

QUESTIONS PEOPLE DON'T LIKE TO ANSWER: WEALTH AND SOCIAL SECURITY NUMBERS

James M. Dahlhamer, Peter S. Meyer, and John R. Pleis
National Center for Health Statistics, 3311 Toledo Road, Hyattsville, Maryland 20782

Abstract¹

The National Health Interview Survey (NHIS), a computer assisted personal interview (CAPI) survey administered annually by the National Center for Health Statistics, Centers for Disease Control and Prevention, collects data on family income, but virtually no data on wealth. Yet research has identified a strong, separate association between wealth and health status. Will people answer questions on wealth? In addition, social security numbers are needed for linkage to administrative databases, but suffer from high item nonresponse. Can we improve reporting? This paper presents early results from a field test designed to address these questions. Though questions on total financial assets, property value, and amount owed on the property suffered from high item nonresponse, they were positively associated with respondent-reported health status, net of other socioeconomic status measures (SES). Additionally, asking for just the last four digits of a social security number improved reporting by roughly 13 percentage points over the current nine-digit format.

Keywords: Sensitive Questions, Familial Wealth, Social Security Numbers, Item Nonresponse, NHIS

1. Introduction

As described in Meyer, Dahlhamer, and Pleis (2007), the National Health Interview Survey (NHIS)² undertook a field test in the second quarter of 2006 to assess the utility of an alternative set of family income questions (see Pleis, Dahlhamer, and Meyer, 2007), a set of family wealth items, and an alternative social security number (SSN) question. The field test was administered to “screened-out” households—households that are designated for inclusion as part of the oversampling of Hispanic, black, and Asian households, but are found to have no Hispanic, black, or Asian members. Under normal procedures, once this determination is made the interview terminates and the household is coded as “screened-out” (out of scope and not included in response rate calculations).³ For the

¹ The findings and conclusions in this paper are those of the author(s) and do not necessarily represent the views of the Department of Health and Human Services, the Centers for Disease Control and Prevention, and the National Center for Health Statistics.

² Covering the civilian, noninstitutionalized population of the United States, the NHIS is a face-to-face (in-person), multipurpose health survey. For more information, please visit the NHIS Web site at www.cdc.gov/nchs/nhis.htm.

³ Screener households that contain a black, Hispanic, or Asian member are “screened-in”: they are considered in-

field test, however, interviews of these “screened-out” households continued with the family core component of the survey. The test questions appeared in the family income section (ninth and final section) of the family core.⁴ A “family respondent”⁵ answered for him/herself (self-report) and all other family members (proxy reports).

This paper presents results on the wealth and alternative SSN questions included in the field test. Because of the nature of the analyses presented below, the screened-out households were treated as a simple random sample, and all analyses were unweighted and performed in SAS (SAS Institute, 2006). Since data from the field test were still being processed at the time this paper was completed, analyses are based on the first eight weeks of quarter 2.⁶

2. Collecting Data on Familial Wealth

While research has consistently shown that socioeconomic status (SES) is strongly related to health status and health outcomes (Robert and House, 1996), the bulk of analyses have relied on a limited set of SES measures such as education and income. More recently, a body of research has identified associations between wealth and a variety of health indicators, net of other, more traditional SES measures (Kahn and Fazio, 2005; Martikainen et al., 2003; von Rueden et al., 2006). Furthermore, while income and education may be good indicators of SES in middle age, wealth may be a better measure of SES for older adults (Matthews et al., 2005; Robert and House, 1996).

Given the highly sensitive nature of the topic (Moore, Stinson, and Welniak, 2000), however, questions on wealth are often excluded from health surveys, as has been the case with the NHIS.⁷ Can we ask about familial wealth on the NHIS? Will respondents provide the information requested? Does knowledge of familial

scope (included in response rate calculations) and receive the full NHIS interview. Approximately 15% of initial sample households are designated as “screeners.”

⁴ Upon reaching the family income section, respondents were randomly sent down one of two paths. The “experimental” path contained a set of test family income questions. The “control” path contained the usual family income questions. The test wealth and SSN questions appeared on both paths.

