Does Encouraging Record Use for Financial Assets Improve Data Accuracy? Evidence from Administrative Data

Jonathan Eggleston and Lori Reeder[1]

April 25, 2017

Abstract

Many surveys ask respondents to consult financial records in order to improve data accuracy. However, the assumption that record use reduces measurement error has not been tested yet with a large-scale comparison to administrative data. In this project, we compare interest, dividend, and rental income in the Survey of Income and Program Participation (SIPP) to administrative IRS 1040 tax data. Our results show that record use is associated with reducing the discrepancy between survey and administrative data by approximately 19 to 43 percent. The effects from records use persists even when controlling for various measures of respondent motivation and precision. In terms of potential costs from encouraging record use, we find that record users spend an extra 4.02 seconds for each asset question, on average, after controlling for their behavior in other parts of the SIPP interview. The extra time per question translates to a 2.6% increase in the total duration of the interview. Finally, we discuss the implications these results have for designing how to ask respondents to consult records and whether these record use requests are beneficial in household surveys.

[1] SEHSD Working Paper Number 2017-34. SIPP Working Paper Number 277. Economist and Survey Statistician (respectively), Social, Economic, and Housing Statistics Division, U.S. Census Bureau. Email: Jonathan.S.Eggleston@census.gov and Lori.Beth.Reeder@census.gov. Mail: 4600 Silver Hill Road, Washington DC, DC 20233. Paper prepared for the 2016 International Conference on Questionnaire Design, Development, Evaluation, and Testing, November 9-13, 2016, Miami, Florida.

1. Introduction

Financial questions on household surveys are invaluable tools for generating estimates of average income and wealth for the general population. However, many respondents have trouble answering these questions because they are unsure about the exact details of their finances at the time of the interview. For example, interviewers for the Survey of Income and Program Participation (SIPP) have left comments in the survey instrument indicating that the respondent's answer was a guess or an approximation, or that another family member manages their finances. Comments like these suggest that such survey responses may be affected by measurement error, which could lead to problems in statistical analyses based on these data.

To mitigate measurement error, household surveys employ a variety of strategies to help respondents retrieve information. For example, respondents may be more aware of their finances from the past year when they are preparing their taxes, so conducting a financial survey around tax season may improve data quality. Another strategy is to encourage respondents to look at financial records (e.g., bank statements or tax returns) during the interview. This strategy is used by many government surveys, such as the SIPP, the Consumer Expenditure Survey (CEX), and the American Housing Survey.

However, there has been little research to determine whether or how much record use is associated with reduced measurement error, and much of the preexisting research has not found benefits of record use for reducing measurement error (e.g., Couper et al. 2013). In this paper, we conduct the first large-scale comparison of survey to administrative data to evaluate whether record use is associated with improved data quality. Specifically, we compare reported interest, dividend, and rental income in the Survey of Income and Program Participation to IRS 1040 tax return data linked at the individual level, and we estimate measurement error by the discrepancy between survey and administrative data. We expect that income from assets is particularly prone to measurement error as respondents are likely more familiar with earnings from employment or social security income than other sources. For example, if a

respondent's dividends are reinvested into their stock portfolio, then the respondent may not be aware of how much income from dividends they are receiving unless they look at tax records or other financial statements. In our regression analyses using the 2001-2014 SIPP panels, we find that record use is associated with a 19 to 43 percent decrease in measurement error.

One potential concern when interpreting our regression analyses is that record use could be associated with other unobserved characteristics that also affect measurement error, such as the respondent's motivation or conscientiousness.[2] To address this problem, we construct person-level variables on the average time spent per question, item nonresponse rates for financial questions, and the average amount of rounding in financial questions. These variables measure how precise the respondent's answers have been in the interview and how much time they spend thinking about each question. As these variables reflect measurement error during the course of the entire interview, they can be used as proxy variables for unobserved characteristics in the asset income measurement error regressions. We find that including these variables has no effect on the record use coefficient.

Although there may be potential benefits from encouraging record use, there could be potential costs in terms of respondent burden and paying interviewers for the time it takes respondents to find and analyze their records. To investigate this, we look at how record use is associated with the average time spent per question in the 2014 panel. In this panel, only the asset section contained an explicit prompt asking respondents to consult records. We analyze both the time spent on asset income and value questions, which are topics to which the record use request pertains, as well as time spent on other questions in the personal interview. When focusing on unconditional means, we find that record use is associated with an increase in time for the asset questions as well as an increase in time in other sections of the interview. In a regression analysis, we find that record use is associated with increased time on asset questions, after controlling for the average time spent on other questions as well as

[2]For example, Hedengren and Stratmann(2012) find that item-nonresponse is associated with conscientiousness.

respondent characteristics.[3] This evidence suggests that while the types of people who use records may be more engaged respondents who spend more time answering questions, record use is still associated with longer interview duration. Overall, we find evidence that record use is associated with lower measurement error but increased interview time. Thus, the decision about whether to encourage respondents to use records depends on weighing the benefits of increased data accuracy against the financial costs of increased interview length.

2. Background

There is a common belief in survey methodology that encouraging respondents to look at financial records will improve data quality (Couper et al. 2013). However, the literature on record use is relatively limited. Most of the work has been non-experimental technical papers which look at the effects of record use for a single survey. For example, Safir and Goldenberg (2008) find that record use is associated with higher total expenditures in the Consumer Expenditure Survey (CEX) and Edgar and Gonzalez (2009) find that data from record users requires less data editing (such as imputing values). Kashihara and Wobus (2006) compare data provided by households to data provided by medical providers in the Medical Expenditure Panel Survey and find that record use was associated with a closer match between the household and medical provider data. In addition, Eggleston (2015) finds that record users in SIPP reported less rounded numbers. That is, record users report numbers that have been rounded to a lesser degree than non-record users.

In these papers, the correlations between record use and measurement error may not be causal, as there could be other unobserved factors, such as respondent engagement, which may also affect data quality. For causal estimates, there are a few papers which conduct an experiment on asset record use. Couper et al. (2013) considers the outcomes of two experiments for the Health and Retirement Study

[3] We control for respondent characteristics including sex, race, Hispanic origin, marital status, average rounding, and item nonresponse rates.

(HRS) Internet Survey designed to encourage respondents to use records for asset questions. The authors found that these experiments resulted in an increase of the proportion of respondents who used records but did not significantly increase the precision of the data, as measured by the amount of rounding in the answers and by item-nonresponse rates. In another study, Maynes (1968) conducted an experiment on a sample of individuals with savings accounts and loans from the Census Federal Credit Union. In this mail survey, some respondents were randomly assigned a condition that instructed them to consult records. He finds that respondents instructed to consult records report more accurate values; nonetheless, the aggregate distribution of report assets was similar between the two experimental conditions. In other words, while non-record users gave less accurate values, these errors, in aggregate, did not decrease the accuracy of the estimated mean and standard deviation. Finally, Murphy et al. (2015) found that their experiment on record use in the Community Advantage Panel Survey increased the proportion of respondents who used records, but had no affect on the precision of reported values or their sample means.

The results from Couper et al. (2013) and Murphy et al. (2015) appear to suggest that record use does not improve data quality. However, there are caveats to these two experimental studies which should be considered. First, these studies use indicators of data quality, such as rounding, but these indicators may or may not be associated with actual measurement error. Second, the increase in the proportion of respondents who use records was relatively small (8 percentage points for Couper et al. (2013) and 7 for Murphy et al. (2015)). Thus, while they found their experimental manipulation increased record use, the effect size may have been too small to detect a significant change in data quality.

To build upon the previous work, we conduct the first large scale comparison of survey to administrative data to determine whether record use is associated with improved data quality. Although Maynes (1968) also used administrative data, our work provides two main benefits: 1) a much larger

sample size, allowing us to obtain more precise estimates; and 2) a sample that is nationally representative, allowing us to estimate the effects of record use for the general population. With this approach, we make the common assumption in the literature that the discrepancy between survey and administrative data represents measurement error in the survey data (e.g. Bound and Krueger 1991). However, there may be other reasons for the discrepancy, such as measurement error in the administrative data (Abowd and Stinson 2013), or a survey respondent being matched to the wrong administrative record. Thus, while the discrepancy between survey and administrative data may be a better indicator of measurement error than rounding or item nonresponse, it may still be an imperfect indicator of measurement error.

Another innovation of our paper is our approach for controlling for confounding factors associated with both record use and measurement error. Because consulting records in a survey is voluntary, it could be the respondents who use financial records are in general more aware of their finances and more accurate in their survey responses regardless of whether they consult records. To control for this effect, we use indicators of how accurate an individual's responses have been throughout the entire survey, as measured by person-level variables for the average time spent per question, item nonresponse rates for financial questions, and the average amount of rounding in financial questions. In other words, we use indicators of measurement error elsewhere in the survey as proxy variables for unobserved respondent characteristics in our regressions for measurement error in asset income. This approach does have limitations compared to an isolated experimental change, as our proxy variables may be imperfect controls for unobserved respondent characteristics. However, in the absence of experimental variation, our approach provides a new way to use additional data on respondent accuracy when investigating measurement error in a survey.

