

## WHO FILLS OUT THE CENSUS FORM?

Theresa J. DeMaio and Nancy A. Bates, U.S. Bureau of the Census  
Theresa J. DeMaio, U.S. Bureau of the Census, Washington, DC 20233

**Keywords:** respondent characteristics, self-administered questionnaires

The decennial census form is designed to capture information about all persons residing within a household as well as facts about the housing unit in which they live. For the most part, the census is self-administered; in 1980 the form was mailed to approximately ninety-five percent of the population with the intent that respondents complete it themselves and mail it back. For the remaining percent of households (and those who failed to mail back a form), census information was obtained by way of an enumerator interview (Bureau of the Census, 1986).

The range of information collected in the census is quite broad, particularly for the long form (sent to approximately 1 in 6 households), which contains questions about each household member's employment, income, past residence, and other various inquiries. It also requires answers to such housing questions as rent, mortgage, and insurance.

The actual completion of the census form involves a complex response task in a situation that is totally unstructured. Once the forms are delivered, anyone (or everyone) in the household can complete them. No instructions are provided about who should fill out the census form, and little is known about who actually completes it. We have some general information about how many respondents participate in the form-filling process for a single household. This is a start, since one person in the household may not have complete information about all the diverse topics included on the form. For example, the family member who knows the most about the financial aspects of the property might not be the best to report on the work histories of other family members. Also, the expansion of "nonfamily" households has significant effects on the amount of information shared between members. Nonfamily households where a householder is not related to any of the persons who share the residence now account for 5 percent of all U.S. households (Bureau of the Census, 1989).

From previous research (DeMaio, 1983), we know that 78 percent of the households in the 1980 census involved one respondent in the form-filling process, and the remaining 22 percent involved two or more persons.<sup>1</sup> Thus, in most cases, a single respondent provides all the information requested. We don't know anything about that respondent, however. Although the name and telephone number of the

respondent are requested on the last page of the census form, this information is used only to contact the household to obtain additional information, and is never captured on the census data file.

In this paper, we present results from one of the pre-1990 census tests in which the identity of the respondent was coded and captured. We use this information to paint a demographic portrait of the census respondent. This information is important for two reasons. First, it may give us some indirect measures of the quality of census data. Differences are likely in the ability of various household members to perform as a household respondent. The identity of the respondent, and such characteristics as age, educational level, and relationship within the household, may provide some indirect information about the quality and accuracy of the data. Factors such as the respondent's knowledge of the housing unit and other household members, and his/her accessibility to information and records may influence how well the respondent completes the form, or how much data he/she can provide about other household members.

The second reason that information about respondent characteristics is important is because this information can help us design a better census. Learning more about the characteristics of persons who fill out the forms can give us a better idea of how to design the form and also suggest where to target our outreach programs.

### METHODOLOGY

This analysis is based on data from the 1988 National Census Test. This test was a nationwide mailout/mailback survey using a sample of 14,448 addresses selected for households from expired rotation groups of the National Crime Survey. The survey was designed to test three experimental long forms that contained slight wording, layout and question order changes from the traditional long form. Although slightly different from each other, all four forms (one control and three experimental forms) contained the same content as the 1988 Dress Rehearsal questionnaire, including a request for the name of the respondent. For the purpose of this research, data from all questionnaire versions are aggregated, as it is assumed that the differences among the four versions did not affect who filled out the form within a household.<sup>2</sup>

Questionnaires were mailed out on May 30, 1988, and each address received a nontargeted mail reminder card the following week. Overall, the

response rate was 45.9 percent, representing 5,875 returns from 12,788 eligible sample units. Post office and respondent returns for vacant housing units, and undeliverable addresses were defined as "ineligible." No information was obtained on nonrespondents to the mailout.

In cases where the respondent recorded his/her name on the back page (approximately 95 percent overall)<sup>3</sup>, coders examined the form to see where this person had placed him/herself in the 100-percent matrix. This refers to the area of the form that includes the basic demographic questions such as relationship to head of household, age, sex, and marital status. The number of the person column where the respondent was found was identified and coded--for example, if the name on the back page matched the name listed in column 2, the form received a code of "2." This enabled a record to be created for the respondent so that characteristics such as age, sex, and relationship could be pinpointed.

