Trends and Correlates of Income Nonresponse: Forty Years of the General Social Survey

Jibum Kim*, Sungkyunkwan University, Korea
Jaesok Son, NORC the University of Chicago, USA
Peter K. Kwok, NORC at the University of Chicago, USA
Jeong-han Kang, Yonsei University, USA
Faith Laken, NORC at the University of Chicago, USA
Jodie Daquilanea, NORC at the University of Chicago, USA
Hee-Choon Shin, NCHS, USA
Tom W. Smith, NORC at the University of Chicago, USA

December 06, 2012

GSS Methodological Report No. 120

JIBUM KIM is an associate professor in the department of Sociology at Sungkyunkwan University, Seoul, Korea, and a research associate at NORC at the University of Chicago, Chicago, IL. USA. JAESOK SON is a research analyst at NORC at the University of Chicago, Chicago, IL. USA. Peter Kwok is a survey statistician at NORC at the University of Chicago, Chicago, IL. USA. JEONG-HAN KANG is an associate professor in the Department of Sociology at Yonsei University. FAITH LAKEN is a graduate research assistant at the University of Chicago, Chicago, IL. USA. Chicago, IL. Jodie Daquilanea is a survey director at NORC at the University of Chicago. HEE-CHOON SHIN is a senior fellow at NCHS, Hyattsville, MD, USA. TOM W. SMITH is a director of the General Social Survey at NORC at the University of Chicago. A preliminary version of this paper was presented at the 2012 meeting of the American Association of Public Opinion Research. Opinions expressed here are those of the authors and do not necessarily reflect those of the sponsors or related organizations.
*Corresponding author. Address: Department of Sociology, Sungkyunkwan University, 25-2 Sungkyunkwan-Ro, Jongno-Gu, Seoul, Korea
E-mail address: jbk7000@gmail.com

Abstract

Using the 40 years of the General Social Survey (GSS), we investigate the long-term trend and the correlates of family and personal income nonresponse. Family and personal income nonresponse has increased slightly by about 5 percentage points from 1974 to 2010 (9% to 13% in family income; 7% to 12% in personal income). While family income nonresponse was equivalently attributed to "Don't Know" and "Refused," personal income nonresponse was mainly attributed to "Refused." We found very similar correlates of family and personal income nonresponse, such as being older, female, married, self-employed, those not answering the number of earners, uncooperative respondents, people living in the East, and those surveyed in recent periods. In addition, based on the interviewer's evaluation, uncooperative respondents are less likely to response "Don't Know" than "Refused" and respondents with poor comprehension are more likely to respond "Don't Know" than "Refused." Our findings suggest that we need to distinguish "Refused" from "Don't Know" if we aim to better understand income nonresponse and to consider paradata to evaluate the cognitive processing of income nonresponse.

Key words: income nonresponse, sensitive question, paradata

1. Introduction

Income nonresponse holds a central place in item nonresponse. Income correlates with many variables social scientists are interested in (Bell, 1984; Micklewright and Schnepf, 2010) and is important for business use, such as segmentation (Peterson and Kerin, 1980) and policy issues (Moore, Stinson and Welniak, Jr., 2000). Income, however, has a higher nonresponse rate than other questions, even more than questions about sexual behaviors (Laumann et al. 1994, p. 60; Tourangeau, Rips and Rasinski, 2000, pp. 263-4). For example, among eighteen standard demographics items in the General Social Survey (GSS), household income shows the highest income nonresponse for 1972-1990 (Smith, 1991). Income is a vital aspect of socioeconomic conditions of individual development, so social scientists, survey specialists, and statisticians have been concerned about, and interested in, designing better income questions (Duncan and Petersen, 2001; Galobardes and Demarest, 2003; Herriot 1977; Hippler and Hippler, 1986; Juster and Smith, 1997; Locander and Burton, 1976), validating income reporting (Kormendi, 1988; Micklewright and Schnepf, 2010; Peterson and Kerin, 1980), and imputing missing income (Allison, 2010).

Multiple factors have been identified for income nonresponse. According to Moore, Stinson, and Welniak (2000), these factors can be placed under the umbrella of cognitive processing of answering income question: respondents need to (1) understand questions, including specific technical terms for various remunerations; then, (2) retrieve proper income information, which may be hindered due to lack of knowledge or recall difficulties; and (3) individuals can consciously omit or misreport income due to mistrust or privacy and confidentiality concerns. This cognitive approach is consistent with findings of previous studies pertaining to income nonresponse or misreporting, as they were for item nonresponse in general

(Beatty and Hermann, 2002). For example, since some members of the household are not knowledgeable about their total household income, family income nonresponse is higher than personal income nonresponse (Sudman and Bradburn, 1982, pp. 194-202; Korrmendi, 1988), and the household income question elicits a response of "Don't Know" or "Refusal" more than the personal income question does (Kormendi, 1988), and households with more adults (Ross and Reynolds, 1996; Smith, 1991) or those who live in complex household structures (Garner and Blanciforti 1994) are less likely to report income. Also, those who mistrust others are more likely not to report household income (Ross and Reynolds, 1996). Compared to those people whose income sources are wages and salaries, people whose income sources are businesses or partnerships are more likely to be silent about income, and this finding could be attributed to their concern for taxation (Turrell, 2000). A related finding is that self-employed respondents are less likely to report income than salaried professionals (Garner and Blanciforti 1994). Separately, or in combination, these factors contribute to understanding income nonresponse.

Previous studies of income nonresponse have mainly focused on socio-demographic factors, and their effects on income nonresponse were mixed except age. Old people are less likely to report income (Bell, 1984; Kim et al., 2007; Ross and Reynolds, 1996; Smith, 1991; Turrell, 2000). In terms of race, Bell (1984) showed that Whites more than non-Whites are also less likely to divulge income information (Bell, 1984). In contrast, Garner and Blanciforti (1994) found that Blacks more than non-Blacks are less likely to report income (Garner and Blanciforti, 1994). While some studies showed no gender differences on income nonresponse (Garner and Blanciforti, 1994; Turrell, 2000), others showed that females are more likely to be income nonrespondents (Smith, 1991). Likewise, the findings of an association between socio-economic status and income nonresponse are mixed. Lower-incomers (Kormendi, 1988) and the less-

educated (Kim et al., 2007) are more likely to be income nonrespondents, but college graduates are less likely to report income than those who are high school graduates (Garner and Blanciforti, 1994). In contrast, higher income groups, those who do not have economic strain (Ross and Reynolds, 1996) and those who reported higher expenditures of consumer unit (Garner and Blanciforti, 1994), are more likely not to report household income. Similarly, those who are higher-level employees (Hippler and Hippler, 1986) and from higher occupational groups (Turrell, 2000) are more likely not to report personal income. In addition, those who live in the Northeast or Midwest are less likely than those who live in the South to report income (Garner and Blanciforti, 1994).

