MEASURING RESPONSE VARIANCE IN CENSUS 2000

J. Joy Sharp, John C. Kindelberger, Bureau of the Census and John M. Bushery, Bureau of Transportation Statistics
J. Joy Sharp, Bureau of the Census, Room 3484/3, Washington D.C. 20233

Key Words: Data Reliability, Data Quality, Reinterview

I. Background

Since 1950, the Census Bureau has conducted a Content Reinterview Survey (CRS) to evaluate the quality of data collected in the decennial census of population and housing. The evaluation of the quality of data collected is important for both data users and census planners. Knowledge of the accuracy and reliability of the data help users make informed decisions about how errors may affect the conclusions they draw from the data. Census planners require similar information to develop and test methods to improve the overall quality of the data produced in future censuses. The quality of the data collected in the census is critical for decisions made by government agencies, policy-makers, social science researchers, and the populace.

The methods used to collect and process census data are complex and subject to error. Response error arises from the erroneous or unreliable reporting of characteristics. This error may introduce bias into the estimate of the population parameter, create variability in the classification of an element over repeated trials, or distort the relationships among the variables. Contributors to response error include, but are not limited to: questionnaire design; interview mode; question wording; inadequate instruction (for enumerators and respondents); interviewer effects; and deliberate falsification by the respondent or interviewer. Sources of procedural error (e.g., data capture and processing errors) can also generate response error.

The CRS has always had two main objectives. First, it has been used to estimate response error for items on the census long form. Second, the reinterview has been used to make historical comparisons to previous studies of census content error.

The CRS of Census 2000 may be the last decennial Content Reinterview Survey conducted. The Census Bureau plans to replace the data collected by the census long form with data collected by the American Community Survey . The Census Bureau expects to fully implement the American Community Survey by 2003, conducting interviews in every county of the United States. The American Community Survey is a monthly survey representing a new approach to collect accurate and timely information the government needs for program and policy decisions. This new approach provides up-to-date profiles of the U.S. population every year. The American Community Survey will provide estimates of demographic, housing, social, and economic characteristics every year for all states, as well as for all cities, counties, metropolitan areas, and population groups of 65,000 people or more.

II. Measures of Response Error

The CRS for Census 2000 aims to measure simple response variance (SRV). SRV is one of the components of mean square error (MSE) in the model for survey error (Forsman and Schreiner, 1991). Another term for SRV is "test-retest reliability." To measure SRV, the reinterview will re-ask virtually the same set of questions and replicate, to the extent possible, the original survey procedures and conditions.

The error model for SRV assumes that if the census could be applied to each unit of the population over repeated trials (t), each unit's response to a census question would consist of the true value plus a random error component. This error component would be drawn from a distribution of errors characteristic of that unit. For any unit, we can write this model as:

$$Response(t) = True\ value + Error(t).$$

Each unit possesses a *True value*, but responds to the t-th application (or trial) of the census with some error, *Error(t)*, yielding *Response(t)*. Census 2000 and the CRS consist of two trials with responses *Response(1)* and *Response(2)*, respectively. This model assumes that *Error(1)* and *Error(2)* come from the same distribution and have the same expected values and variances because the CRS replicates the census. The model also assumes that the errors are not correlated between the census and the CRS.

We use these two responses from a subsample of the long form census to estimate the SRV by,

---

This paper reports the results of research and analysis undertaken by Census Bureau Staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress.

$SRV(hat) = 1/n \Sigma_{i=1} \frac{1}{2} (y_{i1} - y_{i2})^2$, where:

n = the number of units in the CRS,

y_{i1} = *Response(1)* for the i-th unit selected for the CRS,

y_{i2} = *Response(2)* for the i-th unit.

In categorical data the estimate of SRV equals half of the "gross difference rate" (GDR). The GDR is the proportion of units whose responses of in-category or out-of-category status change between the census and the CRS (Forsman and Schreiner, 1991).

