

USING CATI IN THE CURRENT POPULATION SURVEY

J.M. Bushery, L. Weidman, S.M. Briley, S.Ciochetto, U.S. Census Bureau¹
John M. Bushery, U.S. Bureau of the Census, Washington, D.C. 20233

1. INTRODUCTION

Computer assisted telephone interviewing (CATI) has been used by market researchers and academic survey researchers for almost 15 years (Fink 1983). Since its beginnings, CATI has grown rapidly, displaying both advantages and limitations relative to paper and pencil interviewing (Rustemeyer et al. 1978) (Nicholls 1978) (Nicholls 1983) (Groves 1983) (Sudman 1983) (Groves and Mathiowetz 1984) (Nicholls and Groves 1986). The advantages include more efficient survey management, improved data quality, improved interviewer training, reduced data processing time, and more flexibility in questionnaire design.

Governmental agencies now investigating or using CATI include the U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS), the U.S. Bureau of Labor Statistics, the Centers for Disease Control, the U.S. Census Bureau, Statistics Sweden, and Statistics Canada (Nicholls and Groves 1986).

The Census Bureau's interest in CATI began in the late 1970s (Rustemeyer 1978). Census is especially attracted by CATI's promise of improved data quality. However, an important concern is the effect that a change to CATI would have on estimates and established data series. Of the numerous studies carried out comparing CATI with non-CATI methodologies some have been carefully controlled to compare the estimates generated by the competing methodologies, for example, (House 1984) (Coulter 1985) (Tortora 1985). However, there has been little research to evaluate the effects of CATI in a rotating panel survey like the Census Bureau's Current Population Survey (CPS).

Census carried out two tests of CATI in 1982. These were imbedded in the telephone follow-up operations for the 1982 Survey of Scientists and Engineers (Ferrari 1984) and the 1982 Census of Agriculture (Ferrari 1986). In both these studies, CATI was evaluated against a "hard copy" centralized telephone interviewing procedure. Neither test actually compared estimates obtained from the two methodologies, but the Census of Agriculture test compared distributions of completed interviews across categories within several variables and found no significant treatment differences.

In 1985 the Census Bureau opened a 40 station CATI facility in Hagerstown, Maryland. Since then Census has developed and tested CATI systems for the CPS, the National Crime Survey (NCS), and other surveys. In November 1986 Census began a controlled experiment to evaluate estimates and data quality measures under a CATI system, as it would most likely be implemented in the CPS, against the current CPS methodology. Enough sample has been selected to carry out the study from November 1986 through December 1988. One of the major objectives of the study is to estimate the extent that labor force estimates produced by CATI will differ from current CPS estimates.

2. DESIGN

The CATI sample was designed to provide information about the most likely implementation of CATI in the CPS. Our best guess was that CATI would be implemented first in large metropolitan areas with workloads large enough to require more than one interviewer. To approximate this condition, we limited eligibility for the CATI study to CPS self-representing (SR) Metropolitan Statistical Areas (MSAs) having a sample size of 70 or more households.

These requirements yielded a frame of 109 metropolitan areas. From these we selected seven areas with certainty. From the remaining 102 areas we selected 23 with probability proportional to population size.

The total sample size allowed by the CATI budget is 3,000 housing units (HUs) designated for interview each month, 100 in each area. The control sample in these areas, a subset of the CPS production sample, varied from 100 to 1200 HUs. The CATI, or Test group, sample is selected from clusters of HUs neighboring the CPS sample clusters, minimizing differences between the Test and the Control sample designs.

Cost and operational considerations required some compromises in designing the CATI sample. However, the resulting differences between the CATI and production CPS samples are relatively minor. To maintain comparability, the Control sample was adjusted accordingly.

A production implementation of CATI in the CPS might be called a "mixed-mode" CATI system, a system using a mixture of methodologies -- personal visit, CATI, and decentralized non-CATI telephone interviews. The incoming rotation group, always is interviewed in person by the regular Field staff. Of the sample in the remaining three rotation groups, only about 75 percent are assigned to the CATI facility in Hagerstown, Maryland. The remainder will be retained by the Field offices as unsuited for CATI (for example, no telephone). About nine percent of the Hagerstown cases are "recycled" to the field because contact could not be made from Hagerstown before the survey deadline.