Two recent studies on missing income tried to overcome the limitations of previous research on income nonresponse by focusing on a single item income question (Micklewright and Schnepf, 2010) and on the long-term trend of income nonresponse (Yan, Curtin, and Jans, 2010). Micklewright and Schnepf (2010) noted that researchers paid great attention to validate an income question that asked several sources of income, whereas little attention was paid to a single item income question despite being widely used in many well-known cross-national data sets, such as the European Social Survey (ESS) or International Social Survey Programme (ISSP). Using the Office for National Statistics (ONS) Omnibus survey and the British Social Attitudes (BSA) survey, in which income was asked as a single question, they found that household income nonresponse is larger than personal income nonresponse and correlates of household and personal income differ. Women and larger number of adults in the household are positively associated with household income nonresponse, but gender is insignificant and number of adults has small effects in personal income nonresponse.

Yan, Curtin, and Jans (2010, p. 146) noted, "a dynamic, historical view of income nonresponse under the same essential conditions" is lacking for the correlates and explanations of income nonresponse. Using the 1986 to 2005 Survey of Consumers, whose household income question was open-ended followed by a bracketed income question, they found a fluctuating trend of household income nonresponse, decreasing in the 1980s, increasing in the 1990s, and decreasing in the 2000s (nonresponse of initial open-ended and bracketed questions: 23 % vs. 9% in 1986, 20% vs. 5% in 1990, 26% vs. 15% in 2001, and 12% vs. 8% in 2005). Although the authors tried to identify the causes of the up and down of income nonresponse over that period, including study protocols and item wordings, they could not explain what accounts for the down-up-down pattern of income nonresponse over that period. On the other hand, using the March Current Population Survey (CPS) in 1990 and 2000, Atrostic and Kalenkoski (2002) showed the increasing rate of income nonresponse across six sources of income, such as earnings from longest job, interest income, and social security. Given the paucity of prior research on long-term trends of income nonresponse, it definitely poses questions for further research on this overlooked topic in income nonresponse.

In this research, using the 40 years of the General Social Survey (GSS), one of the most important resources in understanding social change in America (e.g., Marsden 2012), we describe the long-term trend of income nonresponse and analyze what individual characteristics distinguish income nonrespondents from income respondents. Our research contributes to the literature on income nonresponse by integrating these two strands of research: long-term trend of nonresponse to a single-item income question (Micklewright and Schnepf, 2010 vs. Yan, Curtin, and Jans 2010). In addition, we tried to overcome the limitation of prior research by analyzing household and personal income from one data source and separating the "Don't Know" from

"Refused" in the analysis because the nonresponse rate of "Refused" and "Don't Know" varies by income items (Kalton, Kasprzyk, and Santos, 1981) and "Refused" and "Don't Know" have relatively different causes, either sensitive question or cognitive efforts, respectively (Shoemaker, Eichholz, and Skewes, 2002). Moreover, to evaluate the cognitive approach of income nonresponse, which seems to be an area of much needed research to understand the causes of the item missing (Groves et al., 2009, pp. 208-9), we incorporated two paradata variables: (1) interviewer's evaluation of interviewee's cooperation and (2) interviewer's evaluation of respondent's understanding of questions.

2. Methods

2.1. Data

Data come from the 1972-2010 cumulative cross-sectional General Social Survey (GSS). The GSS uses a nationally representative sample of English speaking adults aged 18 and over (from 2006 GSS onward, Spanish speaking adults are also included) conducted annually or biennially from 1994. This face-to-face survey was conducted by paper-and-pencil interview until 2000, and from 2002, computer-assisted personal interview was adopted. Due to the GSS sample design, not all respondents were asked to answer all questions. For example, the percentage of "not applicable" for the trust question is 34 percent (18,755 out of 55,087 total cases), while age has no "not applicable." From 1972 to 2010, the response rate decreased from over 75 percent in the 1970s to 70 percent in the 2000s, although it was not a linear decline. More detailed information about the GSS is described in the GSS codebook (Smith et al., 2011).

2.2. Measures

Two income questions— family income and personal income— were asked. The family income question has always been on the questionnaire since the first GSS in 1972, while the question about personal income first appeared in 1974. The main wording of the question about family income has remained the same over years: “In which of these groups did your total family income, from all sources, fall last year—[previous year]—before taxes, that is?” In 1976, one more sentence was added to clarify the meaning of “all sources”: “Total income includes interest or dividends, rent, Social Security, other pensions, alimony or child support, unemployment compensation, public aid (welfare), armed forces or veteran’s allotment.” There has been no more change in wording since then (see Appendix A). Then, respondents choose a response category from the list of income categories presented to them on a show card.

With respect to the personal income question, asked just after family income, the interviewer checks an earlier question about occupation and asks about personal income only if respondents earned any income from the occupation described in the earlier part of the questionnaire. The question wording for personal income, however, has gone through several changes over time, mainly switching back and forth between two versions. This question asked about earning from the job in 1974-86, 1988-89, and 1991-93, but it asked about “total earnings from all sources” in 1987, 1988-89, 1990, and 1991-2010. While the personal income questions were designed to be closed-ended questions, the explicit instruction, “Just tell me the letter [of the income category]”, was given sometimes and omitted at other times on the questionnaire (see Appendix B).

The response categories of income variables have changed over time as well (see Appendix C). In 1973, the GSS questionnaire had twelve valid total family income (INCOME) categories and the top category was “$25,000 or more”. Over time, the rise of income and

inflation made this top category unrealistic, so a new income variable was added (INCOME77) with a top category "$50,000 or more", then to "$60,000 or more" in 1986 (INCOME86), and so on. The most recently added family income variable created in 2006 (INCOME06) uses "$150,000 or over" as its top category. While new income variables have been added, the GSS data set has also included INCOME with the original categories in all years for the convenience of comparison over time. The same changes apply to respondents' personal income (RINCOME).

