

Long Term Retention of Sample Members under Automated Self-Response Data Collection

Richard J. Rosen, Richard L. Clayton, Lynda L. Wolf
Bureau of Labor Statistics

Richard J. Rosen, BLS, 2 Mass. Ave. NE, Rm. 4860, Wash., DC, 20212

KEY WORDS: CASIC, TDE, Establishment Survey

Background: The Current Employment Statistics (CES) survey is a monthly panel survey of 380,000 business establishments conducted by the Bureau of Labor Statistics (BLS). The CES provides monthly estimates of total employment, production or nonsupervisory worker employment, and hours and earnings. These figures are regarded as key economic indicators of the health of the economy. The figures are used by policy-makers such as the Council of Economic Advisors and the Federal Reserve Board as well as businesses throughout the private sector in evaluating plant location decisions and in determining wage rates.

In order to be as timely as possible, only 10-15 days are available to collect, key enter, and edit the data from the 380,000 sample members before publishing preliminary estimates. Traditionally, collection has been via mail. Under mail, each month the sample unit is sent a form to enter the current month's figures. This form is then mailed back to the State Employment Security Agency which has the responsibility to enter and review the figures then transmit them to BLS for production of national estimates.

Under mail collection, response rates for preliminary estimates average 55%. Revised estimates are released after 3 additional weeks of reporting, with a response rate of about 75%, and a final set of estimates are released after 3 more weeks, with a response rate of about 90%. Because of the low response rate for the preliminary estimates, revisions are often larger than desired.

Conversion to Automated Self-response: In an effort to increase response rates for the preliminary estimates and reduce the magnitude of revisions between the preliminary and revised estimates, BLS has introduced several new automated collection methods, including Computer Assisted Telephone Interviewing (CATI) and Touchtone Data Entry (TDE). The BLS strategy for improving response rates is to convert mail units which routinely miss the cutoff for estimates to CATI for a period of about 6 months.

Under CATI, a trained interviewer calls the sample unit on an agreed upon date and time to collect their data. These CATI interviews are used to "educate" the respondent on the timeliness requirements of the CES survey. At the end of the 6-month period, the respondent is converted to TDE self-reporting.

The touchtone phone has gained acceptability for a wide range of uses in the private sector, for everything from banking transactions to shop-at-home services. Its potential use in the statistical community has only begun to be explored.

Under TDE, rather than receiving a survey form in the mail each month, the respondent is provided a reporting form once a year. They record the information on the form, then call a toll-free number and use the number pad on their phone to report their data.

TDE has several advantages over mail. First, it eliminates the delay inherent with mail delivery and check-in procedures. It also eliminates key entry since the data are captured in machine-readable form. It provides for greater control over the sample. It eliminates any uncertainty about whether the sample unit has reported, making early follow-up of nonrespondents possible.

Under mail collection, there is very little follow-up of nonrespondents prior to initial estimates. This is in large part because it is difficult to ascertain whether the unit has already mailed in the data. Under TDE, nonrespondents receive a brief "reminder" telephone call if they do not report by a pre-determined date based on the firms' past reporting histories while on CATI.

Figure 1 summarizes the growth of the TDE portion of the CES sample. Since 1987, BLS has moved over 45,000 mail reporters to TDE. These units comprise 12% of the sample and account for 25% of sample employment. TDE response rates have averaged about 80%, a 30 percentage point improvement over mail. In a typical month, 65% of the sample will self-report prior to receiving the "reminder" call.

Figure 1. TDE Performance

	Research & Development			Implementation			
	1987	1988	1989	1990	1991	1992	1993
Units	400	600	2,000	5,000	11,000	20,000	45,000
Response Rate	78%	80%	84%	82%	79%	80%	80%

Study of Sample Retention: Because this method of data collection is so new, little is known about the long-term aspects of automated self-reporting. This paper examines the TDE reporting behavior of a large group of TDE reporters which were converted to TDE during 1989. Two basic measures are presented; sample retention and response rates for initial estimates.