In addition, our project also sheds light on whether record use is associated with increases in interview length. Edgar (2010) finds that record use was associated with spending more time on various

sections in the Consumer Expenditure Survey interview, even for topics for which respondents likely did not consult records. In contrast, Laurie and Moon (2010) found that record use was not associated with interviewer perceptions of interview length. In our paper, we take a more systematic approach by analyzing how record use is associated with interview length for the relevant asset questions after controlling for the average time spent in other sections of the interview. Thus, our analyses are the first to examine how record use affects interview length after controlling for respondent characteristics.

3. Data

We use the 2001, 2004, 2008, and 2014 panels of the Survey of Income and Program Participation (SIPP) linked to IRS 1040 tax return data from 2000 to 2013. The SIPP is a longitudinal survey which interviews between about 30,000-44,000 households over a four year period.[4] The survey collects information about income, assets, labor market activity, and participation in government programs of U.S. households. Information on a wide variety of assets and debts is collected and includes variables covering financial information about savings accounts, checking accounts, retirement accounts, property values, and credit card debt.

Various changes have been made to the survey over time. In 2004, several changes were made to help respondents report income more accurately. For asset income, the survey gave respondents more flexibility to report income in either a monthly, quarterly, or annual format. This is in contrast to the 2001 panel, in which respondents, for the most part, were asked to add up their asset income from the previous four months. In the 2014 panel, numerous changes were made to the survey. The most substantial change was the introduction of the event history calendar (EHC), which is a visual method of collecting retrospective data on the timing of events, such as the loss of a job or health insurance coverage. In addition, respondents in the 2014 SIPP panel are interviewed less frequently in order to

[4] The sample size varies across panels.

reduce costs (U.S. Census Bureau, 2016). In the 2014 panel, respondents are interviewed once a year, but in earlier panels, respondents were interviewed every four months. For the asset section, questions on new types of assets were added, and the asset questions were rearranged in order to ask about the value and income of an asset consecutively.[5]

The SIPP and IRS data are linked together using the Census Bureau's Person Identification Validation System (PVS), as described in Wagner and Layne (2014). This procedure matches both survey data and administrative data to a master reference file. Individuals who are matched are given an identifier called a Protected Identification Key (PIK), which acts as an anonymized identification number that can be used to link administrative datasets and surveys. Because of incomplete or erroneous information in both the survey and administrative dataset, such as a misspelled name, not all individuals in SIPP and the IRS 1040 file are assigned a PIK. In addition, such erroneous information can also result in the PVS system assigning the wrong PIK to a SIPP respondent or a tax record, which would result in an incorrect match. Such match errors may also affect the match between survey and administrative data. Table 1 lists the proportion of individuals who are assigned a PIK as well as the match rates between SIPP and IRS 1040, conditional on the SIPP respondent having been assigned a PIK. A SIPP respondent may not be matched to a 1040 record if there is either an error in the match process or if the individual did not file taxes in a given year. Overall, about 79 percent of individuals in SIPP over age 15 are assigned a PIK, and among those who are assigned a PIK, about 81 percent are matched to a 1040 record. The 2001 panel is an exception in that the assigned PIK rate is much lower at 56.2 percent. Nonetheless,

[5] See Moore and Griffiths (2003) for more details about changes made in the 2004 SIPP panel and U.S. Census Bureau (2016) for changes made in the 2014 panel. In the 2014 panel, questions on mortgages owned as an asset and royalties were dropped because of low ownership rates. Because of this, our construction of the sum of SIPP asset income variables differs slightly in the 2014 panel.

among those who were assigned a PIK in the 2001 panel, about 82 percent are matched to a 1040 record.[6]

Table 1: PIK and 1040 Match Rates

Panel	Assigned PIK	Matched IRS1040, Given Assigned PIK
2001	56.2	81.9
	(0.57)	(0.23)
2004	80.8	84.0
	(0.18)	(0.11)
2008	88.9	79.0
	(0.15)	(0.16)
2014	91.9	77.4
	(0.17)	(0.24)
Overall	79.2	81.0
	(0.15)	(0.09)

Table 1 describes the proportion SIPP respondents who are matched to the IRS1040 dataset. Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. Sample consists of all individuals over age 15. The cross-sectional unit of analysis in this table is an individual. Standard Errors in Parentheses. Number of observations: 1,133,510.

For this project, the asset income variables we analyze are interest, dividend, and gross rental income. For these three types of income, equivalent variables are found both in SIPP and 1040 datasets. In our main analyses, the unit of observation is a tax-unit (either a single person or a married couple), in a given tax year. Appendix A provides more details about how we clean the datasets to generate a comparable unit of analysis. In particular, Appendix A gives details about how we combine data for married couples and clean data for SIPP respondents who are not in-sample for the full calendar year.

[6] In the 2001 panel, assignment of a PIK was based on respondents providing their social security number, which had high item-nonresponse rates. In the 2004 and subsequent panels, the revised matching algorithm no longer required the respondent to provide their social security number. In addition, the process for obtaining consent from respondents to match them to administrative records was changed from an opt-in to an opt-out procedure.

3.1. Record Use

The explanatory variable of interest is an indicator of whether the respondent used records during the interview. In SIPP, the procedure for requesting that the respondent consult records has changed over time. In the 2001 panel, interviewers read the following prompt before asking the questions on income:

….Since accuracy is important to this survey, it would be very helpful if you could refer to any income records you might have for the next series of questions. I would be happy to wait while you get them. Do you need a moment?

In the 2004 and 2008 panels, the prompt was changed to

…..We want to be both as accurate and as efficient as we can, so it would be very helpful if you could refer to any records you might have.

Unlike in the 2001 panel, interviewers were instructed to encourage respondents to get records at appropriate moments during the subsequent question, rather than right after the initial request for consulting records. In the 2014 SIPP panel, the text of the request remained the same for the most part, but was changed to pertain only to assets, and was placed right before the asset income and value section. In the 2001-2008 panels, the record use request was made before the questions on social assistance and employment income, which preceded the questions on asset income.

Over time, the proportion of SIPP respondents who use records may have been affected by both changes in the survey as well as changes in a respondent's perceived burden from complying with the record use request. To investigate these factors, Figure 1 plots the proportion of respondents who have used records in a given interview over time. Overall, Figure 1 provides evidence that more respondents used records in the 2001 and 2014 panels than in the 2004 and 2008 panels, with a sharp increase in the proportion for the 2014 panel.[7] This suggests that respondents are more likely to use records if they are

[7] Excluding the first and last years of the 2001, 2004 and 2008 panels, the proportion of respondents who used records in the 2001 and 2014 panels is higher than the proportion of respondents who used records in all years of the 2004 and 2008 panels. In addition, the proportion of the respondents who used records in the first wave of the 2014 panel is higher than the proportion in all years of the 2004 and 2008 panels.

either given more time to find records, or if the request has a more narrow focus and is made right before a given set of questions. Record use may also be higher in the 2014 Panel because many interviews are conducted around the time people are preparing their taxes, so they may have tax documents on hand. In addition, the recall period is longer in the 2014 Panel, so using records may reduce the cognitive burden from trying to remember asset income from the previous year.

Figure 1: Record Use Rates over Time

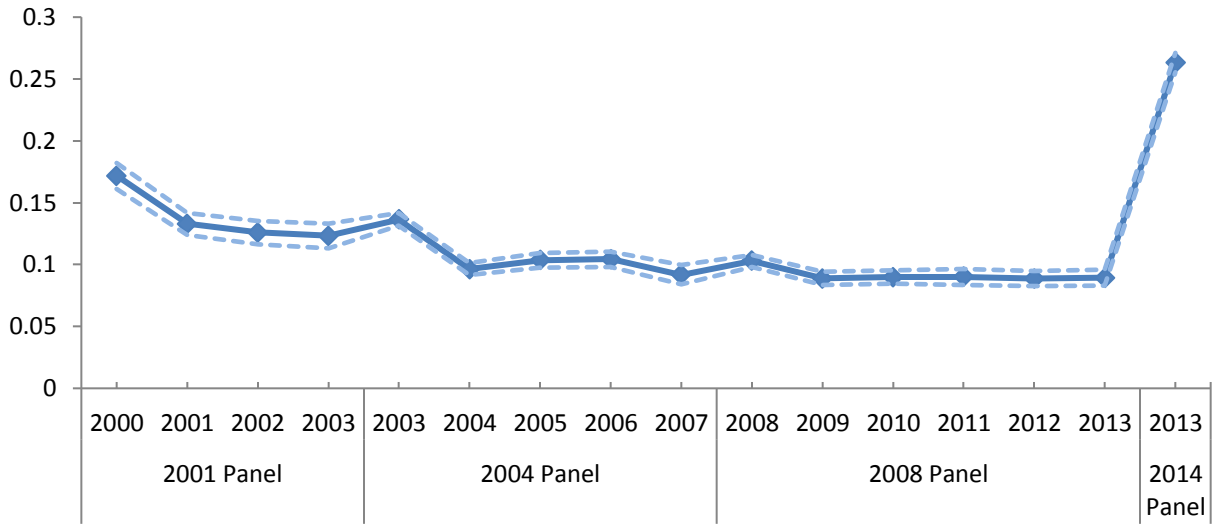


Figure describes the proportion of SIPP respondents who look at financial records for asset questions in a given interview. Source: 2001-2014 SIPP panels. Sample consists of all individuals over age 15 who owned an asset that generated interest, dividend, or rental income. The cross-sectional unit of analysis in this table is an individual. Dashed line shows 95% confidence interval. Number of observations over all years and panels: 607,832.