The Census Bureau also uses a relative measure of SRV, the "index of inconsistency." The index is the ratio of the SRV to the sum of the sampling variance and the SRV (U.S. Census Bureau, 1985). The Census Bureau computes the index of inconsistency for dichotomous data as,

$I = GDR \div (P_1 \cdot Q_2 + P_2 \cdot Q_1)$, where:

P_1 = the proportion in category in trial 1, the census,

P_2 = the proportion in category in trial 2, the CRS,

Q_1 and Q_2 are the corresponding proportions not in category.

The psychometric literature defines the reliability coefficient as $\rho(y_{i1},y_{i2})$ (Crocker and Algina, 1986). When the data perfectly satisfy the reinterview model assumptions of independent errors and replication, the index and the reliability coefficient are related as $I = 1 - \rho(y_{i1},y_{i2})$ (Groves, 1989). The CRS will use $1 - \rho(y_{i1},y_{i2})$ to approximate the index for quantitative variables.

The CRS analysis will report the following estimates, with 90 percent confidence intervals, for each categorical variable:

- Gross difference rate for each category
- Index of inconsistency for each category
- Net difference rate (differences in proportions between census and CRS) for each category
- Aggregate gross difference rate
- Aggregate index of inconsistency

The Census Bureau also will conduct more detailed reliability analyses of particularly important or potentially problematic questions.

III. Methodology

Household Sampling

The universe for the CRS is long form addresses on the Census Bureau's Decennial Master Address File (DMAF) as it existed in January 2000, with certain restrictions. The CRS is limited to units in the fifty states and the District of Columbia that were eligible to mail back a questionnaire, excluding units in block clusters selected for the Accuracy and Coverage Evaluation.

The CRS selected a simple systematic sample of 30,000 units. The sampling interval was the same in each state, so the probability of selection did not vary by state. The probability of a long form household being selected was about 1 in 690.

Within-Household Person Sampling

To reduce respondent burden and increase response rates, the 2000 CRS selected one "sample person" from the household to evaluate the person-level items. We asked a knowledgeable household member to provide reinterview answers about the sample person and about the unit's housing characteristics. Each CRS questionnaire contained a pre-printed sampling table to help the interviewer accurately select the sample person within the household. The sampling tables were randomly and independently generated prior to the questionnaire printing process, and then preprinted on each questionnaire. After the census field representative had completed and verified the roster, he or she used the preprinted sampling table to choose the sample person.

Weighting, Estimation, and Matching

Since the fixed sampling rate did not vary by state and no strata were used, the CRS analysis will use unweighted data to produce estimates. However, the sample is not truly self-weighting, since within household probability of selection varies with household size. Therefore, we also will compute response variance estimates using within-household weights and compare these results with the unweighted estimates.

A CRS case is considered "eligible" if the sampled housing unit is not vacant or demolished and the current household occupied the unit on April 1, 2000. A CRS case will be a considered a "match" if the selected sample person's name also appeared on the Census 2000 roster for the sample address. If the sample person's name is not an exact match, or if more than one household member has the sample person's name, we will use demographic variables to enhance the matching process.

Based on response rates, vacancy rates, and match rates from the 1990 CRS and the Current Population Survey, we expect about 18,000 eligible, valid matches for analysis. With 18,000 cases for analysis, estimates of the Gross Difference Rates (percent inconsistent responses) ranging from one percent to 50 percent will have coefficients of variation ranging from roughly 0.075 to 0.00075.

Questionnaire and Mode of Data Collection

The CRS aims to replicate, as closely as possible, the questions and conditions of Census 2000. The Census 2000 mail-back forms and non-response follow-up forms are different in format and, to some extent, question wording. Operational considerations led us to use an interviewer-administered survey, rather than developing both a mail-back form and an interviewer-administered form. The interviewer-administered form is appropriate for both the original mail-back cases and the enumerator-completed cases. About 54 percent of the respondents who received a long form completed and returned it by mail. Census 2000 enumerators interviewed the remaining long form households.