To determine the long-term effects, we selected all sample units which converted to TDE during 1989 in nine States. In order to provide as direct a comparison as possible, we also selected a parallel sample of units reporting by mail during the same time period. The mail sample was selected at random from the same States and with the same size and industry distribution. Figure 2 compares the size distribution of the TDE units with the size distribution of the overall CES sample. The distributions are similar.

Figure 2. Distribution of TDE Test Panel and Overall CES Sample by Size of Firm

Size:	TDE & Mail Panel	Overall CES
0-49	48%	66%
50-99	13%	14%
100-249	27%	12%
250+	8%	8%
Total	100%	100%
	711	380,000

Analysis: Sample retention and response rates for preliminary estimates are analyzed across two key variables; size of firm and length of pay period. Size of firm is a traditional variable used in analyzing establishment data. Businesses of different sizes often exhibit different reporting patterns. In the CES, small firms generally have higher response rates for the preliminary estimates than do larger firms. However, previous research has suggested that smaller firms are more likely to drop out of the CES survey, perhaps in

part because of the higher rate of business failures among small firms.

The length of pay period has a significant impact on when the establishments have their data available relative to the CES deadline for receipt of data for use in preparing the preliminary estimates. Therefore, analysis of reporting behavior by this characteristic is also important. Establishments with a weekly or semimonthly payroll are most likely to have their data available shortly after the 12th of the month (the reference period for CES data). Bi-weekly payrolls often are not available until a few days before the BLS cut-off for receipt of data, while firms with a monthly payroll often do not have their data available until after this cut-off date.

Attrition: There are three primary sources of attrition:

- Out of business -- the unit ceases to exist in its current form;
- Overt refusal -- the establishment states they will no longer provide information;
- Cease reporting -- the unit does not provide data despite repeated efforts to obtain information.

Clearly, we cannot control sample loss from units going out of business. However, as survey researchers, we would like to believe that the other two components can be controlled or at least minimized; that through improved methods of contact and education we can retain units who might otherwise refuse or cease reporting.

Because the CES survey is voluntary, BLS must rely on the establishment's "good will" to continue reporting. Establishments often cite time constraints, staff shortages, and respondent burden as reasons for refusing to continue to report.

Another factor in attrition is changes in the contact person providing the data. Because of normal staff turnover, the person responsible for compiling the payroll data for CES reporting can change. When this occurs, there is a potential to lose touch with the respondent. For example, there may not be an immediate replacement for the contact. Even if there is an immediate replacement, the previous contact person may not explain the CES reporting procedures to the new contact. As a result, the new person may not be aware of the CES survey or the reporting procedures and may not have the same level of commitment to timely reporting. This can result in a deterioration/discontinuance of reporting.

In a separate study, we re-contacted 100 severely delinquent TDE units in California. These units had failed to report data for 5 or 6 months. The results of this effort highlight various aspects of nonresponse. In almost one-third of the cases, we found a new contact person was responsible for payroll information. In instances where the contact person had changed, one-half of the units were not familiar with how to report, two-thirds did not have the report form or instructions, and 40% did not recall receiving our advance notice postcard in the past few months. Overall, 4% of the units were out of business, 4% had undergone a change in ownership, and 4% could not be contacted because of disconnected or non-working phone numbers. These may be presumed to be out of business as well.

Figure 3 compares the sample loss over a 3-year period for the mail and TDE panels studied. As can be seen, the panel reporting by mail had a higher rate of attrition than the TDE panel. Mail attrition averaged 7.0% per year for the period, compared with 5.0% for the TDE panel. At the end of 3 years, only 79% of the mail panel remained in the sample compared with 85% of the TDE sample. This difference is statistically significant at the .05 level of confidence.