3.2. Description of Asset Measurement Error

In this subsection, we give an overview of the discrepancies between SIPP and IRS asset income. For expositional purposes, we present only the results for interest income in this section. The results for dividend and rental income are comparable, and are presented in Appendix A.8. First, we look at the probabilities of reporting interest income in one dataset but not the other. Table 2 shows the basic cross tabulations. In Table 2, we show that 25.2 percent report income on SIPP but not their tax returns. However, Figure 2 and Figure 3 suggest that much of the discrepancy is from people with low amounts of interest income. For example, the probability of reporting interest income on the IRS 1040 is only 41

percent when SIPP interest income is $20, but the probability jumps to 78 percent when SIPP interest income is $300. Figure 3 shows a similar pattern for reporting any SIPP interest income. While comparison studies between survey and administrative data usually assume the administrative data is more accurate, these results potentially suggest that individuals with low amounts of interest income are more likely to report these amounts in SIPP but fail to report them on their tax returns.[8]

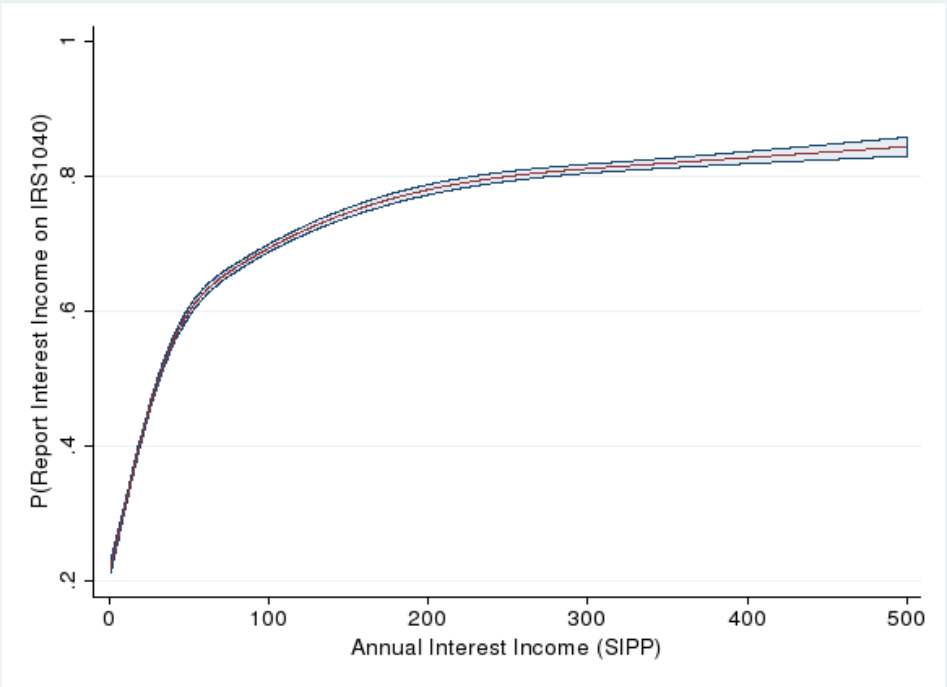
Table 2: Cross Tabulation for Reporting Interest Income

	Interest Income on IRS 1040	
Interest Income in SIPP	Yes	No
Yes	35.1	25.2
	(0.20)	(0.23)
No	7.9	31.8
	(0.12)	(0.23)

Table describes the proportion of SIPP respondents who report any interest income in either SIPP, the IRS 1040, or both. Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. Sample consists of all individuals over age 15. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Standard errors in parenthesis. Number of Observations: 515,732.

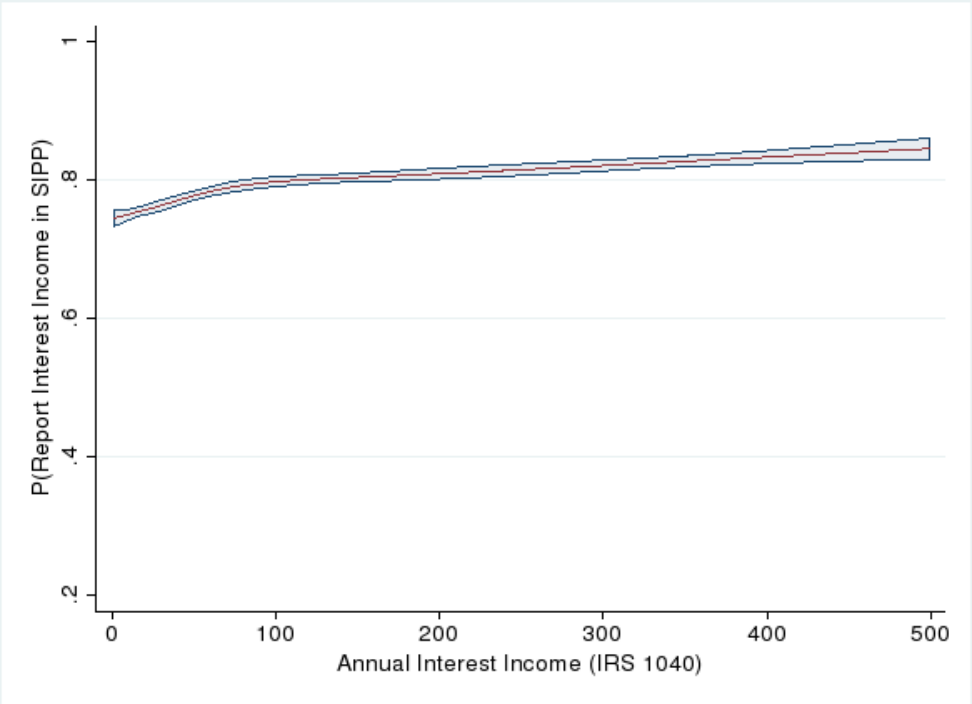
[8]A tax filer is supposed to report any interest income over 50 cents. A financial institution is only required to fill a 1099-INT for a customer with interest income over $10, so some individuals may forget to report interest income or may not know these rules if they do not receive a 1099-INT.

Figure 2: Probability Reporting Interest Income on Tax Return



Graph displays predicted values from a regression of an indicator of reporting interest income in the IRS1040 on a cubic spline of SIPP Interest Income with 5 knots. Source: 2004-2008 SIPP panels and 2004-2013 IRS 1040 datasets. The 2001 and 2014 panels were not used in this graph due to the different number of replicate weights in these panels. See Appendix A.7 for more details on replicate weights. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive interest income in SIPP. Shaded area represents 95% confidence interval. Number of Observations: 210,473.

Figure 3: Probability Reporting Interest Income in SIPP



Graph displays predicted values from a regression of an indicator of reporting interest income in SIPP on a cubic spline of IRS Interest Income with 5 knots. Source: 2004-2008 SIPP panels and 2004-2013 IRS 1040 datasets. The 2001 and 2014 panels were not used in this graph due to the different number of replicate weights in these panels. See Appendix A.7 for more details on replicate weights. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive interest income in IRS 1040. Shaded area represents 95% confidence interval. Number of Observations: 119,277.

Next, we look at the discrepancies between SIPP and IRS interest income, conditional on reporting income in both sources. Figure 4 plots the overall distribution of SIPP and IRS income as reported in 1040 records. This shows that the overall shape of the distributions are similar, but there are more people with low amounts of interest income in SIPP than in the corresponding 1040 records. To investigate discrepancies at the individual level, Table 3 shows percentiles and the mean of the difference in SIPP and IRS 1040 income as well as the absolute value of the difference. The median difference is near zero, but the mean difference is less than zero at negative $2,251, suggesting large underreporting of interest income is more prevalent than large overreporting. The statistics on the absolute value of the difference shows that there is a wide fluctuation in the match between SIPP and

IRS data. For example, the 10th percentile of the absolute difference is only $10 but the 90th percentile is $4,718.

Figure 4: Distribution of SIPP and IRS Interest Income

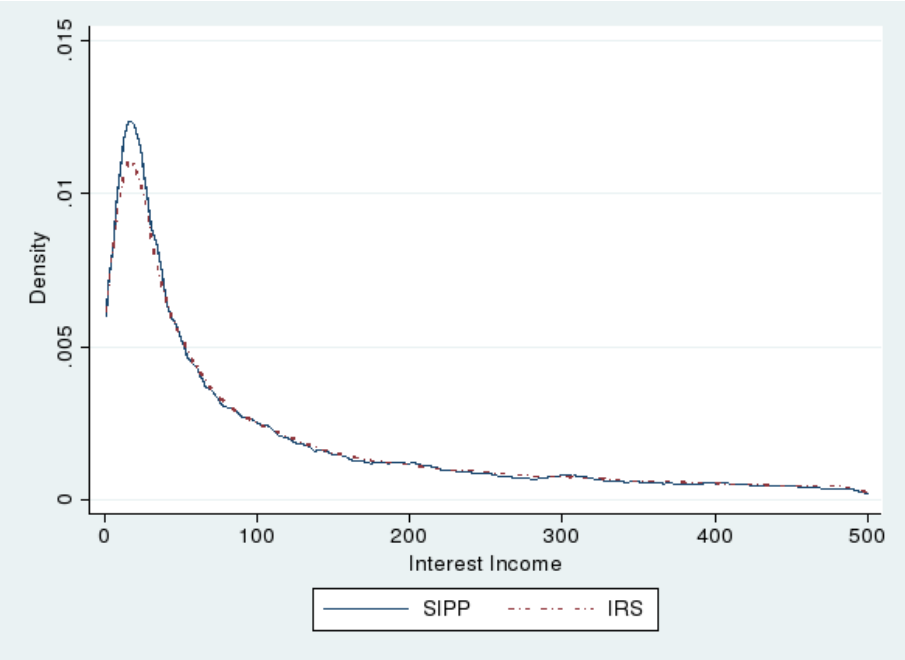


Figure displays a kernel density graph for interest income in SIPP and IRS 1040 data. Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive interest income in SIPP and IRS 1040. Epanechnikov kernel used to generate the density. Number of Observations: 191,232.

