Item Response Rates for Composite Variables

Jonathan Eggleston[1]

November 29, 2017

Abstract

Item response rates frequently serve as indicators of data quality and potential nonresponse bias. However, key variables from surveys, such as total household income or net worth, are often composite variables constructed from several underlying components. Because such composite variables do not have clearly identifiable response rates, inference on the data quality of these key measures is more difficult. In this paper, I propose three new methods for aggregating data on response rates across questions to create a measure of item response for composite variables. To show how these methods can be used to investigate data quality, I analyze item response for net worth in the Survey of Income and Program Participation and the Survey of Consumer Finances. All three of the new measures show that item response rates went up in the SIPP after the redesign in the 2014 Panel, but are still lower than item response rates in the SCF. This comparison between surveys would be difficult without a method for aggregating item response rates. Overall, these new item response rate methods provide a new way of describing data quality for key measures in surveys and for analyzing changes in data quality over time.

[1]Economist, Social, Economic, and Housing Statistics Division, U.S. Census Bureau. Email: Jonathan.S.Eggleston@census.gov. Mail: 4600 Silver Hill Road, Washington DC, 20233. SEHSD Working Paper Number 2017-48. SIPP Working Paper Number 282.

1. Introduction

Unit and item response rates are widely used tools in survey research to measure the potential impact of nonresponse bias (e.g. Bollinger et al. 2015). While low response rates do not necessarily lead to high nonresponse bias (Groves and Peytcheva 2008), constructing an alternative measure based on validation studies is often difficult or infeasible for many surveys. Because of this, response rates are seen as one of the primary indicators of data quality.

Considerable attention has been focused on how to construct unit response rates properly (e.g., American Association for Public Opinion Research 1998) and analyzing item response rates for individual survey questions (e.g. Ferber 1966). However, little attention has been given to item response rates for composite variables which are created from several survey questions.[2] Examples of such variables include net worth from the Survey of Income and Program Participation (SIPP) and the Survey of Consumer Finances (SCF), the unemployment rate from the Current Population Survey (CPS), and household income from the CPS Annual Social and Economic Supplement. While the item response rates are available for the underlying components, this information may be hard for data users to synthesize. The SCF, for example, has over 140 assets and liabilities questions. Without aggregating the response rates in a meaningful way, it is extremely difficult to assess how key estimates from these surveys are impacted by item nonresponse. In addition, these composite variables are often the key measures in surveys, receiving the most attention and scrutiny. Because researchers often look at response rates to gauge the quality of a variable, having such information for composite variables could be useful for individuals evaluating data quality. To the best of my knowledge, there has been no paper which focuses on how to construct item nonresponse rates for composite variables. In a paper on item nonrespone bias, Hokayem et al. (2017) briefly describe total item response rates for household income

[2] Such variables are also called recode or summary variables.

in the CPS Annual Social and Economic Supplement. However, their paper is not focused on item response rate calculations, nor do they discuss alternative ways of constructing such a measure.

In this paper, I propose several new methods for aggregating data on response rates across questions to create a measure of item response for composite variables. These methods provide a useful way to summarize item response rates for key measures, and offer a feasible way of comparing item response rates across surveys and over time. The three proposed measures of response rates are

1. The percent of observations without any missing component,
2. A sum-weighted formula, which is the sum of reported values divided by the sum of all values, and
3. A median-weighted formula, in which the response rate for each question is weighted by the percent of respondents with a non-missing/non-zero value times the median value.

For each method, I discuss their advantages and disadvantages as well as how they differ from the other measures. In these methods, I also incorporate data on partial item nonresponse from questions that contain a range follow-up for respondents who give an initial answer of "Don't Know" or "Refuse." (e.g., respondents who don't know the exact balance of their bank account, but are able to say that the balance was between $500 and $1,000). Similar to the American Association for Public Opinion Research (1998) formulas for unit response rates, these three measures provide a standardized way of comparing item nonresponse rates across surveys.

To show how these measures can be used to analyze item response for composite variables, I analyze item response for net worth in the Survey of Income and Program Participation (SIPP). The U.S. Census Bureau redesigned SIPP starting with the 2014 Panel, at which time numerous changes were made to the survey. The asset section underwent a major revision; new assets were added and asset income and values were asked together rather than in separate sections. With all three of the proposed composite response rate measures, I find that overall item response rates went up for wealth questions

in the new panel. For specific asset and debt categories, item response rates increased for the majority of categories, even for assets for which the question text changed very little between the panels. This increase in item response rates suggests a potential improvement in data quality, although item response rates remain low for a variety of assets.

In addition, I also apply our item response aggregation method to the Survey of Consumer Finances (SCF), which is another survey that contains wealth data. Because SCF has more detailed wealth questions and its interviewers receive more training about types of assets, the SCF has been labeled as the "gold standard" for wealth data, and thus is a useful survey to provide a comparison to SIPP.[3] I find that overall item response rates are lower in SIPP, even after the redesign, than in the SCF. Thus, while item response rates did increase in the 2014 Panel of SIPP, they are still lower than rates for the SCF. My proposed aggregation method for item response rates greatly facilitates such insights. While a researcher could compare item response rates for individual questions between the two surveys, this task would be difficult given the large number of questions and issues with comparability of question text and concepts.

The rest of this paper proceeds as follows. In Section 2, I give a description of the SIPP data and highlight response rates for some key variables. In Section 3, I discuss the new methods for aggregating data on response rates. In Section 4, I present the results of the item response analysis on SIPP and SCF data. Finally, Section 5 concludes.

2. Background

Before I present my method for aggregating item response rates, I first present an overview of the SIPP data and highlight response rates for particular variables. This discussion will show that

[3] National Research Council (2009) is one among many sources that have applied this label in reference to SCF.

comparing question-level response rates over time is difficult when the questionnaire changes, suggesting the need for a different measure for comparing item response rates over time.