Two paradata-derived variables are the interviewers' evaluations of respondents' attitudes toward the interview and of respondents' understanding of questions. The interviewer was asked to answer the question "In general, what was the respondent's attitude toward the interview?" Response categories include "Friendly and interested," "Cooperative but not particularly interested," "Impatient and restless," and "Hostile." The highest scores indicate a hostile attitude toward the survey. Also, the interviewers evaluated respondents' understanding of the questions by answering "good," "fair," or "poor." For both respondents' cooperation and comprehension evaluated by interviewers, highest scores indicate respondents' poor understanding of questions.

The trust question asked "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" Its response categories include "Most people can be trusted," "Can't be too careful," and "Other, Depends." The latter two categories are made into dummies against the first one.

Other variables: Age and education are continuous variables, and education indicates highest year of school completed. We created several dummy variables for gender (female=1), race (dummy variables for black, other race, and Hispanic, with white as a reference group),

marital status (dummy variables for widowed, divorced, separated, and single, with married as a reference group), work status (dummy variables for part-time job, temporary job, unemployed, and other job (retired, school, keeping house, etc.), with full-time job as a reference group), self-employed (yes=1), subjective class (dummy variables for low, middle, and upper class, with working class as a reference group), number of earners in family (dummy variables for zero, two, three, four and more, and "Don't Know," with one person as a reference group). For geographic areas, we include region (dummy variables for Midwest, South, and West, with Northeast as a reference group) and residential geographic information based on the GSS variable, SRCBELT (dummy variables for 13 to 100 ranked among Standard Metropolitan Statistical Area (SMSA), 1 to 12 ranked SMSA suburb, 13 to 100 ranked SMSA suburb, other urban, and other rural, with 1 to 12 ranked SMSA as a reference group). We also include period (dummy variables for 1980s, 1990s, and 2000s, with 1970s as a reference group).

2.3. Analysis

Due to availability of questions as mentioned in the data section above (2.1), the sample size has decreased. First, we show the long-term trend of family and personal income response. Then, based on previous research and the Wald test which rejected the hypothesis that "refused and no answer" and "don't know" can be combined (Long and Freese 2006, pp. 239-240), we used multinomial logistic regression to contrast income "refused and no answer" and income "Don't Know" with the baseline category of income response for family and personal income. We applied a weight variable for multinomial logistic regression and standard errors were adjusted for clustering at the primary sampling unit.

3. Results

<Figure 1 about here>

Figure 1 displays the income response for about forty years. The total percentage of income response for household and personal income is very similar (89% for family income vs. 91% for personal income). Both family and personal income nonresponses have slightly increased from the early 1970s to 2010, from 7% to 13% for family income and 7% to 12% for personal income. Of the 26 occurrences of GSS data collection, in all but four years, family income nonresponse is higher than personal income nonresponse.

With respect to the income nonresponse trend, if we examine “Refused” and “Don’t Know” responses, in the 1970s and 1980s, the response rate of “Don’t Know” is slightly higher than “Refused” for family income, but since the 1990s there have always been more “Refused” than “Don’t Know” responses. For personal income, on the other hand, there have always been more refusals than "Don't Know.” Thus, on average, there is little difference between “Refused” and “Don’t Know” for family income (6% vs. 5%). The difference between them for personal income, however, is large (8% “Refused” vs. 1% “Don’t Know”).

<Table 1 about here>

<Table 2 about here>

Table 1 and Table 2 are results from multinomial logistic regression and list unstandardized coefficients of three comparisons: Models 1(2)-1 and 1(2)-2 compare people who answered “Refused” and “Don’t Know” relative to those who are respondents, and Model 1(2)-3 compares people who answered “Don’t Know” relative to those who answered “Refused”.

Overall, the effects of correlates for income nonresponse in family (Table 1) and personal income (Table 2) are very similar. The common statistically significant correlates of "Refused" and "Don't Know" for both family and personal income nonresponse are gender, self-employment, those who answered "Don't Know" for the question about the number of earners, attitude toward survey, other urban area, West, and period of 2000s.

Concerning the three variables related to cognitive factors, including survey cooperation, understanding of questions, and general trust, uncooperativeness is associated with an increase in "Refused" and "Don't Know" relative to those who answer income questions. Regarding comprehension, poor understanding is associated with an increase in "Don't Know." In terms of general trust, it is interesting that, compared to those who trust others, those who said "Other, Depends" are less likely to report income.

Of socio-demographic variables, age is positively related to income responses of "Refused" for both family and personal income. Age is not, however, statistically significant for responses of "Don't Know" in family income questions but is significant in terms of personal income. Females are less likely than males to report income. Race is not statistically significant. Married people are more likely than non-married people to be in the "Refused" category relative to "Responded", but are less likely than non-married to be in "Don't Know" relative to "Responded". Years of education is associated with an increase in "Refused" although it is not statistically significant, and is associated with a reduction in "Don't Know" relative to "Responded". Also, job status is not significantly related to "Refused" but, compared to "Full-time job", those who occupy another job status are more likely to answer "Don't Know". Concerning subjective class, its effect is not consistent. For example, middle class are more likely to answer "Refused" or "Don't Know" relative to working class, but "Don't Know" is not

significant. For number of earners in the family, the number of earners is associated with an increase in "Don't Know" rather than "Refused" in family income. Also, those who answered zero or "Don't Know" for number of earners are more likely to answer "Refused" or "Don't Know." For geographical variables, people who are not living in ranked 1-12 SMSA are less likely than people who are living in ranked 1-12 SMSA to refuse a response for the income question; and compared to those in the Northeast, people in other regions are less likely to report income. Compared to the 1970s, people are more likely to refuse and answer "Don't Know" for the income question in the recent period. But, the increasing pattern over time is more salient for refusal than for "Don't Know."

