Journal of Official Statistics
Vol. 6. No. 2. 1990. pp. 205–211

Sampling Design for a Monitoring Plan for CATI Interviewing

David W. Chapman and Rachel B. Weinstein[1]

Abstract: The U.S. Bureau of the Census is considering the use of computer assisted telephone interviewing (CATI) to replace decentralized telephone interviewing from interviewers' homes for many of its surveys. The use of a centralized CATI facility permits the monitoring of interviews. A systematic monitoring procedure, intended to be an ongoing activity for all CATI interviewers, has been designed. The method consists of monitoring about 5% of each interviewer's work using interview time segments as sampling units. Here we describe the sampling design of this systematic monitoring procedure.

Key words: Telephone surveys; sample procedures; quality control; time sampling.

1. Introduction

Computer assisted telephone interviewing (CATI) is an automated method of telephone interviewing which eliminates the need for paper and pencil. The survey questionnaire is programmed into a computer and the interviewer conducts interviews from a computer terminal or microcomputer in a centralized facility. Questions are read from the screen to the respondents over the telephone and responses are keyed directly into the terminal by the interviewer.

The U.S. Census Bureau began research in 1980 into the use of CATI to replace the decentralized telephone interviews conducted in its demographic surveys. In 1985, the Census Bureau established a telephone facility, the Hagerstown Telephone Center (HTC), to test and evaluate the use of CATI and random digit dialing for specific surveys. A discussion of the development of CATI at the Census Bureau is given by Nicholls (1983) and an overview of the development and status of CATI is provided by Groves and Nicholls (1986) and Nicholls and Groves (1986). A report on the evaluation of CATI for use in the Current Population Survey (CPS) is provided by the U.S. Bureau of the Census (1988) and for use in the National Crime Survey (NCS) by Hubble and Wilder (1988). Also, a report on an investigation of the use of random digit dialing for the National Health Interview Survey is given by Biemer and Chapman (1985), and for the NCS by Alexander, Sebold, and Pfaff (1986).

Because of the centralization of interview-

[1] Statistical Research Division, U.S. Bureau of the Census, Washington, DC 20233, U.S.A.

Acknowledgements: The authors thank Ms. Karen J. Bryant of Harris Data for her contribution to the preparation of this paper. Before leaving the Census Bureau in February 1988, she helped prepare early drafts of some sections of this paper. The authors also thank the editor and referees for their comments and helpful suggestions.

ing, CATI provides the opportunity for monitoring interviews. Research into monitoring methods for CATI has been underway at the Census Bureau for about three years. Although the emphasis has been in designing and testing monitoring forms, there has been considerable study regarding the sampling aspects of *systematic monitoring* – the procedure for regularly monitoring a portion of each interviewer's work. Such a procedure has been designed, but not yet implemented for the Bureau's CATI experimental surveys. This document describes the sampling design of this monitoring system. (This design could also be used for a centralized telephone facility that does not use CATI).

2. Basic Approach to Systematic Monitoring for CATI

At this time, the primary purpose of systematic monitoring for CATI at the Census Bureau is to provide a general check on interviewer performance. No procedure for deriving formal ratings of interviewers based on monitoring results has been suggested, due primarily to the variability among monitors.

For general purpose monitoring, it seems appropriate to use probability sampling in selecting cases for monitoring and to use a constant selection rate across shifts and interviewers. Specifically, the system has been designed to monitor a random 5% of each interviewer's cases, using time segments as sampling units. There may be some modest advantages to other methods that allow varying sampling rates; however, these would complicate the selection procedure and would probably be perceived as unfair by the interviewers.