3. WEIGHTING AND ESTIMATION

The procedures used to produce weighted estimates for the CATI study are a simplified version of the regular CPS weighting procedures. Sample data are inflated to account for the probability of selection, and a simple adjustment is performed to account for eligible HUs that were not interviewed. The estimates in this paper refer to all 30 CATI sample areas combined, not to the frame of 109 areas.

We estimated variances using a replication technique, the stratified Jackknife (White 1986) (Krewski and Rao 1981). This replication technique also was used to perform the log linear analyses discussed in the next section.

4. ANALYSIS PROCEDURES

Our evaluation made use of two techniques, log linear analysis and direct comparisons (t-tests), to evaluate labor force estimates and coverage under CATI. Software developed at the Census Bureau allowed us to perform log linear analyses despite the complex sample design of the CATI study (Fay 1982, 1983, 1985). This software uses replication techniques to account for the design features and weighting commonly used in multi-stage samples.

Log Linear Models

Hierarchical log linear modeling was used to test for differences in labor force and demographic distributions for the two methodologies. The data from the CATI study was put into a cross-classification consisting of six discrete variables:

- M Methodology: Test, Control
- L Labor Force Status: Employed, Unemployed, Not in Labor Force
- S Sex
- H Hispanic Ethnicity: Hispanic, Not Hispanic
- R Race: Black, White, Other
- A Age: 16-19, 20-24, 25+

This method of analysis provides a means for testing for the existence of multi-way interactions. In this study, we are interested in interactions between methodology and the labor force and demographic variables. Tests for two- and three- way interactions involving methodology provide useful information about differential coverage and reporting among subpopulations between the Test and Control methodologies.

Consider the arbitrary model:

$$\ln(F_{ijkl}) = u + u_i^I + u_j^J + u_k^K + u_l^L + u_{ij}^{IJ}$$

We use the shorthand notation [IJ][K][L] to describe this model. The term [IJ] represents the two-way u_{ij}^{IJ} parameters and the u_i^I and u_j^J parameters as well. The [K] and [L] terms represent the u_k^K and u_l^L parameters, respectively. We can test for the [IJ] interaction by comparing the Likelihood Ratio Test statistic for the first model, $G^2(\text{model 1})$, with the corresponding statistic for a second, nested model, [I][J][K][L]. The difference $G^2(\text{model 1}) - G^2(\text{model 2})$ can be compared with the critical values for a Chi-Squared variate with $k(1) - k(2)$ degrees of freedom to determine whether the interaction represented by the u_{ij}^{IJ} parameters is statistically significant. Testing for three-way interactions can be performed in an analogous manner. We used the basic model [LSHRA][ML][MS][MH][MR][MA] in our hypothesis testing.

For the examples given above, simple random sampling was assumed. There are a number of books on log linear modeling under simple random sampling, (Feinberg 1979) (Bishop, Fienberg, and Holland 1975) (Goodman 1977). However as noted earlier, the complex sample design of the CATI study requires a theory that accounts for clustered sampling and weighted estimates. The software available at the Census Bureau allows for log linear analyses that proceed completely analogously to those for simple random sampling assumptions.

5. FINDINGS

The data presented here were collected and averaged from November 1986 through June 1987, about one-third the planned duration of the CATI study. The findings discussed below should be viewed as preliminary, and as suggesting relationships between the Test and the Control methodology, rather than clearly demonstrating that such relationships exist. Although the data suggest that some differences may exist between the Test and the Control, these results do not mean that CATI is not viable for the CPS. At present the data provide only weak evidence of differences in labor force estimates. At this time our results simply underscore the need for caution when considering a major methodological change in the survey, as would be involved in implementing CATI on a large scale.

Coverage and Labor Force Estimates

There is some evidence of an overall coverage difference between the Test and the Control methodology. The Test estimate of 25,119,000 eligible housing units (HUs) is about 4.1% lower than the corresponding Control estimate of 26,192,000 HUs. The total population estimate in the Test group, 51,122,000 persons, is 4.6% lower than the Control group's estimate of 53,617,000 (alpha = 0.10 for both comparisons). There is no evidence that within household coverage, based on the estimates of persons per HU, differs by methodology. The 2.04 persons per HU in the Test group and the 2.05 in the Control were not significantly different.