The SIPP is a longitudinal survey that collects information about the income, assets, labor market activity, and participation in government programs of U.S. households. Information on a wide variety of assets and debt is collected and includes financial variables such as savings accounts, checking accounts, retirement accounts, property values, and credit card debt. SIPP interviews households for a period of about 2.5 to 5 years (depending on the panel), with each panel containing a new set of households. In order to improve estimates of receipt from government programs, SIPP oversamples low-income areas.

In the 2014 panel, the U.S. Census Bureau made many changes to the survey. One substantial change is that SIPP now interviews respondents less frequently in order to reduce costs (U.S. Census Bureau, 2016). In the 2014 panel, interviews occurred once per year, but in earlier panels, interviews occurred every four months. However, interviews with questions on asset values occurred about once a year in previous panels, with some gaps across time. Therefore, SIPP collects wealth data with roughly the same frequency in the 2014 Panel as before. Another 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.

The asset section also underwent a major revision. SIPP now collects data on additional assets that were not asked about previously, such as annuities, trusts, student loans, and education savings accounts. Moreover, questions on asset income and asset values are now asked concurrently rather than in separate sections of the interview. Wording for many questions changed as well. Because so much of the survey was modified, the quality of the net worth data may have changed substantially. To investigate changes in data quality, Eggleston and Gideon (2017) compared SIPP wealth estimates to estimates from the SCF. They found that the discrepancy between SIPP and SCF wealth estimates, such

as median net worth, decreased in the 2014 Panel, suggesting an improvement in SIPP data quality after the redesign.

Table 1: Response Rates for Select Value Variables

SIPP 2014 Wave 1 (Calendar Year 2013)				SIPP 2008 Wave 4 (Fall 2009)			
Value Variable	Reported	Imputed without range	Imputed within range	Value Variable	Reported	Imputed without range	Imputed within range
Primary Residence	87.1 (0.28)	12.9 (0.28)		Primary Residence	75.1 (0.48)	24.9 (0.48)	
401k/Thrift Retirement Account	56.0 (0.56)	19.5 (0.41)	24.5 (0.39)	401k/Thrift Retirement Account	46.2 (0.57)	27.3 (0.53)	26.5 (0.54)
IRA/Keogh Retirement Account	59.2 (0.67)	19.2 (0.46)	21.6 (0.53)	IRA Retirement Account	49.6 (0.64)	27.8 (0.55)	22.6 (0.50)
				Keogh Retirement Account	28.5 (1.87)	42.6 (2.43)	28.9 (1.94)
Savings Account (Own Name)	63.2 (0.50)	23.0 (0.46)	13.8 (0.31)	Interest-Earning Bank Accounts (Own Name)	54.2 (0.53)	25.9 (0.46)	23.1 (0.34)
Interest Checking (Own Name)	62.5 (0.70)	22.0 (0.62)	15.5 (0.48)				
Education Savings Account (1st Account)	72.1 (1.55)	12.4 (1.30)	15.5 (1.15)				
Trusts	47.2 (2.58)	31.4 (2.87)	21.4 (2.47)				

Table presents allocation rates for a small set of value variables (variables which give the market value of an asset or debt) from the 2008 and 2014 SIPP Panels. Replicate weights were used to construct standard errors, which are shows in parenthesis. "Imputed within range" means that the respondent had item nonresponse for the original question, but responded to the question which asked if the asset or debt value was in a particular interval (e.g. $1,000 to $2,000). The difference in the imputation rates between the 2014 and 2008 Panel for Primary Residence and 401k/Thrift Retirement Accounts is significant at the one percent level.
Source: 2008 SIPP Panel (Wave 4) and 2014 SIPP Panel (Wave 1).

Table 1 presents item response rates for some select variables in 2008 and 2014 SIPP panels. I define item nonresponse as the proportion of people who give an answer of "don't know" or "refuse" to a question, or who drop out of the survey before reaching the particular asset question. Households that have unit nonresponse are dropped from the sample, so they are not included in the item response rate calculations.[4] In addition, for initial item nonresponders, some questions in SIPP have a follow-up question which asks whether the value of an asset is within a given range (e.g. $1,000 to $2,000). Because this creates varying degrees of item response, I make this distinction when presenting response rates in tables.

For some variables in Table 1, such as the value of primary residences and 401k/Thrift retirement accounts, there is a directly comparable variable in both the 2008 and the 2014 Panels. This table shows that item response rates went up for these variables, with the reporting rate for primary residences values increasing from 75.1 percent to 87.1 percent, and the reporting rate for 401k/Thrift retirement accounts values increasing from 46.2 percent to 56.0. These changes are surprising, given the question texts changed very little for these variables (Eggleston and Gideon 2017).

For other SIPP variables, comparisons between the 2008 and 2014 Panels are not clear cut. For example, in the 2008 Panel, there was one combined question on the value of interest-earning checking, savings, CDs, or money market accounts. In the 2014 Panel, there are separate questions for these different types of bank accounts. With differences in the question design, how to make direct comparisons of the response rates across panels is unclear. Similarly, questions on IRA and Keogh retirement account balances are now combined for the 2014 Panel, so the old questions are not directly comparable. In addition, questions on educational savings accounts and trusts are new to the 2014

[4]In SIPP, the goal is to interview everyone in the household, either through an in-person interview, or through a proxy interview. If no one in the household responds, then the household is considered to be a unit nonresponder. In some other households, one person is interviewed but other people are unable to be interviewed. These noninterviewed people are treated as item nonresponders, so they are included in the item response rate calculations.

Panel. The trusts item, for example, has a low item response rate of 52.8 percent. Because these new questions have no analogues from earlier panels, evaluating the effects of these new variables on data quality needs to be addressed in a different way than variable-to-variable comparisons in response rates over time.

3. Equations for Item Nonresponse for Composite Variables

Because changes in a survey often involve adding or combining questions, there needs to be a way of aggregating item non-response data across variables in order to generate statistics which can be used to compare response rates over time. In this section, I present three different ways of aggregating response rates over several variables. I focus on continuous numeric variables in this section, such as data on income and wealth, but I discuss at the end how the results can be generalized to other types of variables.