The Model 1(2)-3 column (Don't know/Refused) in Table 1 and 2 indicates why it is worthwhile to distinguish "Refused" from "Don't Know". Again, the common significant correlates of "Don't Know" relative to "Refused" are age, marital status, part-time, zero, two, and three earners, survey cooperation, comprehension, and the periods of the 1990s and 2000s. Regarding cognitive factors, negative attitude toward survey participation is associated with a decrease in "Don't Know" relative to "Refused", but poor understanding is associated with an increase in "Don't Know" relative to "Refused". Although not statistically significant, it is interesting that those who do not trust others are less likely than those who trust to be "Don't Know" relative to "Refused" in family income nonresponse; but those who do not trust are more likely than those who trust to be "Don't Know" relative to "Refused" in personal income response.

Regarding socio-demographic variables, age is associated with a reduction in "Don't Know" relative to "Refused" in income response. Married rather than non-married are less likely to answer "Don't Know" relative to "Refused". Those who are unemployed are less likely than

those who are full-time to answer “Don’t Know” relative to “Refused”. Also, those whose number of earners is zero are less likely than those whose number of earners is one to answer “Don’t Know” relative to “Refused”, but those whose number of earners is two or three are more likely than those whose number of earners is one to answer “Don’t Know” relative to “Refused”. For periods, people who were surveyed in the 1990s or 2000s are less likely than those who were surveyed in the 1970s to report “Don’t Know” relative to “Refused”. Differential effects of those who are in other job category on family and personal income are worth noting. People who are in other job categories than full-time are more likely than those who report full-time to answer “Don’t Know” relative to “Refused” in family income, but people who are in other jobs are less likely than those who are in full-time jobs to answer “Don’t Know” relative to “Refused” in family income.

4. Conclusion

Despite the fact that income nonresponse spawns a considerable research due to the high rate of income nonresponse and its significant effect on the analysis of data, studies of the trend and correlates of income nonresponse whose income question is based on a single item were rare. To fill this gap, this study examines trends and correlates of missing income response for one single-item income question. Our investigation has four distinctive features: (1) we examined the trend of income for four decades, (2) we used a single-item income question, (3) we incorporated both family and personal income nonresponse for our analysis, and (4) we differentiated “Don’t Know” from “Refused” for the analysis of income nonresponse. Family and personal income nonresponse has increased slightly by about 5 percentage points from 1974 to 2010 (9% to 13%

in family income; 7% to 12% in personal income). The percentage of income response is similar between family and personal income, although, in general, personal income response is higher than family income response. We also found very similar correlates of family and personal income nonresponse, and showed that those who refuse in the income question were different from those who answered "Don't Know". As expected, the percentage difference between "Refused" and "Don't Know" is much larger for personal income than family income. Thus, it is meaningful to distinguish "Refused" from "Don't Know" if we aim to better understand those who did not report income.

The trend of income nonresponse in our study suggests another trend for income nonresponse. If we compared our finding with previous results, in general, our findings are not consistent with Yan, Curtin, and Jans's study of a decreasing trend of family income nonresponse. They showed a fluctuating trend of initial open-ended family income nonresponse and the final family income, which reflects those who do not provide an open-ended response but answer the following income bracket question. Unlike Yan, Curtin, and Jans's study, the GSS family income nonresponse has slightly increased, 10% in 1972, 8% in 1980, 12% in 1990, and 13% in 2010. Also, Atrostic and Kalenkoski's March CPS study (2002) showed rapid increasing nonresponse from 1990 to 2000, which varies from 6 to 21% depending on the source of income while the GSS personal income nonresponse shows a slow upward trend of 8% in 1990, 10% in 2000, and 12% in 2010. Thus, compared to previous studies, our study showed a slowly increasing trend of income nonresponse for both family and personal income, and we can identify that increasing income nonresponse is driven more by "Refused" than "Don't know."

Due to the fact that limited attention has been paid to a single item income question (Micklewright and Schnepf, 2010), few studies are comparable to our findings. The exception is

Micklewright and Schnepf's study based on the BSA and the ONS (2010). Both studies show similar family income nonresponse rates and that the old and females are less likely to report family income. But, while they found that the self-employed are less likely to report only personal income in the ONS, our studies consistently showed that the self-employed are less likely to report both family and personal income. The inconsistent effects of self-employed on family and personal income nonresponse by Micklewright and Schnepf might be attributed to the use of two different datasets.

In terms of studying the correlates of family and personal income nonresponse, unlike Micklewright and Schnepf's study of a single item income question based on two different datasets (2010), we used one data set, the GSS. Our findings suggested overall similarity of correlates for both family and personal income nonresponse. Old people, female, married, self-employed, those who do not answer the number of earners, uncooperative respondents, people living in the East, and those surveyed in recent periods are more likely to be missing income respondents. However, we did not find racial and socio-economic status differences on missing income respondents, which contradict some previous studies (Kim et al., 2007). The greater the overall similarity of correlates of family and personal income is, the more confidence we may have on our study's important implication regarding the relationship between unit nonresponse and item nonresponse. Our finding that hostile respondents are more likely to be missing income respondents is especially consistent with the finding that inclusion of uncooperative respondents may decrease the quality of data by increasing the item nonresponse (Yan, Curtin, and Jans, 2010).

Our analysis confirms the need to separate "Refused" and "Don't Know" (Kalton, Kasprzyk, and Santos, 1981; Shoemaker, Eichholz, and Skewes, 2002). We have shown that

while in family income nonresponse was equivalently attributed to "Don't Know" and "Refused," in personal income, nonresponse was mainly attributed to "Refused." It is consistent with the previous finding that questions that require cognitive efforts are more likely to generate "Don't Know" than "Refused" (Shoemaker, Eichholz, and Skewes, 2002). Naturally, we found that respondents with poor comprehension are more likely to respond "Don't Know" rather than "Refused" in both family and personal income question.

There are several limitations. First, since our study was guided by the cognitive theoretical framework of income nonresponse, we have included both the interviewer's evaluation of the respondent's attitude toward the survey and of their level of understanding. Although respondents' attitudes toward the survey seem to indicate willingness to answer questions in a broad sense, we cannot exclude the possibility that respondents' nonresponse on the income question may affect the interviewers' overall evaluation. Second, with regard to different trends or correlates of income nonresponse, we cannot explain why different patterns exist because of the many different survey dimensions, such as mode of data collection, survey organization, incentives, and income question wording, across these studies. Finally, in order to use consistent correlates for forty years, we did not use variables that were available only for recent years, such as such as interviewer characteristics or incentives.