Generally it is acknowledged that the CPS does not provide complete coverage of all eligible units (or persons), so the lower coverage exhibited by the Test methodology is a cause for concern. At this time we are unable to explain the coverage differences to our satisfaction. We performed some exploratory analyses which suggest that the Test group's lower housing estimates may be caused by higher vacancy rates. We are conducting a special reinterview of vacant units in the Test group to examine this possibility more closely. Also, analysis is underway to determine whether the apparent undercoverage occurs primarily during the first month-in-sample (MIS) interview or in the later MIS interviews.

We tested for two-way interactions between methodology and each of the four major demographic variables to determine whether the Test methodology exhibited differential coverage across these broad demographic groups, relative to the Control, but found not even weak evidence of statistical significance. Nor did the two-way methodology-labor force interaction test statistically significant. Only one interaction, the three-way interaction among methodology, labor force and sex, [M L S], was statistically significant (alpha < 0.10). This suggests that the distribution of persons by labor force and sex for the Test group differs from that for the Control.

Two of the primary statistics produced by the CPS are the civilian labor force participation rate, or CLF rate, and the unemployment rate. Table 1 compares these rates by methodology for the major demographic groups in the study. There

is no evidence that the CLF rate is reported differently under the two methodologies. Nor is there evidence except for women, that the unemployment rate is reported differently by methodology. The data provide some evidence that women report a higher unemployment rate in the Test group ($\alpha = 0.10$). The evidence for this difference is weak, since it is the only comparison to prove significant, among several that were tested. The problems involved in testing numerous comparisons are discussed more completely in the section **LIMITATIONS OF THE DATA**.

There also is evidence that fewer Not in Labor Force (NILF) persons and fewer NILF males are reported in the Test group than in the Control

($\alpha < 0.05$). Finally, the estimated number of males reported in the CLF is lower for the Test group ($\alpha < 0.10$). We suspect the lower Test group NILF and CLF estimates result from that group's lower population estimates, rather than from differential reporting of labor force status. This suspicion is supported by the lack of evidence that CLF rates differ between treatments.

Noninterview Rates

About one tenth of all addresses assigned to be interviewed are found to be vacant units or otherwise out of scope: type B noninterviews. Another one percent turn out not to be housing