First, I discuss the mathematical notation used for these equations. Let individual i's response to question q be denoted by $r_{i,q}$, where $r_{i,q} = 1$ if the individual responds to the numeric question, $r_{i,q} = 1/2$ if the individual does not answer the numeric question but gives an answer to a range follow-up question, $r_{i,q} = 0$ if the individual does not respond to either the numeric question or the range follow-up question, and $r_{i,q} = NIU$ (not in universe) if the question is not applicable to the respondent. The value a respondent has for question q, which is either the actual response or the imputed value, is denoted by $x_{i,q}$. From these Q underlying variables, the value of the c^{th} composite variable $y_{i,c}$ is the sum of a set of variables minus the sum of a different set of variables. For example, an individual's or household's net worth is the sum of all assets minus all debts. Letting $d_{c,q}$ =1,0,-1 denote the indicator of the variables that are added, ignored, or subtracted, respectively, for the composite variable, the equation for $y_{i,c}$ is given by $\sum_{q=1}^{Q} d_{c,q} x_{i,q}$.

For the item response rate for question q, $p_{q,k}$ denotes the weighted proportion of individuals with response outcome $r_{i,q} = k$. Using w_i to denote the sample weight for individual i, the formula for $p_{q,k}$ is given by the equation

$$p_{q,k} = \sum_{i=1}^{N} w_i 1(r_{i,q} = k) \; / \; \sum_{i=1}^{N} w_i \, 1(r_{i,q} \neq NIU) \tag{1}$$

.

I use survey weights in order to account for over-sampling of certain populations, allowing $p_{q,k}$ to reflect the non-response propensity of a random person from the general population.

3.1. Percent with No Missing Value

The first aggregation method I present is the proportion of individuals who give a response to every question used to create the composite variable c, given by the equation

$$a_c = 1 - \sum_{i=1}^{N} w_i \max\left(|d_{c,1}|1(r_{i,1} \in \{0,1/2\}), \dots, |d_{c,Q}|1(r_{i,Q} \in \{0,1/2\})\right) / \tag{2}$$

$$\sum_{i=1}^{N} w_i \max\left(|d_{c,1}|1(r_{i,1} \neq NIU), \dots, |d_{c,Q}|1(r_{i,Q} \neq NIU)\right).$$

One advantage of this formula is that it is straightforward to explain and understand. In addition, as some researchers drop imputed values from their analyses, this statistic gives an indicator of the proportion of the sample that would be kept if imputed values are excluded. The main disadvantage of this formula is that it does not capture varying degrees of item response which may be meaningful. For example, this formula does not describe whether households tended to give an answer of "Don't Know" or "Refuse" for a small number of questions. Also, the more questions that are used to create the composite variable, the more likely it is that the household will not give an answer to at least one question, which may create a false sense of inaccuracy for composite variables that are created for a large set of detailed questions.

3.2. Sum-Weighted Response Rates

The second aggregation method is the weighted sum of values with a given response type (e.g. gave a range follow-up response) divided by the weighted sum of all values. This is given by the equation

$$s_{c,k} = \sum_{i=1}^{N} w_i \left(\sum_{\mathrm{q}=1}^{\mathrm{Q}} |d_{c,q}| 1(r_{i,q} = k) x_{i,q} \right) / \sum_{i=1}^{N} w_i \left(\sum_{\mathrm{q}=1}^{\mathrm{Q}} |d_{c,q}| x_{i,q} \right), \quad (3)$$

where $s_{c,1}$, for example, represents the percent of all values for composite variable c that consist of reported values. Hokayem et al. (2017) use such a formula to describe response rates for household income in the CPS Annual Social and Economic Supplement. One advantage of this formula (Equation 3) over the previous one (Equation 2) is that it gives more weight to assets that have larger values, on average. For example, if home equity constitutes a larger proportion of household net worth than bank accounts, then the sum-weighted formula would put more weight on imputed home values than bank account values. Equation 2 gives equal weight to home value and bank account values if the household owns both assets, which may be undesirable if a researcher is concerned about assets that constitute a larger proportion of wealth portfolios.

One potential disadvantage of the sum-weighted formula is that respondents with larger values are given more weight than other respondents. For example, if high-wealth households have lower response rates than other respondents, then this facet of the data will cause $s_{c,k}$ to be higher.[5] This feature may be desirable when looking at statistics based on a mean or sum, such as aggregate income or net worth, as such statistics are influenced heavily by outliers. However, if a researcher is instead focused on median net worth, then this item response statistic may not properly reflect the behavior of respondents in the middle of the distribution.

[5] Lillard et al. (1986), for example, find that high-income respondents are less likely to answer the wage question in the CPS.

3.3. Median-Weighted Response Rates

The final formula is a weighted average of response rates from each question. The weight is the weighted percent of individuals who have a non-missing and non-zero value for question q, denoted by o_q, times the weighted median of non-missing and non-zero responses, denoted by $med(x_{i,q})$. For assets, o_q represents the ownership rate for a given asset while for income, o_q represents the proportion of people who are receiving a particular source of income.[6] With this notation, the median-weighted item response rate for composite variable c is denoted by

$$m_{c,k} = \sum_{q=1}^{Q} |d_{c,q}|\, o_q med(x_{i,q}) p_{q,k} - \sum_{q=1}^{Q} |d_{c,q}|\, o_q med(x_{i,q}), \quad (4)$$

In which $p_{q,k}$ is as defined in Equation (1). In this formula, more weight is given to variables with a higher ownership (non-missing and non-zero) rate, and higher median values. Similar to the second formula, this formula gives more weight to more "important" variables that constitute a larger proportion of the composite variable. However, one benefit of this formula over the second is that the statistic is less influenced by outliers, which could be useful in order to gauge the nonresponse behavior of the "typical" respondent, or for analyzing the impact of item nonresponse on median estimates.

[6] In SIPP, people are asked if they owned a particular asset at any time during the reference period, but are asked to give the asset's value as of the end of the reference period. If they sold the asset, they are instructed to report a value of zero. Because of this technicality, I construct the ownership rate as the proportion of people with both a non-missing and non-zero value.

