THE EFFECT OF COMPUTER ASSISTED TELEPHONE INTERVIEWING ON RESPONSE INCONSISTENCY

Carol L. Kosary and John Paul Sommers
Statistical Methods Division
Office of Prices and Living Conditions
Bureau of Labor Statistics

INTRODUCTION

According to Groves and Nicholls (1986) there has been little empirical evidence demonstrating the impact of Computer Assisted Telephone Interviewing (CATI) on data quality to date. The expectations have generally been that CATI will lead to quality improvements in the data collected (Nicholls, 1978; Groves, 1983) but these assumptions have not as yet been demonstrated empirically and evidence of quality improvements may, in fact, prove to be specific to each survey studied. It is felt that the use of CATI, especially centralized CATI, will generally lead to improvements in survey data quality due to its strict enforcement of question wording and survey procedures and it ability to support frequent supervisory monitoring of calls. It should also be noted, though, that improvements in data quality due to this type of rigid standardization of wording and procedures must go hand in hand with other areas of overall survey design, such as the testing of question wording and procedures. CATI has the capability of increasing the reliability of survey data but without care also being taken in the areas of question wording and procedures, especially to any aspects unique to telephone interviewing, CATI cannot guarantee the validity of the data collected.

The following paper presents the first in a series of studies designed to demonstrate the effect of CATI on the data quality of one particular survey, the Consumer Price Index (CPI) Housing Survey. The present CPI Housing Survey was developed as a part of the CPI revision of 1987 and has been used in index calculations since January of 1987, when it replaced the previously fielded version of the survey. As a part of the CPI revision extensive work was undertaken to test both the wording of questions and the survey procedures (Kosary, Sommers, Branscome, and McCarthy, 1986). Additional work was also done prior to this study on the wording of standard introductions.

BACKGROUND

The Consumer Price Index Housing Survey measures the Rent and Rental Equivalency components of the Consumer Price Index. The survey consists of a rotating panel of approximately 40,000 renter occupied units one-sixth of which are surveyed every six month and 20,000 owner occupied units one-twelfth priced every month. Data is collected using one data collection instrument (BLS Series 222R) with different sections of the instrument pertaining to either owner or renter occupied housing units. Data from renter occupied units can be collected from either an occupant or the unit manager, with the choice of respondent left to the data collector. The data collector may also elect to collect the data from a different respondent every six months.

The purpose of the data collected in the case of renter occupied housing units is to provide a measure of rent change. The owner occupied units are matched, on the basis of collected characteristics, to similar renter units and a measure of the change in the owner's shelter cost is calculated. Data is collected in 94 Primary Sampling Units (PSU's) of various size throughout the U.S.. The data is presently collected mainly by personal visit with increasing emphasis on telephone collection.

TEST DESIGN

Data analyzed consisted of two sources, a CATI Pilot test consisting of a sample of CATI reinterviews of interviews initially conducted in the field by BLS data collectors and the 1986 CPI Housing Survey Process Audit consisting of a sample of reinterviews of interviews originally conducted in the field. Both of these reinterview efforts took place in the summer and fall of 1986. Both samples were selected from the same set of PSU's, in the Eastern and Central time zones. This was do. ... scheduling pur- poses on the CATI reinterviews and in order to obtain a representative sample of the CPI Housing Survey sample since over 75% of the total Housing Survey sample are found in these PSU's.

CATI Pilot

The CPI Housing Survey CATI pilot test was conducted from August-December, 1986 utilizing a CATI version of the current CPI Housing Survey data collection. All telephone calls were made from BLS's Washington, D.C. office by members of the Washington staff. The sample was selected from 24 PSU's in the Eastern and Central time zone. All units selected had been scheduled for data collection by the field staff in the previous month so that the lag between interview and reinterview was approximately one to three weeks. Units were eligible for selection if they had at least one reported telephone number. This requirement led to the loss of approximately 4% of the units. So as to collect information on contact rates for various time periods, selected units were randomly assigned to one of nine time periods (weekday, 9am-11am, 11am-1pm, 1pm-3pm, 3pm-5pm, 5-7pm, 7pm-9pm, or weekend 11am-1pm, 1pm-3pm, 3pm-5pm). A total of 1250 cases were attempted during the course of the pilot. 516 completed interviews were obtained. As could be expected, contact rates varied across time periods, from the low 40%'s in the late morning and afternoon time slots to the mid 90%'s in the evening and on weekend. The completed cases were then matched to their corresponding original interviews. When this was completed a total of 334 renter occupied and 152 owner occupied housing units remained.