TABLE 1. CPS CATI PHASE II LABOR FORCE ESTIMATES TEST GROUP VERSUS CONTROL GROUP
NOVEMBER 1986 TO JUNE 1987

NUMBERS IN 1000s	TOTAL PERSONS 16+	CIVILIAN LABOR FORCE	EMPLOYED	UNEM- PLOYED	NOT IN LABOR FORCE	CLF RATE	UE RATE
TOTAL PERSONS							
TEST GROUP	51,122	34,105	31,803	2,303	17,016	66.71	6.75
CONTROL GROUP	53,617	35,572	33,311	2,262	18,045	66.35	6.36
DIFFERENCE	-2,496*	-1,467	-1,508	41	-1,028**	0.37	0.39
S.E. (DIFF.)	1,426	1,161	1,119	172	497	0.78	0.48
MALE							
TEST GROUP	23,606	18,377	17,217	1,159	5,229	77.85	6.31
CONTROL GROUP	25,154	19,417	18,139	1,278	5,737	77.19	6.58
DIFFERENCE	-1,548**	-1,040*	-922	-119	-508**	0.65	-0.27
S.E. (DIFF.)	654	592	576	87	200	0.78	0.46
FEMALE							
TEST GROUP	27,516	15,729	14,585	1,143	11,788	57.16	7.27
CONTROL GROUP	28,464	16,155	15,172	983	12,308	56.76	6.09
DIFFERENCE	-947	-427	-587	160	-521	0.40	1.18*
S.E. (DIFF.)	867	674	646	116	383	1.11	0.70
WHITE							
TEST GROUP	42,449	28,540	27,095	1,444	13,909	67.23	5.06
CONTROL GROUP	44,474	29,684	28,139	1,545	14,791	66.74	5.20
DIFFERENCE	-2,026	-1,144	-1,044	-101	-881	0.49	-0.14
S.E. (DIFF.)	1,581	1,245	1,183	134	549	0.93	0.40
MALE							
TEST GROUP	19,970	15,725	14,937	788	4,245	78.74	5.01
CONTROL GROUP	21,151	16,544	15,622	921	4,607	78.22	5.57
DIFFERENCE	-1,181	-819	-685	-134*	-362	0.52	-0.56
S.E. (DIFF.)	761	661	635	70	223	0.92	0.40
FEMALE							
TEST GROUP	22,478	12,814	12,158	656	9,664	57.01	5.12
CONTROL GROUP	23,323	13,140	12,517	623	10,183	56.34	4.74
DIFFERENCE	-845	-326	-358	33	-520	0.67	0.38
S.E. (DIFF.)	896	674	641	83	395	1.31	0.58
BLACK							
TEST GROUP	6,845	4,319	3,572	747	2,525	63.10	17.29
CONTROL GROUP	7,025	4,445	3,827	618	2,580	63.27	13.90
DIFFERENCE	-181	-126	-255	129	-55	-0.17	3.40
S.E. (DIFF.)	799	539	422	149	320	2.24	2.07
MALE							
TEST GROUP	2,740	1,988	1,665	322	752	72.55	16.22
CONTROL GROUP	3,004	2,115	1,820	295	890	70.39	13.95
DIFFERENCE	-264	-127	-154	27	-138	2.16	2.27
S.E. (DIFF.)	351	257	210	66	134	3.01	2.15
FEMALE							
TEST GROUP	4,105	2,331	1,907	425	1,773	56.80	18.21
CONTROL GROUP	4,021	2,330	2,008	323	1,691	57.95	13.85
DIFFERENCE	84	1	-101	102	83	-1.16	4.36
S.E. (DIFF.)	486	308	237	95	230	2.81	2.67

* Significant at $\alpha < 0.10$

** Significant at $\alpha < 0.05$

units at all: type C noninterviews. Among eligible housing units (HUs), those for which an interview can't be completed are called type A noninterviews. Average response figures for November 1986 through June 1987 are given below.

November 1986 through June 1987 are given below.

ITEM COMPLETION RATES AND REFUSAL RATES
TEST VERSUS CONTROL
AVERAGE: NOVEMBER 1986 - JUNE 1987

CPS CATI PHASE II NONINTERVIEW DATA				
TEST VS CONTROL				
AVERAGE: NOVEMBER 1986 TO JUNE 1987				
	TEST	CONTROL	DIFFERENCE	S.E.
HUs ASSIGNED	2964	9310	-	-
HUs ELIGIBLE	2586	8232	-	-
TYPE A RATE	5.14	5.50	-0.36	0.27
REFUSAL RATE	3.09	3.10	-0.01	0.22
TYPE B/C RATE	12.79	11.59	1.20	0.97
TYPE B RATE	11.81	10.46	1.35	0.96

	COMPLETION RATE		CONTROL	TEST MINUS CONTROL DIFFERENCE	SE
20A	98.29	97.38	99.77	-1.48 ****	0.13
20C	97.10	96.13	98.75	-1.65 ****	0.35
22A	98.57	97.78	99.49	-0.92 **	0.46
22E	98.50	98.35	98.95	-0.15	0.42
25C	78.64	78.28	86.72	-8.08 ****	1.21
25C ¹	7.97	5.96	10.63	-2.66 ***	0.86

¹ Refusal rate.

** Significant at alpha < 0.05

*** Significant at alpha < 0.01

**** Significant at alpha < 0.001

The data provide no evidence, that type A, B, or C noninterviews or refusals occur at different rates between treatment groups, at alpha < 0.10.

Item Completion Rates

Higher item completion rates usually reflect higher data quality. Unanswered items must be imputed, which is unlikely to be as accurate as the data obtained from the respondent. We compared completion rates for the key CPS questionnaire items listed below. Items 20A, 20C, 22A, and 22E are used in determining labor force status, with items 22A and 22E particularly important.