⁵ A “family respondent” can be any responsible adult age 18 or older, an adult less than 18 if ever married, or an emancipated minor.

⁶ Once available, all analyses will be replicated with the full quarter 2, 2006 test file.

⁷ The NHIS does include a question on home ownership.

wealth enhance our analyses of health outcomes, net of currently collected SES measures such as education and income? To address these questions, we added four wealth items to the quarter 2, 2006 field test.

2.1 Wealth Questions Tested

The wealth questions included in the field test appeared immediately following items on total family income and housing tenure, and preceding items on program participation. An introduction screen was provided as a transition, and read as follows:

- “The next few questions are about the value of any financial assets that [you own/your family owns]. By assets, we are referring to things like bank accounts, retirement accounts, investments such as stocks and bonds, and other financial assets listed on this card. Financial assets are also important in analyzing the health information we collect. Please be assured that, like all other information you have provided, these answers will be kept strictly confidential.”

The intro screen was accompanied by a card, handed to the family respondent, which listed the various types of financial assets to be considered when answering these wealth questions. The intro was then followed by a question on total financial assets:

- [if own home] “Not including the value of this residence, what is the total value of all financial assets that [you own/your family owns]?”
- [else] “What is the total value of all financial assets that [you own/your family owns]?”

The question was designed to elicit an actual dollar amount. The next two questions were reserved for home owners. Like total financial assets, both were designed to capture an actual dollar amount: the value of the respondent’s property and the amount still owed on the property.

- “Earlier you mentioned that [you own or are buying/your family owns or is buying] this residence. Approximately, how much would this property sell for today?”
- “How much [do you/does your family] still owe on this property? Please include all mortgages or home equity loans [you/your family] might have.”⁸

Finally, all respondents received a question on car ownership:

- “[Do you/Does anyone in your family] own a car?”

With the exception of home ownership, the NHIS does not currently include questions on wealth. The test items, therefore, were asked of all screened-out cases.

2.2 Item Response Rates for the Wealth Questions

Our initial research question was whether or not respondents were willing to provide or capable of providing the requested information. Table 1 presents item response rates for the four wealth questions. Consistent with rates for income questions (Dahlhamer et al., 2004), response rates for the three exact amount questions were fairly low (first column of rates).⁹ Only 51.2% of family respondents provided an amount in response to the question on total financial assets. Of the nonresponders, the bulk provided a refusal response (31.4%), suggesting a highly sensitive item. Response rates were somewhat better, though still poor, for the items on home ownership. Not quite 71% of family respondents provided a dollar amount for the value of their property. Don’t know and refusal rates were fairly similar (13.6% and 15.7% respectively). Just over 65% of family respondents provided a figure for the amount still owed on the property, with refusal (24.0%) responses more than doubling don’t know responses (10.9%). Again, this question appeared to heighten respondent sensitivity. Finally, a yes or no response was secured from just over 93% of family respondents when asked about car ownership.

2.3 Associations between Familial Wealth and Self-Reported Health Status

The second question about the test was whether or not the addition of such items added any value to the NHIS. For example, do these items enhance our analyses of health status, health outcomes, and health care service utilization? To address this, we explored associations with respondent-reported health status, as measured by the following question:

- “Would you say [your/person’s] health in general is excellent, very good, good, fair, or poor?”

The five health status outcomes were assigned values from 1 (poor) to 5 (excellent).

We first looked at the relationship between total financial assets and respondent-reported health status for three different age groups: 18-44, 45-64, and 65 or older. All persons, not just family respondents, age 18 or older were included in the analysis. Thus, the total of the wealth

⁸ The intent of this question was to capture the amount owed on the property if paid off today.

⁹ Response rates are also provided in Table 1 (second and third columns of rates) by the type of income questions received (experimental or control). Similar breakouts are provided in Table 2 and Figures 4 and 5.

question amounts reported by the family respondent was attributed to all family members. Total financial assets was broken into quartiles and, given the high level of nonresponse to this item, a “missing” category. Quartile 1 refers to the lowest amount of financial assets, while quartile 4 represents the highest amount of financial assets.¹⁰ For each quartile (and the missing category) within each age group, the average respondent-reported health status was computed. So, for example, persons age 18 to 44 in quartile 1 of assets had an average respondent-reported health status of 3.78 (between good and excellent).