No attempt was made to contact the previous re-

spondent. In the case of owner occupied units the telephone calls were made to the occupant's telephone number. When contacting renters, the first attempt made was to the occupant, if an occupant's telephone number was available. Where no number was available or a no contact was obtained the interviewer was branched to the manager's telephone number and a call was attempted. The number of attempts were limited to 5, spread over 5 separate days. Attempts were made during the first three weeks of each month beginning on the first Monday. No information on previous data was provided.

Process Audit

As part of the 1986 Process Audit of the CPI Housing Survey reinterviews were conducted for a sample of housing units. In order to achieve a sample comparable to that of the CATI pilot, this sample was selected from the same 24 PSU's reinterviewed in the CATI pilot. Reinterviews were conducted from October, 1986 to January, 1987. All reinterviews were conducted in person by members of the Washington, D.C. based CPI Process Audit staff, many of whom also participated in the CATI pilot. No attempt was made to contact the previous respondent. The same information given to the original data collector on back data was provided to the process audit. A total sample of 788 housing units were selected for reinterview. Of these 588 were completed by the Process Auditors. The completed cases were then matched to their corresponding original interviews. When this was completed a total of 408 renter occupied and 162 owner occupied housing units remained.

ANALYSIS METHODOLOGY

For the purpose of this paper 6 of the questions which appear on the CPI Housing Survey data collection form have been selected for analysis and discussion (for the exact wording of these question see Appendix A). The questions under investigation were selected due to their importance to the survey results (such as the question on amount of rent paid) or to illustrate typical questions asked in the survey

After the data were collected a variable was created to indicate whether the reinterview had been conducted with the same respondent as the original or with a different respondent.

In order to assess the difference between the CATI reinterviews when compared to the reinterviews conducted by the Process Auditors (PA) a number of indicators were examined. These included 1) the gross discrepancy rate, 2) simple response variance, 3) mean, mean difference, and mean absolute difference for several survey estimates, and 4) the refusal rate for the income question.

Gross Discrepancy Rates

The gross discrepancy rate was calculated as the number of times a reinterview response differed from the original interview over the total num-

ber of interviews. The gross discrepancy rate has a strong intuitive appeal as a measure of reliability and can also be used as a measure of simple response variance (Lord and Novick, 1974; O'Muircheartaigh, 1986). Rates were calculated for CATI and PA reinterviews and same and different reinterview respondents. These rates were examined using multidimensional contingency table analysis (Fienberg, 1981) with method of reinterview data collection (CATI vs PA) and reinterview respondent (same vs different) as effects. If CATI offers an improvement over traditional data collection methods it would be expected that CATI will show a smaller gross discrepancy rate, both overall and when compared with reinterview respondent type, than will PA.

When the same respondent reports on both the original interview and reinterview the possibility exists that she will remember and give the original response, regardless of whether this response is in fact correct. The gross discrepancy rates observed should therefore be smaller when the reinterview respondent is the same as opposed to when the respondent is different. For this reason, in order to demonstrate that CATI has a positive effect on gross discrepancy rates both respondent types must be taken into consideration.

Simple Response Variance

The simple response variance (Hansen, Hurwitz, and Pritzker, 1964; O'Muircheartaigh, 1986) was calculated both overall and separately for same and different reinterview respondents for both the CATI and PA. If CATI has a positive effect on the data collected, the simple response variances obtained for the independent questions will be lower than the simple response variances obtained for the PA reinterviews, both overall and for same and different reinterview respondents.

Survey Estimates

In several instances the responses given to the survey questions are combined in order to determine additional estimates. IN the CPI Housing Survey two of these estimates are normalized rent and total number of rooms. Normalized rent is calculated by combining the responses given for the amount of rent plus any reductions in the rent and/or extra charges paid by the tenant to the landlord and are based on monthly rent, regardless of the actual rental period. Normalized rent is the estimate of the monthly economic rent for the particular rental unit being priced. Total number of rooms is the number of bedrooms plus the number of other rooms. Since some respondents may have a bedroom which is used for some other purpose, such as a den or office, there may be a fluctuation in the reported number of bedrooms and other rooms. Total number of rooms may, therefore provide a better estimator.

