

Respondent Acceptance of Web and E-mail Data Reporting for an Establishment Survey

Louis Harrell, Hong Yu, Richard Rosen,
Bureau of Labor Statistics

Keywords: response rate, web, e-mail, Current Employment Statistics

1.0 Introduction

This paper reports on recent efforts to evaluate respondent acceptance of e-mail reporting in the Current Employment Statistics (CES) program. The CES also uses a large enterprise-wide web site for collecting data. The e-mail results are compared to those obtained using the web site.

1. Background on the Current Employment Statistics Survey

The Current Employment Statistics (CES) survey is a monthly survey of employment, payroll, and hours. It is conducted in cooperation with State governments. The sample consists of a probability sample design of 300,000 business establishments. Estimates are published after only 10 to 14 days of collection. The CES is the source of key economic indicators, including employment by industry, state, and area; average hourly earnings, and average weekly hours.

2. Use of E-mail Data Collection in the Current Employment Statistics Survey

The Current Employment Statistics program began experimenting with E-mail data collection starting in October, 2006. The CES program incorporated E-mail as one of the survey's collection methods in February of 2007. The CES program is continually investigating different methods of collecting CES reports. Emphasis is placed on methods that use "off the shelf" software and have the potential to reduce respondent burden. For a production test of e-mail collection, the CES program used a package that had been successfully used for internal BLS employee surveys, Survey Tracker.

Some of the advantages of using e-mail for data collection include the ubiquity of e-mail and the low cost of internet service, the use of off the shelf software, the very low marginal cost of sending an e-mail, and the ability to use the existing network infrastructure. Some of the disadvantages include the

lack of built-in data edits and the possibility that spam blocking software may treat the survey as "spam."

Survey Tracker version 4.5 from Training Technologies Inc., was used to design, distribute, and collect the Current Employment Statistics data by e-mail. Survey Tracker has proven to be sufficient for a production test of e-mail data collection. However, a decision has been made to find different collection software for long term use, as CES requires more comprehensive database support than that provided by Survey Tracker. Full scale CES production requires the ability to easily collect and process thousands of reports and the scalability of the software proved to be problematic during our production test.

2.2 E-mail Data Collection Features

E-mail data collection does not require its users to enter either a user name or password. This feature is highly desirable, as previous research for CES web data collection has shown that security protocols can have a significant impact on response rates.

E-mail data collection is similar to web collection in that each system has a similar screen for data entry. Instead of requiring reporters to logon to a web site, e-mail reporters receive a form embedded in an e-mail and need to fill out the data items.

The reporters receive e-mails with an embedded form in their inbox. (See Figure 1.) When the reporter presses the submit button, an HTTPS transaction is initiated with a BLS web server, which in turn, will re-route the response to a collection machine using a Simple Mail Transfer Protocol. Using an HTTPS connection requires the user to have a live internet connection when entering and submitting data but also allows us to use 128-bit Secure Socket Layer encryption to ensure the security of the sensitive reporter data.

The reporter must use the Microsoft Internet Explorer web browser and Microsoft Outlook e-mail reader.

Figure 1. CES form embedded E-mail

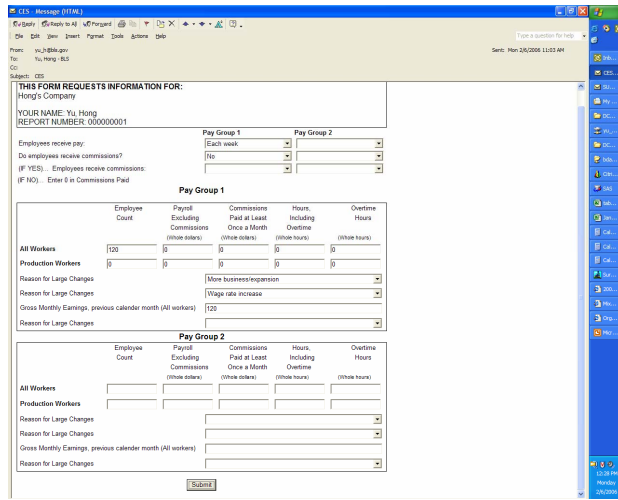


Table 3 shows the changes in collection rate between 1st and 2nd closing. More TDE reporters report after 1st closing than both E-mail and Web reporters. TDE reporters receive non-response prompts by fax and reminder calls and both E-mail and Web reporters receive non-response prompts by e-mails.

Table 3. Changes in Collection Rate Between 1st and 2nd Closing

	TDE	Web	E-mail	Overall
1st Average	69.2%	61.9%	66.2%	66.4%
2nd Average	83.3%	72.0%	79.5%	84.7%
Increase	14.1%	10.2%	13.3%	18.3%

2.3 Collection Rate Comparisons

Table 1 shows the comparisons of 1st closing data collection rate between Touchtone Data Entry (TDE), Web, E-mail, and overall. TDE has the highest average 1st closing collection rate from February of 2007 to July of 2007 at 69.2% and E-mail and Web are 3% and 7.3% lower than TDE respectively.