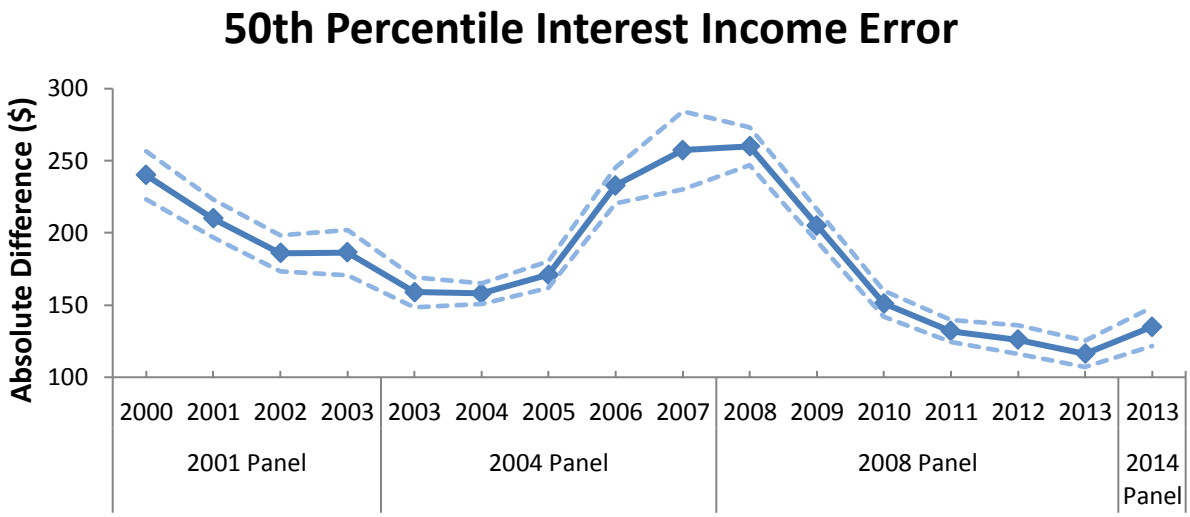
Table 3: Percentiles and Mean of Difference between SIPP and IRS Interest Income (SIPP-IRS)

	Percentiles					Mean
	10th	25th	50th	75th	90th	
Difference	-3,325	-458	-19	68	554	-2,251
	(92.7)	(11.3)	(1.3)	(2.8)	(19.3)	(562.5)
Absolute Difference	10	35	181	1,039	4,718	3,439
	(0.1)	(1.0)	(3.2)	(19.2)	(99.6)	(568.4)

Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive interest income in SIPP and IRS 1040. Standard Errors in Parenthesis. Number of Observations: 191,232.

To show how the discrepancy between survey and administrative data has changed over time in SIPP, Figure 5 shows the median of the absolute difference by panel and year. The figure shows fluctuations in accuracy from year to year, but no overall time trend and no notable differences between the SIPP panels. The values for the 50th percentile is higher between 2006 and 2008, but this is likely due to the higher interest rates resulting in increasing asset income around this time. In other words, SIPP respondents may have had higher interest income between 2006 and 2008, which could lead to larger reporting error in terms of absolute value, even if the reporting error in terms of percentages of IRS income remained the same. Thus, while there were numerous changes made to the SIPP over time, including changes to the record use request, it does not appear that these changes affected the match between SIPP and IRS interest income.

Figure 5: 50th Percentile in Absolute Difference of Interest Income over Time



Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive interest income in SIPP and IRS 1040. Dashed line represents 95% confidence interval. Number of Observations over all years and panels: 191,232.

4. Regression Analysis for Measurement Error

To investigate whether record use is associated with a reduction in measurement error, we conduct a regression analysis to control for other observed factors. Our research question is whether the use of financial records for asset questions reduces the absolute value of the difference between SIPP and IRS asset income for asset a and tax unit i in year t, denoted by $abs(I_{a,i,t}^{SIPP} - I_{a,i,t}^{IRS})$. Note from Table 3 that the distribution of the absolute difference contains many outliers with very large values. To both reduce the effect of outliers and include people with a zero difference between SIPP and IRS income, our dependent variable is the inverse hyperbolic sine (IHS) of the absolute difference, denoted by

$$\sinh^{-1}\left(abs(I_{a,i,t}^{SIPP} - I_{a,i,t}^{IRS})\right), \text{where } \sinh^{-1}(x) = ln(x + \sqrt{x\text{^}2 + 1}). \quad (1)$$

The $\sinh^{-1}$ function approximately equals the natural log function for most values greater than zero, but has the added benefit of being defined at zero. From this setup, our equation of interest is

$$\sinh^{-1}\left(abs(I_{a,i,t}^{SIPP} - I_{a,i,t}^{IRS})\right) = \sinh^{-1}(y_{a,i,t}) = \beta_{a,0} + \beta_{a,1}R_{i,t} + \beta_{a,2}X_{a,i,t} + \epsilon_{i,t}. \quad (2)$$

In which $R_{i,t}$ is an indicator of record use and $X_{a,i,t}$ is a vector of other explanatory variables. Other explanatory variables include the respondent's IRS asset income; adjusted gross income; indicators of imputed values; percent of SIPP asset income which is imputed; indicator of a proxy interview;[9] demographic variables such as age, race, Hispanic origin, and education; and employment income reported in SIPP. Note that since $\sinh^{-1}$ approximately equals the natural log function for most values, each β coefficient for continuous variables roughly represents the average percent change in the dependent variable from a small change in the explanatory variable, based on taking the derivative. For the record use coefficient, the average percent change in the dependent variables from using records is $e^{\beta_{a,1}} - 1$, as

$$\left(y(R_{i,t} = 1) - y(R_{i,t} = 0)\right)/y(R_{i,t} = 0) \approx$$

$$(exp(\beta_{a,0} + \beta_{a,1} + \beta_{a,2}X_{a,i,t} + \epsilon_i) - \exp(\beta_{a,0} + \beta_{a,2}X_{a,i,t} + \epsilon_i))/exp(\beta_{a,0} + \beta_{a,2}X_{a,i,t} + \epsilon_i) =$$

$$exp(\beta_{a,0} + \beta_{a,1} + \beta_{a,2}X_{a,i,t} + \epsilon_i - (\beta_{a,0} + \beta_{a,2}X_{a,i,t} + \epsilon_i)) - 1 = \exp(\beta_{a,1}) - 1,$$

and the average percent change for other discrete variables takes a similar form.

Table 4: Regression for Asset Income Measurement Error

	Inverse Hyperbolic Sine (IHS) of Absolute Difference in Asset Income		
	Interest	**Dividend**	**Rental**
Record Use	***-0.257	***-0.219	***-0.569
	(0.019)	(0.032)	(0.077)
IHS IRS Income	***0.756	***0.648	***0.817
	(0.003)	(0.005)	(0.016)

[9]A proxy interview is where another household members provides the information about a respondent.

IHS AGI	***0.019	***0.057	-0.007
	(0.003)	(0.006)	(0.008)
Any Ownership	-0.011	0.020	***0.256
Indicators Imputed	(0.013)	(0.025)	(0.046)
Percent Amounts Imputed	***0.412	***0.498	***1.095
	(0.009)	(0.016)	(0.032)
Proxy	0.015	0.026	***0.223
	(0.012)	(0.025)	(0.059)
Male, Not Married	***0.069	***0.098	0.080
	(0.015)	(0.035)	(0.084)
Married	***0.079	***0.119	0.101
	(0.016)	(0.033)	(0.084)
White	**0.042	0.024	*0.179
	(0.020)	(0.039)	(0.096)
Black	-0.024	***0.236	***0.369
	(0.032)	(0.081)	(0.124)
Latino	***-0.094	*-0.132	0.215
	(0.029)	(0.072)	(0.132)
High School	0.029	-0.067	-0.118
	(0.018)	(0.053)	(0.088)
Some College	***0.096	-0.027	0.006
	(0.016)	(0.052)	(0.079)
College	***0.265	*0.086	*0.139
	(0.019)	(0.051)	(0.078)
Age	*0.029	0.051	0.027
	(0.016)	(0.036)	(0.111)
Age Squared	***0.006	0.004	0.006
	(0.001)	(0.003)	(0.010)
Any SIPP Labor Income	***-0.670	**-0.256	0.077
	(0.052)	(0.100)	(0.178)
IHS SIPP Labor Income	***0.050	*0.017	0.000
	(0.005)	(0.009)	(0.015)
Constant	***0.589	***1.420	***-1.272
	(0.053)	(0.116)	(0.367)
Number of Observations	191,232	59,970	19,408

Table present the results of three regression in which the dependent variable is the inverse of the hyperbolic sine of interest, dividend, and rental income. Source:2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive asset income in SIPP and IRS 1040. Year fixed effects included but not shown. See Appendixes A.2-A.5 for more details on the construction of the explanatory variables. Standard Errors in Parenthesis. Significance asterisks: *** p<.01 ** p<.05 * p<.1.