Overall, this study shows slowly growing nonresponse on the single-item income question and the need to differentiate between "Refused" and "Don't Know" to better comprehend income nonresponse. Future research can be benefited if the data encompass items relevant to the cognitive approach of income nonresponse to better understand income nonresponse. As Tourangeau, Rips and Rasinski (2000: 259) noted, accessing how income nonresponse is directly related to indicators or measures of privacy and confidentiality may sort

out the relative importance of the components of questions that are regarded as sensitive topics. Furthermore, given the increasing use of paradata for survey analysis, a thorough understanding of what determines interviewers' evaluations of respondents' attitudes toward the survey is very useful.

Acknowledgements

We would like to thank Lauren Doerr and Steven Pedlow for helpful comments, and Ting Yan for providing us the figures of income nonresponse of her paper.

References

Allison, P.D., 2010. Missing data. In: Marsden, P.V., Wright, J.D. Handbook of Survey Research. (2nd Ed.), Emerald Group Publishing Limited, Bingley, WA, UK, pp. 631-57.

Atrostic, B. K., Kalenkoski, C., 2002. Item response rates: One indicator of how well we measure income. In: American Statistical Association. New York: NY. < http://www.amstat.org/sections/srms/Proceedings/y2002/Files/JSM2002-000785.pdf> (accessed Nov. 12, 2012).

Beatty, P., Herrmann, D., 2002. To answer or not to answer: Decision processes related to survey item nonresponse. In: Groves, R.M., Dillman, D.A., Eltinge, J.L., Survey Nonresponse. John Wiley & Sons., NY, pp. 71-85.

Bell, R. 1984. Item nonresponse in telephone surveys - An analysis of who fails to report income. Social Science Quarterly 65, 207-215.

Duncan, G.J., Petersen, E., 2001. The long and short of asking questions about income, wealth, and labor supply. Social Science Research 30, 248-263.

Galobardes, B., Demarest, S., 2003. Asking sensitive information: an example with income. Sozial-Und Praventivmedizin 48, 70-72.

Garner, T.I., Blanciforty, L.A. 1994. Household income reporting: An analysis of U.S. consumer expenditure survey data. Journal of Official Statistics 10, 69-91.

Groves, R.M., Fowler Jr., F.J., Couper, M.P., Lepkowski, J.M., Singer, E. Tourangeau, R. 2009. Survey Methodology. (2nd Ed.) John Wiley & Sons, Hoboken, NJ.

Herriot, R.A. 1977. Collecting income data on sample-surveys - Evidence from split-panel studies. Journal of Marketing Research 14, 322-329.

Hippler, H-J., Hippler, G. 1986. Reducing refusal rates in the case of threatening questions: The 'Door-in-Face' technique. Journal of Official Statistics 2, 25-33.

Juster, F. T., Smith, J.P., 1997. Improving the quality of economic data: Lessons from the HRS and AHEAD. Journal of the American Statistical Association 92, 1268-1278.

Kalton, G., Kasprzyk, D., Santos, R. 1981. Issues of nonresponse and imputation in the survey of income and program participation. In: Krewski, D., Platek, R., Rao, J.N.K., Current Topics in Survey Sampling. Academic Press, NY, pp. 450-480.

Kim, S., Egerter, S., Cubbin, C., Takahashi, E., Braveman, P. 2007. Potential implications of missing income data in population-based surveys: An example from a postpartum survey in California. Public Health Reports 122, 753-763.

Kormendi, E., 1988. The quality of income information in telephone and face to face surveys. In: Groves, R.M., Biemer, P.P., Lyberg, L.E, Massey, J.T., Nicholos II, W.L., Waksberg, J., Telephone Survey Methodology. John Wiley & Sons, New York, pp. 341-356

Laumann, E.O., Gagnon, J.H., Michael, R.T., Michaels, S., 1994. The Social Organization of Sexuality: Sexual Practices in the United States. University of Chicago, Chicago.

Locander, W.B, Burton, J.P., 1976. Effect of question form on gathering income data by telephone. Journal of Marketing Research 13, 189-192.

Long, S.J., Freese, J., 2006. Regression Models for Categorical Dependent Variables Using Stata. Stata Press, College Station

Marsden, P.V., 2012. Social Trends in American Life: Findings from the General Social Survey since 1972. Ed. Princeton University, Princeton.

Micklewright, J., Schnepf, S.V., 2010. How reliable are income data collected with a single question? Journal of the Royal Statistical Society Series a-Statistics in Society 173, 409-429.

Moore, J.C., Stinson, L.L., Welniak Jr., E.J., 2000. Income measurement error in surveys: A review. Journal of Official Statistics 16, 331-361.

Peterson, R.A., Kerin, R.A., 1980. Household income data reports in mail surveys. Journal of Business Research 8, 301-313.

Ross, C.E., Reynolds, J.R., 1996. The effects of power, knowledge, and trust on income disclosure in surveys. Social Science Quarterly 77, 899-911.

Shoemaker, P.J., Eichholz, M., Skewes, E.A., 2002. Item nonresponse: Distinguishing between don't know and refuse. International Journal of Public Opinion Research 14, 193-201.

Smith, T. W. 1991. An analysis of missing income information on the General Social Surveys. NORC, Chicago.

Smith, T. W., Marsden, P., Hout, M., Kim, J., 2011. General Social Surveys, 1972-2010[machine-readable data file], National Opinion Research Center [producer], Chicago; The Roper Center for Public Opinion Research, University of Connecticut [distributor], Storrs, CT.

Sudman, S., Bradburn, N.M., 1982. Asking Questions. Jossey-Bass, San Francisco.

Tourangeau, R., Rips, L.J., Rasinski, K., 2000. The Psychology of Survey Response. University of Cambridge, Cambridge, UK.

Turrell, G. 2000. Income non-reporting: implications for health inequalities research. Journal of Epidemiology and Community Health 54, 207-214.

Yan, T., Curtin, R., Jans M., 2010. Trends in income nonresponse over two decades. Journal of Official Statistics 26, 145-164.