It should be noted that assigning values 1 to 5 to the health status variable outcomes assumes the outcomes are equally spaced. This is not necessarily true. However, the type of analysis described above has been done elsewhere (for example, see Robert and House, 1996). In addition, graphs of the percentage of adults with very good or excellent health broken out by the same total financial asset (and net property value quartiles—see below) and age measures show virtually identical patterns as Figures 1 and 2.

As shown in Figure 1, persons in families with the greatest amount of financial assets (quartile 4) had better health than persons in families with fewer financial assets. While this relationship emerged across all age groups, the differences in reported health by financial asset quartile were greatest for persons age 65 or older. Within this group a full scale-point difference in health exists between persons with the highest (quartile 4) and lowest (quartile 1) financial assets, suggesting that total financial assets may be a more meaningful predictor of health status (and possibly proxy of SES) among the elderly, compared to younger persons.

Figure 2 presents a similar graph, except total financial assets has been replaced by net property value. The latter was created by subtracting the amount still owed on the property from the value of the property. Once again, the resulting distribution was broken into quartiles, and a “missing” category was included in the analysis.¹¹ Similar to the results for total financial assets, persons in families with higher net property values (quartile 4) reported better health than persons in families with lower net property values. Again, this relationship was observed for all three age groups. And similar to the previous results, the greatest range in health status scores across the net property value quartiles was observed for persons aged 65 or older.

¹⁰ The quartiles break down as follows: quartile 1 (\$0-\$400), quartile 2 (\$401-\$14,000), quartile 3 (\$14,001-\$83,000) and quartile 4 (\$83,001-\$999,995).

¹¹ The quartiles break down as follows: quartile 1 (-\$160,000-\$30,000), quartile 2 (\$30,001-\$80,000), quartile 3 (\$80,001 - \$180,000), and quartile 4 (\$180,001-\$999,995).

The results clearly suggest a relationship between familial wealth and respondent-reported health status, with the association being somewhat stronger among older persons. The question still remains, however, as to whether familial wealth adds value to our analyses of health status and health outcomes, net of other SES measures (e.g., income, education, employment status). To assess this question we performed two logistic regressions, with respondent-reported health status dichotomized as follows: 0=poor/fair/good and 1=very good/excellent. The first model included the quartile measure of total financial assets (along with a “missing” category) and controlled for age, sex, education, employment status, and total family income. For the second model, total financial assets was replaced by net property value (quartile measure with a “missing” category), and age, sex, education, employment status, and total family income were entered as controls.

The results from both logistic regressions were consistent with findings from the descriptive analyses.¹² In the first model, persons in quartiles 3 and 4 (highest amount) of total financial assets, compared to persons in quartile 1 (lowest amount), were more likely to report very good/excellent health, net of other SES measures. With the second model, persons in quartile 4 (highest amount) of net property value were more likely to report very good/excellent health than persons in quartile 1 (lowest amount). Again, these findings held while controlling for other SES influences.

2.4 Wealth Summary

In sum, the wealth questions appear to provide additional explanatory power when predicting self-reported health status, net of other SES measures. We would expect similar relationships to emerge with other health outcomes. In addition, the results suggest that models of health outcomes that utilize more traditional measures of SES (e.g., education, income) in the NHIS may be underspecified. This would especially appear to be the case when focusing on older adults.

However, before the inclusion of such measures in the NHIS can be considered, the high item nonresponse must be addressed. Note that we did not include interval-based or bracketing follow-up questions for nonrespondents to the exact amount questions. This was largely due to insufficient knowledge of appropriate dollar intervals to be employed. To aid in the development of follow-ups, distributions of the wealth items will be reviewed using the full quarter test file. And finally, more emphasis will be placed on educating interviewers on the utility of asking wealth-based questions. Post-test focus groups revealed considerable interviewer sensitivity to these items.