Item	Description
20 A	hours worked
20 C	reason worked < 35 hours
22 A	doing what to find work
22 E	could have taken job if offered?
25 C	earnings per hour (and refusal rate)

Item 25C is subject to low completion rates, because it deals with a difficult topic, earnings. The household respondent may not have precise information on the earnings of other household members, or may be unwilling to provide this information. Because of these difficulties, refusals may be recorded for item 25C without penalty to the interviewer in the regular CPS.

The completion rate for an item is the percent of cases receiving an entry from the interviewer and requiring no allocation, among all cases requiring a response to the item. An item not completed is considered "blank." The CPS field interviewers' performance ratings are hurt by blank items. In a regular field interview, blank items occur for the following reasons: interviewer skip pattern error; respondent unable to provide the information (Don't Know - DK); respondent refuses to provide the information.

Blank items occur in a CATI interview only for "Don't Know" (DK) responses or refusals. Barring software failures, the computer assures that the skip patterns are followed correctly. The refusal rate for an item is simply the percent of cases requiring a response, which the respondent refused to answer. Average completion rates for

One might expect CATI to produce higher completion rates by eliminating skip pattern refusals. On the other hand, by providing DK and refusal options for every item, CATI reduces the motivation for an interviewer to probe for an answer, which would tend to reduce the CATI completion rate.

As can be seen there is fairly strong evidence that the Test procedure results in lower item completion rates than the Control. Items 20A, 20C, and 25C all have fewer completions in the Test group (alpha < 0.01). This leads us to infer that CATI interviewers are taking advantage of the DK and refusal options. There is no evidence that Test group completion rate for the crucial item 22E differs from the Control group rate. The Test group refusal rate for item 25C is significantly lower than the Control's, but this is most likely because the CATI interviewers are making more use of the DK option.

We display the Test group completion data separately for cases interviewed by CATI from Hagerstown (HTC). The following observations are without the benefit of statistical testing. The HTC cases seem to be the reason for the Test group's lower completion rates. For all items the observed HTC rates are lower than the Test group rates. The observed HTC refusal rate for item 25C is lower than the overall test group rate, probably reflecting the CATI interviewers' ability to use the DK option.

Several features of the CATI system used for this study contributed to the lower Test group completion rates. These features can be changed easily, and weaken these results for predicting how CATI will perform in production in the CPS. A primary feature is that the CATI case management system did not allow callbacks to obtain missing data, except for mid-interview break offs. Field interviewers, on the other hand, are encouraged to make callbacks to reach a knowledgeable respondent. Another important factor is that the CATI interviewers in Hagerstown did not receive the same feedback as the Field interviewers. The Field interviewers receive monthly reports on their completion rates, which are used in their performance ratings. The CATI interviewers did not receive this kind of feedback emphasizing high completion rates. Finally, the CATI instru-

ment might have included explicit probes after DK responses. Such probes would underscore the importance of obtaining complete responses. Steps are being taken to implement explicit probes in the CPS CATI system.

Industry and Occupation Coding

The ability to record responses to open questions is important. To evaluate this, we coded industry and occupation data from CATI for three months, November 1986 through January 1987. Log linear analysis revealed no differences between the Test and the Control group in the distributions of employed persons across the major industry and occupation groups. There also was no evidence that either methodology was more subject to missing industry and occupation codes, but there were too few cases with missing codes for accurate hypothesis testing.

<u>Industry Groups</u>	<u>Occupation Groups</u>
Manufacturing	Managers and Professionals
Wholesale / Retail Trade	Technical, Sales, and Administrative Support
Finance / Services	Operators, Fabricators, and Laborers
Other	Service and Other

Response Variance

Response variance is another important factor in data quality. The usual method of measuring simple response variance is to conduct an independent reinterview using the procedures identical to the original interview. At the present time, the CATI study does not include a reinterview program, so we decided to use the percent of "gross change" in labor force status as a proxy variable for response variance. We use the number of persons in sample in two consecutive months as the base of the gross change percentage. The number of these persons whose labor force status changes from one month to the next is the numerator of the percentage.

