

RESULTS OF THE NONSAMPLING ERROR MEASUREMENT SURVEY OF FIELD INTERVIEWERS

Susan A. Lavin, U.S. Bureau of the Census
Field Division, Room 3027-3, Washington, D.C. 20233

KEY WORDS: Nonsampling error,
interviewer perspective

1. Background

Total survey error is made up of two components, sampling error and nonsampling error. For many years, sampling error estimates have been routinely calculated and reported for all statistics published by the U.S. Census Bureau. However, there is no agreed-upon method for estimating the effect of nonsampling error on the data. Important statistics can be seriously biased by errors due to misleading or ambiguous questions, improper interviewing, respondent misunderstanding or memory failure, and coding, editing, and processing errors. The Census Bureau presently has no standard way of calculating the total effect of these nonsampling errors on the sample survey estimates it publishes.

The Census Bureau has conducted research on nonsampling error for over 40 years and has gained a great deal of knowledge about the causes of error (Hansen, Hurwitz and Bershad, 1961; Brooks and Bailar, 1978; Staff, Bureau of the Census, 1974). The Bureau has taken steps to reduce nonsampling error by improving interviewer training, testing questionnaires, and using prior interview data. However, much work remains to be done. The potential for serious error remains, and we lack full understanding of the causes and consequences of nonsampling error in the data. The Census Bureau has a continuing commitment to research and evaluation of nonsampling error.

At the Second Annual Research conference in March 1986, sponsored by the Bureau of the Census, Dr. John G. Keane outlined an eight point program to advance the understanding and measurement of nonsampling error in Census Bureau data (Keane, 1986). One goal of this program was to learn more about the interviewers' opinions of the causes of error in the data they collect. Interviewers, as a group, have more direct experience with the surveys than anyone, and it makes sense to take advantage of their knowledge and experience to gain a fresh, "close-up" perspective on the reasons for errors in the data. This report summarizes the results of a survey conducted of the field staff for the purpose of identifying the major causes of error as interviewers perceive them.

2. Methodology

Designing the survey questionnaire was not an easy task. The field staff is not generally familiar with the terminology "nonsampling error". Additionally, many of the known causes of nonsampling error are interviewer-related. For example, interviewers not asking questions as worded, or using biased probes, are known causes of nonsampling error. However, it is unlikely that interviewers would willingly confirm these activities since this would be openly admitting to not performing an aspect of their jobs according to established guidelines. Instead it was necessary to break down the causes of nonsampling error further, i.e., what conditions would cause interviewers to not ask questions as worded, or use biased probes? As a result, the survey questions were worded to ask about conditions in surveys that "cause errors and reduce the quality of the data".

In mid-November, 1987 the questionnaire was sent to 810 interviewers and 97 SFRs. Interviewers were selected at a sampling rate of 1 in 3 and SFRs were selected at a rate of 1 in 2 without stratifying by survey. The sample was selected from lists that were arranged by regional office, then alphabetically within each regional office (separate lists for SFRs and interviewers). The final response rate achieved was 86.2%. No compensation was made in the analysis of the survey data for questionnaire or item nonresponse. The observations in the report are the opinions only of the responding field staff. Of this group of employees, 54.5% had worked for the Census Bureau for 5 or more years.

3. Results

The results of the survey should be examined with consideration of several factors. This study was not intended to yield a quantitative measure of nonsampling error in our surveys. Rather, it was designed to gather qualitative data on the causes of survey error from the field staff perspective, and the results should be used in that context only. The field staff is in a unique position to provide subjective insights on aspects of problems with our surveys that are not available through

standard measurement techniques. It is evident that a great deal of thought went into the responses provided by the responding field staff.

Data tables from the survey which are referred to in the report can be found in Appendix 1.