¹² Logistic regression results available upon request.

3. Collecting Social Security Numbers

Currently, the NHIS attempts to collect social security numbers (SSN) from selected, consenting adults 18 or older. This is done to enable NCHS to link survey data with administrative data sources. The SSN question currently employed in the NHIS is designed to capture all nine digits and reads as follows:

- “We also need your Social Security Number. The National Center for Health Statistics will use your Social Security Number to conduct health-related research by linking your survey data with vital statistics and other records. We may also use it if we need to recontact you or your family. Except for these purposes, the National Center for Health Statistics will not release your Social Security Number to anyone, including any government agency. Providing this information is voluntary and is collected under the authority of the Public Health Service Act. There will be no effect on your benefits if you do not provide it.”

While respondents may be linked via other identifying information (e.g., name, date of birth), some matching algorithms require an SSN. More importantly, respondents who refuse to provide an SSN are ineligible for linkage. Figure 3 presents item nonresponse rates to the SSN question, for adults 18 or older, from 1993-2004. The top black line represents the total nonresponse rate (don't know and refusals combined), which has climbed steadily from just under 30% in 1993 to over 50% by 1999. This high item nonresponse rate has been driven primarily by respondent refusals (red line). The refusal rate exceeded 50% in 2003 and 2004.

To ameliorate the high refusal rate, it was decided, in consultation with NCHS data linkage staff, to test an alternative question that asks for just the last four digits of one's SSN. Designed to reduce respondent sensitivity, this alternative has performed well in other tests (see Bates, 2004). The test question read as follows:

- “Finally, we would like the last four digits of your Social Security Number. The National Center for Health Statistics will use the last four digits of your Social Security Number to conduct health-related research by linking your survey data with vital statistics and other records. Otherwise, the National Center for Health Statistics will not release the last four digits of your Social Security Number to anyone, including any government agency. Providing this information is voluntary and is collected under the authority of the Public health Service Act. There will be no effect on your benefits if you do not provide it.”¹³

¹³ Compared to the 9-digit question, the alternative item was slightly shorter due to the deletion of a sentence on recontact and the streamlined wording of a subsequent sentence.

This alternative SSN question was included at the very end of the family questionnaire (immediately following the family income section) and was asked of family respondents in all screened-out cases.

3.1 Comparing 4-digit and 9-digit SSN Responses

To assess the utility of the alternative SSN question, we compared item nonresponse rates for the two versions overall and by a set of selected family respondent characteristics (correlates of SSN nonresponse in earlier analyses). Note, however, that the comparisons are not based on a split-ballot or experimental design. The decision to test an SSN alternative occurred very late in the development of the test instrument, prohibiting the programming of a randomized administration of the questions. Thus, the decision was made to ask the 4-digit alternative of family respondents in all screened-out cases (similar to the wealth items). To perform comparisons of responses to the 4-digit and 9-digit SSN questions, we identified a set of participating, non-screener (non-test) households from quarter 2 with similar racial/ethnic compositions to our screened-out households. Since these were in-scope, participating households, respondents received the 9-digit SSN question. Understanding that the two groups may differ on characteristics other than race/ethnicity, we first compared them on a set of family respondent characteristics. With the exception of housing tenure and marital status, the groups were very similar in their sociodemographic compositions.¹⁴ This bolstered our confidence that any differences we observed in SSN reporting were real and not simply artifacts of the group compositions. But again, the SSN response comparisons should be viewed with some caution.

3.2 SSN Results

Table 2 presents item response rates for the two types of SSN questions. Within the 9-digit sub-sample, only 40.2% of family respondents provided an SSN or indicated they did not have an SSN.¹⁵ Nearly 59% refused to provide the information. Conversely, 53.7% of family respondents in screened-out households (all 4-digit cases) reported the last four digits of their SSN or indicated they did not have an SSN. About 45% refused to provide the requested information. In sum, a roughly 13 percentage point increase in reporting was achieved with the 4-digit alternative, largely due to a concomitant decrease in refusals. As hypothesized, the 4-digit alternative appears to reduce respondent sensitivity.