Table 2: Ownership Rates and Median Values for Select Variables (2014 SIPP)

Variable	Ownership Rate	Median Value	Weight (in Thousands)	Weight as a Percentage of the Primary Residence Weight
Primary Residence	58.9	180,000	10,602	100
401k/Thrift Retirement Account	40.2	45,300	1,821	17.2
IRA/Keogh Retirement Account	28.1	40,000	1,124	10.6
Savings Account (Own Name)	48.4	1,400	68	0.6
Interest Checking (Own Name)	22.5	2,000	45	0.4
Education Savings Account (1st Account)	4.0	10,000	40	0.4
Trusts	1.5	100,000	150	1.4

Table presents ownership rates and median values for select variables from SIPP. In addition, to help explain the median-weighted formula for allocation rates presented in Equation (4), this table also shows the weight of the ownership rates times median value, and the percentage the weight is of the weight for primary residences. Because SIPP is a household survey, I construct the ownership rates and the median values at the household level, even for questions asked to every adult in the household.
Source: 2014 SIPP Panel (Wave 1).

To help provide context for the median-weighted formula, Table 2 presents ownership rates and median values in 2014 SIPP for select variables, as well as the weight for these variables (before they are normalized to sum to one). This table shows that primary residences are given a large weight because 58.9 percent own a home, and the median home value is $180,000, which generates a weight of $o_q * med(x_{i,q})$= 188,000*58.9= 10,602,000.[7] The weight for 401k and thrift accounts is 17.2 percent of the weight for primary residence, as the ownership rates are lower and the median value is about $45,000.

[7] These statistics exclude mobile homes, which are captured through a separate variable.

IRAs follow a similar pattern with a weight that is 10.6 percent of the weight for primary residence. Other assets presented in this table have a much lower weight. Because savings accounts held in someone's own name have a median value of only $1,400, the weight is 0.6 percent of the weight for primary residences. Even through trusts have a median value of $100,000, the weight for trust is 1.4 percent of the weight for primary residences since the ownership rate is 1.5 percent. In summary, this table shows that when applying the median-weighted formula for all assets and debt in 2014 SIPP, the item response rates for home values and retirement account balances are given a relatively high weight, while bank accounts and trusts have a much lower weight.

3.4. Non-Numeric Variables

The previous formulas have focused on numeric variables such a household income or wealth. However, some key measures, such as the unemployment rate, are based on a series of yes/no and discrete-choice questions, rather than on numeric variables. For such variables, aggregated measures of item response can also be constructed by modifying some of the formulas already presented in this section. The statistic of having no imputed values is easy to extend to non-numeric variables, as Equation (2) does not rely on the underlying variables being numeric. Specifically, if $d_{c,q}$ now denotes whether question q was used to construct the composite variable c, the measure for having any imputed value is exactly the same as in Equation 2. In addition, Equation 4 for the median-weighted response rate can be modified to remove the median function from the weight to generate a statistic for non-numeric variables. In this revised formula, more weight is given to questions that are on-path/in-universe for a larger set of respondents, which similarly gives more weight to more "important" questions. Thus, this method of aggregating response rates across questions is applicable to a variety of key measures, including measures based on non-numeric variables.

4. Analyses on SIPP and SCF Data

To show how the formulas presented in Section 3 can be used to analyze item response rates, I investigate how item response rates for wealth questions changed in SIPP after the 2014 redesign. Tables 3-5 present the results from applying the three formulas presented in Section 3. In these tables, I organize assets into broad categories, such as bank accounts. Details about these categories are presented in Appendix A. Because SIPP is a household survey and some of the composite variables are at the household level, I let the indicator of having no imputed values equal zero if anyone in the household has an imputed value. In addition, for the median-weighted response rate, I construct the ownership rates and the median values at the household level, even for questions asked to every adult in the household.[8]

In these tables, I construct standard errors for the response rates. I do this to account for sampling error that could cause one of the panels to have more respondents who item respond, even if the questionnaire remained the same. I use the SIPP replicate weights for the standard errors, which are created using Fay's method. See U.S. Census Bureau (2016) for details about the construction of standard errors.

[8] In SIPP, some questions, such as home values, are asked to only one member in the household, while other questions are asked to every individual over 15. For financial assets and rental property that married couples own jointly, only one spouse is asked the value of the asset.

Table 3: SIPP Results for No Imputed Values

Category	Percent of Households with no Value Imputed		
	SIPP 2014	SIPP 2008	P-value Difference
All Assets and Debt	27.7 (0.29)	29.7 (0.36)	<.001
Retirement Assets	52.4 (0.52)	42.8 (0.51)	<.001
All Financial Assets not in Retirement Accounts	42.4 (0.38)	29.2 (0.46)	<.001
Bank accounts	56.2 (0.37)	52.2 (0.49)	<.001
Stocks	45.8 (0.81)	40.2 (0.68)	<.001
Bonds	49.8 (1.34)	51.8 (0.85)	0.215
Real Estate	75.7 (0.34)	64.9 (0.49)	<.001
Vehicles	63.3 (0.39)	62.4 (0.45)	0.106
Other Assets	42.4 (0.50)	29.2 (0.64)	<.001
Business	52.5 (0.82)	23.7 (0.69)	<.001
Unsecured Debt	68.7 (0.42)	71.9 (0.46)	<.001

Table presents the percentage of households with any imputed value within an aggregated wealth category. Replicate weights were used to construct standard errors, which are shown in parenthesis.
Source: 2008 SIPP Panel (Wave 4) and 2014 SIPP Panel (Wave 1).