In effect, we treated the second month's interview as a substitute for the independent reinterview. Our assumption was that a higher level of response variance in one of the methodologies would inflate the gross change estimate for that treatment. The data, averaged from November 1986 through June 1987, show the Test group's gross change rate, 7.74% is higher than the Control's 6.68% rate ($\alpha < 0.05$).

After seeing the data, we realized that factors other than simple response variance might affect estimates of gross change. For example, interviewer variance may be showing up in a different way, because of the change from field interviewers to CATI interviewers after the first monthly interview. However, the consensus opinion is that the CPS already overstates gross change (Fuller and Chua 1986), so any increase in gross change is likely to represent a decrease in data quality. In any event, this is yet another indication that the Test procedure may produce results different from the current CPS methodology.

Increased response variance, reflected in the greater gross change estimates, might be caused by greater turnover in the respondents actually

contacted at identical households from one month to the next. However, we found no evidence that this proportion varied by treatment.

6. LIMITATIONS OF THE STUDY

Caution should be exercised when using these findings to predict what would happen if CATI were used in the production CPS. The Test and Control group samples are comparable, but both differ slightly from the full CPS design:

- 1) The CATI study includes no HUs built after the 1980 Decennial Census.
- 2) The CATI rotation pattern is simpler than the CPS. Only four rotation groups are used rather than the eight in the CPS. Only the first four MIS interviews are used in the Control group. As a result, we have no information on the effects of CATI on data for MIS 5 and later.
- 3) The Test group cases in this study are identified as different from the regular CPS cases and may be treated differently by the field interviewers. This may accentuate or mask differences which would occur in a production setting.
- 4) Only 30 metropolitan areas are represented in the study. It is possible that different results may occur in a broader implementation of CATI.
- 5) This CATI system only approximates the system which finally will be implemented for production in the CPS. Some changes already have been made to the system. As we learn more, other changes no doubt will be made.

In a production CATI system, most of the interviews which can be completed by telephone will be handled by the CATI facility. Only the initial interviews and more difficult cases will be handled by the field staff, usually in personal visit (PV) interviews. This means the composition of field assignments will shift substantially to PV interviews. Shifting the easier cases to the CATI facility may affect interviewer performance, but we cannot evaluate such an effect in this study.

Sample sizes for this study are fairly small, particularly for the Test group, making the estimates obtained highly variable. The current accumulated sample provides only a 30 percent chance of detecting a real difference of 0.5 percentage points in the overall unemployment rate.

We made a large number of direct comparisons of coverage and labor force estimates in this analysis. On the average, about ten percent of these can be expected to test significant through chance alone at the 0.10 alpha level, in the absence of methodological effects. The number of significant comparisons observed here is close to that expected ten percent.

7. CONCLUSIONS

We found several indications that the mixed-mode implementation of CATI tested in this study may produce lower coverage than current CPS interview methodology. However, there is no conclusive evidence that labor force estimates will be different.

Data quality measures, such as item completion

rates and gross change estimates reveal lower quality in the CATI implementation used in this study. Rates of household noninterviews, on the other hand, show no evidence of treatment differences. Some actions already are underway to improve CATI data quality; explicit probes are being inserted into the CATI instrument.

At this point we feel that CATI should not be considered as perfectly interchangeable with the current CPS interview methodology. These results do not mean that CATI is not viable or desirable for the CPS. Rather, they serve to underscore the principle that when implementing any new methodology we must proceed with caution.

We will continue to collect and analyze CATI data, at least through March 1988. We will continue to refine the precision of our original estimates, and seek explanations for the apparent coverage differences. We also are trying to determine whether the field interviewers are treating CATI cases differently from the regular CPS cases, thus confounding the study.

This paper discusses only data comparability and data quality issues. Other work is being done to evaluate the costs of CATI relative to the current CPS methodology (Bryant and Weidman 1987). Another study is attempting to measure more intangible characteristics, such as the effects CATI will have on field interviewer morale when the more desirable cases are transferred to Hagerstown, and what complications CATI might cause the field offices. That study also will evaluate CATI's effect on interviewer turnover.

REFERENCES

Bishop, Yvonne M., Stephen E. Feinberg, and Paul W. Holland (1975), Discrete Multivariate Analysis, Cambridge, MA: MIT Press.