Next, nonresponse rates to the two SSN questions were compared on a set of family respondent characteristics.

¹⁴ Family respondents in screened-out households were less likely to be married and to own their residence.

¹⁵ In Table 2 and Figures 4 and 5, the participating, non-screener cases used in the comparisons are referred to as the “9-digit sub-sample” cases.

We report on two of those comparisons here. Figure 4 presents item nonresponse rates for the two questions by family respondent age. Focusing on the first set of bars, it appears that non-reporting of 9-digit SSNs was highest among middle-aged followed by younger respondents. Surprisingly, family respondents age 65 or older had the lowest level of item nonresponse to the 9-digit question. A different pattern emerges, however, among family respondents receiving the 4-digit question (second set of bars). Younger followed by middle-aged family respondents were more receptive to this question, with elderly respondents producing the highest level of item nonresponse.

Figure 5 presents item nonresponse rates for the two SSN questions by family respondent education. A clear pattern emerges among the 9-digit sub-sample whereby more educated respondents have much higher levels of item nonresponse than respondents with a high school education or less. However, among those receiving the 4-digit alternative, there is a leveling of item nonresponse across educational categories. Respondents with less than a high school education still have the best reporting rates, but the differences among the remaining categories are minimal.

3.3 SSN Summary

Early results suggest that the 4-digit SSN question is a clear alternative to asking for the full nine digits. Again, a roughly 13 percentage point increase in reporting was achieved using the 4-digit version.

We also observed variation in response to the 4-digit question by certain family respondent characteristics. While response rates (or item nonresponse rates) were fairly consistent across educational categories, differences emerged by respondent age. Thus, while reporting rates are improved using the 4-digit question, considerable item nonresponse remains and the resulting sample of SSN respondents may not be representative of all respondents or the larger population. The implications for linkage should be addressed.

Finally, barring dramatic changes to these outcomes upon analyzing the full quarter file, plans are underway to replace the 9-digit SSN question with a 4-digit SSN question in the NHIS sometime in 2007-08. However, given the considerable item nonresponse still present with the 4-digit version, additional investigations are warranted including the feasibility of moving toward more SSN-less matching algorithms.

References

- Bates, N. 2004. "Quantitative Testing of Informed Consent Questions in an RDD Survey: Summary Report." Washington, DC: Demographic Surveys Division, U.S. Census Bureau.
- Dahlhamer, J., J. Dixon, P. Doyle, J. Eargle, D. Griffin, and P. McGovern. 2004. "Quality at the Item Level: Terms, Methods, and Guidelines for Cross-Survey Comparisons." Proceedings of the 2003 Federal Committee on Statistical Methods Conference.
- Kahn, J. R. and E. M. Fazio. 2005. "Economic Status Over the Life Course and Racial Disparities in Health." *Journal of Gerontology: SERIES B*, 60B: 76-84.
- Martikainen, P., J. Adda, J. E. Ferrie, G. Davey Smith, and M. Marmot. 2003. "Effects of Income and Wealth on GHQ Depression and Poor Self-Rated Health in White Collar Women and Men in the Whitehall II Study." *Journal of Epidemiology and Community Health*, 57: 718-723.
- Matthews, R. J., L. K. Smith, R. M. Hancock, C. Jagger, and N. A. Spiers. 2005. "Socioeconomic Factors Associated with the Onset of Disability in Older Age: A Longitudinal Study of People Aged 75 Years and Over." *Social Science and Medicine*, 61: 1567-1575.
- Meyer, P. S., J. M. Dahlhamer, and J. R. Pleis. 2007. "Developing New Methods and Questions for Improving Response and Measurement on Sensitive Questions on the National Health Interview Survey." Proceedings of the 2006 Joint Statistical Meetings.
- Moore, J. C., L. L. Stinson, E. J. Welniak, Jr. 2000. "Income Measurement Error in Surveys: A Review." *Journal of Official Statistics*, 16(4): 331-361.
- Pleis, J. R., J. M. Dahlhamer, and P. S. Meyer. 2007. "Unfolding the Answers? Income Brackets and Income Nonresponse in the National Health Interview Survey (NHIS)." Proceedings of the 2006 Joint Statistical Meetings.
- Robert, S. and J. S. House. 1996. "SES Differentials in Health by Age and Alternative Indicators of SES." *Journal of Aging and Health*, 8(3): 359-388.
- SAS Institute Inc. 2006. SAS Version 9. Cary, NC.
- von Rueden, U., A. Gosch, L. Rajmil, C. Bisegger, U. Ravens-Sieberer, and the European KIDSCREEN group. 2006. "Socioeconomic Determinants of Health Related Quality of Life in Childhood and Adolescence: Results from a European Study." *Journal of Epidemiology and Community Health*, 60: 130-135.