Permanent Census Bureau interviewers used both telephone and personal visit modes to conduct the CRS. The interviewers first attempted to interview a CRS household by telephone, but then used personal visit interviews as a last resort. Interviewers made six telephone attempts to contact a household, on various days of the week and at different times of the day, before planning a visit. Personal visits were also necessary when a working telephone number could not be obtained, or because the household was otherwise not reachable by telephone.

The wording of items on the CRS questionnaire was designed to replicate the wording of the census long form as closely as possible. The CRS preserved the Census 2000 wording with only a few exceptions. For example, some Census 2000 questions implicitly refered to April 1, that is, the official Census Day. Because the CRS interviews began in late June, the corresponding CRS questions refered to April 1, explicitly.

The CRS omitted a few questions asked in Census 2000. Recall and memory effects caused by the time lag between census and the CRS made replication impossible and would confound the analysis. The omitted questions included those asking about the person's work and commuting "last week."

Finally, the CRS added several items to determine who provided the CRS and census responses. This information will help us analyze the effect that different respondents in the two interviews have on the census-CRS discrepancies.

The Census Bureau used DocuPrint technology to print the CRS questionnaire, including sample address contact information. DocuPrinting allowed us to include household-specific information on each questionnaire during printing, and also offered a much quicker turnaround in printing questionnaires after sample units had been enumerated by the census.

We printed the CRS questionnaires only after the Census 2000 control system recorded that the census form was checked-in and data were captured from the form. The Census Bureau's National Processing Center printed questionnaires on a flow, approximately every two to four weeks, and sent them to the Census Bureau's 12 regional offices who distributed them to the appropriate field interviewers.

Data Capture and Processing

Data collected on the CRS questionnaire were captured and processed using procedures similar to those used for the census forms. Regional offices sent completed questionnaires to the Census Bureau's National Processing Center where the data were captured through imaging and scanning operations. Questionnaires were first sent through an imaging process whereby "pictures" of each page of the questionnaire were taken. Processing staff then used optical readers to capture marked category boxes as well as written text from these images. When questionnaires were damaged or otherwise difficult to scan, clerical staff keyed the responses directly from the image or, if necessary, the paper form.

Preparation of Matched Analysis File

Following the completion of the data capture process, we will match the CRS household records to unedited census long form records. Then we will obtain the CRS sample person's record from the census household. The first step will be to verify the match of CRS and census households, and then to match the sample person for CRS with the same person listed on the census long form roster.

IV. Differences Between the 1990 and 2000 CRS

Technological advances allowed rapid access to Census 2000 data, permitting more timely turnaround of CRS cases. Other improvements over earlier content reinterviews follow.

Better estimates of question reliability -- Reduced turnaround between the census and the CRS should result in more accurate estimates of response variance. We reduced the time lag from the census enumeration to the CRS interviewer receiving the CRS questionnaire from as much as nine months (in 1990) to approximately two to three months. This improvement should substantially reduce memory effects, real changes in status, census respondents moving out of the unit, and other confounding factors associated with lengthy delays between the original interview and reinterview.

Evaluation of more questions -- the CRS questionnaire re-asked almost every item that was asked on the census long form. Only a few time-sensitive questions were omitted from the CRS. The 2000 CRS is the only content reinterview to evaluate so many of the long form questions.

Reduced respondent burden -- The CRS collected person-level data for only one person per household. This change significantly reduced the reporting burden on the public. We also expect it improved the CRS response rate. DocuPrint technology allowed us to print an independently generated sampling table on each questionnaire. This table helped to minimize field errors in selecting the CRS sample person. Despite re-asking almost all the census questions, substantial burden savings remained because we collected the data from only one person per household.

V. Limitations

The CRS measures simple response variance only. Earlier content reinterviews attempted to measure bias in some of the long form questions. We decided not to attempt bias estimation because the model assumptions are much harder to satisfy. To estimate bias, we would have to assume that the CRS answer is "truth".