Coulter, Richard (1985), "A Comparison of CATI and NonCATI on a Nebraska Hog Survey." Staff Report No. 85, Statistical Research Division, Statistical Reporting Service, U.S. Department of Agriculture.

Fay, Robert E. (1982), "Contingency Table analysis for Complex Designs: CPLX," Proceedings of the Section on Survey Research Methods, American Statistical Association, p. 44-53.

Fay, Robert E. (1983), "CPLX -- Contingency Table analysis for Complex Designs, Program Documentation", unpublished report, U.S. Bureau of the Census.

Fay, Robert E. (1985), "A Jackknifed Chi-Squared Test for Complex Samples," Journal of the American Statistical Association, March 1985, Vol. 80, No. 389, p. 148-157.

Feinberg, Stephen E. (1979), The Analysis of Cross-Classified Categorical Data, Cambridge, MA: MIT Press.

Ferrari, Pamela W. (1984), "Preliminary Results from the Evaluation of the CATI Test for the 1982 National Survey of Natural Scientists and Engineers." Unpublished research report, U.S. Bureau of the Census.

Ferrari, Pamela W. (1986), "An Evaluation of Computer Assisted Telephone Interviewing Used During the 1982 Census of Agriculture." Unpublished evaluation report, Agriculture Division, U.S. Bureau of the Census.

Fink, James C. (1983), "CATI's First Decade: The Chilton Experience." Sociological Methods and

Research, Vol. 12, No. 2, p. 153-168.

Fuller, Wayne and Tin Chiu Chua (1986), "Response Error Models for Changes in Variables." Proceedings of the Second Annual Research Conference, U.S. Bureau of the Census, p. 425-441.

Goodman, Leo A. (1970), Analyzing Qualitative/Categorical Data, Cambridge, MA: Abt Associates.

Groves, Robert M. (1983), "Implications of CATI: Costs, Errors, and Organization of Telephone Survey Research." Sociological Methods and Research, Vol. 12, No. 2, p. 199-215.

Groves, Robert M. and Nancy A. Mathiowetz (1984), "Computer Assisted Telephone Interviewing: Effects on Interviewers and Respondents." Public Opinion Quarterly, Vol. 48, p. 356-369.

Groves, Robert M. and William L. Nicholls, II (1986), "The Status of Computer Assisted Telephone Interviewing: Part II - Data Quality Issues." Journal of Official Statistics, Vol. 2, No. 2, p. 117-134.

House, Carol C. (1984), "Computer Assisted Telephone Interviewing on the Cattle Multiple Frame Survey." Staff Report No. 82, Statistical Research Division, Statistical Reporting Service, U.S. Department of Agriculture.

Krewski, D. and J.N.K. Rao (1981), "Inference from Stratified Samples: Properties of the Linearization, Jackknife and Balanced Replicate Methods." The Annals of Statistics, Vol. 9, No. 5, p. 1010-1019.

Nicholls, William L., II (1978), "Experiences with CATI in a Large-Scale Survey." Proceedings of the American Statistical Association, Section on Survey Research Methods, p. 9-17.

Nicholls, William L., II (1983), "CATI Research and Development at the Census Bureau." Sociological Methods and Research, Vol. 12, No. 2, p. 191-197.

Nicholls, II, William L. and Robert M. Groves (1986), "The Status of Computer Assisted Telephone Interviewing: Part I - Introduction and Impact on Cost and Timeliness of Survey Data." Journal of Official Statistics, Vol. 2, No. 2, p. 93-116.

Rustemeyer, Anita, Gerald H. Shure, Miles S. Rogers, and Robert J. Meeker (1978), "Computer Assisted Telephone Interviewing: Design Considerations." Proceedings of the American Statistical Association, Section on Survey Research Methods, p. 1-8.

Sudman, Seymour (1983), "Survey Research and Technological Change." Sociological Methods and Research, Vol. 12, No. 2, p. 217-230.

Tortora, Robert D. (1985), "CATI in an Agricultural Statistical Agency." Journal of Official Statistics, Vol. 1, No. 2, p. 301-314.

White, Glenn D. (1986), "CATI Phase II - Variance Estimation," Unpublished Census Bureau memorandum.

1 This paper reports the preliminary results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.