Table 4: SIPP Results for Sum-Weighted Response Rates

Category	SIPP 2014			SIPP 2008			P-value Difference
	Reported	Imputed without range	Imputed within range	Reported	Imputed without range	Imputed within range	
All Assets and Debt	69.9	18.2	11.9	63.9	28.6	7.5	<.001
	(1.66)	(0.90)	(1.09)	(1.41)	(1.52)	(1.59)	
Retirement Assets	66.3	11.4	22.2	51.0	25.9	23.1	0.017
	(0.98)	(0.44)	(0.73)	(5.39)	(7.39)	(6.44)	
All Financial Assets not in Retirement Accounts	49.7	24.6	25.7	55.4	28.0	16.6	<.001
	(1.60)	(1.15)	(1.58)	(4.23)	(3.60)	(1.82)	
Bank accounts	52.6	22.0	25.4	59.5	26.4	14.1	<.001
	(1.85)	(1.13)	(1.92)	(5.16)	(3.38)	(1.95)	
Stocks	48.2	26.3	25.4	50.1	31.2	18.8	0.153
	(2.33)	(1.98)	(2.17)	(6.71)	(5.99)	(2.94)	
Bonds	42.9	28.0	29.1	72.5	14.6	12.9	<.001
	(4.00)	(3.77)	(3.24)	(4.38)	(2.63)	(2.85)	
Real Estate	83.2	14.5	2.2	73.8	24.6	1.7	<.001
	(0.61)	(0.54)	(0.39)	(0.74)	(0.57)	(0.42)	
Vehicles	78.0	22.0		74.1	25.9		<.001
	(0.33)	(0.33)		(0.41)	(0.41)		
Other Assets	50.7	26.2	23.1	32.0	66.8	1.3	<.001
	(7.21)	(4.12)	(4.72)	(2.73)	(2.72)	(0.22)	
Business	52.1	24.2	23.7	27.9	72.1		<.001
	(10.00)	(5.56)	(6.34)	(3.15)	(3.15)		
Unsecured Debt	76.1	23.9		75.7	24.3		0.727
	(0.91)	(0.91)		(1.03)	(1.03)		

Table presents allocation rates across aggregated wealth categories using the sum-weighted formula presented in Equation (3). The allocation rates are the ratio of the sum of all values with a given allocation flag divided by the sum of all values for a given wealth category. Replicate weights were used to construct standard errors, which are shown in parenthesis. For the statistical test comparing the allocation rates between the 2008 and 2014 Panels, a Z-test was used for categories without a range follow-up option in either panel, and a Chi-squared test was used for all the other categories. "Imputed within range" means that the respondent had item nonresponse for the original question, but responded to the question which asked if the asset or debt value was in a particular interval (e.g., $1,000 to $2,000).
Source: 2008 SIPP Panel (Wave 4) and 2014 SIPP Panel (Wave 1).

Table 5: SIPP Results for Median-Weighted Response Rates

Category	SIPP 2014 Reported	SIPP 2014 Imputed without range	SIPP 2014 Imputed within range	SIPP 2008 Reported	SIPP 2008 Imputed without range	SIPP 2008 Imputed within range	P-value Difference
All Assets and Debt	77.3	17.3	5.3	68.8	27.7	3.5	<.001
	(0.30)	(0.26)	(0.12)	(0.43)	(0.41)	(0.11)	
Retirement Assets	57.2	19.4	23.4	47.3	27.7	25.0	<.001
	(0.53)	(0.35)	(0.37)	(0.53)	(0.48)	(0.46)	
All Financial Assets not in Retirement Accounts	53.4	25.1	21.5	48.6	32.0	19.4	<.001
	(0.72)	(0.62)	(0.51)	(0.61)	(0.60)	(0.50)	
Bank accounts	60.2	22.0	17.8	57.6	29.3	13.1	<.001
	(0.63)	(0.60)	(0.46)	(0.51)	(0.43)	(0.31)	
Stocks	47.5	27.7	24.8	43.0	34.9	22.1	<.001
	(0.93)	(0.83)	(0.69)	(0.72)	(0.75)	(0.61)	
Bonds	47.5	28.1	24.3	55.7	23.5	20.7	0.032
	(2.59)	(1.81)	(1.84)	(1.76)	(1.60)	(1.46)	
Real Estate	83.8	15.1	1.1	72.5	26.6	0.9	<.001
	(0.29)	(0.29)	(0.07)	(0.46)	(0.45)	(0.08)	
Vehicles	79.3	20.7		73.5	26.5		<.001
	(0.28)	(0.28)		(0.39)	(0.39)		
Other Assets	52.3	31.0	16.7	34.4	61.9	3.8	<.001
	(0.99)	(1.10)	(0.79)	(0.84)	(0.95)	(0.42)	
Business	54.7	22.2	23.1	27.1	72.9		<.001
	(1.48)	(1.36)	(1.18)	(0.77)	(0.77)		
Unsecured Debt	75.1	24.9		74.0	26.0		0.129
	(0.45)	(0.45)		(0.56)	(0.56)		

Table presents allocation rates across aggregated wealth categories using the median-weighted formula presented in Equation (4). The allocation rates are a weighted average of the allocation rates from the underlying variables, where the weights are the ownership rate times the median value conditional on ownership. Replicate weights used to construct standard errors. For the statistical test comparing the allocation rates between the 2008 and 2014 Panels, a Z-test was used for categories without a range follow-up option in either panel, and a Chi-squared test was used for all the other categories. "Imputed within range" means that the respondent had item nonresponse for the original question, but responded to the question which asked if the asset or debt value was in a particular interval (e.g., $1,000 to $2,000). Source: 2008 SIPP Panel (Wave 4) and 2014 SIPP Panel (Wave 1).

4.1. Percent with no Missing Value

Table 3 presents the results for having no imputed values. When looking at any asset or debt variable, 29.7 percent of households in the 2008 Panel have no imputed values, while 27.7 percent of household in the 2014 panel have no imputed value. Rates vary across assets, with the rates being high for real estate (75.7 percent in the 2014 Panel), but lower for stocks (45.8 percent in the 2014 Panel). When comparing across panels, the 2008 Panel has lower response rates for most categories. However, the response rates for unsecured debt and all assets and debt are lower in 2014 SIPP, and the difference for bonds is insignificant. Overall, the results for Table 3 show the rate of having no missing values is low in both panels, with the 2008 Panel having lower rates for many assets.