Table 4 presents the results of this regression using data from the 2001-2014 SIPP panels. After the $e^{\beta_{a,1}} - 1$ transformation discussed above, the coefficient for record use shows that record users have a 22.6 percent lower discrepancy between their SIPP and IRS interest income. At the median value of measurement error in interest income, this coefficient translates into decreasing the difference between SIPP and IRS data from $181 to $140. Record use is also associated with a 19.7 percent lower discrepancy in dividend income and a 43.4 percent lower discrepancy in rental income across the two sources. In terms of other explanatory variables in the interest income regression, a 1 percentage point increase in the share of interest income that is imputed increases measurement error by 0.412 percent, married couples have an 8.2 percent higher measurement error relative to single females, and college educated respondents have 30.3 percent higher measurement error than respondents without a high school education. These results show that the effect of record use on measurement error is relatively strong compared to other variables in the model.

The previous analyses potentially suggest that record use reduces measurement error. However, it could be that these results are explained by record use being correlated with other unobserved respondent characteristics, and that the types of respondents who use records are more motivated to give accurate answers, regardless of whether they consulted records. To address this problem, we run our regressions again on just the 2014 panel including variables on the average time spent per question, item nonresponse rates for financial questions, and the average amount of rounding in financial questions as proxies for unobserved characteristics which also affect measurement error in asset income.[10] We include the average time per question because Eggleston et al. (2016) shows that similar measures are associated with reporting numbers that have been rounded to a lesser degree,

[10] In 2014 the SIPP changed from CASES to BLAISE survey software. The BLAISE survey software automatically generates key stroke and timestamp file which can be used to create paradata on the time spent per question and person-level data on item nonresponse rates and average amount of rounding. Because such files were absent in the CASES software, the variables of interest are only available for the 2014 panel.

suggesting that this measure is associated with the respondent's motivation in providing accurate answers. We also include the interviewer's average time per question over all interviews to separate out how long it takes a respondent to answer a question from how fast an interviewer reads questions. Our measure of rounding is taken from Gideon et al. (2015), as the number of trailing zeros in an answer divided by the number of potential trailing zeros. For example, a reported value of $1,034 has a rounding value of 0 and a reported value of $1,000 has a rounding value of 1. Gideon et al. (2015) found that rounding behavior was consistent with satisficing strategies, suggesting that rounding is related to respondent motivation. More details on the construction of these variables can be found in Appendix A.6.

Our key assumption is that these three proxy variables account for all other unobserved factors that are correlated with both measurement error in asset income and record use in the asset section, which allows for the estimate of the record use coefficient to represent the causal effect of record use on measurement error. While these proxy variables serve as indicators of overall measurement error in a SIPP respondent's data, and thus should be useful proxy variables, these variables may not capture all additional factors which affect measurement error in asset income. For example, a respondent may be interested in the content of other sections of the SIPP interview, but may be uninterested in the asset questions. Such a respondent may provide accurate answers for other questions in SIPP, but may provide inaccurate answers for the asset questions and refuse to consult records. In this case, record use would still be correlated with other unobserved respondent characteristics, even after controlling for the respondent's behavior elsewhere in the interview. Despite this potential limitation, our approach with proxy variables provides a new way to use additional data on respondent accuracy for investigating measurement error in a survey when an experimental manipulation is unavailable.

Table 5: Asset Income Regression with 2014 SIPP Data and Respondent Engagement Variables

	Inverse Hyperbolic Sine (IHS) of Absolute Difference in Asset Income					
	Interest		Dividend		Rental	
Record Use	***-0.319	***-0.185	***-0.511	***-0.333	***-0.834	**-0.572
	(0.049)	(0.053)	(0.103)	(0.105)	(0.259)	(0.271)
Time Per Question		-0.003		-0.006		-0.002
		(0.005)		(0.009)		(0.016)
Interviewer's Time		0.003		-0.012		-0.050
Per Question		(0.007)		(0.013)		(0.036)
Item Nonresponse Rate		**0.262		0.187		0.133
		(0.118)		(0.189)		(0.562)
Rounding Average		***1.033		***1.063		***2.277
		(0.142)		(0.250)		(0.816)
Respondent Engagement Variables	No	Yes	No	Yes	No	Yes
Number of Observations	7761	7480	2953	2894	1278	1260

Table 5 presents estimates of Equation (2) using only SIPP 2014. Source:2014 SIPP panel and 2013 IRS 1040 dataset. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive asset income in SIPP and IRS 1040. For the regression with the respondent engagement variables, I exclude household members who were not interviewed, either in-person or through a proxy, or who were interviewed but did not give an answer to any financial question (because I have no measure of average rounding for such respondents). Additional variables included but not shown are respondents IRS Asset Income; Adjusted Gross Income; Indicators of Imputed Values; Percent of SIPP interest income which is imputed; indicator of a proxy interview; demographics variables such as age, race, and education, and employment income reported in SIPP. See Appendixes A.2-A.6 for more details on the construction of the explanatory variables. Standard Errors in Parenthesis. Significance asterisks: *** p<.01 ** p<.05 * p<.1

Table 5 shows the results for the 2014 panel with and without our three measures of respondent engagement. These results show that the record use coefficient is still significant and negative with the inclusion of these variables into the model. The time per question coefficients and item nonresponse rate coefficients are not significant for any assets. However, the rounding coefficients are significant and large in magnitude in all models, showing that a 1 percentage point increase in average rounding is associated with a 1 to 2.3 percent increase in measurement error. In summary, this regression analysis provides evidence which supports the proposition that record use reduces measurement error in interest income. While there could be other characteristics of the respondent

which explain the relationship between record use and measurement error, this result is mitigated by including a variety of additional variables from administrative data, survey data, and paradata which account for both socio-economic factors and respondent engagement.

5. Interview Length

Table 6: Mean Time per Question by Record Use in Seconds

Section of Personal Interview	Record Use	
	Yes	No
Asset Value and Income	19.04	11.64
	(0.13)	(0.05)
All Other Sections	14.52	12.30
	(0.07)	(0.04)

Source: 2014 SIPP panel. The cross-sectional unit of analysis in this table is an individual. Sample consists of all individuals over age 15 who owned any assets, excluding household members who were not interviewed, either in-person or through a proxy, or individuals who dropped out of the survey before the asset section. See Appendixes A.4 and A.6 for more details on the construction of the variables. Standard Errors in Parenthesis. In this table, the difference between sections is statistically significant for record users and for non-record users. Number of Observations: 54,399

While encouraging record use may offer potential benefits in terms of data accuracy, this request could increase respondent burden through the time it takes to find and analyze financial records, although having records on hand could make survey questions easier to answer and reduce cognitive burden. In addition, there could be increased financial costs of collecting the survey responses due to paying interviewers for the additional time it takes respondents to find and retrieve financial records. To show how record use may affect interview length, we first present in Table 6 the average amount of time respondents spend in both the section on asset income and values as well as other sections of the 2014 SIPP interview, separating out respondents who use records from those who do not. This table shows that for respondents who do not use records, the average amount of time per question is about 11.6 seconds per question in the asset section and 12.3 seconds in other sections of the interview. Respondents who use records spend about 19 seconds per question on each asset question and about 14.5 seconds per question in other sections. These means suggest that respondents who use records in

the asset section may be more diligent in other sections on the interview, but consulting records in the asset section may result in them taking even longer to respond to asset questions.

Table 7: Regression for Seconds Spent Per Question for the Asset Income and Value Section

	Time Per Question in Asset Income and Value Section (Seconds)
Record Use	***4.023
	(0.103)
Time Per Question in Other Sections (Seconds)	***0.589
	(0.015)
Interviewer's Time Per Question in Other Sections (Seconds)	***0.329
	(0.018)
Item Nonresponse Rate	***-2.603
	(0.193)
Rounding Average	***-3.451
	(0.183)

Source:2014 SIPP panel. The cross-sectional unit of analysis in this table is an individual. Sample consists of all individuals over age 15 who owned any assets, excluding household members who were not interviewed, either in-person or through a proxy, individuals who dropped out of the survey before the asset section, or individuals who were interviewed but did not give an answer to any financial question (because I have no measure of average rounding for such respondents). See Appendixes A.2-A.6 for more details on the construction of the time per question variables. Additional variables included but not shown are age, race, and education, and employment income reported in SIPP. Standard Errors in Parenthesis. Significance asterisks: *** p<.01 ** p<.05 * p<.1. Number of observations: 48,677.