Table 1. Item Response Rates for Wealth Questions (Family Respondents): NHIS, 2006, Quarter 2 (Weeks 1-8)

Item	All Screened-Out Cases	Control Path	Experimental Path
Total Value of Family Assets			
Valid	51.2%	51.9%	50.5%
Don't know	17.5	15.8	19.1
Refused	31.4	32.3	30.4
	(n=1,339)	(n=665)	(n=674)
Value of Property (if own home)			
Valid	70.8%	73.8%	67.7%
Don't know	13.6	9.6	17.7
Refused	15.7	16.6	14.7
	(n=869)	(n=439)	(n=430)
Owe on Property (if own home)			
Valid	65.1%	65.3%	64.9%
Don't know	10.9	9.4	12.6
Refused	24.0	25.3	22.6
	(n=868)	(n=438)	(n=430)
Car Ownership			
Valid	93.3%	92.6%	93.9%
Don't know	0.3	0.3	0.3
Refused	6.4	7.1	5.8
	(n=1,338)	(n=664)	(n=674)

Table 2. Item Response Rates for SSN Questions (Family Respondents): NHIS, 2006, Quarter 2 (Weeks 1-8)

Item	Provided SSN/		
	No SSN	Refused	Don't know
9-Digit Sub-Sample (n=913)	40.2%	58.5%	1.3%
All Screened-Out Cases—4-Digit (n=1,335)	53.7	44.8	1.7
4-Digit—Control Path (n=662)	53.8%	44.0%	2.3%
4-Digit—Experimental Path (n=673)	53.3	45.6	1.0

Figure 1. Mean Respondent-Reported Health Status by Total Family Financial Assets (Quartiles) and Age: NHIS, 2006, Quarter 2 (Weeks 1-8)