### FINDINGS

The first characteristic we looked at was the relationship of the respondent to the householder--that is, the person listed in column 1 of the 100-percent matrix. The Census Bureau defines the householder as the person in whose name the house/apartment is owned, being bought, or rented. Without any information about the census respondent, the Census Bureau tends to assume that the person listed in column 1 is the householder and is also the respondent. There is no information available to validate whether the person listed in column 1 is actually the householder; however, we wanted to investigate the assumption that the person listed in column 1 is the respondent.

In addition, relationship is important in terms of household knowledge. We know that each household member does not fill out his/her own portion of the form. In terms of completeness of information, the next best thing is a proxy with a high level of knowledge about each person.

Table 1 contains information about how the person who filled out the test form related themselves to the householder. A large majority (69 percent) of respondents actually did not answer the relationship question about themselves. They placed themselves first in the matrix, making them "person 1" or the householder. This percentage sheds some light on how reliable the assumption is that the householder is the person who completes the form. As Table 1 reveals, this assumption would be correct the majority of the time, but would still prove incorrect in over 30 percent of all cases. Later, we will examine whether significant differences exist between the demographic characteristics of the householder and those of the

person who completed the form.

Slightly less than 30 percent of the persons completing the form listed themselves as the householder's husband/wife. In total, 97 percent of the forms mailed back were completed by either the self-identified householder or by the spouse of the householder. This would indicate that proxy knowledge levels are high in most cases.

Looking at the remainder who defined their relationship as something other than householder or spouse, we see that the householder's children have the next highest tendency to complete the form. Persons identifying themselves as sons and/or daughters of the householder completed 2 percent of the forms. A smaller percentage of respondents (1 percent) considered themselves to be "unmarried partners" of person 1. The other categories (sibling, parent, grandchild, other relative, roomer, roommate, employee, other nonrelative) represented less than one percent of those who completed the form. It is noteworthy that 1.1 percent of the names on the back page did not match to the person columns because the person was a non-household member. Several of these forms were examined and found to be completed by neighbors and/or non-household relatives who filled out the form for an elderly person(s).

The next variable we examined was the age of the respondent. This has few direct implications for the quality of the data reported, except in the case of the elderly. Older respondents are more likely than younger ones to suffer from visual and other impairments, which could affect their ability to read and complete the form correctly.

Table 2 presents a breakdown by age categories. Almost all respondents reported being over 18 years of age, with the largest category reporting themselves in the 35-44 age category (20 percent). Almost one-quarter of all forms were completed by persons aged 65 or older. Most importantly, nine percent of the respondents were aged 75 or older. In some cases, as noted above, the elderly may enlist the assistance of non-household members to complete the census form. However, a sizable percentage of persons completing the form are in an age group that might have difficulty reading, understanding, and filling it out. This would seem to have implications for how the census form is designed.

The next respondent characteristic, education, is of extreme importance when considering the design and content of the census form. Because the form cuts across all socioeconomic boundaries, it must be designed so that persons from a variety of educational backgrounds can successfully comprehend and complete it. One expert who examined the word frequency and the grade level of the vocabulary found that the census form is

comprehensible to persons with a 5th or 7th grade education level, depending on the word frequency reference used (Farstrup, 1987). Another study that evaluated the reading and comprehension level of the introductory letter and selected blocks of questions indicated that anywhere from three to twelve years of education is required (Williamson, 1989). It should be pointed out, however, that this analysis is based on the use of readability formulas developed to evaluate written prose, which are not necessarily valid for census questions (see DeMaio, 1985).

Table 3 reveals the education levels of persons completing the form. It suggests that respondents are fairly well-educated, with 83 percent having completed high school. Fewer than 1 percent reported less than a fifth grade education. This percentage reflects, of course, only those persons mailing back a questionnaire; no education information was obtained about the 54 percent who failed to respond.

Table 4 presents the distribution of levels of education completed for householders who did not complete the census form and for the member of the household who did. In cases where the householder was not the respondent, the family member who completed the form tended to have more education than the householder. For example, non-householders completing the form were more likely to have a high school education than their householder counterparts (37 percent vs. 31 percent). This suggests that in cases where the householder is less educated than another member of the family, the more educated member tends to complete the form. This, of course, is likely to have positive consequences for the quality of data reported in the census.