3.1 Items not reported as causes of survey error

Before presenting the survey results regarding those items identified as major contributors to survey error, it is interesting to note which items were not identified. Those items that dealt mainly with the training of the field staff, as well as survey instructions and procedures, were notably missing from the list of major contributors to survey error. In fact, the single most positive feature of Census surveys mentioned by the field staff was good training. This is gratifying since a great deal of effort has been made in the past several years to upgrade the quality of interviewer training and instructions.

3.2 Items reported as causes of survey error

The factors identified by the majority of the field staff as major contributors to survey error can be grouped into three general categories:

1. Factors related to the external environment
2. Factors related to survey work in general
3. Factors inherent to a specific survey

3.2.1 Factors related to the external environment

Factors which exist in the external environment that affect surveys, and which were identified as major contributors to survey error are:

- * Respondents do not trust the government
- * Public awareness of Census Bureau surveys is low

Interpretation of Results

Trust in the government varies depending on the political climate of the time and the circumstances of the individual respondents. We cannot control either situation. One interviewer, in providing the remedy to this problem, noted that "there is a need to restore integrity and ethics of elected officials". It is clear the Census Bureau has no power in this respect. Nor can we affect the circumstances of our respondents, and the lack of trust they have in the government. However, we must do everything we can to convince the public that the Census Bureau is not a threat, and that information they provide to us truly is confidential, and will only be

used for statistical purposes.

Respondents must also be convinced of the value of our surveys, and shown how information gathered from Census surveys does affect their lives. To be truly effective, this cannot be done only with respondents. Every resident of the United States is a potential respondent. The public, for the most part, is unaware of the vast amount of information collected by the Census Bureau. Yet statistics from our data collection activity are featured daily in both the print and broadcast media. However, rarely is the Census Bureau credited as the source of the information. We must convince our sponsors to give us that credit. Interviewers identified lack of public knowledge of Census Bureau surveys as the single greatest cause of nonsampling error. This is primarily due to the fact that, if the public had greater awareness of Census Bureau surveys, the negative effect of many other causes of nonsampling error would diminish. For example, response rates should improve and respondents would be more willing to give their time, recognizing that our data are necessary. Recognizing the value of our data could convince respondents of the need to be thorough with their responses. Increasing public awareness of Census Bureau surveys would benefit both the sponsor and the Census Bureau.

3.2.2 Factors related to survey work in general

The factor which was identified as a major contributor to survey error and which is related to survey work in general is:

- * There is too much pressure to maintain a certain response rate

Interpretation of Results

The Census Bureau takes great pride in providing quality data products. Part of that measure of quality is the very high response rates that are achieved on our surveys. The field staff reported that "there is too much pressure to maintain a certain response rate".

There is no question that the Census Bureau puts a great deal of emphasis on interviewers achieving high response rates. In order to stress common goals in a field organization where interviewers work from dispersed locations and are supervised from regional offices, we have established performance standards, or "guidelines" for each survey. These guidelines help interviewers know what they are expected to achieve, and they help supervisors to provide feedback on the quality of interviewers' work. Although interviewers may perceive guidelines to be rigid, they are flexible in that program supervisors can provide feedback that is appropriate for the individual

circumstances of each interviewer. We are presently testing a revision to our performance standard system that specifies response rate guidelines that differ by area; for example, interviewers who work in our most difficult central cities will be expected to perform to a less strict standard than those who work in rural areas. We feel that this new system continues to emphasize the need for interviewers to produce the very best that they can while assuring that field supervisors recognize the difficulties inherent in individual assignment areas.

3.2.3 Factors inherent to a specific survey

Factors which were identified as major contributors to survey error and which fall under this category are listed in Table 1. Those causes of survey specific error which were identified by a majority of the interviewers as being a problem on the specific survey are identified by an "X".