4.2. Sum-Weighted Response Rates

Table 4 shows a similar pattern with the sum-weighted formula. This table shows that the 2008 Panel has lower response rates, although the results are somewhat mixed. When combining data on all asset and debt questions, the sum-weighted response rate is 69.9 percent for the 2014 Panel, meaning that about 70 percent of the estimate of aggregate wealth in SIPP consists of reported values rather than imputed values. In the 2008 Panel, the sum-weighted response rate is lower at 63.9 percent. The pattern is similar for many other asset groupings, with 66.3 percent of the aggregate value of retirement assets in the 2014 Panel consisting of reported values, versus only 51.0 percent in the 2008 Panel. However, for bank accounts, item response rates are higher in the 2008 Panel, with a reporting rate of 59.5 percent for the 2008 Panel but 52.6 percent for the 2014 Panel.

Table 6: Response Rates and Percent of Aggregate Assets by Net Worth Quintile

Net Worth Quintile	Percent Wealth Reported, Sum Weighted	Percent Aggregate Assets
1	83.0	2.1
	(0.91)	(0.14)
3	77.1	8.4
	(0.63)	(0.41)
5	67.52	89.5
	(2.40)	(0.53)

Table presents allocation rates across all wealth variables by net worth quintile, and the percent of aggregate assets held by each net worth quintile. The allocation rates are the ratio of the sum of all values with a given allocation flag divided by the sum of all values for a given wealth category. Replicate weights were used to construct standard errors, which are shown in parenthesis.
Source: 2014 SIPP Panel (Wave 1).

One potential characteristic of the sum-weighted formula is that more weight may be given to observations with larger values. To illustrate this, Table 6 shows how the overall reporting rate for asset and debt items varies by net worth quintile. This table shows that for observations in the bottom 5th of the wealth distribution, the overall weighted response rate is 83 percent, but the response rate drops to 67.5 percent for the top 5th of the wealth distribution. This result could be driven either by high-wealth individual being less likely to respond for any given question, or because high wealth households hold assets like trusts, which have a lower item response rates. To describe the relative weight each quintile has on the sum-weighted response rate, Table 6 also shows the percent of aggregate assets that are held by each quintile of the wealth distribution. Aggregate assets are shown because unlike net worth, assets are greater than or equal to zero for every household, which prevents the percent aggregate statistics from being negative for any quintile. These results show that the top quintile of the wealth distribution holds 89.5 percent of total assets, suggesting that the item response behavior of the top quintile is given a large weight when looking at item response rates for the entire sample.

Table 7: Decomposition of Median-Weighted Formula

Category	Percentage Total Weight
Retirement Assets	12.0 (0.33)
All Financial Assets not in Retirement Accounts	5.5 (0.16)
Bank accounts	2.5 (0.09)
Stocks	2.7 (0.13)
Bonds	0.3 (0.06)
Real Estate	70.3 (0.46)
Vehicles	6.0 (0.07)
Other Assets	4.0 (0.35)
Business	1.3 (0.30)
Unsecured Debt	2.2 (0.07)

Table presents the percentage contribution each group's item response rate has when constructing the median-weighted response rate in Equation (4). Replicate weights were used to construct standard errors, which are shown in parenthesis.
Source: 2014 SIPP Panel (Wave 1).

4.3. Median-Weighted Response Rates

To reduce the influence of large values, I also present the median-weighted response rates in Table 5. In this table, the reporting rate from aggregating all asset and debt variables is 77.3 percent in the 2014 Panel, compared with 68.8 percent in the 2008 Panel. These median-weighted results are similar to Table 4 for the sum-weighted formula, although the levels are higher for each panel. When looking at other asset and debt categories, Table 5 also shows that response rates are higher in the 2014 Panel,

with a response rate of 60.2 percent for bank accounts in the 2014 Panel, but 57.6 percent in the 2008 Panel.

In the median-weighted formula, some assets and debt are given more weight than others. For example, Table 2 in Section 3.3 shows that home value and retirement accounts are given a relatively high weight. To expound on this detail, Table 7 presents the relative weight each of the asset and debt categories have when calculating the median-weight response rate for all asset and debt items. As suggested by earlier results, Table 7 shows that the response rate for real estate (which includes home values) make up 70.3 percent of the overall response rate, and retirement accounts make up 12.0 percent. Bonds, which are a more uncommon asset, only make up 0.3 percent of the overall response rates. The weight real estate is given in the formula is reflective of how home equity constitutes a large proportion of household net worth. For example, in 2014 SIPP (Wave 1) median net worth is $80,039, but when home equity is excluded, the median drops by 68.6 percent to $25,116 (U.S. Census Bureau 2017). Overall, this table shows that assets that have lower response rates have less weight in the median-weighted response rate formula. Thus, while there are some assets like trusts that have a response rate of under 50 percent, many of the assets which constitute a larger proportion of many people's net worth, such as home values, have a much higher response rates. These assets have more influence in the median-weighted formula.

Despite the differences in levels, all three formulas show a similar pattern of the 2014 Panel having higher item response rates. The reasons for this are unclear. While question text did change for a variety of assets, the response rate is also higher for assets and asset groupings in which the question text changed very little between the panels. For example, in Table 1 in Section 2, the item response rate for primary residences is higher in the 2014 Panel, even though the question text is almost identical between the panels (Eggleston and Gideon 2017).

One difference in the wealth data between the panels is that wealth data is not collected until the 4th wave of the 2008 Panel, but wealth data is collected in the 1st wave of the 2014 Panel. Because of this, either attrition or the effect of repeated interviewing may have some impact on the response rates in the 2008 Panel. In waves after the initial interview, households are removed from the dataset in a given wave if the household either attrites from the panel or has unit nonresponse in that wave but is interviewed in later waves. Because of this, respondents who are less engaged and have item nonresponse in the initial interview may attrite from the SIPP panel, which would increase item response rates in later waves. On the other hand, respondents may learn that the interview goes faster if they have item nonresponse, which would decrease item response rates in later waves.