In order to assess the estimates of these variables the mean of the estimate for both the original and reinterview and the difference and absolute difference between the two estimates

were calculated. These estimates were then tested to determine whether any differences encountered were significant.

Refusal Rates

It has been been observed that on most surveys the most sensitive question is the income. In order to assess the possible impact of CATI on refusals of this question the refusal rates for the CATI reinterview were compared to that of its matched set of original interviews and also to the refusal rates from the Process Audit reinterviews and their matched set of original interviews using Chi-square tests.

RESULTS

There were a total of 486 CATI reinterviews paired to an original data collection and 570 PA reinterviews paired to an original data collection.

Gross Discrepancy Rates

Gross discrepancy rates (GDR's) were calculated for the 6 questions examined. These rates were then analyzed using multidimensional contingency table analysis in order to determine the significance of type of reinterview respondent (same vs different) method of reinterview collection (CATI vs PA) and the interaction between respondent and collection. The results are presented in Appendix B for both those questions where interviewing was conducted in an independent manner and those which were dependent.

The results indicate that the overall GDR for the CATI reinterviews are not significantly greater than the PA reinterviews and in some instances are, in fact, significantly lower. This relationship also holds true for same and different reinterview respondent.

Simple Response Variance

The results for the CATI and PA reinterviews are presented in Appendix C. The simple response variance for the CATI reinterviews are smaller than those for the PA reinterviews for the independent questions. These differences are particularly striking on Q.17 - Rent, Normalized rent, Q.37 - number of bedrooms, and Q.39 - number of other rooms.

As expected, simple response variance is higher when the reinterview respondent is different than when the respondent is the same.

Survey Estimates

The results for the CATI and PA reinterviews are presented in Appendix D. No significant differences were found for the estimates of normalized rent. The estimate of total rooms obtained from the PA are significantly lower than obtained from the original interview. Mean difference and mean absolute difference between the original and reinterview are significantly lower for CATI indicating that the CATI estimates vary less from the original than the PA estimates.

Refusal Rates

The refusal rates for the original interviews and their corresponding CATI and Process Audit reinterview are given below. The CATI refusal rate is significantly lower than the original interview while no differences exist between the Process Audit reinterviews and their matched originals. There are several possible explanations for this result. It is possible that CATI is more impersonal and respondents are more willing to provide sensitive information in such an environment. Another explanation pertains to CATI's forced structure. It has been observed during data collection conducted by personal visit on the CPI Housing Survey that data collectors are very nervous asking this question and frequently begin with an apology to the respondent and a reassurance that the respondent is free to refuse if they are offended. Since the CATI data collectors were instructed not too deviate from the question wording as it appeared on the screen and since this wording did not include an apology or option to refuse, respondents may have just answered the question without a thought.

Q.42 Income

Process Audit

	Reinterview	Original
% refusals	14.9	16.1

CHI SQUARE = 0.16
DF = 1
PROB = .6854

CATI

	Reinterview	Original
% refusals	4.9	14.2

CHI SQUARE = 20.82
DF = 1
PROB = .0000

CONCLUSIONS

The results indicate that the use of CATI has potential to increase data quality in terms of item gross discrepancy rate, simple response variance, and stability of survey estimates. Indications are that CATI can also be an aid in dealing with questions of a sensitive nature, such as income, in terms of decreased refusal rates. The data suggests that due to its enforced consistency of procedures and question wording CATI is an aid to obtaining more reliable survey results. This does not mean that the answers obtained by either CATI or personal visit are indeed the answers intended by the survey designers. Such questions are in fact better addressed through other means, such as cognitive laboratory experiments, at the time in which the survey instrument is developed. What the results do infer is that given a well developed survey instrument increased data consistency and reliability can be obtained through the use of CATI.

REFERENCES

Bishop Yvonne M, S. E. Fienberg, and P. W. Holland, (1975). *Discrete Multivariate Analysis*, Cambridge, MA:MIT Press.

Fienberg, S.E. (1981). *The Analysis of Cross-Classified Categorical Data*, Cambridge,MA:MIT Press.

Groves, R.M. (1983) Implications of CATI: Costs, Errors, and Organization of Telephone Survey Research. *Sociological Methods and Research*, vol. 12(2), pp.199-215.