A second concern about the comprehensibility of the census form involves households where English may be a second language. In 1980, approximately 11 percent of the U.S. population aged 5 and older reported speaking a language other than English in the home. Of this population, approximately 44 percent reported speaking English less than "very well" (Bureau of the Census, 1983). English proficiency, then, becomes an issue, especially considering one evaluation of the proposed 1990 forms which described them as "nearly insurmountable to the reader for whom English may be a secondary language" (Farstrup, 1987).

In the census, households are given a toll-free number to call and request a Spanish language form. For the 1988 National Census Test, no alternative to the English version was provided. It is possible, however, that in households where a second language is primary to adults, a child with stronger English skills may be able to intercede and fill out the form for the household.

We tested this hypothesis by examining households where a second language was spoken. Among this group of households (N = 412), 68 percent of the respondents were householders while 25 percent reported they were the spouse of the householder. Six percent, however, indicated they were children of the householder. This is a larger percentage (by 4 percent) than we found earlier with the overall population. Thus, apparently, the census form is sometimes delegated to the more English-fluent member within second-language households. This would seem to have implications in terms of our outreach programs.

The last demographic variable we examined was the sex of the respondent. Table 5 presents gender frequencies for both the respondent and the householder. A divergence between the characteristics of the householder and the respondent is evident for this variable. If one assumes that the householder completes the form, we would be inclined to believe that the large majority (71 percent) of respondents are male. As Table 5 shows, this is not the case; more females than males completed the test census questionnaire (56 percent vs. 44 percent). Thus, it appears that the feminism movement notwithstanding, women completing the form tend to list their husbands rather than themselves as the householder.

We reported earlier that for the most part, the respondent was either the householder or his/her spouse. We would expect either of these respondents to be relatively knowledgeable about the kinds of information requested in the census. However, there have been studies suggesting that answers to questions may differ significantly depending on the sex of the respondent. A study comparing the responses of husbands and wives (Anderson and Silver, 1987) showed that substantial percentages of the respondents gave different answers to such items as the number of rooms in a housing unit and the listing of household members. In a study examining household income reporting (Hutcheson and Prather, 1978), the percentage of cases in which the respondent refused and/or was unable to provide an answer varied by the sex of the person being interviewed. While females were found to be less likely to refuse household income questions, they were more likely to lack the knowledge to answer them.

The items on the census long form cover a wide range of information levels, from those that are more or less obvious (e.g., sex, marital status) to those that require detailed knowledge or looking in records (e.g., income amounts). We chose to examine some of the more difficult items to see if response to these items differed by the sex of the respondent. We used the presence or absence of a response to an item as a measure of respondent

knowledgeability.

We first examined nonresponse differences by sex for selected housing items. The long form contains a series of questions for homeowners that concentrates on property value and mortgage payments. We tested for sex differences by examining rates of item nonresponse to questions on property value, real estate taxes, insurance payments, presence of mortgage and mortgage amount. Item nonresponse was defined as either leaving the question blank or writing in a "don't know" response.

Table 6 shows that of the five items included in this analysis, significant differences in item nonresponse were found in two cases. Females were just as likely to answer questions concerning property value and mortgage; however, they were significantly less likely to answer the questions regarding real estate taxes and property insurance. Perhaps this occurred because a check containing the mortgage amount is written each month, while the real estate taxes and insurance payments are either paid less frequently or are included in the mortgage payment. Females may have enough information on the former to answer--perhaps they are regularly contributing to the payments. Conversely, they may lack the knowledge necessary to answer the latter because the (annual or semi-annual) task of payment is not their responsibility or they are unfamiliar with the records that contain the information.

We also examined nonresponse differences to household income questions by the sex of the respondent. This part of the questionnaire asks the respondent to report income earnings for each member of the household aged 14 and over from a variety of sources: wages, self-employment, farm self-employment, interest, social security, supplemental security, pensions, other income, and total income. Nonresponse was defined as leaving both portions of the question blank (yes/no and amount), or reporting "don't know," or indicating "yes" to an income category but leaving the amount earned blank. As Panel 1 of Table 7 reveals, income item nonresponse is higher when females complete the form. The differences between male and female respondents in the level of item nonresponse for any household member averaged approximately 3 percent and were found to be statistically significant for every kind of income. When males completed the test form, for example, a household member's wage income data was missing 16 percent of the time. This is comparable to the significantly higher nonresponse rate of 20 percent when a female provided wage income information for the household.