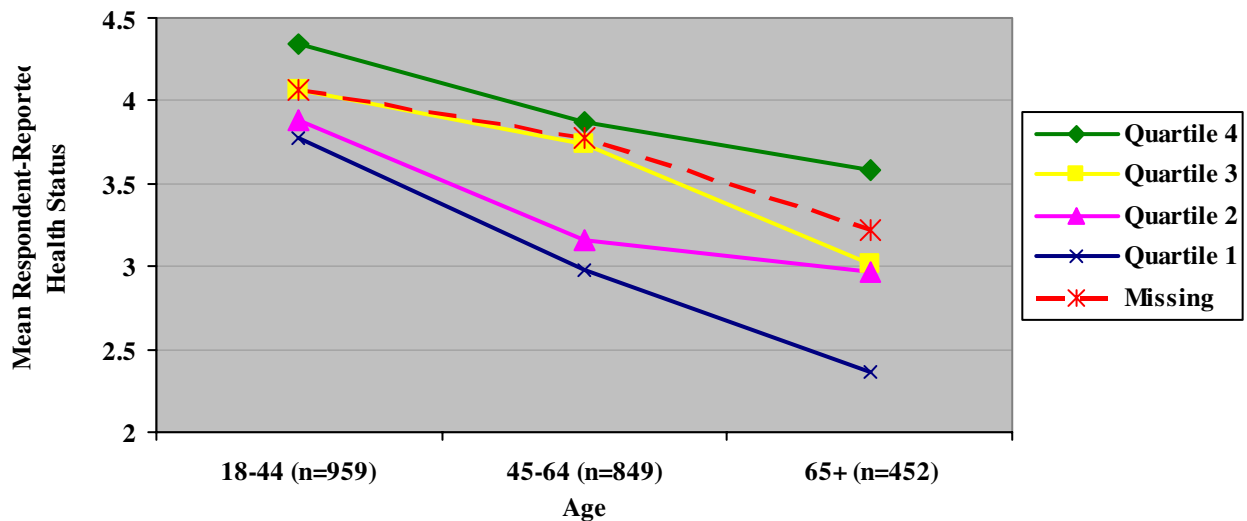


Figure 2. Mean Respondent-Reported Health Status by Net Property Value (Quartiles) and Age: NHIS, 2006, Quarter 2 (Weeks 1-8)

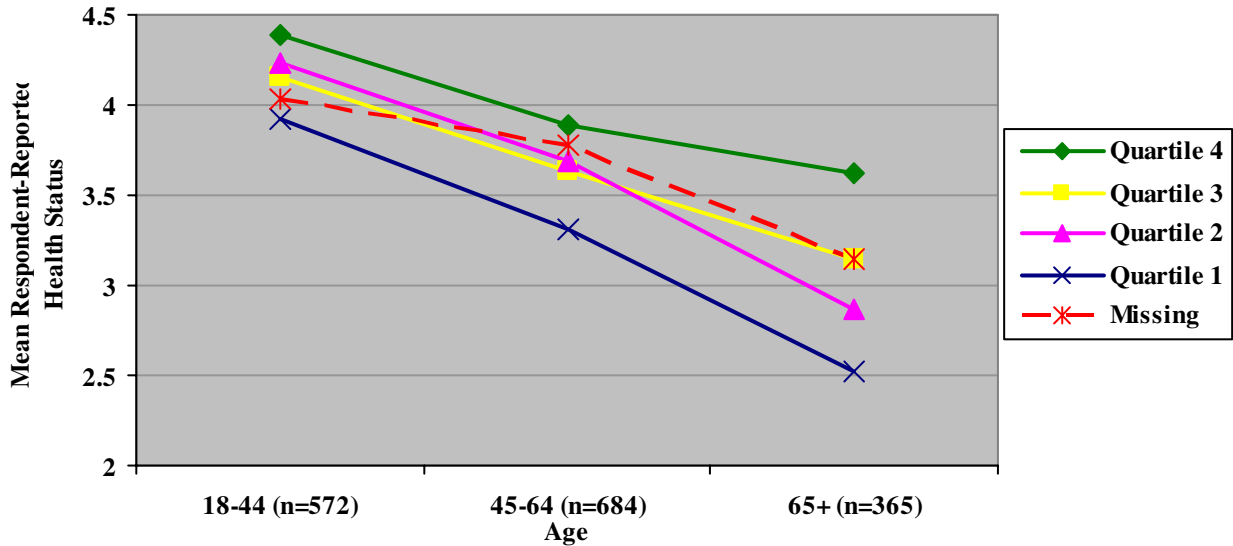
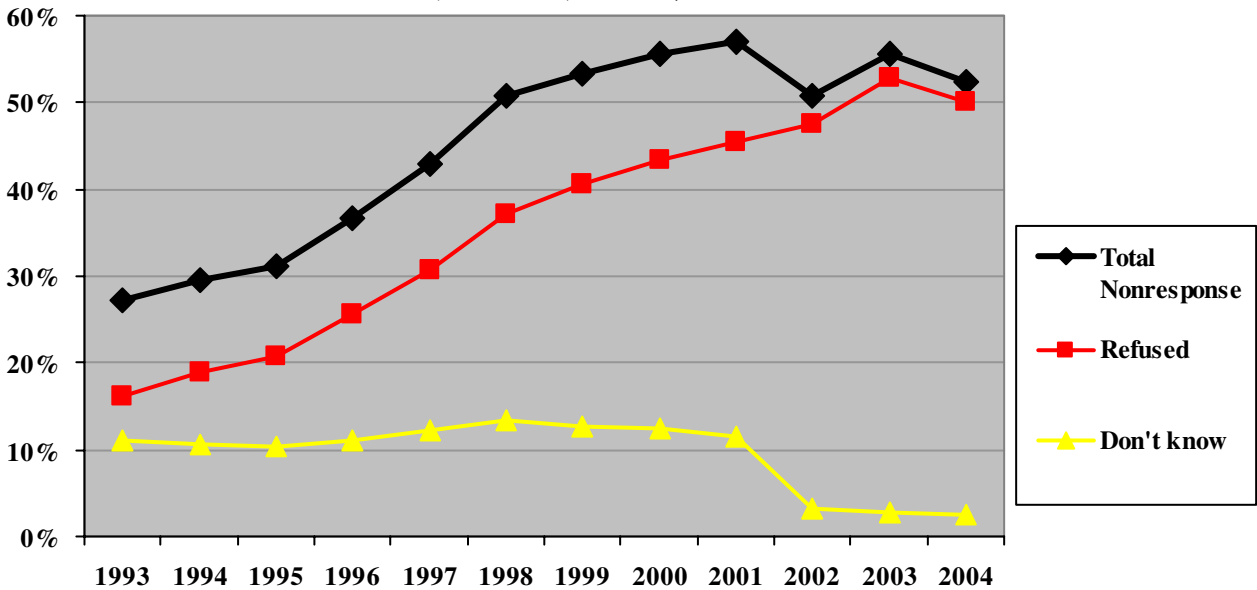