To more formally control for record users' behavior in other sections of the SIPP Interview, we conduct a regression analysis in Table 7 in which the dependent variable is the average number of seconds spent on each asset value and income question. Our main explanatory variables are whether the respondent used records, the average time the respondent spent on other SIPP questions, and the average amount of time the respondent's interviewer spent on other SIPP questions. We include the interviewer-level variable to separate out the time it takes a respondent to answer a question from the

speed in which an interviewer reads the survey questions.[11] We find that record use is associated with spending an extra 4.02 seconds on each asset question. To put this number into context, the average number of questions a respondent answers in the asset income and value section is 16.02, and the average amount of time for an individual's personal interview is 41.65 minutes. Thus, at the average values, using records increase the amount of time spend in the asset section by 1.07 minutes, which translates to a 2.6% increase in the duration of the interview. We also find that spending one second longer for other questions is associated with spending an extra 0.59 seconds on each asset question, and the interviewer taking one extra second in other interviews is associated with an extra 0.33 seconds in the asset questions. Together, the evidence from Table 6 and Table 7 suggests that while the types of people who use records may be more engaged respondents who spend longer on questions, record use is still associated with longer interview duration.

6. Conclusion

In this paper, we investigate whether consulting financial records to answer asset questions affects measurement error and interview length. We find that record use is associated with a 19 to 43 percent decrease in asset income measurement error, and is associated with spending an extra 4.03 seconds answering each asset question, on average. Our results have several implications for the design of record use requests. First, our comparison between the 2001 and 2004 panels suggests that prompting respondents to take their time to find financial records may increase the proportion who use records while answering questions in the asset section. Our comparison between the 2008 and 2014 panels suggest that if the record use request is made only once, respondents may either forget about consulting records over the course of the survey or may stop doing so after consulting records for a few

[11]When constructing the interviewer-level average for a given respondent, we exclude the respondent's time-per-question to make sure the respondent's behavior does not affect the calculation.

questions. In terms of whether record use requests are beneficial in surveys, our results show that there is a cost-benefit trade off between increased interview length and reduced measurement error. While respondents who abide to the request may be more engaged and find the request less burdensome, our results suggest that record use still increases interview length for such respondents. Thus, while record use may be an effective tool for improving data accuracy, it may come at a financial cost of interviewer compensation.

References

Abowd, John M., and Martha H. Stinson. "Estimating measurement error in annual job earnings: A comparison of survey and administrative data." Review of Economics and Statistics 95, no. 5 (2013): 1451-1467.

Bond, Brittany, J David Brown, Adela Luque, and Amy O'Hara. 2013. "The Nature of the Bias When Studying Only Linkable Person Records: Evidence from the American Community Survey." *Proceedings of the 2013 Federal Committee on Statistical Methodology (FCSM) Research Conference*.

Bound, John, and Alan B. Krueger. "The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?." Journal of Labor Economics (1991): 1-24.

Couper, Mick P, Mary Beth Ofstedal, and Sunghee Lee. 2013. "Encouraging Record Use for Financial Asset Questions in a Web Survey." *Journal of Survey Statistics and Methodology* 1 (2):171-182.

Edgar, Jennifer. 2010. "Respondent Record Use in the US Consumer Expenditure Survey." *Annual Meeting of the American Association for Public Opinion Research*.

Edgar, Jennifer, and Jeffrey M Gonzalez. 2009. "Correlates of Data Quality in the Consumer Expenditure Quarterly Interview Survey." *Joint Statistical Meetings*.

Eggleston, Jonathan. 2015. "Asset Record Use and Measurement Error." *Annual Meeting of the Southern Economic Association*.

Eggleston, Jonathan, Michael Gideon, Joanne W. Hsu, and Brooke Helppie. 2016. "Heaping at Round Numbers on Financial Questions: Satisficing and Fatigue." *American Association for Public Opinion Research Annual Conference*.

Gideon, Michael, Joanne Hsu, and Brooke Helppie McFall. 2015. "Heaping at Round Numbers on Financial Questions: The Role of Satisficing." *Federal Committee on Statistical Methodology Research Conference*.

Hedengren, David, and Thomas Stratmann. 2012. "The Dog that Didn't Bark: What Item Nonresponse Shows about Cognitive and Non-Cognitive Ability." *Available at SSRN 2194373*.

Kashihara, David, and Diana Wobus. 2006. "Accuracy of Household-Reported Expenditure Data in the Medical Expenditure Panel Survey." *Proceedings of the American Statistical Association, Alexandria, VA: American Statistical Association*:3193-3200.

Laurie, Heather, and Nick. Moon. 2010. "Is It Worth the Effort? Respondent Record Use on the British Household Panel Study." *Annual Meeting of the American Association for Public Opinion Research*.

Layne, Mary, Deborah Wagner, and Cynthia Rothhaas. 2014. "Estimating Record Linkage False Match Rate for the Person Identification Validation System." *Center for Administrative Records Research and Applications Working Paper* 2.

Maynes, E Scott. 1968. "Minimizing Response Errors in Financial Data: The Possibilities." *Journal of the American Statistical Association* 63 (321):214-227.

Moore, Jeffrey C, and Julia Klein Griffiths. 2003. "Asset Ownership, Program Participation, and Asset and Program Income: Improving Reporting in the Survey of Income and Program Participation." *Joint Statistical Meetings*.

Murphy, Joe, Jeffrey Rosen, Ashley Richards, Sarah Riley, Andy Peytchev, and Mark Lindblad. 2015. "Financial Record Checking in Surveys: Do Prompts Improve Data Quality?" *Field Methods*. doi: 10.1177/1525822x15616346.

Safir, Adam, and Karen L Goldenberg. 2008. "Mode Effects in a Survey of Consumer Expenditures." *Proceedings of the Survey Research Methods Section, American Statistical Association*.

U.S. Census Bureau. 2016. *Survey of Income and Program Participation 2014 Panel Users' Guide*. Washington, D.C..

A. Appendix

In this Appendix, we give more details about sample selection, construction of our variables, and weighting. We also present additional results for dividend and rental income which follow a similar pattern to the results for interest income presented in the main body of the paper.

A.1. Sample Selection

From SIPP we include all individuals who were over 15 years old and in the sample during at least one month in a calendar year. This criteria includes people who leave or join the sample in the middle of the year. For the analyses that use tax data, we include only individuals who were matched to the IRS 1040 as either a primary filer or as the spouse of a primary filer. In our administrative data, we have a small number of cases in which the PIK of the spouse is actually the PIK of the primary filer's child. In these cases, we reclassify the child as someone who was not matched to the 1040, unless the child is an adult and is matched to their own separate 1040 record.

A.2. IRS Asset Income

The Individual Master File of all IRS 1040 tax returns contains data on interest, dividend and gross rental income. Interest income corresponds to taxable interest and tax-exempt interest from lines 8a and 8b of

the 1040 form. Dividend income consists of ordinary and qualified dividends from lines 9a and 9b. Gross rental income consists of rental and royalty income from lines 3 and 4 of the Schedule E form. For most observations, the IRS variable is taken from one matched 1040 record. For a small number of married couples, each spouse is matched to a different 1040 record. We've concluded that in most of these cases, the married partners are filing separate tax returns. Because of this, we add up the values from the two different returns when creating the combined income totals for a married couple.

A.3. SIPP Asset Income

For the 2001-2008 SIPP panels, asset income data is given at the month level. To construct an annual income variable, we add up the values from all months in a given calendar year. For individuals who are only interviewed in some months, which can happen frequently in the years in which a SIPP panel both begins and ends, we inflate the sum of the monthly amounts by 12/(the number of interviewed months) to generate an annual variable. For the 2014 panel, interest income is reported at the calendar-year level, so no summation is necessary. For married couples, we add their annual total together to generate one observation per married couple.[12] Because marital status can change throughout the year, we define a person as married if they were married in December of the tax year (or in the last month they were observed in sample), as the IRS criteria for married filing jointly is based on marital status as of December 31st.

The SIPP contains multiple interest, dividend, and rental income variables. The interest income variables include money received from savings accounts, interest checking accounts, money market accounts, certificates of deposit, municipal and corporate bonds, U.S. Savings Bonds, U. S. Government Securities, and mortgages owned as an asset. Mortgages owned as an asset are uncommon, and were

[12] For a bank account jointly owned by a married couple, only one spouse is asked about the interest income and balance from that account. The final values in the SIPP dataset are divided by two to prevent double counting.

asked directly about in the 2008 panel and prior panels. Because of very low ownership rates, this direct question was dropped in 2014 SIPP and was instead included in a catch-all question. Because of this, interest from mortgages owned as an asset is not used for the 2014 panel but is used for other panels. Dividend income consists of dividends from stocks and mutual funds. SIPP collects data on both gross and net rental income, but we only use gross rental income for this project, as we only have administrative data on gross rental income. In our SIPP rental income variable, we also include data on royalty income, as the IRS variable consists of rental and royalty income. Because of low rates of receipt for royalty income, the question on royalty income was dropped from the 2014 SIPP panel. For each of these asset variables, income is collected separately for assets owned jointly and for assets owned only by the respondent.

In addition, for our rental income analyses, we exclude individuals in the 2001-2008 Panels who report owning rental property jointly with someone besides their spouse. These respondents were only asked a question about the net rental income from their property, while our administrative variable is the gross rental income from the property. Because we do have the gross rental income from such properties, we excluded these individuals from our analyses.

