CONTROLLING NONRESPONSE IN AN ESTABLISHMENT SURVEY

Richard J. Rosen, BLS; Richard L. Clayton, BLS; Thomas B. Rubino, Jr., BLS
Richard Rosen, Bureau of Labor Statistics, 441 G St NW, Rm 2089, Washington, DC 20212

Key Words: CASIC, CATI, Touchtone Data Entry, Automated Collection

Introduction:

Timely collection of data is a priority in all surveys. As a result, most surveys have procedures to follow-up on nonrespondents. These procedures generally involve setting a fixed "cutoff" for receipt of data. Sample units which do not respond by this date are recontacted.

This paper describes development of a model to maximize response rates and minimize workload, costs, and respondent burden in an automated self-response collection environment.

Background:

The Current Employment Statistics Survey (CES), conducted by the Bureau of Labor Statistics (BLS), requires collection over a very narrow 10-15 day time frame. This heightens the demands for prompt reporting and makes effective nonresponse follow-up difficult.

The CES survey collects monthly data on employment, payroll, and hours from a panel of over 350,000 business establishments. Data are collected based on the pay period which includes the 12th of each month. In order to release the data on the first Friday of the following month, preliminary estimates are produced based on only 10-15 days of collection. Collection is decentralized, with each State Employment Security Agency responsible for collection within their State. The data are key entered, edited, and then transmitted to BLS-Washington for estimation.

Since its inception more than 50 years ago, mail has been the primary mode of data collection. The response rate for preliminary estimates is generally 50-55 percent. As a result of this relatively low response rate, subsequent estimates, based on more complete data, can be significantly different from the initial figures. These initial estimates are closely watched by policy-makers, the financial and business communities, and others concerned with the health of the U.S. economy. Thus, major revisions are viewed with concern.

Under mail collection, it is difficult to obtain precise information on sample status. Apparent nonrespondents may in fact have reported; however, the form may still be in the mail, awaiting key entry, or simply not "run" into the data base. Prompting these units would likely be wasteful and irritating. For this reason, and because of the volume of nonresponse, almost all nonresponse activity under mail currently occurs after the cut-off for initial estimates. Thus, the mail process relies entirely on what respondents are willing and able to report, and the efficiency of the mail delivery system.

Since the mid-1980's, BLS has evaluated various automated collection techniques to speed-up and control the collection process. Automated collection methods offer the potential for vast improvements in both response rates and the survey process. Portions of the CES Survey currently report via Computer Assisted Telephone Interviewing (CATI), while others report directly to a computer using their phone.

The BLS Touchtone Date Entry (TDE) system allows respondents to enter data using the touchtone pad on their phone. Data are entered directly into the computer, and the system reads each item back for verification. The BLS Voice Recognition (VR) system allows respondents to report data over any phone by "speaking" the digits in a conversational manner. *[Werking and Clayton, 1990; Clayton and Winter, 1990]*

Collection via CATI eliminates nonresponse prompting, since dates for collection are set in advance with the respondent based on their data availability, and the CATI interviewer initiates the call each month on the agreed upon date. Additional calls are scheduled and made as needed.

Automated collection methods such as CATI, TDE, and VR, offer many advantages over mail. These methods:
1) eliminate the timing lag associated with mail;
2) capture the data in machine readable form;
3) provide greater control over the collection process.

In addition to the advantages cited above, TDE and VR self-response offer the additional advantage of lower survey costs. Because most respondents self-report, collection costs are reduced in comparison to either CATI or mail. *[Clayton and Harrell, 1989]* However, like mail, these methods rely on self-response on the part of sample

members--respondents must remember to report on their own.

Current TDE Nonresponse Procedures:

Currently, all nonresponse activity for TDE is undertaken three days prior to the cut-off for preliminary estimates. In the absence of more specific information on when to begin, this practice allows the maximum time for units to self-report and sufficient time for a small number units to be contacted prior to the cut-off date.

Figure 1 shows the TDE response rates achieved under these procedures. Prior to prompts, the response rate averages about 70%, with two more days of collection remaining. Thus, a 30% nonresponse rate becomes the benchmark against which to compare alternative models. After nonresponse follow-up, the response rate averages nearly 85%.