The choice of a monitoring rate of 5% has to be somewhat arbitrary since there are no specific measurements generated from monitoring sessions that could provide a precision criterion. Based primarily on cost considerations, a monitoring rate of 2.5% was initially proposed. However, a small, informal survey of some survey organizations revealed that, although telephone monitoring rates vary substantially, there may be an industry standard emerging for Federal surveys of 10%. These surveys generally did not include any reinterviewing of samples of respondents as a check of the quality of interviewing. Since the Census Bureau plans to do both reinterviewing and monitoring, it was decided that 5% would be a reasonable monitoring rate to begin with.

The choice of time segments, rather than cases, as sampling units for monitoring is dictated by the basic operation of a CATI facility. First, the number of cases per interviewer is not fixed, making it difficult to determine how many cases per interviewer should be monitored to achieve a 5% rate. Also, there is considerable variety of interviewer work schedules at the HTC which could make it difficult to schedule monitoring of a specific interviewer's cases. Furthermore, there is no easy way at this time for a monitor to find the beginning of an interview attempt (i.e., the beginning of a telephone call) other than when an interviewer first begins work each day. The use of time segments for monitoring sessions allows for a 5% rate to be achieved and for relatively easy scheduling of monitoring sessions.

The number of monitoring time segments scheduled for a shift is based on the anticipated number of interviewers for the shift. These time segments are assigned randomly to time slots throughout the shift. The interviewers working the shift are assigned randomly to the selected monitoring sessions, hour by hour, except that controls have been established to increase the selection

probabilities of those interviewers who have been monitored fewer times per hour worked than others.

The specific features of the sample design for monitoring are given in Sections 3–7. These include (a) the length of monitoring time segments, (b) the selection of monitoring segments for a shift, (c) the assignment of interviewers to monitoring sessions, and (d) the assignment of staff to conduct the monitoring sessions.

3. Definition of Monitoring Time Segments

Monitoring time segments of various lengths have been considered. Initially, segments of 15 and 20 minutes were proposed. Based on some trial monitoring sessions for CPS and NCS, longer segments were recommended because of the increased likelihood that a monitoring session will contain an entire interview for CPS or NCS, two important surveys for implementing CATI.

The interview length varies between CPS and NCS. The average length of a regular CPS interview is 10–15 minutes. Supplemental questions, which usually add about five minutes to the length of an interview, are added to the CPS questionnaire in about half the months. However, the March income supplement adds approximately 30 minutes. The length of the NCS interview is about 10–30 minutes, depending on whether there are any crime incidents reported. Therefore, it has been decided to use either 25- or 50-minute monitoring sessions: 25-minute sessions for CPS interviews (except March) and 50-minute sessions for NCS and March CPS interviews.

Hourly time slots have been defined for the 25- and 50-minute monitoring segment lengths, as follows:

1. One 50-minute time segment: 10–60 minutes past the hour.
2. Two 25-minute segments with a scheduled 5 minute break between them: Session 1: 5–30 minutes past the hour. Session 2: 35–60 minutes past the hour.

4. Determination of the Number of Monitoring Segments to Assign for a Given Shift

The target number of monitoring minutes for each shift is 5% of the anticipated total number of interviewer minutes for the shift. The actual number of assigned monitoring segments for a shift is the number which best approximates the target number of monitoring minutes. For instance, if 23 interviewers are anticipated to work a five-hour shift, then the number of interviewer minutes expected to be worked is 6900 and the target monitoring time is 345 minutes (5% of 6900). For a 50-minute monitoring segment, the multiple of 50 closest to the target is 350 minutes, or seven sessions of monitoring.

The number of 50-minute monitoring segments assigned for five hour shifts is specified in Table 1 for 1, 2, 10, 25, 50, and 75 interviewers. In some cases the target monitoring time falls exactly between two multiples of 50 (e.g., when there are 15 or 25 interviewers). In these instances the choice of the lower or higher multiple alternates, starting with the lower multiple.