A.4. Record Use Variable

For the 2001-2008 SIPP panels, a respondent could consult records in one interview but not the other. We create a calendar-year level variable for these panels which equals to the average number of interviews a respondent consulted records in. For example, if a respondent used records in two out of three interviews for 2012, and the reference period for these interviews were completely contained within 2012, then their record use variable would have a value of two-thirds. For reference periods that spans over two calendar years, we weight by the proportion of the reference period contain within the year. For example, if a respondent used records in one interview, and the reference period was between

October 2011-January 2012 (4 months), then the record use indicator for this interview would be given a weight of three-fourths for the 2011 calculation and a weight on one-fourth for the 2012 calculation.

A.5. Other explanatory variables

For married couples, we use the spouse with the highest level of education to construct the couple's education level. For race, we have the White indicator equal to one if either spouse is White non-Hispanic. If this is not true, then we have the Black indicator equal to one if either spouse is Black non-Hispanic. If this is not true, then we have the Hispanic indicator equal to one if either spouse is Latino/Hispanic. For other indicator variables, such as having any imputed values, a proxy interview, or having labor market income, we set these indicator variables equal to one if they equal one for any spouse.

A.6. Respondent Engagement Variables for 2014 Panel

Our respondent engagement variables are constructed by a program which reads in and analyses the BLAISE audit trail (key-stroke files) from the SIPP interview. The file contains the date and time an interviewer started reading a question, the date and time when the interviewer moved onto a new question, and the respondent's answer to the question. From this information, we can construct the average time spent per question as well as the percent of financial questions in which the response was "don't know" or "refuse." When constructing the time per question variables, we account for interviews that are stopped in the middle and then completed later. Our set of financial questions consists of questions that ask respondents to give a dollar amount for some source of income, expenses, or wealth. We use the item nonresponse rate for financial questions rather than for all questions in the survey to reflect better the respondent's behavior in answering questions similar to the asset income questions.

Our measure of rounding is taken from Gideon, Hsu, and McFall (2015) as the number of trailing zeros in an answer divided by the number of potential trailing zeros. For example, a reported value of

$1,034 has a rounding value of 0 and a reported value of $1,000 has a rounding value of 1. In our rounding calculations, we remove single digit responses (e.g. $5), as these types of responses have no opportunity to display rounding behavior. A small number of respondent's give an answer in cents (e.g., $20.21). We assign these responses a rounding value of 0. For our final variable, we then take the average rounding value over all financial questions in which the respondent gave an answer and the given response was not in the single digits.

A.7. Weighting and Variance Estimation

We use sample weights for our point estimates and replicate weights to construct standard errors to account for complex sampling. The 2001, 2004-2008, and 2014 panels all have a different number of replicate weights. To construct pooled estimates from all the panels, we run separate analyses for the 2001, 2004-2008, and 2014 panels, and combine the point estimates and standard errors as a weighted average of the estimates from the three sets of panels. For married couples, we take the average of their two weights in our estimates.

For our analyses that use administrative data, we also make other adjustments to our weights. For a small number of observations, one SIPP respondent is matched to multiple 1040 records. In this case, we include this SIPP respondent multiple times in our data, and divide their weight by the number of 1040 records they are matched to. In addition, we run a probit model to generate the predicted probability a SIPP respondent has a PIK, and we then multiply the respondent's weight by the inverse of the predicted probability. As noted by Bond et al. (2013), the probability someone is matched to administrative records depends on age, ethnicity, and other observed characteristics, so adjusting our weight by the PIK probability accounts for non-random selection based on observed characteristics.

A.8. Additional Results for Dividend and Rental Income

The following tables and figures present additional results for dividend and rental income, which correspond to the various tables and figures for interest income presented in the main body of the paper.

Table 8: Cross Tabulation for Reporting Dividend Income

	Interest Income on IRS 1040	
Interest Income in SIPP	Yes	No
Yes	10.9	4.2
	(0.13)	(0.19)
No	11.0	73.9
	(0.12)	(0.19)

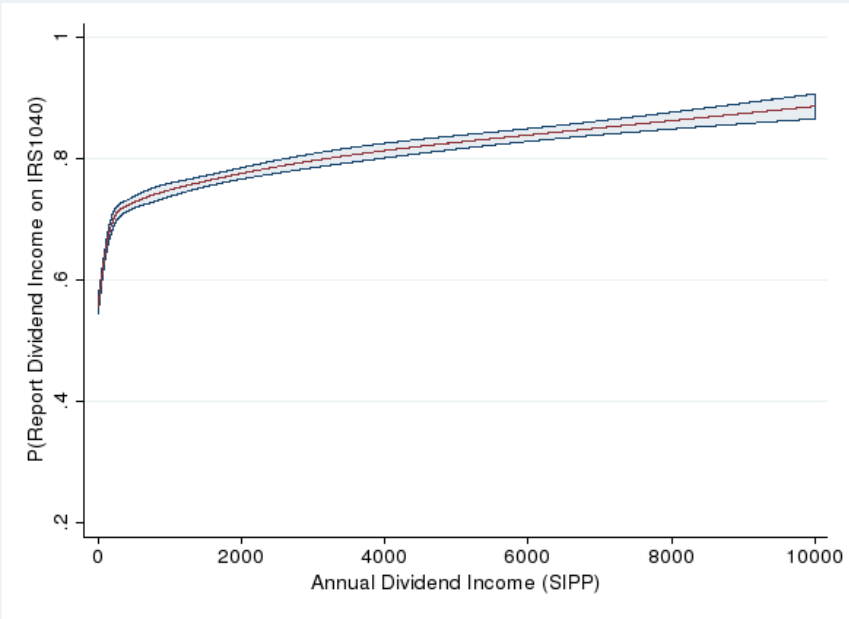
Table describes the proportion of SIPP respondents who report any dividend income in either SIPP, the IRS 1040, or both. Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. Sample consists of all individuals over age 15. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Standard errors in parenthesis. Number of Observations: 515,734.

Table 9: Cross Tabulation for Reporting Rental Income

	Interest Income on IRS 1040	
Interest Income in SIPP	Yes	No
Yes	3.5	1.2
	(0.07)	(0.11)
No	3.7	91.6
	(0.07)	(0.11)

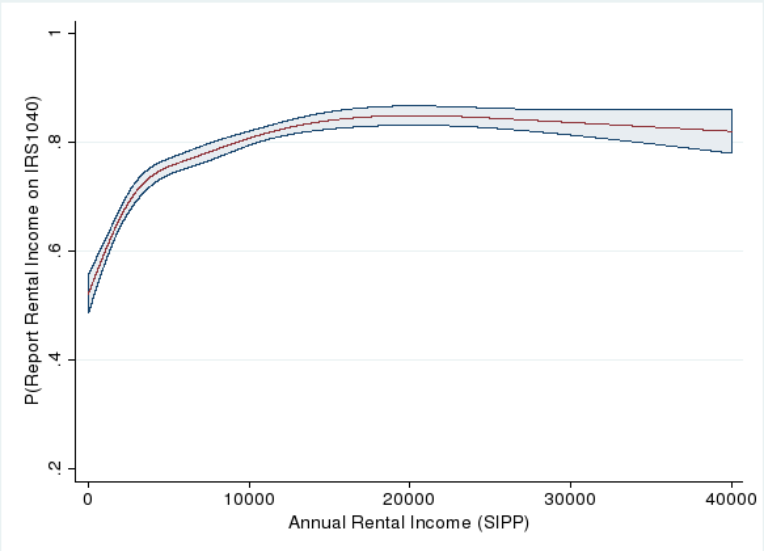
Table describes the proportion of SIPP respondents who report any rental income in either SIPP, the IRS 1040, or both. Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. Sample consists of all individuals over age 15. In the 2001-2008 Panels, we exclude individuals who report owning rental property jointly with someone besides their spouse. These respondents were only asked a question about the net rental income from their property, while our administrative variable is the gross rental income from the property. Because we do have the gross rental income from such properties, we excluded these individuals from our analyses. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Standard errors in parenthesis. Number of Observations: 512,723.

Figure 6: Probability Reporting Dividend Income on Tax Return



Graph displays predicted values from a regression of an indicator of reporting dividend income in the IRS1040 on a cubic spline of SIPP dividend income with 5 knots. Source: 2004-2008 SIPP panels and 2004-2013 IRS 1040 datasets. The 2001 and 2014 panels were not used in this graph due to the different number of replicate weights in these panels. See Appendix A.7 for more details on replicate weights. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive dividend income in SIPP. Shaded area represents 95% confidence interval. Number of Observations: 59,691.