To measure simple response variance accurately the CRS must satisfy the reinterview model assumptions of replication and independence. However, due to the operational constraints and budget and resource limitations, the CRS does not perfectly satisfy these assumptions. The CRS does not fully replicate the census conditions because of differences in mode, respondent burden, and memory effects.

Differences in interviewers and interview mode -- Permanent Census Bureau interviewers conducted the CRS, either by telephone or in person. Census responses came from mail returned questionnaires or interviews conducted by temporary Census 2000 enumerators. The effect of mode differences on estimates of census SRV is likely strongest among the 54 percent of Census 2000 long form questionnaires returned by mail. Research suggests that self-enumeration by mail yields more reliable responses than data collected by telephone (Bushery, et al., 1996). If this relationship holds for mail versus personal visit interviews, the CRS will overstate the SRV to some extent.

Mode difference also may affect the CRS data collected by the telephone versus the census data collected by personal visit. We anticipate that the data collected by permanent Census Bureau interviewers will be more reliable than the census data collected by temporary enumerators. This effect would generate a tendency to understate the census SRV.

Differences in respondent burden -- Respondent burden in the CRS was lower than in Census 2000 because the CRS asked slightly fewer questions and collected person-level data from only one person per household. The average burden estimate for the CRS interview was 20 minutes, compared to 38 minutes for the census long form. This lowered burden may cause the CRS responses to be more reliable than Census 2000 responses. However, because the CRS is another contact, respondent resistance may be higher during the CRS. Respondents may be more likely to "satisfice"- provide a "top of the head" answer just to get the interview over with (Krosnick 1990).

Differential memory effects -- The two-month lag between the CRS and Census 2000 interviews probably caused more severe memory effects in the CRS for items referring to April 1. These effects would cause a tendency to overstate the Census SRV.

Wording effects -- The CRS revised several census questions to account for collecting person-level data from only one household member. The enumerator questionnaire for the Census 2000 used a topic-based approach, asking the basic demographic questions on age, race, sex, and Hispanic origin for all household members at the same time. Because the CRS collected responses for only one person, the topic-based approach was not applicable. And, as mentioned earlier, the CRS explicitly anchored all questions to April 1.

Independent errors -- The CRS also may not fully satisfy the assumption of independence of errors between Census and CRS. It is possible that respondents remembered their original responses. However, the two- to three-month lag between the Census and the CRS most likely was long enough that the respondents were able to answer the CRS independently of their census responses.

VI. Schedule

Interviewing for the Content Reinterview Survey began on June 26, 2000 and continued until November 6, 2000. Cases were sent to field staff on a flow basis, corresponding to the completion of census enumeration. Key dates for activities follow on the next page.

CRS ACTIVITY SCHEDULE

Activity	Date
Census forms mailed to households	March 13-15
Official Census Day	April 1
Conduct census nonresponse activities	April 27- July 7
Conduct CRS interviews	June 26 - Nov. 6
Conduct match operation of census and CRS data	March - April, 2001
Perform data analysis	April - Oct. 2001
Prepare final report	December 2001

References

Bushery, J., Brick, J. Severynse, J. and McGuinness, R. (1996), "How Interview Mode Affects Data Reliability," Joint Statistical Meetings.

Crocker, L. and Algina, J. (1986). *Introduction to Classical and Modern Test Theory*, Holt, Rinehart, and Winston, Inc.

Forsman, G. and Schreiner, I. (1991). "The design and analysis of reinterview: an overview," in P. Biemer et al. (Ed.). *Measurement Errors in Surveys*, John Wiley and Sons, Inc.

Groves, R. (1989). *Survey Errors and Survey Costs*, John Wiley and Sons, Inc.

Krosnick, J.A., "The Impact of Satisficing on Survey Data Quality," Proceedings of the 1990 Annual Research Conference, U.S. Bureau of the Census, March, 1990, pp. 835-845.

U.S. Census Bureau (1985). *Evaluating Census of Population and Housing*, Statistical Training Document ISP-TR-5.