Another potential reason for the change is due to differences in unit response rates. Unit response rates (AAPOR RR6) are 80.6 percent in the 2008 Panel but only 68.8 percent in the 2014 Panel. Yan et al. (2010) find that decreasing unit response rates in the Survey of Consumers has been associated with higher income item response rates, potentially suggesting that people who unit nonresponse are more likely to have item nonresponse at well. This finding would suggest that part of the reason item response rates are higher in the 2014 Panel is because unit response rates are lower.

4.4. SIPP and SCF Comparison

To further investigate data quality in the 2014 Panel, I also analyze item response for wealth questions in the Survey of Consumer Finances (SCF) and compare these numbers to SIPP. The SCF is a survey on the income and wealth of U.S. families sponsored by the Federal Reserve. This survey is primarily cross-sectional and is conducted once every three years.[9] To estimate the income and wealth of high-wealth families, the SCF oversamples high-income individuals based on data from the IRS. This is in contrast with SIPP, which oversamples low-income areas. In addition, the SCF has a much smaller sample size than SIPP. The 2013 SCF contains about 6,000 families, while wave 1 of the 2014 SIPP Panel

[9] The SCF has only reinterviewed families once after the 2007 SCF in response to the Great Recession.

contains about 30,000 households, and the first wave of the 2008 Panel contains about 42,000 households.[10]

Table 8: SIPP and SCF Results for Median-Weighted Response Rates

Category	SIPP 2014			SCF 2013			P-value Difference
	Reported	Imputed without range	Imputed within range	Reported	Imputed without range	Imputed within range	
All Assets and Debt	77.3	17.3	5.3	85.0	3.4	11.7	<.001
	(0.30)	(0.26)	(0.12)	(0.72)	(0.23)	(0.73)	
Retirement Assets	57.2	19.4	23.4	73.6	7.9	18.5	<.001
	(0.53)	(0.35)	(0.37)	(1.72)	(0.80)	(1.41)	
All Financial Assets not in Retirement Accounts	53.4	25.1	21.5	75.6	7.2	17.2	<.001
	(0.72)	(0.62)	(0.51)	(1.70)	(1.18)	(1.59)	
Bank accounts	60.2	22.0	17.8	80.8	4.0	15.2	<.001
	(0.63)	(0.60)	(0.46)	(1.34)	(0.63)	(1.26)	
Stocks	47.5	27.7	24.8	73.6	7.5	18.8	<.001
	(0.93)	(0.83)	(0.69)	(2.48)	(1.57)	(2.17)	
Bonds	47.5	28.1	24.3	71.7	17.1	11.2	0.032
	(2.59)	(1.81)	(1.84)	(7.82)	(6.87)	(4.41)	
Real Estate	83.8	15.1	1.1	88.8	1.4	9.7	<.001
	(0.29)	(0.29)	(0.07)	(0.67)	(0.23)	(0.72)	
Vehicles	79.3	20.7		90.8	1.3	7.9	<.001
	(0.28)	(0.28)		(0.73)	(0.19)	(0.74)	
Other Assets	52.3	31.0	16.7	73.7	8.9	17.4	<.001
	(0.99)	(1.10)	(0.79)	(3.89)	(2.35)	(3.39)	
Business	54.7	22.2	23.1	72.5	6.3	21.2	<.001
	(1.48)	(1.36)	(1.18)	(3.35)	(1.40)	(3.32)	
Unsecured Debt	75.1	24.9		85.5	3.3	11.2	0.129
	(0.45)	(0.45)		(1.48)	(0.58)	(1.49)	

Table compares allocation rates between SIPP and SCF aggregated wealth categories using the median-weighted formula in Equation (4). The allocation rates are a weighted average of the allocation rates from the underlying variables, where the weights are the ownership rate times the median value conditional on ownership. Replicate weights used to construct standard errors in both surveys, and imputation implicates were used to construct the standard error for SCF. Standard errors shown in parenthesis. For the statistical test comparing the allocation rates between 2014 SIPP and 2013 SCF, a Z-test was used for categories without a range follow-up option in either panel, and a Chi-squared test was used for all the other categories. "Imputed within range" means that the respondent had item nonresponse for the original question, but responded to the question which asked if the asset or debt value was in a particular interval (e.g., $1,000 to $2,000).
Source: 2014 SIPP Panel (Wave 1) and 2013 SCF.

[10] The 2013 SCF has a response rate (AAPOR RR1) of about 70 percent for the general population sample, and a rate of about one-third for the high wealth oversample. See Bricker et al. (2014) for more details.

The SCF has been labeled the "gold standard" for wealth data because it has more detailed wealth questions and its interviewers receive more training about types of assets. Because of this, SCF is a useful survey to compare with SIPP. In Table 8, I present results comparing median-weighted response rates between the 2014 SIPP and 2013 SCF.[11] This table shows that item response rates are higher in the SCF than in SIPP for all wealth categories. For bank accounts, SCF has a weighted response rate of 80.8 percent, but the response rate for the SIPP is only 60.2 percent. This trend holds when looking at all asset and debt variables, in which the response rate for SCF is 85 percent but is 77.3 percent in SIPP. Of all respondents with some degree of item nonresponse, SCF also has more individuals whose values are imputed from a range than SIPP. For bank accounts, 44.7 percent of SIPP respondents who didn't respond to the initial bank account question have a value that is imputed from a range, but this rate is 79.2 percent for SCF respondents.[12]

The reason for SCF having higher item response rates is unclear. The surveys do have some differences in question text that could affect response rates. For example, SCF lets respondents list the balance of different checking accounts separately, while the checking question in SIPP requires that respondents add up the balance of different accounts when reporting their answer (Eggleston and Gideon 2017). Given this mathematical calculation might be more cognitively demanding, this requirement in SIPP could decrease item response rates for checking accounts. However, differences in question text might not entirely explain the differences between SIPP and SCF. For example, item response rates are lower in SIPP for real estate, even though the question text for primary residences is similar between the two surveys (Eggleston and Gideon 2017). One overall difference between the

[11] While not shown in this paper, the analogue analysis between 2014 SIPP and 2013 SCF is comparable when looking at the rates of any imputed values or sum-weighted response rates.