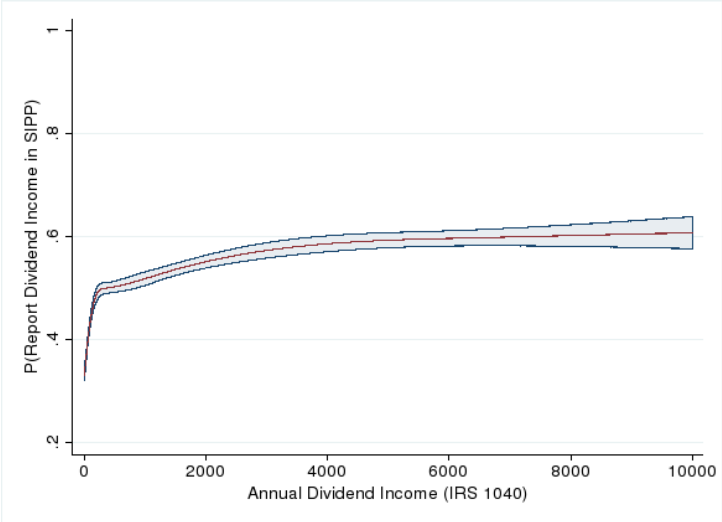
Figure 7: Probability Reporting Rental Income on Tax Return



Graph displays predicted values from a regression of an indicator of reporting rental income in the IRS1040 on a cubic spline of SIPP dividend income with 5 knots. Source: 2004-2008 SIPP panels and 2004-2013 IRS 1040 datasets. The 2001 and 2014 panels were not used in this graph due to the different number of replicate weights in these panels. See Appendix A.7 for more details

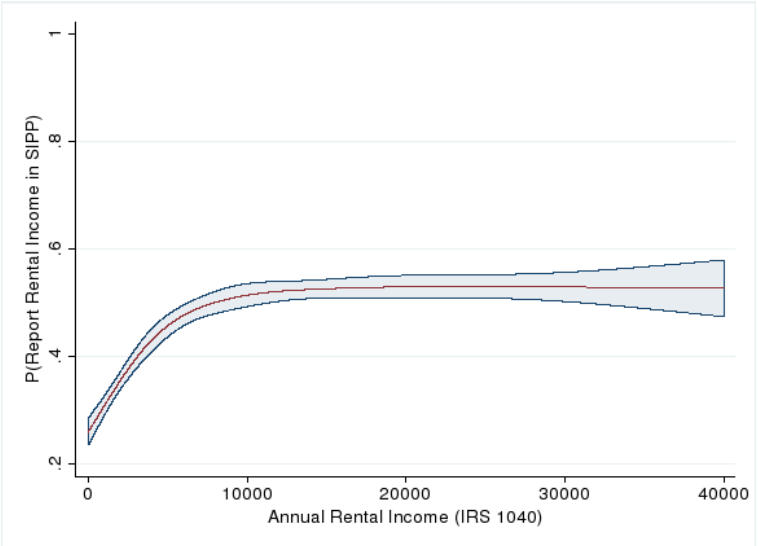
on replicate weights. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive dividend income in SIPP. Shaded area represents 95% confidence interval. Number of Observations: 19,159.

Figure 8: Probability Reporting Dividend Income in SIPP



Graph displays predicted values from a regression of an indicator of reporting dividend income in SIPP on a cubic spline of IRS Interest Income with 5 knots. Source: 2004-2008 SIPP panels and 2004-2013 IRS 1040 datasets. The 2001 and 2014 panels were not used in this graph due to the different number of replicate weights in these panels. See Appendix A.7 for more details on replicate weights. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive interest income in IRS 1040. Shaded area represents 95% confidence interval. Number of Observations: 85,769

Figure 9: Probability Reporting Rental Income in SIPP



Graph displays predicted values from a regression of an indicator of reporting rental income in SIPP on a cubic spline of IRS Interest Income with 5 knots. Source: 2004-2008 SIPP panels and 2004-2013 IRS 1040 datasets. The 2001 and 2014 panels were not used in this graph due to the different number of replicate weights in these panels. See Appendix A.7 for more details on replicate weights. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive interest income in IRS 1040. Shaded area represents 95% confidence interval. Number of Observations: 28,142.

Table 10: Percentiles of Difference between SIPP and IRS Dividend Income (SIPP – IRS)

	Percentiles					Mean
	10th	25th	50th	75th	90th	
Difference	-5,787	-1,099	-20	652	4,515	795
	(204.9)	(51.8)	(3.6)	(32.3)	(169.9)	(623.3)
Absolute Difference	34	164	856	3,655	11,859	7,446
	(1.4)	(4.9)	(22.0)	(115.9)	(381.8)	(630.3)

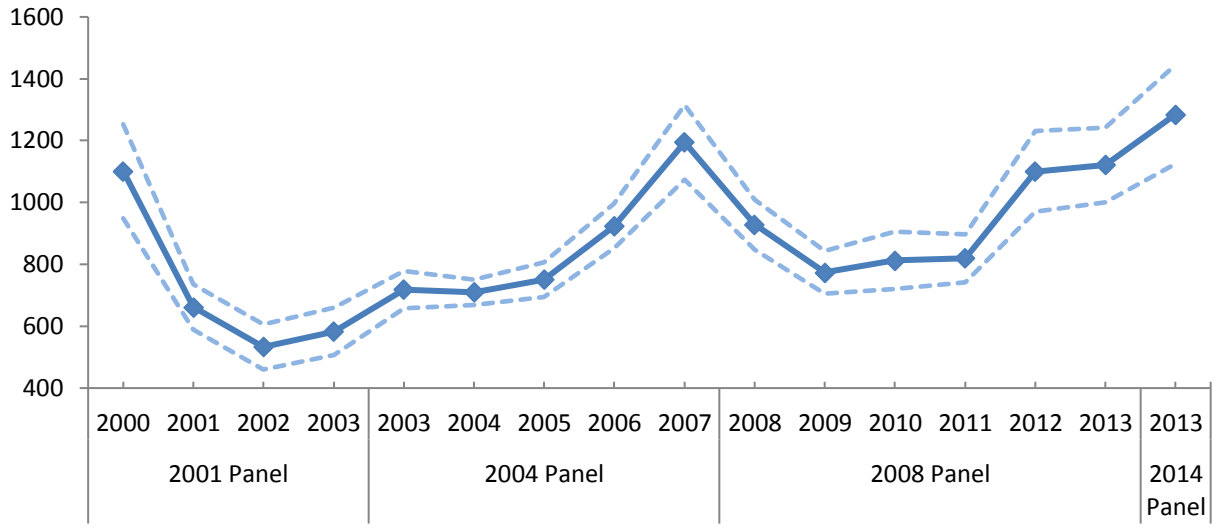
Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive dividend income in SIPP and IRS 1040. Standard Errors in Parenthesis. Number of Observations: 59,970.

Table 11: Percentiles of Difference between SIPP and IRS Rental Income (SIPP – IRS)

	Percentiles					Mean
	10th	25th	50th	75th	90th	
Difference	-25,943	-8,274	-1,578	669	6,365	-8,826
	(1598.2)	(392.2)	(121.7)	(93.5)	(480.8)	(1171.3)
Absolute Difference	192	1,095	4,085	12,032	33,389	16,844
	(15.3)	(62.9)	(148.2)	(392.9)	(1896.1)	(1178.5)

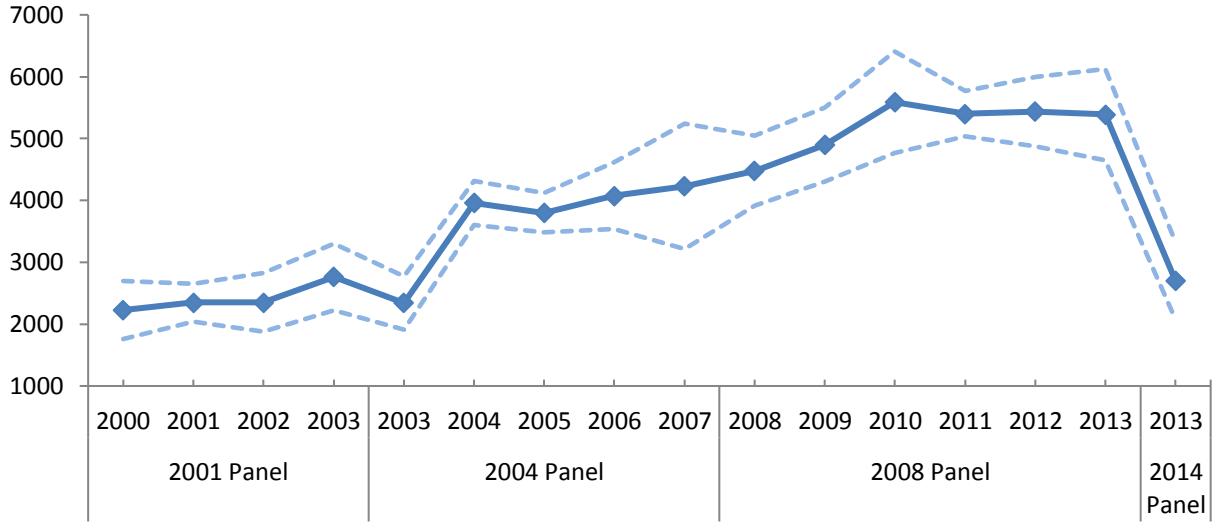
Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive rental income in SIPP and IRS 1040. Standard Errors in Parenthesis. Number of Observations: 19,408.

Figure 10: 50th Percentile Dividend Income Error over Time



Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive dividend income in SIPP and IRS 1040. Dashed line represents 95% confidence interval. Number of Observations: 59,970.

Figure 11: 50th Percentile Rental Income Error over Time



Source: 2001-2014 SIPP panels and 2000-2013 IRS 1040 datasets. The cross-sectional unit of analysis in this table is a tax unit (which is either a single person or a married couple). Sample consists of all individuals over age 15 and tax units with positive rental income in SIPP and IRS 1040. Dashed line represents 95% confidence interval. Number of Observations: 19,408.