Looking at these survey specific results provides several insights:

- o Respondent fatigue and interviewer efforts to speed up the interview seem to be reported on surveys of one hour or more in-household time.
- o On multiple-visit surveys, where the same basic set of questions are asked each time and the interview is lengthy, respondent conditioning effects are reported on subsequent visits.
- o Respondent conditioning during the interview, where respondents are perceived to withhold information to shorten the interview, seems to be suggested on surveys which are designed so that a set of detailed questions are repeatedly asked immediately following positive replies to screener questions.
- o Interviewers believe respondents are not knowledgeable, or are reluctant to disclose information on income and expenditures.

TABLE 1

SUMMARY OF MAJOR CONTRIBUTORS TO SURVEY SPECIFIC ERROR

	CE	SIPP	HIS	NCS	CPS	SOC
The questionnaire is too long	X	X	X			
Questions are too repetitious		X		X		
Questions are wordy, or use language that is too complex	X		X			
Respondents often do not know the answers to the questions	X	X				
Sometimes it is necessary to reword questions, or not ask all of them, to avoid a break-off or refusal	X	X				
Interviews often must be rushed	X	X	X	X		
Respondents knowingly withhold information or report incomplete information	X	X				
In subsequent interviews, respondents learn not to give certain answers, in order to shorten the interview	X			X		
Maps are inadequate			X			
There is not enough time to complete the assignments			X			

Interpretation of Results

Consumer Expenditure Survey

According to the responses, the Consumer Expenditure Survey (CE) has the greatest vulnerability to survey error. Seven of the top ten survey specific contributors to error were identified by the majority of the field staff as a problem on CE. This survey is undoubtedly one of the most complex surveys conducted by the Census Bureau, and, of our continuing surveys, is the one that places the greatest demands on the respondents. The average CE interview lasts approximately 2 hours, with each household being interviewed 5 times over the course of 15 months (or once every 3 months). The problems identified on CE are predictive of the concern over the length and complexity of the CE interview and its subject matter. It is also evident that at times respondents and/or interviewers knowingly take steps to shorten the interview.

Survey of Income and Program Participation

The Survey of Income and Program Participation (SIPP) is also quite complex and places great demands on our respondents. The average SIPP interview lasts approximately 1 hour, with each household being interviewed 8 times over the course of 2 1/2 years (or once every 4 months). Six survey specific causes of error were identified by a majority of the field staff as a problem on SIPP. As with CE, the problems identified reflect the concern over the length of the interview and its subject matter. Additionally, the repetitive nature of the questions was indicated as a problem. This survey asks the same set of questions separately about each adult household number. It is interesting that respondent conditioning effects on subsequent interviews was not reported as a problem on SIPP. Unlike other lengthy surveys requiring multiple visits, the SIPP does not utilize the same set of questions on each visit as does the CE and the NCS.

Health Interview Survey

On the Health Interview Survey (HIS) five major contributors to survey error were reported. Three of these contributors dealt with survey length (the questionnaire is too long, interviews often must be rushed, questions are wordy, or use language that is too complex). Each household in this survey is visited only once. However, this survey is in a state of transition, with supplemental questions of increasing length and complexity being asked each year. In 1986 the average HIS interview lasted approximately 1 hour. In 1987 and 1988 that figure increased to 1 1/2 hours. Meanwhile, response rates, traditionally the best of all of our

ongoing surveys, have declined.

The fourth contributor to survey specific error identified on HIS reflects the increasing demands on the interviewer to complete the assignment within a one week time frame (there is not enough time to complete the assignments). It is interesting to note that this is the only survey where this was identified as a major cause of survey error. The sample design selected for the HIS anticipated a weekly workload for an interviewer to be approximately 16 cases. The average time per case for all interviewer-related activity to complete a case was 112 minutes. This resulted in a weekly assignment requiring approximately 30 hours. By 1987, the time frame of this survey, the average assignment size had grown to 20 cases, with the average time to complete each case being 156 minutes. Thus, the average time required to complete an average weekly assignment in 1987 was 52 hours. With the growing need to complete assignments in the evenings and on Saturdays, it is no surprise the HIS interviewers are concerned about the effect this has on obtaining quality data.