Groves, R. M. and W.L. Nicholls II (1986). The Status of Computer Assisted Telephone Interviewing: Part II – Data Quality Issues. *Journal of Official Statistics*, vol 2(2), pp. 117-134.

Hansen, M. H., W.N. Hurwitz, and L. Pritzker (1964). The Estimation and Interpretation of Gross Differences and the Simple Response Variance. in C.R. Rao, ed. *Contributions to Statistics*, Calcutta:Pergamon Press Ltd.

Kosary, C.L., J.P. Sommers, J.M. Branscome, and W.D. McCarthy (1986). Quality Assurance of the Consumer Price Index Housing Survey. *Proceedings of the American Statistical Association, Section of Survey Methodology*, pp. 556-561.

Lord, F.M. and Novick M.R. (1974). *Statistical Theories of Mental Test Scores*, Reading,MA:Addison Wesley.

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

O'Muircheartaigh, C. A. (1986). Correlates of Reinterview Response Inconsistency in the Current Population Survey. *Proceedings of the Second Annual Research Conference*, U.S. Bureau of the Census, March 23-26, 1986 Reston, VA.

Pafford, B.V. (1987) Use of Historic Data in a Current Interview Situation: Response error Analyses and Application to Computer Assisted Telephone Interviewing. *Proceedings of the Third Annual Research Conference*, U.S. Bureau of the Census, March 29 – April 1, 1987 Baltimore, MD.

APPENDIX A – Analyzed Questions

Complete for all renter occupied units

Q.17 (After the rent has been lowered) How much is the rent for this (apartment/house) for the current time period?

Complete for all occupied units

Q.36A Which of these fuels is the main fuel used for heating?

1 electricity
2 natural gas
3 lp gas
6 firewood/charcoal
7 solar
8 other, specify

4 fuel oil/kerosene 9 none
5 coal

Q.37 How many bedrooms are there in this (apartment/house)?

Q.39 Excluding bedrooms, bathrooms, and any unfinished rooms, how many other rooms are there?

Q.41 How many people live in this (apartment/house)?

Q.42 Which group letter best describes the total combined yearly income of all of the people living in this (apartment/house)?

A $ 0 - $ 4999   G $30000 - $39999
B $ 5000 - $ 9999   H $40000 - $49999
C $10000 - $14999   I $50000 - $74999
D $15000 - $19999   J $75000 and above
E $20000 - $24999   K refusal
F $25000 - $29999

APPENDIX B – Results of analysis of gross discrepancy rates for CATI and PA reinterviews

Q.17 – Current Rent

*respondent collected	Different		Same	
	CATI	PA	CATI	PA
% discrp	32.4	37.2	18.3	28.1

*collected	CATI	PA
% discrepant	23.6	32.5

*respondent	Different	Same
% discrepant	35.9	24.5

---

* Significant Results

Effect	DF	Chi Square	Prob
Respondent	1	7.79	.0193
Collected	1	3.81	.0510
Respondent X Collected	1	5.47	.0193

Reinterview respondent, method of reinterview data collection and the interaction between respondent and mode of collection are significant. The GDR is lower when the reinterview respondent was the same. The GDR is lower when CATI was the reinterview data collection method. When the reinterview respondent was the same the GDR from the CATI reinterview is lower than the PA reinterviews. When the reinterview respondent was different this relation is also evident.

Q.36A - Main heating fuel

*respondent collected % discrp	Different CATI	PA	Same CATI	PA
	17.0	20.2	7.7	12.5

collected % discrepant	CATI 11.6	PA 16.4

*respondent % discrepant	Different 18.3	Same 9.3

*Significant Results

Effect	DF	Chi Square	Prob
Respondent	1	8.52	.0035
Respondent X Collected	1	3.06	.0804

Reinterview respondent and the interaction between respondent and the reinterview data collection method are significant. The GDR is lower when the reinterview respondent is the same. When the reinterview respondent was the same the GDR from the CATI reinterview is lower then the PA reinterview. When the reinterview respondent was different this relationship is also evident.

