiles. RDI edit checks have been used to query changes in earnings in the Survey of Labour and Income Dynamics (SLID). Hale and Michaud (1995) reported that 8.3% of respondents reported

earnings that differed by more than $\pm 10\%$ and were then asked an RDI edit check. Two thirds of respondents confirmed the change as true. The experimental study we use in this article also included a test of RDI edit checks for earnings. In this study, 59% of respondents reported a change in earnings larger than $\pm 10\%$ and were asked the edit check. All but one confirmed that the change was true (Jäckle 2009). It is not clear from these studies however what the effect of the edit check for amounts of labour income on estimates of household income is.

Second, our validation study is limited in that it contains only data on state cash transfers (and not on the other non-labour income sources for which RDI is used in the BHPS), and in that the study contrasted only INDI and RDI (and did not use within-wave edit checks). To check whether it is reasonable to assume that the findings from the validation study also apply to the BHPS, to the other income sources, and maybe also to the within-wave edit checks, we have carried out some further analyses (see Lugtig and Jäckle 2011). We compared 1) the effects of RDI in the BHPS versus the validation survey, 2) the effects of RDI on reporting of state cash transfers versus other non-labour income sources in the BHPS, and 3) the effects of RDI versus within-wave edit checks in the BHPS. The results suggest that the effects of the questioning method are similar for these three comparisons. We therefore assume that the effects of RDI on measurement error in the validation study are also likely to apply to the BHPS data. We further assume that for those income sources for which we have no validation data, the changes in responses also reflect a reduction in measurement error. Finally we assume that the changes in responses due to the within-wave edit checks also reflect reduced measurement error. As a result, we assume that the changes we find in estimates related to household income represent improvements in data accuracy.

Third, in this article we do not investigate which specific types of income sources are most likely to be misreported, or which types of respondents are most likely to misreport. Lynn et al. (2012) examined the same validation study we use here and reported error rates for the six most common state cash transfer programmes. Among these, underreporting rates were highest for Incapacity Benefit (50%), followed by Tax Credit (29%), Child Benefit (23%), Housing Benefit (17%), Income Support (11%), and lowest for Retirement Pension (0%). Overreporting rates were between 0% and 2% for all sources (rates derived from Table 3). Our own analyses for the same income sources on the BHPS data suggested that the extent of underreporting was generally lower and differences between individual sources were much smaller: underreporting was highest for Housing Benefit (17%), followed by Income Support (8%), Incapacity Benefit (7%), Tax Credit (5%), and lowest for Child Benefit (4%), and Retirement Pension (4%).

Lynn et al. (2006), using the validation data, also examined which types of respondents were most sensitive to RDI edit checks. Respondents who reported income sources only in response to the edit check were less likely to be retired (or born before 1943) or living with a spouse or partner, but more likely to be registered disabled than respondents who reported receipt in response to the independent questions. The authors found no differences by gender, whether in paid work, children in the household, qualifications, general health, duration living at the address, regular car, or mobile phone use (Table 6 in Lynn et al. 2006).

4.3. Future Research

There are a number of issues, regarding both the effects of RDI and within-wave edit checks and the mechanisms through which these methods work, which in our view warrant further attention. Reactive and proactive DI have rarely been compared. The reason why RDI was implemented in the BHPS was that this made it possible to maintain comparability with the previous 15 waves of data collection, in which independent interviewing was used. The responses given to the independent question can still be identified and for comparisons with previous waves the responses to the reactive follow-up can be ignored.

Our ability to compare the effects of within-wave edit checks and RDI were limited by the fact that they were always used in combination, with the edit checks always preceding the RDI checks. With this design, the RDI edits seemed more effective at increasing the reporting of income sources than the within-wave edit checks: depending on the wave and income type, 1-2% of total income sources were reported in response to the within-wave check, while 5-12% were reported in response to the RDI check. Their relative effects may be quite different if compared individually or in different order. Since the edit checks do not require feeding forward information from previous interviews, they can be used in cross-sectional surveys and are cheaper to implement than RDI. Their use is however restricted to income sources for which there are questions earlier in the questionnaire that are good predictors of eligibility.