The second panel of Table 7 shows that this is true for self-reports as well. In every category, male

respondents were more likely to answer questions concerning their own income than were female respondents. This finding seems at odds with other research involving a personal visit survey (Spiers et al, 1978), which suggests that item nonresponse for males (regardless of who the respondent is) is higher than that for females. One explanation is that the item nonresponse for women includes unemployed women who leave the question blank rather than indicate that they earned no income. However, a nonresponse/refusal cannot be distinguished from an implied entry of "none." If this is the case, the finding is not a consequence of lack of knowledge on the part of females, but rather a misunderstanding of how to properly respond to the question. It is also important to note here that we have not controlled for household characteristics, so differential characteristics of households that have a female respondent may also contribute to the observed differences.

To further examine the effect of potential levels of knowledge by the sex of the respondent, we limited the analysis to all household members for whom proxy reports were made. Here, as the third panel of Table 7 shows, we see no significant differences between the item nonresponse rates on forms completed by males and females. Apparently, females are just as likely to answer questions about other household members' income as male respondents are; differences seen earlier are a result of female respondents' tendency to have higher self-nonresponse rates. Thus, male and female respondents tend to be equally knowledgeable in terms of answering income items for household members.

## SUMMARY

In this paper we have attempted to paint a descriptive picture of who the census respondent within a household is. Data from the 1988 National Census Test reveal that respondents are most likely to be listed either in column 1 or column 2 of the census form and to be defined as the householder or the spouse of the householder. However, in cases where a second language is spoken in the home, the percentage of children completing the form (presumably with a better understanding of English) is higher than average. This suggests that in areas with a prevalence of foreign-language speakers, targeting outreach efforts to children (either through the schools or outside of the schools) would have positive benefits.

Examination of the demographic characteristics of the respondent shows that nine percent were aged 75 or older, posing potential response problems due to physical infirmity. Perhaps enlarging the size of the print on the census form would improve response among this age group.

Most respondents are fairly well-educated, and almost all have completed at least fifth grade, which is the education level required according to one reading specialist who evaluated the census form. In terms of gender, respondents are fairly evenly split between males and females.

Analysis of item nonresponse rates to selected long form items requesting financial information shows that there are some differences in the knowledgeability of male and female respondents, but the differences are not overwhelming. Female respondents were less likely than males to answer questions regarding tax and insurance payments and were less likely to provide their own income. However, there were no differences in ability to report amounts of mortgage payments or income information for other household members.

#### NOTES

<sup>1</sup> These data from the Applied Behavior Analysis Survey are not included in PERM #61.

<sup>2</sup> No interaction was found between form type and location of respondent in the 100-percent matrix.

<sup>3</sup> The three experimental forms had a significantly higher response rate to back page items compared to the control form.

#### REFERENCES

Anderson, B. A. and B. D. Silver (1987) "The Validity of Survey Responses: Insights from Interviews of Married Couples in a Survey of Soviet Emigrants," Social Forces, Vol. 66, No. 2, 537- 554.

DeMaio, T. J. (1983) "Results from the Applied Behavior Analysis Survey," (Preliminary Evaluation Results Memorandum #61). Internal Report, U.S. Bureau of the Census.

DeMaio, T. J. (1985) "Information About the Reading Level of the 1980 Census Form," Internal Census Bureau Memorandum for Paulette Lichtman-Panzer. March 5, 1985.

Farstrup, A. (1987) Testimony before a Joint Oversight Hearing of the Senate Subcommittee on Federal Services, Post Office, and Civil Service and the House Subcommittee on Census and Populations on U.S. Bureau of the Census Questionnaire Content for the 1990 Census. May 14, 1987.

Hutcheson, J. D. and J. E. Prather (1978) "Maximizing Response to Household Income Items: Respondent Selection." Proceedings of the American Statistical Association, Section on Survey Research Methods: 294-296.