Figure 1. Family Income (left) vs. Personal Income (right) Response/Nonresponse Percent (GSS 1972-2010)

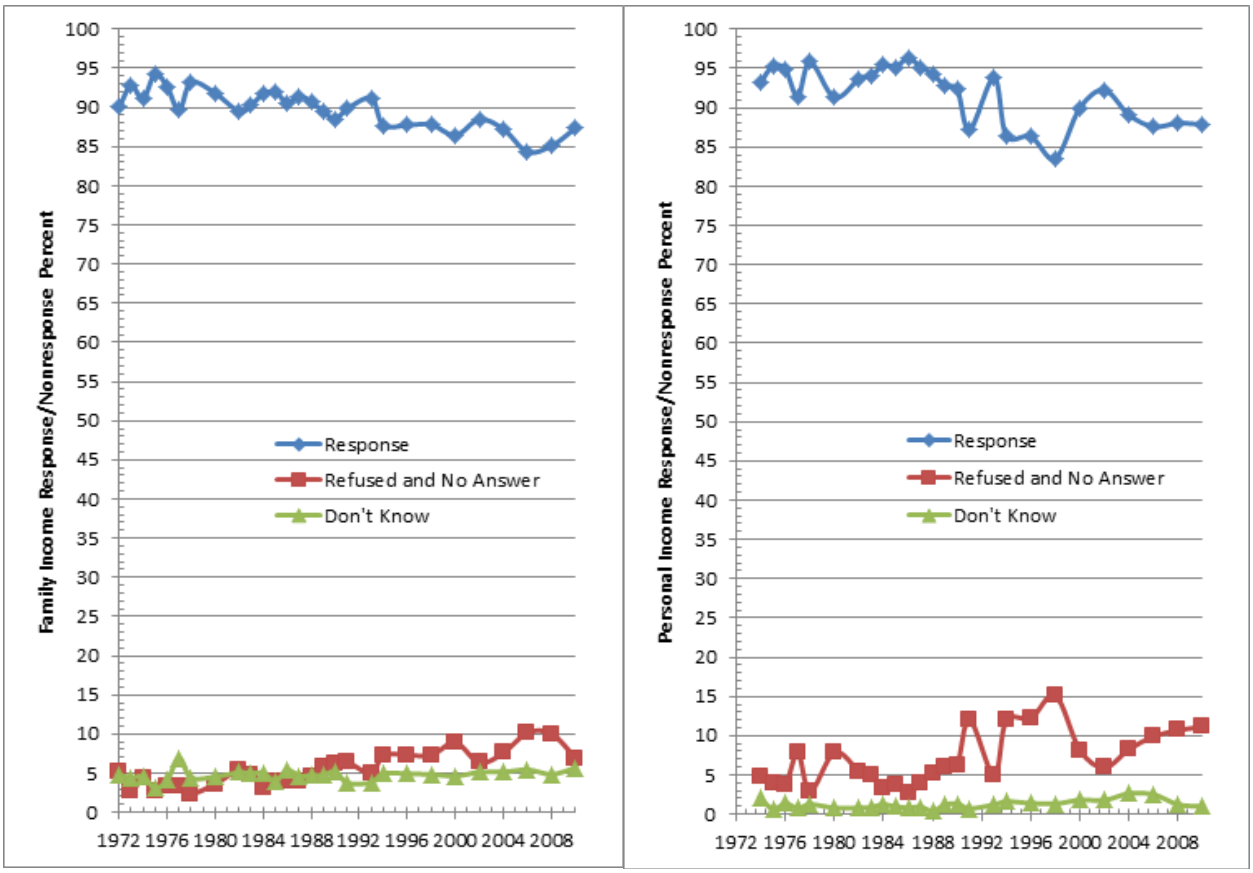

Table 1. Unstandardized Multinomial Logistic Coefficients from the Regression of Family Income Nonresponse (Unweighted N=29,637)

	Model 1-1 Refused/Responded		Model 1-2 Don't Know/Responded		Model 1-3 Don't know/Refused	
Age	0.021 (0.002)	***	0.003 (0.002)		-0.017 (0.003)	***
Female	0.240 (0.066)	***	0.469 (0.077)	***	0.229 (0.096)	*
Black	-0.098 (0.105)		-0.083 (0.119)		0.015 (0.151)	
Hispanic	-0.172 (0.198)		-0.213 (0.218)		-0.041 (0.271)	
Other race	0.048 (0.201)		0.275 (0.229)		0.227 (0.277)	
Married	0.291 (0.070)	***	-0.915 (0.091)	***	-1.206 (0.112)	***
Yrs. Of educ	0.014 (0.012)		-0.089 (0.014)	***	-0.104 (0.017)	***
Part-time	-0.040 (0.106)		0.655 (0.119)	***	0.695 (0.148)	***
Temp. no job	0.038 (0.221)		0.705 (0.250)	**	0.667 (0.344)	
Unemployed	-0.358 (0.200)		0.891 (0.179)	***	1.249 (0.268)	***
Other job	-0.118 (0.088)		1.176 (0.097)	***	1.295 (0.125)	***
Self-employed	0.318 (0.087)	***	0.286 (0.109)	**	-0.033 (0.135)	
Lower class	-0.212 (0.175)		0.186 (0.139)		0.398 (0.215)	
Middle class	0.186 (0.072)	**	0.192 (0.100)		0.006 (0.126)	
Upper class	0.244 (0.142)		0.113 (0.261)		-0.131 (0.288)	
# of earners: 0	0.460 (0.095)	***	-0.147 (0.133)		-0.607 (0.157)	***
# of earners: 2	-0.090 (0.084)		0.582 (0.105)	***	0.672 (0.132)	***
# of earners: 3-4	0.009 (0.107)		1.322 (0.120)	***	1.313 (0.165)	***
# of earners: 5 & +	0.151 (0.360)		1.369 (0.239)	***	1.218 (0.428)	**
# of earners: DK	2.383 (0.215)	***	1.368 (0.335)	***	-1.014 (0.359)	**
Trust: cannot	0.171 (0.069)	*	0.032 (0.094)		-0.140 (0.111)	
Trust: depends	0.580 (0.112)	***	0.317 (0.174)		-0.263 (0.200)	
Cooperation (1 - 4: hostile)	0.904 (0.050)	***	0.345 (0.062)	***	-0.559 (0.076)	***
Comprehension (1 - 4: poor)	-0.096 (0.074)		0.499 (0.075)	***	0.594 (0.105)	***
13-100 SMSA	-0.474 (0.154)	**	-0.299 (0.178)		0.175 (0.198)	
1-12 SMSA suburb	-0.496 (0.164)	**	-0.179 (0.182)		0.317 (0.220)	
13-100 SMSA suburb	-0.250 (0.149)		-0.406 (0.199)	*	-0.155 (0.215)	
Other urban	-0.536 (0.133)	***	-0.443 (0.161)	**	0.093 (0.174)	
Other rural	-0.414 (0.179)	*	-0.321 (0.218)		0.093 (0.228)	
Midwest	-0.413 (0.109)	***	-0.224 (0.126)		0.190 (0.143)	
South	-0.233 (0.108)	*	-0.042 (0.117)		0.191 (0.142)	
West	-0.903 (0.132)	***	-0.410 (0.115)	***	0.494 (0.164)	**
Period: 1980s	0.361 (0.144)	*	0.296 (0.144)	*	-0.065 (0.190)	
Period: 1990s	0.893 (0.135)	***	0.393 (0.137)	**	-0.500 (0.188)	**
Period: 2000s	1.236 (0.144)	***	0.562 (0.138)	***	-0.675 (0.188)	***