The limitation of the current prompting scheme is that, since all prompting activities are undertaken in a single day, resources will become strained as the prompting workload grows. Arbitrarily beginning earlier risks "wasting" calls to units whose data are not yet available or would have called on their own given more time. In addition, respondent cooperation may be jeopardized by calling too early--before their data are available.

Factors Affecting Nonresponse:

Much has been written on the subject of nonresponse, however most of this research has been focused on household surveys. *[Groves, 1988; Groves, 1989; Kasprezyk, 1989]* In this context, nonresponse is analyzed by such demographic characteristics as age, sex, and race. Similarly, in establishment surveys, one can look at establishment-specific characteristics related to nonresponse.

Response/nonresponse in the CES, regardless of collection method, is affected by three key factors:

1) the length of the establishment's pay period;
2) the size of the firm; and
3) the number of collection days available.

Figure 2 shows response rates for mail and automated collection for each of these factors.

The length of the pay period relates to how often employees are paid and directly determines how early or late in the CES collection cycle the respondent can report. For example, some firms pay employees every week, while others issue payroll checks once each month. Establishments with a weekly payroll generally have their data available early in the cycle, while monthly payrolls often end on the last day of the month, right at the cut-off for preliminary estimates.

Small firms generally have their data available earlier than large firms. This may be because of the time lag for summary information to reach the person reporting the CES figures. Large firms are more likely to have separate departments responsible for various aspects of the payroll process, whereas in small firms one person may be responsible for all aspects of the process.

The number of collection days represents the number of working days from the 12th of the month (the first potential collection day) to the last Friday of the month (the cut-off for initial estimates). The number of collection days can vary from 10 to 15 days. While this factor does not influence the nonresponse prompting process, it is nonetheless a significant factor in explaining monthly differences in response rates. Both mail and automated collection show lower response rates for shorter collection periods.

Figure 3 illustrates the TDE reporting process and the various factors which influence reporting. The horizontal line represents time during the collection cycle. Along the "time line" are the normal events which occur before a respondent reports. These include the ending of the pay cycle, compiling summary information, completing the CES form, and reporting the data.

Below the time line are events which can negatively impact reporting--respondent workload or technical problems with the TDE system.

Above the time line are the positive procedural steps that encourage timely reporting. First, TDE respondents are provided a set of suggested reporting dates on the form which they retain. Second, reporters are sent an "advance notice" postcard in the mail at about the time their data should be available each month. Finally, units which do not self-report prior to the initial cut-off date are called by phone and reminded to please call-in their data.

Developing a TDE Non-Response Prompt Model:

The goal was to develop a TDE nonresponse prompting (NRP) model that would improve on the current method.

The specific objectives were to:

1) maximize the response rate for initial estimates;
2) minimize the total NRP workload;
3) spread out the NRP workload to fit available staff;
4) minimize "wasted" calls; and
5) minimize respondent frustration from early prompts.

Some of these objectives are naturally conflicting. For example, the best way to minimize "wasted" calls, calls to

respondents who's data are not yet available or would have reported on their own, is to wait until the last possible moment. This would also minimize the total NRP workload since respondents would be given the maximum amount of time to self-report. However, this would not satisfy the objective of spreading out the work, and could impact the overall response rate if respondents were not given sufficient time to report after being prompted. Providing sufficient time and staff resources for the NRP call is critical to large-scale implementation of TDE. For example, with 1,000 units reporting on TDE, a 30% nonresponse rate translates into 300 prompting calls.

Thus, the optimal model must balance these conflicting objectives. By using information on "data availability" and other known respondent characteristics such as the length of the pay period and size of firm, an effective prompting scheme was developed that would both spread the calls over a greater number of days, yet reduce the total number of prompting calls.

Methodology:

Various models of nonresponse were developed using actual CATI and TDE data. The sample contained approximately 3,000 CES reporters from eight States. These units had generally reported for at least 5 months on CATI, followed by at least 5 months on TDE.

The following variables, captured during CATI calls, were used to analyze reporting under CATI:

1) first call date;
2) last call date;
3) length of pay period (LP); and
4) size of firm.