Figure 3. Attrition: TDE vs Mail

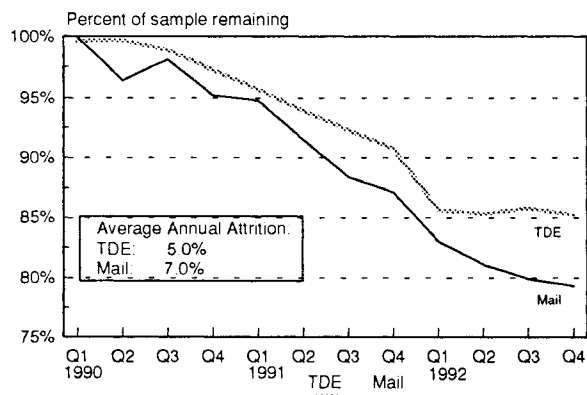


Figure 4 shows the three components of attrition. The out of business rate is estimated at 2.2% per year. This rate is based on direct measurement of the out of business rate experienced for over 10,000 CES units while on CATI, where we obtain direct information from sample members when they go out of business. This rate can also be applied to the mail component, since there is no reason to suspect that the out of business rate would be affected by the mode of data collection.

The overt refusal rate is estimated at 1.3% per year for TDE units. This is based on actual refusals while

conducting nonresponse follow-up calls to late TDE units. While we do not have a direct measure of this component for mail units, we have applied this rate to the mail panel as a rough proxy. Units which overtly refuse to continue participation generally cite such reasons as time constraints and company policy. TDE respondents often cite the convenience of TDE as a benefit. Thus, the overt refusal rate for mail may be somewhat higher than we have estimated.

After deducting these two components, we can estimate the cease reporting rate as the residual. For TDE units the cease reporting rate is estimated at 1.3% per year; for mail, it is estimated at 3.3%.

Figure 4. Attrition: TDE vs Mail

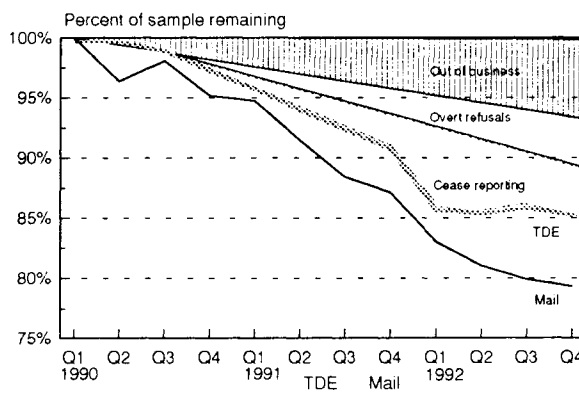
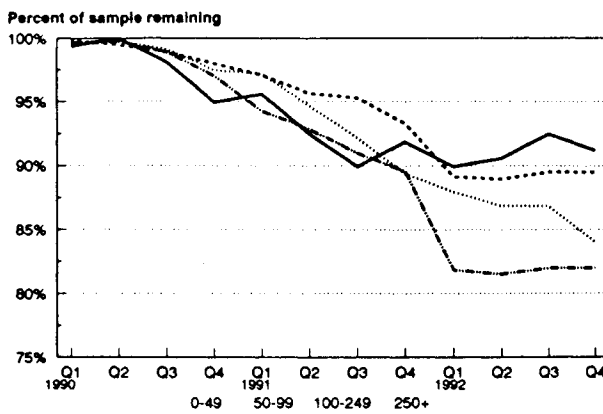


Figure 5 illustrates how attrition by size class shows the expected variation. For the smaller size class (units with employment under 50), only 82% of the original sample remained active after three years. By contrast, for units with employment of 250 or more, 91% of the establishments were still reporting. These results parallel the findings of previous studies of CES sample loss [Werking, Clayton, Rosen, and Winter, 1989].