_cons	-5.755	(0.294)	***	-4.163	(0.351)	***	1.593 (0.427) ***

Note: Income respondents serve as baseline category. Sample weights and cluster membership were considered for estimates. Number in parentheses are standard errors. Omitted category for race is white; unmarried for marital status; full-time job for work status; not self-employed for self-employment; working class for class; earner=1 for earners; "yes" for trust; 1-12 SMSA for place; Northeast for region, 1970s for periods. *<.05; **<.01; ***<.001

Pseudo-R^2 =.123 for family income; Pseudo-R2 =.194 for personal income

Table 2. Unstandardized Multinomial Logistic Coefficients from the regression of Personal Income Nonresponse (Unweighted N=21,123)

	Model 2-1 Refused/Responded		Model 2-2 Don't Know/Responded		Model 2-3 Don't know/Refused	
Age	0.028 (0.003)	***	0.014 (0.006)	*	-0.014 (0.006)	*
Female	0.246 (0.076)	**	0.385 (0.152)	*	0.139 (0.162)	
Black	0.070 (0.105)		-0.342 (0.229)		-0.413 (0.227)	
Hispanic	-0.151 (0.175)		0.004 (0.273)		0.156 (0.293)	
Other race	0.079 (0.189)		0.504 (0.377)		0.425 (0.403)	
Married	0.456 (0.081)	***	-0.206 (0.180)		-0.662 (0.192)	**
Yrs. of educ	-0.022 (0.014)		-0.060 (0.029)	*	-0.038 (0.033)	
Part-time	-0.190 (0.099)		0.416 (0.181)	*	0.606 (0.197)	**
Temp. no job	0.046 (0.202)		0.246 (0.439)		0.200 (0.487)	
Unemployed	-0.592 (0.206)	**	0.351 (0.348)		0.943 (0.396)	*
Other job	1.175 (0.094)	***	0.601 (0.230)	**	-0.574 (0.241)	*
Self-employed	0.433 (0.091)	***	0.759 (0.176)	***	0.326 (0.196)	
Lower class	0.253 (0.165)		0.768 (0.272)	**	0.515 (0.303)	
Middle class	0.169 (0.078)	*	0.071 (0.171)		-0.099 (0.192)	
Upper class	0.285 (0.170)		0.316 (0.382)		0.031 (0.411)	
# of earners: 0	2.089 (0.153)	***	0.808 (0.398)	*	-1.280 (0.404)	**
# of earners: 2	-0.290 (0.084)	**	0.303 (0.195)		0.593 (0.214)	**
# of earners: 3-4	-0.282 (0.107)	**	1.023 (0.225)	***	1.304 (0.245)	***
# of earners: 5 & +	-0.044 (0.298)		0.877 (0.497)		0.920 (0.573)	
# of earners: DK	2.053 (0.252)	***	1.604 (0.679)	*	-0.449 (0.716)	
Trust: cannot	0.052 (0.077)		0.128 (0.179)		0.076 (0.192)	
Trust: depends	0.529 (0.136)	***	0.733 (0.330)	*	0.204 (0.339)	
Cooperation (1 - 4: hostile)	0.785 (0.071)	***	0.408 (0.122)	**	-0.378 (0.138)	**
Comprehension (1 - 4: poor)	-0.065 (0.088)		0.599 (0.152)	***	0.664 (0.167)	***
13-100 SMSA	-0.294 (0.175)		-0.499 (0.319)		-0.205 (0.346)	
1-12 SMSA suburb	-0.333 (0.162)	*	-0.518 (0.362)		-0.185 (0.412)	
13-100 SMSA suburb	-0.271 (0.148)		-0.950 (0.396)	*	-0.679 (0.399)	
Other urban	-0.448 (0.139)	**	-0.696 (0.279)	*	-0.248 (0.294)	
Other rural	-0.375 (0.178)	*	-0.666 (0.325)	*	-0.292 (0.342)	
Midwest	-0.323 (0.116)	**	-0.316 (0.206)		0.007 (0.220)	
South	-0.188 (0.104)		0.167 (0.200)		0.356 (0.213)	
West	-0.677 (0.121)	***	-0.525 (0.255)	*	0.152 (0.280)	
Period: 1980s	0.444 (0.183)	*	-0.144 (0.274)		-0.589 (0.320)	
Period: 1990s	1.310 (0.166)	***	0.306 (0.248)		-1.003 (0.283)	***
Period: 2000s	1.256 (0.174)	***	0.607 (0.256)	*	-0.649 (0.294)	*

_cons	-5.609	(0.328)	***	-5.983	(0.732)	***	-0.374 (0.758)