[12] In this statistic on bank accounts and range follow-ups, the 44.7 estimate for SIPP is 17.8, the imputed within range estimate, dividend by (100-60.2), which is the reported rate. The SCF estimate is constructed analogously. In addition, SCF allows respondents to give their own bounds for a range follow-up response, which may result in SCF having more people in this response category.

surveys is that wealth questions are asked near the end of the SIPP interview, while SCF has more wealth questions earlier on in the survey. Because of this, some respondents may have either dropped out of the SIPP interview or had become fatigued from the interview length by the time they reach the section, which would decrease response rates in SIPP.

5. Conclusion

In this paper, I present new methods for aggregating data on item response across questions in order to generate statistics of item response for key measures, such as household income and wealth. Similar to American Association for Public Opinion Research (1998) for unit response rates, these new methods provide a standardized way of comparing item response rates across surveys. After applying these methods to wealth data, I find item response rates went up in SIPP after the 2014 redesign, but the rates are still lower than in the SCF. This comparison in item response rates amongst surveys is greatly facilitated by my methods for aggregating item response rates. The SCF, for example, has over 140 assets and liabilities questions, so analyzing item nonresponse rates for each of these questions would be tedious and burdensome for researchers and data users. In addition, many of these questions do not have a direct correspondence in SIPP, so direct comparison by question is infeasible. These aggregation methods potentially alleviate these difficulties by combining all asset and debt questions into categories which are comparable across surveys. These methods can be applied to other key measures, such as the unemployment rate, allowing for a new way of evaluating data quality in key measures.

References

The American Association for Public Opinion Research. 1998. "Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. 1st edition." AAPOR.

Bollinger, C.R., Hirsch, B.T., Hokayem, C.M. and Ziliak, J.P., 2015. "Trouble in the Tails? What We Know about Earnings Nonresponse Thirty Years after Lillard, Smith, and Welch."

Bricker, Jesse, Lisa J. Dettling, Alice Henriques, Joanne W. Hsu, Kevin B. Moore, John Sabelhaus, Jeffrey Thompson, and Richard A. Windle. 2014. "Changes in U.S. Family Finances from 2010 to 2013: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, vol. 100, no. 4.

Eggleston, Jonathan, and Michael Gideon. 2017. "Evaluating Wealth Data Quality in the Redesigned 2014 Panel of the Survey of Income and Program Participation." U.S. Census Bureau.

Ferber, Robert. 1966. "Item nonresponse in a consumer survey." *Public Opinion Quarterly* 30.3: 399-415.

Gottschalck, Alfred O. and Jeffrey C. Moore. 2007. "Evaluation of Questionnaire Design Changes on Life Insurance Policy Data". *Research Report Series, Survey Methodology #2007–14*, U.S. Census Bureau.

Groves, Robert M., and Emilia Peytcheva. 2008. "The impact of nonresponse rates on nonresponse bias a meta-analysis." *Public Opinion Quarterly* 72.2: 167-189.U.S. Census Bureau. 2016. *Survey of Income and Program Participation 2014 Panel Users' Guide*. Washington, D.C..

Hokayem, Charles, Trivellore Raghunathan, and Jonathan Rothbaum. 2017. "Match Bias or Nonignorable Nonresponse? Improved Imputation and Administrative Data in the CPS ASEC." U.S. Census Bureau.

U.S. Census Bureau. 2017. *Wealth, Asset Ownership, & Debt of Households Detailed Tables: 2013*. Washington, D.C.

Yan, Ting, Richard Curtin, and Matthew Jans. 2010. "Trends in Income Nonresponse Over Two Decades." *Journal of Official Statistics* 26.1: 145-164.

Appendix A

In this Appendix, I describe the asset and debt groupings presented in the tables.

1. **Bank Accounts**: Consists of money in checking, savings, money market accounts, and Certificates of Deposit (CDs).

2. **Bonds**: Consists of U.S. Treasury securities, municipal bonds, and corporate bonds held outside of retirement accounts, as well as U.S. savings bonds. For SCF, this category also includes foreign and mortgage-backed bonds.
3. **Stocks**: Consists of shares of stocks and mutual funds held outside of retirement accounts.
4. **Financial Assets**: Consists of all assets in the bank account, bonds, and stock categories.
5. **Business**: Consists of the value and debt of businesses. SIPP asks respondents the percent of the business that they own. I use this variable to construct the business value for the household, but I do not incorporate the item nonresponse status of the percent owned variable when calculating the response rate for businesses.
6. **Other Assets**: For both 2014 SIPP Panel and SCF, this consists of the cash value of life insurance policies, annuities, trusts, and the value of all other assets captured in a catchall question. For 2008 SIPP, the measure consists of only values from a catchall question, as the annuity and trust questions were added in the 2014 Panel. The 2008 measures also excludes the cash value of life insurance policies, as this variable was excluded from net worth calculations because many respondents conflated cash value and face value of life insurance (Gottschalck and Moore 2007). The 2013 SCF measure also includes money owed to the respondent by friends, family, or businesses.
7. **Retirement Assets**: Consists on money in Individual Retirement Accounts (IRAs), Keogh accounts, and 401k/thrift accounts.
8. **Real Estate**: Consists of the value of primary residences; rental property; and other real estate, such as timeshares and vacation properties. The SCF has a question about the percent of the other real estate the respondent owns. I use this variable to construct the real estate value for the household, but I do not incorporate the item nonresponse status of the percent owned variable when calculating the response rates for real estate.

9. **Vehicles**: Consists of cars, trucks, SUVs, and recreational vehicles such as motorcycles, boats, and RVs.
10. **Unsecured Debt**: For 2014 SIPP, this consists of credit cards, student loans, and other debt. For 2008 SIPP, there was no separate question on student loans, but student loans should be included with "other debt." The SCF measure includes everything collected in the 2014 SIPP as well as data on other consumer loans and lines of credit.