The first and last call dates provide observable parameters around "data availability", and serve as a proxy for when the respondent should be able to report. The date of first call is the date the respondent agreed to be contacted at the conclusion of their report for the prior month. It is unlikely the respondent would agree to a date that was not realistic. The date of last call represents the latest call during the collection period--when data were actually collected. If data were collected on the first call, then the first and last dates would be the same. Since about half the units require one or more callbacks, the last contact date is often later than the first contact date.

The TDE system captures the date the call was entered into the system, making it possible to determine exactly when each respondent reported each month.

Three types of models were tested: 1) firm specific; 2) group specific; and 3) a combination of firm and group models. For the firm specific models, the first and last call dates were used. For the group specific models, the length of pay period and size were used. The combination model used aspects of both the firm and group models.

Firm-specific: A number of firm-specific models were tested using both the first and last call dates. These included using the mean value for each unit across the 5 CATI observations, as well as using the standard deviation for each unit. However, the model which provided the best results was the "latest" first call date across all observations.

The "latest" date represents the date closest to the cut-off for preliminary estimates across all months the unit reported on CATI. Thus, this date provides the absolute outside limits on data availability--the latest date the unit scheduled a first call. As will be shown later, this date still fell within the cutoff for first estimates and provided sufficient spread of the workload.

Group-specific: To develop the group-specific model, we analyzed the scheduling pattern of CATI calls across length of payroll and size. In general, units with weekly payrolls are scheduled relatively early in the collection cycle, as are units with semi-monthly payrolls. By contrast, biweekly units are scheduled throughout the cycle, and monthly units are scheduled relatively late.

Based on this pattern the following nonresponse calling rules were developed:

Days Prior to cutoff:

5 - Weekly units with less than 100 employees;
4 - Weekly units with 100 or more employees;
3 - Semi-monthly units;
2 - Biweekly units;
1 - Monthly units.

Units with a weekly payroll were split among two days because they comprise the largest portion of the sample, over 50%. Scheduling these units in a single day would pose workload problems. The split by size takes into account the fact that larger units generally report later than smaller units.

Combination: This model used aspects of both the firm-specific and group models. The group-specific date was used, except were the call date fell later in the collection cycle. Thus, units who report after the generic LP/size criteria would be given the required time to report. Units where the call date fell before the generic LP/size date would be allowed extra time to report. The only exception to this rule was for units where the call date fell after the cutoff for preliminary estimates, but had call dates in other months that

were prior to the cutoff. These units were assigned for prompt one day before the cutoff.

This model embodies the best of both the firm-specific and group models. It schedules units who have call dates later than the generic LP/size model into a more appropriate time slot (thus avoiding wasted calls), and provides additional reporting time for units who would otherwise be scheduled for a nonresponse prompt call earlier than the generic date.

Results:

Simulations were performed using the three models. Using the CATI data, a projected TDE reporting date was determined for each model. The projected "report by" date for each unit was compared with the actual call-in date. If the sample unit did not report by the projected call-in date, then the unit would be "prompted". The TDE call-in date was then compared against the cutoff date for preliminary estimates. If the unit called prior to the cutoff but was scheduled for a prompt, then this was considered a "wasted" prompt, as the unit would have called in on their own.

Figure 4 shows the results of these simulations, along with a comparison with the current method.

The expected response rate was determined by counting all units who called prior to the prompt, plus 60% of the prompted units (the 60% figures was determined based on actual results achieved under the current method).

In comparing the three models, the combination model has a slight advantage in that somewhat fewer units are prompted and the expected response rate is somewhat higher.

The primary weakness of the latest scheduled date model is that 4% of the units would not be prompted until after the first cutoff. While not excessively high, it is considerably higher than the other models. Since no early prompting occurs for these units, there is little sample control.

The relatively simple LP/size model performs remarkably well, with slightly more prompts and a very small differential in the response rate in comparison with the other two models.

Another major objective was to spread out the collection. Figure 5 provides estimates of the number of prompting calls which would be made each day prior to the first cutoff. The estimates assume a sample of 1,000 units (about the average expected for a State), with the same length of payroll/size distribution as the test sample.