Figure 5. TDE Attrition by Size of Establishment

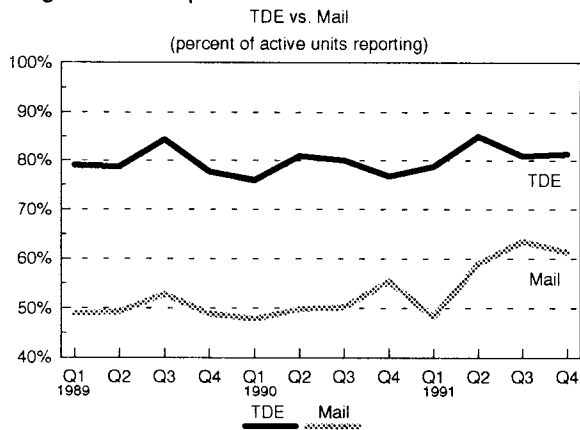


The differences across size class in part reflect the long-standing trend toward higher levels of business failures among small firms and the increased reporting burden placed on small businesses who may not have sufficient staff, time, or resources to continue reporting over a long period of time. Since CES is a voluntary survey, establishments can decline further participation. Larger firms are more likely to have personnel dedicated to completing various government forms and appear less likely to refuse.

There were no remarkable differences in attrition for units with different lengths of pay.

Response Rates: Figure 6 compares the response rates for the mail and TDE panels. Response rates are calculated by dividing the number of units which reported for the initial cut-off by the current "active" sample. Thus, units which have ceased to report for a period of time are eliminated from the base.

Figure 6. Response Rates for Initial Estimates



The response rate for the TDE panel holds steady at about 80% throughout the 3-year time span, while the mail response rate stays in the 50% range. The TDE response rates by length of pay period showed the expected differences. Weekly and semi-monthly payrolls had response rates in the 85-90% range while establishments with biweekly payrolls had response rates of about 70%, and establishments with monthly payrolls had response rates under 50%.

In an effort to understand the dynamics of TDE attrition and response, we classified each unit in the TDE panel according to their performance during their first three months on TDE. Thus, the best performers were units which reported for our preliminary estimates in all three months, while the worst performers were units which failed to report for the preliminary estimates in any of their first 3 months on TDE.

We then looked at attrition and response rates for each of these groups to determine whether units which were timely during their initial months on TDE maintained this level of performance over time, and whether there was a relationship between this timely response and attrition.

Figure 7 provides the results of this analysis. In terms of attrition, there appears to be a correlation between timely reporting during the first months on TDE and remaining in the sample. Eighty five percent of TDE units which reported in time for the preliminary estimates in 2 out of their first 3 months on TDE remained in the sample 3 years later. By contrast, only 35% of the units which did not report for the preliminary estimates remained in the survey after 3 years. However, it should be noted that these poor performers comprise less than 15% of the sample, while the good performers comprise over 75% of the sample. Still, the sample loss among this group is disturbing and suggests that more attention must be paid to these early poor performers.

Figure 7. Comparison of Attrition & Response Rate by Initial Performance on TDE

Initial Performance on TDE Made Preliminary Estimates	Sample	Performance, Oct. - Dec. '92	
		Attrition % Remaining	Response Rate
3	384	88%	87%
2	164	85%	81%
1	72	76%	67%
0	91	35%	60%
Total	711	85%	81%

A comparison of response rates for each group 3 years later shows some positive results. Sample units which were timely during their initial months on TDE remain timely. There is also improvement in timely response for units whose initial performance was poor. Fully 60% of TDE units which did not report for the initial estimates when they began TDE reported for the initial estimates at the end of 3 years.

Summary: Automated self-reporting via TDE appears to be a viable method of reporting. It combines the advantages of high response rates and relatively low attrition. TDE also provides significant cost savings over both mail and CATI [Clayton and Harrell, 1989]. Response rates for active TDE units do not

decline over time and attrition is reduced compared to units reporting by mail.