When monitoring segments are 25 minutes, the assigned monitoring time is the multiple of 25 closest to the target monitoring time, except when the anticipated number of interviewers for a shift is only two. In this case, 50 minutes of monitoring will be assigned instead of 25 minutes, in order to provide a minimum monitoring workload of 50 minutes for a shift. The number of 25-minute monitoring segments assigned for five hour shifts is specified in

Table 1 50-minute monitoring segments for five hour shifts

Projected number of interviewers	No. of interviewer minutes	Target monitoring time	Assigned monitoring time	No. of monitoring segments	Projected monitoring rate
1	300	15	0	0	0
2	600	30	50	1	.083
10	3000	150	150	3	.050
25	7500	375	400	8	.053
50	15000	750	750	15	.050
75	22500	1125	1100	22	.051

Table 2 for 1, 2, 10, 25, 50, and 75 interviewers.

5. Assignment of Monitoring Segments to Time Slots

Once the number of monitoring segments (m) assigned to a shift is determined, segments are assigned to time slots. Letting s denote the number of time slots in a shift, the number of monitoring segments to be assigned can be expressed as an integral multiple of s, plus a remainder (if any) that is less than s:

$$m = qs + r, \tag{1}$$

where

$$q \geqslant 0, 0 \leqslant r \leqslant s - 1.$$

Initially, each time slot is assigned q monitoring sessions. If $r = 0$, no other assignments are required. If $r > 0$, the remaining r segments are randomly assigned to the s time slots. After this assignment, each time slot will ultimately have either q or $q + 1$ monitoring sessions assigned to it.

For example, if 23 interviewers are anticipated to work a five-hour shift and monitoring segments are 50 minutes, then it was determined in the example in the previous section that the number of monitoring sessions to be scheduled is 7. Following equation (1), with $s = 5$ this number is expressed as

$$7 = 1(5) + 2.$$

Therefore, each time slot is initially assigned one segment and the remaining two segments are randomly assigned to two of the five time slots.

Table 2 25-minute monitoring segments for five hour shifts

Projected number of interviewers	No. of interview minutes	Target monitoring time	Assigned monitoring time	No. of monitoring segments	Projected monitoring rate
1	300	15	0	0	0
2	600	30	50	2	.083
10	3000	150	150	6	.050
25	7500	375	375	15	.050
50	15000	750	750	30	.050
75	22500	1125	1125	45	.050

6. Assignment of Interviewers to Monitoring Sessions

Once the time slots in a shift are assigned a specific number of monitoring sessions, interviewers are assigned to monitoring sessions. Assignments are made hourly, rather than for an entire shift, because interviewer work schedules often vary from strict shift definitions.

First, the interviewers working a specific hour are listed in ascending order, based on the priority-to-be-monitored index, I, defined as

$$I = \frac{\text{No. of time segments monitored}}{\text{No. of hours worked} + 1} .$$

This index is based on the rate at which interviewers have already been monitored. Interviewers with the lower index values have the highest priority. Any ties are randomly ordered except zeroes which are listed in descending order by number of hours worked. The "1" in the denominator ensures that division by zero does not occur when an interviewer has not worked any hours in the month. Also, if two interviewers have the same ratio of number of segments monitored to number of hours worked, the "1" lowers the index more for the interviewer that has worked less. Ideally, the index would be updated each hour. However, it may be feasible to update the index only once per shift or per day.

If there are u monitoring sessions assigned to a time slot contained in a specific hour, the first u interviewers listed on the priority-to-be-monitored list are monitored that time slot. Interviewers are assigned to specific sessions in the order listed.

The priority index, I, could be modified to provide differential monitoring rates for interviewers, For example, to decrease the rate for experienced interviewers, a constant could be added to the numerator of I, or could be inserted as a multiplier, based on the amount of experience an interviewer has.

This procedure of assigning interviewers to monitoring sessions is not significantly altered by the fact that interviewer shifts may overlap. The appropriate number of monitoring sessions for a shift is calculated without regard to overlapping shifts. Then for any hour contained in two overlapping shifts, the assigned numbers of monitoring sessions for each time slot are added for the two shifts. Also, in assigning interviewers to monitoring sessions, a combined list of interviewers from the two shifts is used.