All three models provide sufficient spread over the collection period. The combination model has the largest number of calls on a single day, 77 on Thursday (two days prior to the initial cutoff). Even this number is reasonable, requiring about 3 hours of staff time, since these calls generally take less than 2 minutes. This is in contrast to the 265 calls in a single day using current procedures.

Conclusions:

By analyzing the prior reporting pattern of sample units and other characteristics, it is possible to develop a model for targeting nonresponse. This has obvious advantages over more traditional nonresponse methods which simply wait unil a fixed cutoff time. These advantages include reducing the proportion of the sample which require follow-up, lowering collection costs, and providing more efficient utilization of staff resources. Since nonresponse is spread out over a longer time period, the peak associated with nonresponse is reduced and staff can be redirected to other activities.

The research indicates that even a relatively simple model based on known properties of reporting can provide significant improvements in nonresponse priorities. For CES, a simple calling scheme based on the length of the pay period and size of firm performed quite adequately.

Automated collection simplifies nonresponse activities, providing the survey manager with instantaneous access to nonresponse information. Under mail collection, response status is less certain since the survey form may be in the mail or awaiting processing.

While these models were developed specifically for TDE reporting, the approach can be applied to other surveys and collection methods. The key variables related to reporting and distributed lags can be estimated and a nonresponse follow-up scheme developed accordingly. Targeted nonresponse has the advantage of speeding up the collection process, improving respondent cooperation, and ensuring that nonrespondents are followed up in a timely and manner.

Implementation:

The combination model is being implemented for TDE collection in the CES. By determining a precise date when the respondent should have their data available, we will be better able to target the suggested reporting dates which are provided to each TDE respondent and to use the same dates on the advance notice postcard. This will provide for consistency of dates across these vehicles and reinforce the timeliness factor with the respondent.

References:

Werking George S. and Clayton, Richard L., *"Enhancing the Quality of Time Critical Estimates Through the use of Mixed Mode CATI/CASIC Collection"*, Proceedings of Statistics Canada Symposium 90, Measurement and Improvement of Data Quality, October 1990.

Clayton, Richard L., and Winter, Debbie, L.S., *"Speech Data Entry: Results of the First Test of Voice Recognition for Data Collection"*, Proceedings of the American Statistical Association, Survey Methods Section, August 1990.

Clayton, Richard L. and Harrell, Louis J. *"Developing a Cost Model for Alternative Data Collection Methods: Mail, CATI and TDE"*, Proceedings of the American Statistical Association, Survey Methodology Section, August 1989.

Groves, Robert M. et. al., eds. Telephone Survey Methodology, New York: John Wiley & Sons, 1988. Section C.

Kasprezyk, Daniel, et. al., eds. Panel Surveys, New York: John Wiley & Sons, 1989. Part one.

Robert M. Groves, Survey Errors and Survey Costs, New York: John Wiley & Sons. 1989. Chapters 4 and 5.

Figure 1. TDE Response Rates

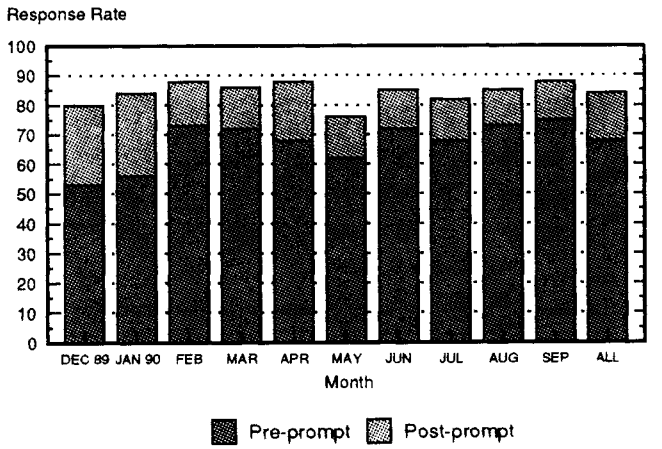

Figure 2. Comparison of Response Rates by Length of Pay, Size, and Collection Days