The long-term effects of RDI have not been assessed. The ability of RDI to reduce underreporting is limited by the fact that the respondent can only be reminded of income sources reported in the past. Since RDI reduces underreporting in a given wave, this means that in the following wave a larger proportion of recipients can be reminded of income sources they have received in the past, increasing the effectiveness of RDI. As a result, over time RDI may decrease measurement error more than it does in a single wave. The effects of RDI across waves have not been assessed, as most previous studies have focused on the wave when dependent interviewing was first introduced. The effectiveness of the reminders depends on the quality of the first report and there is concern that dependent interviewing may lead measurement error to be fed forward into future waves. In the case of reporting of income sources, where the questions are about whether or not a source was received (yes/no), this risk is somewhat reduced. If the previous wave report was wrong and the respondent had underreported receipt, the RDI follow-up is not triggered in the following wave. That is, RDI simply has no effect. If the previous wave report is wrong because the respondent reported a source they had not actually received, the RDI follow-up would be triggered, leading to a risk that the respondent may continue to overreport receipt. The results of our validation study (see Table 6) however suggest that overreporting hardly occurs. Therefore we would conclude that underreporting in the previous wave reduces the effectiveness of RDI because the RDI check is not asked when it should be, but that overreporting in the previous wave does not impact the effectiveness of RDI.

The extent of measurement error in independent survey questions is presumably affected by the question format. The shortcut method of using showcards instead of separate yes/no questions about the receipt of all potential income sources presumably leads to more underreporting. On the other hand, the shorter interview time reduces respondent burden, which could lead to less measurement error using the showcards. This trade-off between cost savings in terms of questionnaire time and measurement error has not been assessed to our knowledge.

Finally, we have not touched on the question through which mechanisms RDI and edit checks work, that is, which types of sources are most likely to be misreported, by which types of respondents, and how the edit checks work for these different groups (see Lynn et al. 2012; Pascale et al. 2009). We have also not touched on the question how these methods could further be improved. Improvements could focus further on the reduction of underreporting, but also on capturing new receipt. This could be done by extending the use of within-wave edit checks by incorporating more factual questions into earlier sections of the questionnaire that predict eligibility for income receipt. Measurement error in household income was reduced by our study design, but there is room for further reductions in error with potentially greater impact on substantive conclusions.

5.   References

AAPOR (2011). Standard Definitions: Final Dispositions of Case Codes and Outcome Reports for Surveys, (7th edition). Lenexa, KS: American Association for Public Opinion Research. Available at: http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions2&Template=/CM/ContentDisplay.cfm&ContentID=3156 (accessed October 30, 2013).

Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement Error in Survey Data. In Handbook of Econometrics Vol. 5, J.J. Heckman and E. Leamer (eds). Amsterdam: Elsevier, 3705–3843.

Brewer, M., Muriel, A., Philips, D., and Sibieta, L. (2009). Poverty and Inequality in the UK: 2009. The institute for fiscal studies, London. Available at: http://www.ifs.org.uk/comms/c109.pdf (accessed October 30, 2013).

Böheim, R. and Jenkins, S.P. (2006). A Comparison of Current and Annual Measures of Income in the British Household Panel Study. Journal of Official Statistics, 22, 733–758.

Hale, A. and Michaud, S. (1995). Dependent Interviewing: Impact on Recall and on Labour Market Transitions. SLID Research Paper Series No. 95-06. Ottawa: Statistics Canada.