7. Assignment of Supervisors to Monitoring Sessions

For this initial stage of development, the Census Bureau has decided to use shift supervisors to do the monitoring. It is important that the number of shift supervisors present for each hour include at least one more supervisor than the number of monitors needed during the hour.

There are several ways that supervisors can be assigned to monitoring sessions. It would probably be best to have a small number of supervisors assigned to monitor for an entire shift so that they could focus their attention on monitoring. The monitoring function for a given shift would probably be rotated among supervisors from day to day because of the tedious nature of monitoring. In fact, it may turn out that monitoring is so tedious that the monitoring work load for each shift has to be shared among all supervisors present.

8. Discussion

Although the procedures described in this document were designed primarily for two

Census Bureau surveys, CPS and NCS, they could be modified, if necessary, for application to other surveys. It would be straightforward to adjust the procedures to incorporate a different monitoring rate, a different interviewer shift length, or a different length of monitoring segment.

To help supervisors administer the monitoring procedures, some computer programs have been written: one program to derive each interviewer's priority-to-be-monitored index, I, and two programs for selecting monitoring time segments for each shift. In future applications it might be helpful to write programs that would combine these two functions.

One of the major concerns regarding the implementation of systematic monitoring is whether or not the supervisors assigned to monitoring during a given shift will be able to adhere to the monitoring schedule since they may have other responsibilities. For the plan to work, it may be necessary to form a special staff of dedicated monitors, an approach already used by some survey organizations. Some additional testing of the procedures is needed.

Another issue is how well the time segments will work as sampling units. With time segments, monitoring sessions will often begin and end in the middle of an interview. Some consideration should be given to alternate methods that would allow monitoring of entire interviews. For example, one approach would be to monitor all cases that *originate* during a given time segment. However, this would involve lost time at the beginning of a time segment, waiting for a new case, and would make it difficult to schedule a monitor's time due to the uncertainty of when a session would end. It might also be useful to investigate the possibility of using interviewer cases, rather than time segments, as sampling units. For this to work efficiently, interviewers would have to keep fixed work schedules and the number of cases per interviewer would have to be reasonably stable.

Finally, there is unexplored potential in the automation of the monitoring operation with a CATI system. It might be possible to eliminate the need for paper monitoring forms, with monitoring notes being entered directly into the computer. There could perhaps be a split screen arrangement, half the screen showing the interview responses and half being used for monitor entries.

9. References

Alexander, C., Sebold, J., and Pfaff, P. (1986). Some Results of an Experiment with Telephone Sampling for the U.S. National Crime Survey. Proceedings of the Survey Research Methods Section, American Statistical Association, 351–356.

Biemer, P.P. and Chapman, D.W., eds. (1985). Results of the 1984 NHIS/RDD Feasibility Study: Final Report. Internal U.S. Census Bureau report, February.

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

Hubble, D.L. and Wilder, B.E. (1988). Preliminary Results from the National Crime Survey (NCS) CATI Experiment. Proceedings of the Survey Research Methods Section, American Statistical Association 558–563.

Nicholls II, W.L. (1983). Development of CATI at the U.S. Census Bureau. Proceedings of the Survey Research Methods Section, American Statistical Association, 642–647.

Nicholls II, W.L. and Groves, R.M, (1986).

The Status of Computer-Assisted Telephone Interviewing: Part I – Introduction and Impact on Cost and Timeliness of Survey Data. Journal of Official Statistics, 2, 93–115.

U.S. Bureau of the Census (1988). Evaluation of CATI Data Quality and Cost in the Current Population Survey. CATI Research Report No. CPS-4, Computer Assisted Interviewing Research and Analysis Subcommittee. Internal U.S. Census Bureau report, October.

Received January 1989
Revised December 1989