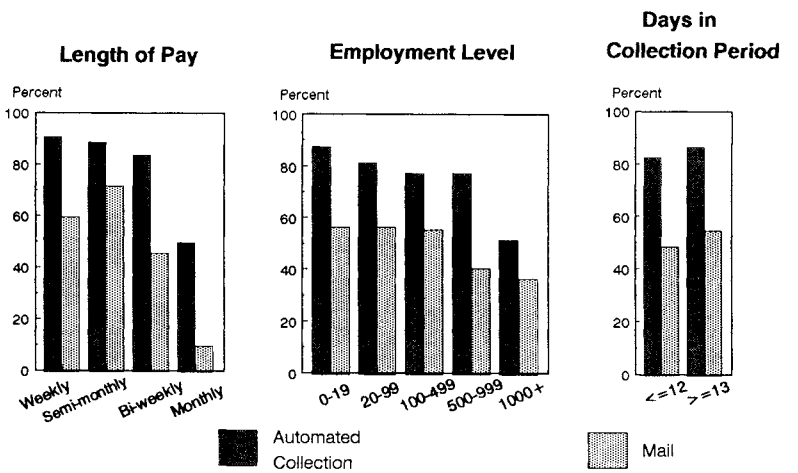

Figure 3. CES Payroll Reporting Process Under TDE

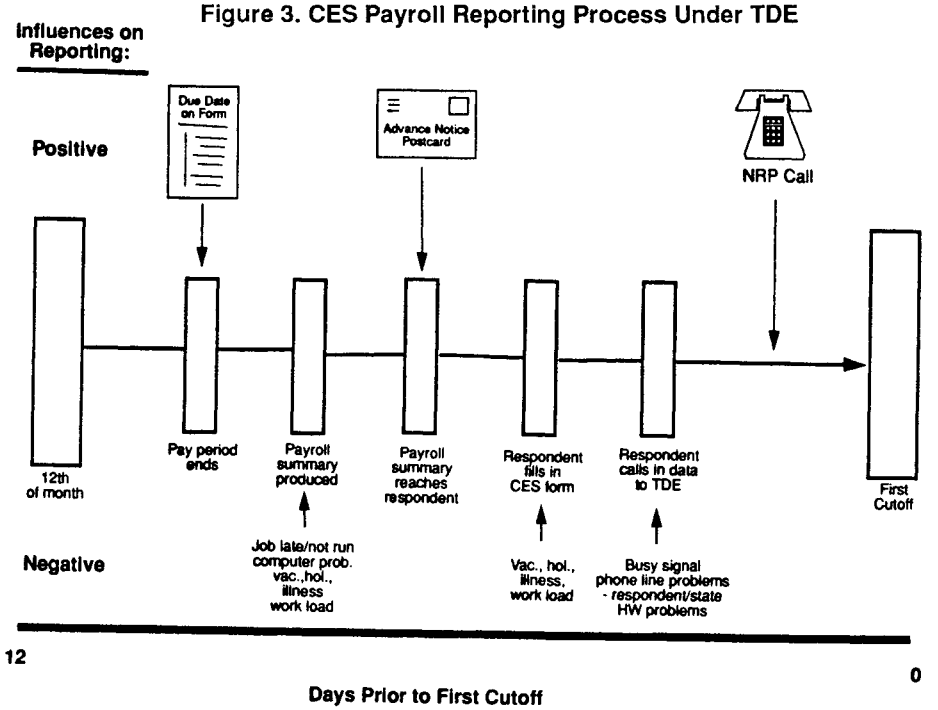

Figure 4. Summary of Nonresponse Model Results

Model:	Prompted	Wasted	Predicted After First Cutoff	Expected Response Rate (1)
Current Method	26.5%	(2)	0.0%	89.4%
Firm Specific	18.0	32.5	4.0	92.1
Group Specific	20.5	23.4	0.0	91.8
Combination	17.3	24.3	1.1	93.1

(1) Assumes 60 percent of the prompts are Successful.
(2) There is no way to predict the excess prompts for the current method.

Figure 5. Number Of Prompting Calls By Day
(Based on a 1000 Unit Sample)