The success of TDE may lie in part with its multi-contact approach. All units receive a monthly "advance" notice postcard. This helps improve self-reporting. Delinquent units are contacted by phone. Thus we are able to maintain some contact with respondents. A recent study of CES TDE reporters in Texas showed that 83% of the units received 1 or more prompting calls during a 12-month period. Thus, almost everyone receives a call some time during the year. However, most reporters did not require continuous prompting--only 13% of the units required a prompt 9 or more times during the year [Clayton, Rosen, and Rubino, 1993].

Future Research: Efforts should be undertaken to further solidify reporting and reduce attrition among TDE reporters. Special attention must be paid to units during their first few months on TDE in order to resolve any reporting problems. In addition, more rigorous contact of delinquent units should be undertaken. For example, units which do not respond for 2-3 months could receive a structured call to ascertain the reason they are not reporting. This call should focus on identifying the reasons for nonresponse, provide information about the importance of the CES survey, and respond to any questions/problems raised by the sample unit.

Another method of contact could be via FAX. FAX provides a relatively inexpensive, fast, automated means of contacting sample members. Most businesses have FAX machines and are willing to receive messages [Rosen and Clayton, 1992]. Special messages could be designed and sent periodically to all sample units. These messages could provide information which may be of use to them (such as recent trends in earnings for their industry), point out instances where their data were cited in major publications or used in policy decisions, etc. Such messages could help solidify ongoing timely reporting.

REFERENCES

- Clayton, R.L. and L.J. Harrell, Jr. (1989), "Developing a Cost Model of Alternative Data Collection Methods: MAIL, CATI, and TDE," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 264-269.
- Clayton, R.L., R.J. Rosen, and T.B. Rubino, Jr., (1993) "Nonresponse Prompting Behavior in an Establishment Survey," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, in print.
- Groves, R., P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicholls II and J. Waksberg (1988), *Telephone Survey Methodology*, New York: Wiley, Sect. D & E.
- Lyberg, L. and D. Kasprzyk (1991), "Data Collection Methods and Measurement Error: An Overview," in P.P. Biemer, et al., (eds.) *Measurement Errors in Surveys*, New York: Wiley, pp. 237-257.
- Moore, D.L., and R.K. Baxter (1993), "Increasing Mail Questionnaire Response Rates for Surveys of Businesses: A Telephone Follow-up Procedure as an Element of the Total Design Method," Paper presented at the International Conference on Establishment Surveys, Buffalo, New York.
- Ponikowski, C.H., and S.A. Meily (1988), "Use of Touchtone Recognition Technology in Establishment Survey Data Collection," paper presented at the First Annual Field Technologies Conference, St. Petersburg, Florida.
- Rosen, R.J., R.L. Clayton, and T.B. Rubino (1991), "Controlling Nonresponse in an Establishment Survey," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 587-591.
- Rosen, R.J., and R.L. Clayton (1992), "An Operational Test of FAX for Data Collection," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, in print.
- Statistical Policy Working Paper 15 (1988); *Quality in Establishment Surveys*, Office of Management and Budget.
- Statistical Policy Working Paper 19 (1990); *Computer Assisted Survey Information Collection*, Office of Management and Budget.
- Werking, G.S., A.R. Tupek and R.L. Clayton (1988), "CATI and Touchtone Self-Response Applications for Establishment Surveys," *Journal of Official Statistics*, Vol. 4, No. 4, 1988, pp. 349-362.
- Werking, G.S., and R.L. Clayton, "Enhancing Data Quality Through the Use of Mixed Mode Collection," *Survey Methodology*, June 1991, 17, No. 1, pp. 3-14.
- Werking, G.S., R.L. Clayton, R.J. Rosen and D.L.S. Winter (1988), "Conversion from Mail to CATI in the Current Employment Statistics Program," *Proceedings of the Section on Survey Research Methods, American Statistical Association*, pp. 431-436.