Q.37 - Number bedrooms

*respondent collected % discrp	Different CATI	PA	Same CATI	PA
	13.5	22.0	9.3	36.5

*collected % discrepant	CATI 11.0	PA 29.5

respondent % discrepant	Different 17.4	Same 20.0

*Significant Results

Effect	DF	Chi Square	Prob
Collected	1	32.04	.0000
Respondent X Collected	1	3.01	.0828

Reinterview data collection and the interaction between respondent and reinterview data collection method are significant. The GDR is lower when CATI was the reinterview data collection method. When the reinterview respondent was the same the GDR from the CATI reinterview is lower than the PA reinterview. When the reinterview respondent was different this relationship is also evident.

Q.39 - Number other rooms

*respondent collected % discrp	Different CATI	PA	Same CATI	PA
	43.6	50.9	38.1	58.4

*collected % discrepant	CATI 40.4	PA 54.8

respondent % discrepant	Different 46.9	Same 46.1

*Significant Results

Effect	DF	Chi Square	Prob
Collected	1	11.58	.0007

Reinterview data collection method is significant. The GDR is lower when CATI was the reinterview data collection method. When the reinterview respondent was the same the GDR from the CATI reinterview is lower tan the PA reinterview.

Q.41 Number persons

*respondent collected % discrp	Different CATI	PA	Same CATI	PA
	25.0	38.3	17.5	18.1

*collected % discrepant	CATI 20.7	PA 28.5

*respondent % discrepant	Different 30.5	Same 17.7

*Significant Results

Effect	DF	Chi Square	Prob
Respondent	1	4.56	.0327
Collected	1	10.84	.0010
Respondent X Collected	1	3.2	.0717

Reinterview respondent, reinterview data collection method and the interaction between respondent and reinterview data collection method are significant. The GDR is lower when the reinterview respondent was the same. The GDR is lower when CATI was the reinterview data collection method. When the reinterview respondent was different the GDR from the CATI reinterview is lower than the PA reinterview. When the reinterview respondent was the same there is little difference between the GDR's for the CATI and PA reinterviews.

APPENDIX C - Simple response variance for CATI and PA reinterviews

Q.17 Rent

	CATI	PA
Overall	354.47	1485.15
Same respondent	250.37	639.50
Different resp	529.47	2394.45

Normalized rent

	CATI	PA
Overall	461.89	1145.39
Same respondent	437.81	613.04
Different resp	508.66	1801.71

Q.36A <u>Heating fuel</u>

	<u>CATI</u>	<u>PA</u>
Overall	.058	.082
Same respondent	.039	.063
Different resp	.085	.101

Q.37 <u>Number bedrooms</u>

	<u>CATI</u>	<u>PA</u>
Overall	.112	.613
Same respondent	.066	.897
Different resp	.173	.309

Q.39 <u>Number other rooms</u>

	<u>CATI</u>	<u>PA</u>
Overall	.449	1.06
Same respondent	.388	1.46
Different resp	.533	.634

<u>Total number rooms</u>

	<u>CATI</u>	<u>PA</u>
Overall	.786	3.07
Same respondent	.871	1.83
Different resp	.492	4.24

Q.41 <u>Number of persons</u>

	<u>CATI</u>	<u>PA</u>
Overall	.199	.337
Same respondent	.197	.223
Different resp	.208	.444

APPENDIX D – Survey estimates for CATI and PA reinterviews

<u>Normalized Rents</u>

OVERALL

	CATI	Field	PA	Field
Mean	314.12	314.40	335.24	333.30
Mean diff	.27		-1.95	
Mean abs difference	11.11		15.71	

SAME RESPONDENT

	CATI	Field	PA	Field
Mean	302.98	304.32	328.15	326.05
Mean diff	1.34		-2.10	
Mean abs difference	9.81		12.51	

DIFFERENT RESPONDENT

	CATI	Field	PA	Field
Mean	335.77	333.98	343.98	342.23
Mean diff	-1.79		-1.76	
Mean abs difference	13.64		19.66	

Total Rooms

OVERALL

	CATI	Field	PA	Field
Mean	5.55	5.56	3.81 *	5.00
Mean diff	-0.12		*	0.86
Mean abs difference	0.58		*	1.42

SAME RESPONDENT

	CATI	Field	PA	Field
Mean	5.40	5.41	3.90 *	5.21
Mean diff	-0.18		*	1.20
Mean abs difference	0.52		*	1.74

DIFFERENT RESPONDENT

	CATI	Field	PA	Field
Mean	5.78	5.73	3.87 *	4.73
Mean diff	-0.18		*	0.50
Mean abs difference	0.66		*	1.08

* Significant difference