The final cause of error reported by a majority of the HIS interviewers was that maps are inadequate. Having inadequate maps has been a concern of the field staff for years. This complaint is generally confined to the maps used in area segments, which comprise approximately 25% of the sample on current surveys. However, the bulk of the HIS sample consists of block segments, which are treated like area segments. Map complaints generally deal with difficulty determining segment boundaries, especially when a physical feature has changed, or when a political boundary, such as a city limit, is used as a boundary. Also of concern is the amount of detail provided on the maps (roadnames clearly marked, landmarks indicated) and the age of the maps. Additionally, the adequacy of the hand-drawn maps, done by the field staff to pinpoint all housing units in a defined area, is variable.

National Crime Survey

On the National Crime Survey (NCS) three of the major contributors to survey error were reported. NCS households are interviewed 7 times over the course of 3 years (every 6 months). The average length of interview is 20 minutes. The first problem cited was that the questions were too repetitious. After some initial questions are asked of a household respondent, each household member 12 years of age or older is asked a set of screening questions to determine whether any crime incidents had been experienced. It is this set of questions that are perceived to be repetitious. Interestingly, a new questionnaire has

been developed on NCS by an independent research organization under contract to the sponsor and is currently being tested. This new questionnaire expands the set of screening questions. Preliminary results of the testing indicate that this expanded set of screening questions, which could be seen as even more repetitious than the original set of questions, may produce more reporting of crime incidents, thus at least partially correcting the underreporting of crimes that internal analysis shows is present in the current questionnaire.

The second problem noted was that interviews often must be rushed. This was consistently reported as a problem for surveys whose average interview length was greater than 15 minutes.

The final problem dealt with respondent conditioning effects on subsequent visits. Again, this is consistent with the conclusion that respondent conditioning effects occur when the same set of questions are administered at each visit, and the interview is considered lengthy.

Current Population Survey and Survey of Construction

The other two surveys included in the study had no instances of a survey specific cause of nonsampling error being reported by a majority of the staff working on these surveys. These surveys were the Current Population Survey (CPS) and the Survey of Construction (SOC). The Current Population Survey is the Census Bureau's largest ongoing survey, with approximately 1500 interviewers contacting 67,300 households each month. Each household is interviewed 8 times over the course of 16 months. The average interview lasts about 10-15 minutes. The Survey of Construction is much smaller in scope. This survey has two basic parts - sampling of new buildings and subsequent interviews with owners or builders of those buildings. In that respect it is unlike the other Census surveys included in the study that rely on household data collection.

4. Conclusions

We can draw several conclusions from the results of this survey. Among the more important of these is that those surveyed believe that one of the major causes of nonsampling error is the public's lack of knowledge about all the different kinds of work that we do and the importance of it. We need to pursue ways to sell the importance of our surveys to our respondents, so they will be willing to give their time freely and answer the questions we ask to the best of their ability. This must include, but not be limited to, increasing general public awareness of the full range of activities of the Census Bureau. Those surveyed also believe that long interview surveys, particularly those that return to the same address for several interviews and seem repetitive in nature, are subject to a high degree of nonsampling error. In this regard, dependent interviewing, utilizing information gathered in previous interviews, should be examined for feasibility on surveys which gather similar information each visit. This is true especially for items that are perceived to have relatively little change over time. Questions should be written to be short, simple and direct, and questionnaire designers should be concerned about the ability of the elderly and the less educated to understand the language of the questions. We need to try to avoid situations where respondents learn that positive responses lead to additional questions. Finally, with the increasing need to conduct interviews during the evening and Saturday hours, we must ensure that there is enough time in the survey period to actually complete an assignment.

The speculations regarding possible sources of survey error provided herein should guide the Census Bureau into more carefully constructed statistical research to measure the amount of error attributable to the various sources. Additionally, those persons who design survey instruments and procedures should carefully consider these results.