Spiers, E., J. Coder and M. Ono (1971) "Characteristics of Income Nonrespondents in the Current Population Survey." Proceedings of the American Statistical Association, Social Statistics Section: 369-374.

U. S. Bureau of the Census (1989) Household, Families, Marital Status and Living Arrangement: March 1989 (Advance Report). Current Population Report, Series P-20, No. 441.

U.S. Bureau of the Census (1986) 1980 Census of Population and Housing: History Part A. PHC80-R-2A, September, 1986.

U.S. Bureau of the Census (1983) Census of Population: General Social and Economic Statistics.

Williamson, C. (1987) "Readability Formulas and the 1990 Census." Internal Report, U.S. Bureau of the Census, Director's Office. August 11, 1987.

**Table 1**  
**Relationship of Respondent to Householder**

Cases in which respondent was:	<u>Percent</u>
Householder	69%
Husband/Wife of Householder	28
Child of Householder	2
Unmarried Partner of Householder	1
All Other Relationships	**
	N=5,452

Note: \*\* = less than one percent

**Table 2**  
**Frequency Distribution of Respondent's Age**

Cases in which respondent's age was:	<u>Percent</u>
14 - 18	**%
19 - 24	3
25 - 34	18
35 - 44	20
45 - 55	16
56 - 64	18
65 - 74	15
75 +	9

N=5,302

Note: \*\* = less than 1 percent

**Table 3**  
**Frequency Distribution of Respondent's Education**

Cases in which respondent completed:	<u>Percent</u>
0-4th grade	**%
5th-8th grade	6
9th-11th grade	8
12th grade, no diploma	2
High School	29
Some College	22
Associate Degree	6
Bachelor's Degree	16
Master's Degree	7
Professional Degree	2
Doctorate Degree	1
	N=5,342

Note: \*\* = less than one percent

**Table 4**  
**Distribution of Levels of Education of Respondents and Householders for Cases Where the Householder Did Not Complete the Form**

Cases in which person completed:	<u>Respondent</u>	<u>Householder</u>
0-4th grade	**%	2%
5th-8th grade	3	9
9th-11th grade	7	11
12th grade, no diploma	2	3
High school	37	31
More than high school	50	44
	N=1,704	N=1,710
Overall $X^2 = 105.9$ , d.f.=5, $p < .0001$		
Note: ** = less than one percent		

**Table 5**  
**Distribution of Sex of Respondents and Householders**

Cases in which person was:	<u>Respondent</u>	<u>Householder</u>
Male	44%	71%
Female	56	29
	N=5,359	N=5,713

**Table 6**  
**Nonresponse to Selected Housing Items by Sex of Respondent**

Cases in which housing item was:	Item Nonresponse Rate When Respondent was:	
	<u>Male</u>	<u>Female</u>
Property Value	6%	7%
Base N	(1,869)	(2,102)
Real Estate Taxes*	10	14
Base N	(1,869)	(2,102)
Insurance Amount**	11	16
Base N	(1,869)	(2,102)
Presence of Mortgage	7	8
Base N	(1,869)	(2,102)
Mortgage Amount	3	3
Base N	(1,056)	(1,142)
* $X^2 = 19.0$ , d.f.=1, $p < .001$		
** $X^2 = 18.6$ , d.f.=1, $p < .001$		

**Table 7**  
**Nonresponse to Income Items by Sex of Respondent**

Cases in which Income Item was:	Item Nonresponse For All HH Members*		Item Nonresponse For Self*		Item Nonresponse for Proxy HH Members <sup>+</sup>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Wages	16%	20%	14%	21%	17%	18%
Self-Employment	31	34	34	29	32	33
Farm Self-Empl.	33	34	31	38	34	34
Interest	27	30	22	31	31	30
Social Security	28	30	24	30	30	30
SSI	32	35	31	36	34	34
Pensions	30	33	26	34	33	32
Other Income	32	34	30	36	34	33
Total Income	12	16	10	17	14	14
	N=5,054	N=5,864	N=2,355	N=2,927	N=2,699	N=2,937

\* All  $X^2$  values between male and female nonresponse rates are significant at .05 level, d.f.=1. These tests have not yet been adjusted for within household clustering.

<sup>+</sup>  $X^2$  values between male and female item nonresponse not significant at .05 level, d.f.=1.