Jenkins, S.P., Cappellari, L., Lynn, P., Jäckle, A., and Sala, E. (2006). Patterns of Consent: Evidence from a General Household Survey. Journal of the Royal Statistical Society Series A, 169, 701–722. DOI: http://www.dx.doi.org/10.1111/j.1467-985X.2006.00417.x

Jenkins, S.P., Lynn, P., Jäckle, A., and Sala, E. (2008). The Feasibility of Linking Household Survey and Administrative Record Data: New Evidence from Britain. International Journal of Social Research Methodology, 11, 29–43. DOI: http://www.dx.doi.org/10.1080/13645570701401602

Jäckle, A. (2008). Measurement Error and Data Collection Methods: Effects on Estimates from Event History Data. ISER Working Paper 2008-13. Colchester: University of Essex. Available at: http://www.iser.essex.ac.uk/publications/working-papers/iser/2008-13.pdf (accessed October 30, 2013).

Jäckle, A. (2009). Dependent Interviewing: A Framework and Application to Current Research. In Methodology of Longitudinal Surveys, P. Lynn (ed.). Chichester: Wiley, 92–112.

Jäckle, H., Laurie, H., and Uhrig, S.C.N. (2007). The Introduction of Dependent Interviewing on the British Household Panel Survey. ISER Working Paper 2007-7. Colchester: University of Essex. Available at: http://www.iser.essex.ac.uk/publications/working-papers/iser/2007-07.pdf (accessed October 30, 2013).

Jäckle, A., Sala, E., Jenkins, S.P., and Lynn, P. (2004). Validation of Survey Data and Employment: the ISMIE Experience. ISER Working Paper 2004-14. Colcheste: University of Essex. Available at: http://www.iser.essex.ac.uk/publications/working-papers/iser/2004-14 (accessed October 30, 2013).

Lynn, P., Jäckle, A., Jenkins, S.P., and Sala, E. (2006). The Effects of Dependent Interviewing on Responses to Questions on Income Sources. Journal of Official Statistics, 22, 357–384.

Lynn, P., Jäckle, A., Jenkins, S.P., Sala, E. (2012). The Impact of Interviewing 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, 289-308. DOI: http://www.dx.doi.org/10.1111/j.1467-985X.2011.00717.x

Lugtig, P. and Jäckle, A. (2011). Can I just check. . . ? Effects of edit check questions on measurement error and survey estimates, ISER Working Paper 2011-23. Colchester: University of Essex. Available at: http://www.iser.essex.ac.uk/publications/working-papers/iser/2011-23 (accessed October 30, 2013).

Mathiowetz, N.A. and McGonagle, K.A. (2000). An Assessment of the Current State of Dependent Interviewing in Household Surveys. Journal of Official Statistics, 16, 401–418.

Moore, J., Bates, N., Pascale, J., and Okon, A. (2009). Tackling Seam Bias through Questionnaire Design. In Methodology of Longitudinal Surveys, P. Lynn (ed.). Chichester: Wiley, 72–92.

Pascale, J., Roemer, M.I., and Resnick, M. (2009). Medicaid Underreporting in the CPS. Results from a Record Check Study. Public Opinion Quarterly, 73, 497–520. DOI: http://www.dx.doi.org/10.1093/poq/nfp028

Pennell, S.G. (1993). Cross-Sectional Imputation and Longitudinal Editing Procedures in the Survey of Income and Program Participation. Washington, DC: US Bureau of the Census. Available at: http://www.census.gov/sipp/workpapr/wp9314.pdf (accessed October 30, 2013).

Sala, E., Uhrig, S.C.N., and Lynn, P. (2009). It is Time Computers Do Clever Things! The Impact of Dependent Interviewing on Interviewer Burden. Field Methods, 23, 3–23. DOI: http://www.dx.doi.org/10.1177/1525822X10384087

Taylor, M.F., Brice, J., Buck, N., and Prentice-Lane, E. (2009). British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices. Colchester: University of Essex. Available at: http://www.iser.essex.ac.uk/bhps/documentation/pdf_versions/volumes/bhpsvola.pdf (accessed October 30, 2013).