Note: Income respondents serve as baseline category. Sample weights and cluster membership were considered for estimates. Number in parentheses are standard errors. Omitted category for race is white; unmarried for marital status; full-time job for work status; not self-employed for self-employment; working class for class; earner=1 for earners; "yes" for trust; 1-12 SMSA for place; Northeast for region, 1970s for periods. *<.05; **<.01; ***<.001
Pseudo-R2 =.123 for family income; Pseudo-R2 =.194 for personal income

Appendix A. Question wording: Family Income

GSS year	Family Income question wording	Note
1972	In which of these groups did your total family income, from all sources, fall last year--[previous year]--before taxes, that is?	
1973-75	In which of these groups did your total family income, from all sources, fall last year--[previous year]--before taxes, that is? Just tell me the letter.	Added: "Just tell me…"
1976 - present	In which of these groups did your total family income, from all sources, fall last year--[previous year]--before taxes, that is? Just tell me the letter. Total income includes interest or dividends, rent, Social Security, other pensions, alimony or child support, unemployment compensation, public aid (welfare), armed forces or veterans allotment.	Box added: "Total income includes…"

Appendix B. Question Wording: Personal Income

GSS year	Ballot/Version	Personal Income question wording	Note
1974-86	All	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your earnings from (JOB IN Q_), for last year--[previous year]--fall? That is, before taxes or other deductions. Just tell me the letter.	
1987	All	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_) from all sources, fall last year--[previous year]--before taxes, that is. Just tell me the letter.	Added: "from all sources"
1988-89	Ballot A & B	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your earnings from (JOB IN Q_), for last year--[previous year]--fall? That is, before taxes or other deductions. Just tell me the letter.	Dropped: "from all sources"
1988-89	Ballot C	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_) from all sources, fall last year--[previous year]--before taxes, that is.	Added: "from all sources" Dropped: "Just tell me…"
1990	Ballot A & B	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your earnings from (JOB IN Q_), for last year--[previous year]--fall? That is, before taxes or other deductions. Just tell me the letter.	
1990	Ballot C	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_) from all sources, fall last year--[previous year]--before taxes, that is. Total income includes interest or dividends, rent, Social Security, other pensions, alimony or child support, unemployment compensation, public aid (welfare), armed forces or veterans allotment.	Dropped: "Just tell me…" Added: "Total income…"

1991-93	Ballot A & B	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your earnings from (JOB IN Q_), for last year--[previous year]--fall? That is, before taxes or other deductions. Just tell me the letter.	
1991-93	Ballot C	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_) from all sources, fall last year--[previous year]--before taxes, that is.	Dropped: "Just tell me…" Dropped: "Total income…"
1994	Version 1-5, 7-9	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_), from all sources for [previous year] fall? That is, before taxes or other deductions. Just tell me the letter.	Added: "Just tell me…"
1994	Version 6	IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_), from all sources for [previous year] fall? That is, before taxes or other deductions.	Dropped: "Just tell me…"
1996-2004	Version 1-2, 4-6	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_), from all sources for [previous year] fall? That is, before taxes or other deductions. Just tell me the letter.	Added: "Just tell me…"
1996-2004	Version 3	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_), from all sources for [previous year] fall? That is, before taxes or other deductions.	Dropped: "Just tell me…"
2006-2010	All	Did you earn any income from (JOB DESCRIBED IN Q_) in [previous year]? IF YES: In which of these groups did your total earnings from (OCCUPATION IN Q_), from all sources for [previous year] fall? That is, before taxes or other deductions. Just tell me the letter.	Added: "Just tell me…"

Appendix C. Response Categories of Income question

Variable Name	INCOME	INCOME77	INCOME82	INCOME86	INCOME91	INCOME98	INCOME06
GSS Year	1973-2010	1977-80	1982-85	1986-90	1991-96	1998-2004	2006-10
UNDER $1,000	Y	Y	Y	Y	Y	Y	Y
$1,000 to $2,999	Y	Y	Y	Y	Y	Y	Y
$3,000 to 3,999	Y	Y	Y	Y	Y	Y	Y
$4,000 to 4,999	Y	Y	Y	Y	Y	Y	Y
$5,000 to 5,999	Y	Y	Y	Y	Y	Y	Y
$6,000 to 6,999	Y	Y	Y	Y	Y	Y	Y
$7,000 to 7,999	Y	Y	Y	Y	Y	Y	Y
$8,000 to 9,999	Y	Y	Y	Y	Y	Y	Y
$10,000 to 12,499	$10,000 - 14999	Y	Y	Y	Y	Y	Y
$12,500 to 14,999	$15,000 – 19,999	Y	Y	Y	Y	Y	Y
$15,000 to 17,499	$20,000 – 24,999	Y	Y	Y	Y	Y	Y
$17,500 to 19,999	$25,000+	Y	Y	Y	Y	Y	Y
$20,000 to 22,499		Y	Y	Y	Y	Y	Y
$22,500 to 24,999		Y	Y	Y	Y	Y	Y
$25,000 to 29,999		$25,000-49,999	$25,000-34,999	Y	Y	Y	Y
$30,000 to 34,999		$50,000+	$35,000-49,999	Y	Y	Y	Y
$35,000 to 39,999			$50,000+	Y	Y	Y	Y
$40,000 to 49,999				Y	Y	Y	Y
$50,000 to 59,999				Y	Y	Y	Y
$60,000 to 74,999				$60,000+	Y	Y	Y
$75,000 to $89,999					$75,000+	Y	Y
$90,000 to $109,999						Y	Y
$110,000 to $129,999						$110,000+	Y
$130,000 to $149,999							Y
$150,000 or over							Y
REFUSED	Y	Y	Y	Y	Y	Y	Y
DK	Y	Y	Y	Y	Y	Y	Y
NA	Y	Y	Y	Y	Y	Y	Y
IAP	Y	Y	Y	Y	Y	Y	Y

Notes: "Y" denotes that the category existed on the question in that GSS year.

INCOME is a variable for total family income, which uses the same categories as RINCOME, a variable for personal income.

GSS 1972 used income categories very different from the ones used afterward, and they are omitted in this table:

less than $2,000, $2,000-3,999, $4,000-5,999, $6,000-7,999, $8,000-9,999, $10,000-12,499, $12,500-14,999, $15,000-17,499, $17,500-19,999, $20,000-24,999, $25,000-29,999, and $30,000 or more.