Figure 3. Item Nonresponse Rates for 9-Digit Social Security Number Question (adults 18+)*: NHIS, 1993-2004



* Asked of all adults prior to 2002 (included proxy response); asked only of consented adults beginning in 2002 (family respondent, sample adult respondent, sample child respondent).

Figure 4. Item Nonresponse Rates for the 4-Digit and 9-Digit SSN Questions by Age of Family Respondent: NHIS, 2006, Quarter 2 (Weeks 1-8)

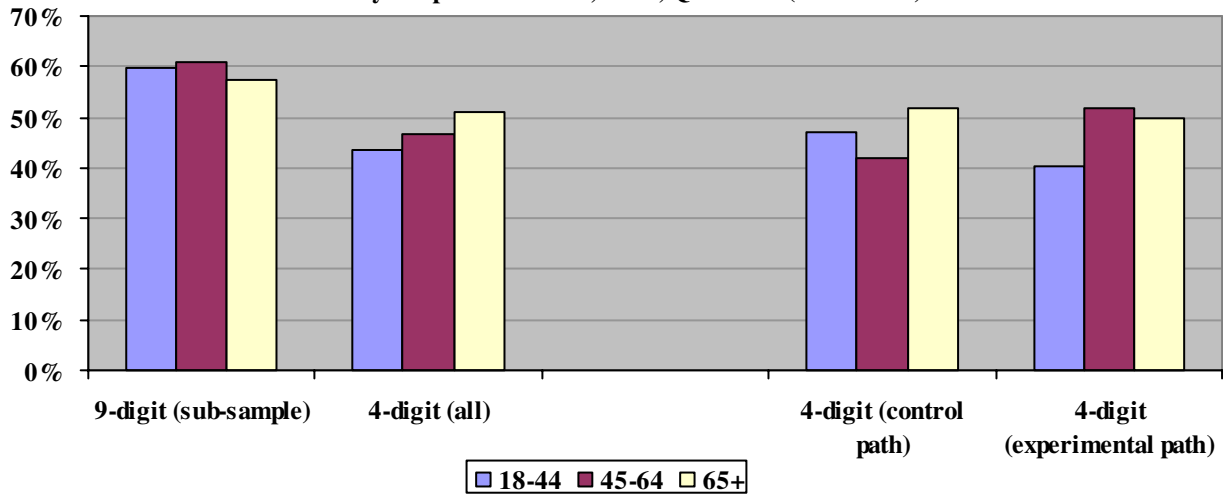


Figure 5. Item Nonresponse Rates for the 4-Digit and 9-Digit SSN Questions by Education of Family Respondent: NHIS, 2006, Quarter 2 (Weeks 1-8)