REFERENCES

1. Hansen, H.H., Hurwitz, W.N., and Bershad, M.A., (1961): Measurement Errors in Censuses and Surveys. Bulletin of the International Statistical Institute, 38, pp. 359-374.
2. Brooks, C.A. and Bailar, B.A. (1978): An Error Profile: Employment as Measured by the Current Population Survey. Statistical Policy Working Paper 3, U.S. Department of Commerce, Washington, D.C.
3. Staff, Bureau of the Census, (1974); Indexes to Survey Methodology Literature, Technical Paper 34.
4. Keane, John G. (1986): Have We Erred in Nonsampling Error? Proceedings from the Bureau of the Census' Second Annual Research Conference, pp. 183-189.

APPENDIX 1

EVALUATION OF PROBLEMS WITH INDIVIDUAL SURVEYS

[Question 7: Please answer questions a through z for each survey you have worked on during the past 6 months. Based on your experience with the

survey, circle Y or N beside each statement to indicate if you feel it is a problem for that survey, in terms of causing errors and reducing the quality of the data. Please answer questions a through z for one survey before going on to the next survey.]

	HIS n=94	CPS n=538	NCS n=241	SIPP n=176	CE n=239	SOC n=118
a. The questionnaire is too long.	67%	4%	30%	51%	77%	4%
b. Skip patterns are too complex and difficult to follow.	8%	4%	14%	11%	28%	1%
c. Questions are repetitious.	45%	23%	65%	54%	48%	3%
d. Questions are wordy, or use language that is too complex.	52%	11%	31%	38%	74%	14%
e. The reasons why the information is needed are not clear (to either you or the respondent).	21%	16%	18%	39%	20%	10%
f. Respondents often do not know the answers to the questions.	23%	17%	10%	51%	52%	32%
g. Some questions are insulting or offensive to respondents.	13%	26%	14%	37%	24%	10%
h. Sometimes it is necessary to reword questions, or not ask all of them, to avoid a breakoff or refusal.	43%	43%	39%	56%	55%	26%
i. Interviews often must be rushed.	78%	44%	51%	66%	79%	42%
j. Respondents do not trust the government.	56%	59%	44%	65%	56%	40%
k. Respondents do not want to give complete information about who is living in the household.	27%	31%	29%	34%	22%	0%
l. Respondents knowingly withhold information or report incomplete information.	23%	39%	38%	64%	55%	28%
m. Respondents exaggerate or make up information.	10%	18%	19%	22%	20%	14%
n. In subsequent interviews, respondents learn not to give certain answers, in order to shorten the interview.	10%	18%	51%	35%	59%	9%
o. Public awareness of Census Bureau surveys is low.	94%	88%	87%	90%	92%	81%
p. The interviewer's manual does not provide enough guidance in resolving problem situations.	18%	14%	12%	20%	37%	23%
q. Sample units are difficult to locate.	7%	9%	13%	6%	4%	13%
r. Listing/coverage definitions or rules are not clear.	11%	15%	17%	10%	10%	1%
s. Listing procedures are not easy to understand and follow.	16%	20%	18%	15%	8%	3%
t. Maps are inadequate.	61%	39%	39%	31%	34%	27%
u. There is not enough time to complete the assignments.	55%	19%	8%	9%	10%	16%
v. There is too much pressure to maintain a certain response rate.	70%	58%	44%	49%	58%	13%
w. There is too much change and substitutions of interviewer assignments.	7%	23%	12%	17%	18%	1%
x. Procedures vary too much from survey to survey.	24%	23%	28%	29%	24%	6%
y. Survey procedures are changed without really informing interviewers.	7%	11%	8%	8%	5%	4%
z. Training does not adequately explain the purpose, concepts, and procedures of the survey.	13%	14%	12%	17%	12%	14%

n=number of field staff working on the specific survey