Table 1. 1st Closing Collection Rate Comparisons

	TDE	Web	E-mail	Overall
02/2007	73.3%	65.9%	74.3%	73.4%
03/2007	73.1%	63.7%	66.7%	67.6%
04/2007	69.2%	59.5%	69.7%	69.3%
05/2007	63.7%	57.0%	55.2%	62.8%
06/2007	68.9%	62.9%	64.3%	63.7%
07/2007	66.8%	62.1%	67.1%	61.6%
Average	69.2%	61.9%	66.2%	66.4%

Table 2 shows the comparisons of 2nd closing data collection rate between TDE, Web, E-mail, and overall. TDE has the highest average 2nd closing collection rate from February of 2007 to June of 2007 at 83.3% and E-mail and Web are 3.8% and 11.3% lower than TDE respectively.

Table 2. 2nd Closing Collection Rate Comparisons

	TDE	Web	E-mail	Overall
02/2007	84.3%	74.2%	82.8%	83.3%
03/2007	84.1%	72.2%	76.8%	82.5%
04/2007	82.8%	69.3%	79.3%	85.9%
05/2007	81.7%	71.7%	78.0%	85.9%
06/2007	83.6%	72.8%	80.5%	86.1%
Average	83.3%	72.0%	79.5%	84.7%

Table 4 shows the data item response rate of Web, E-mail, and TDE from February 2007 to June 2007. The data item response rates appear similar between different collection methods. The E-mail reporters should exhibit similar patterns as the TDE reporters as most of the E-mail reporters have been converted from TDE.

Table 4. Data Item Response Rate Comparisons

	Women Workers %	All Employee Payroll %	All Employee Hour %
Web	84.1%	82.0%	79.8%
Email	88.6%	81.7%	74.5%
TDE	88.7%	76.8%	74.3%

	Gross Monthly Earnings %	Production Worker Payroll %	Production Worker Hours %
Web	76.8%	65.6%	64.9%
Email	75.8%	65.6%	65.2%
TDE	68.2%	67.9%	67.0%

3. Use of Web Data Collection in the Current Employment Statistics Survey

The Current Employment Statistics program began collecting data through the Internet in 1996 (Harrell, 1996; Manning, 1998). It was the first federal survey to collect data via the Web. The CES web collection effort was moved to the BLS enterprise level data collection facility in 2004.

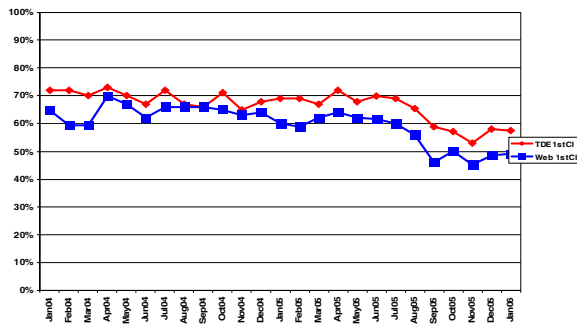
There were several benefits to moving to the enterprise level environment. First, there is only one system to manage for security and continuity of operations. Second, uniform standards for web pages can be used.

Finally, an individual respondent can report for different surveys during one session.

3.1 Evolution of the CES Web Collection Instrument

The original CES web instrument used a complex security paradigm that lowered response rates. In addition, the requirement to use account numbers and passwords or digital certificates generated substantial numbers of help desk inquiries regarding forgotten passwords. (Harrell, 2006). Chart 1 shows a comparison of Touchtone Data Entry and original web response rates. The chart shows that TDE response rates exceeded Web response rates every month for a 2 year period.

Chart 1. Comparison of TDE and Web Response Rates, January 2004-2006

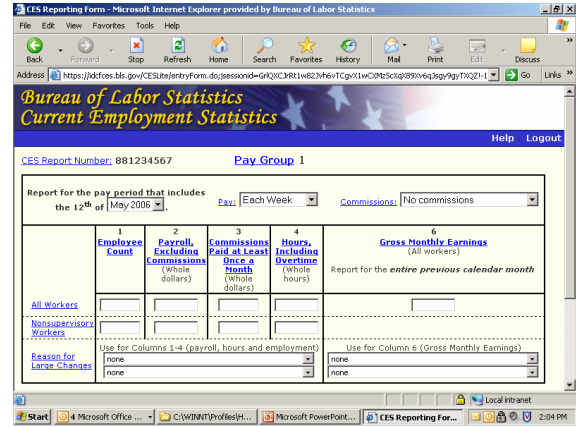


To overcome some of the difficulties encountered with the original web site, a new site, “Web-Lite”, was implemented. Unlike the original site, Web-Lite does not use temporary or permanent account numbers or passwords, and digital certificates are not used. The CES report number, which is printed on the form provided to the respondent, is used to access the site. The goal of the design changes is to make the site more like the Touchtone Data Entry system where the entry point is also the CES report number.

Web-Lite did require some design compromises to accommodate security considerations. Since the login procedure was less complex, BLS security rules required that access to contact information and previously submitted data be eliminated. No respondent identifying information can be displayed and contact information cannot be updated on-line. No historical data is shown. The lack of historical data reduces the visual continuity of reporting. Respondents can no longer visually compare their current data to the previously submitted data. Also, over-the-month change edits cannot be used. Finally,

Web-Lite users cannot directly link to other BLS web surveys. Figure 2 shows the data collection screen.

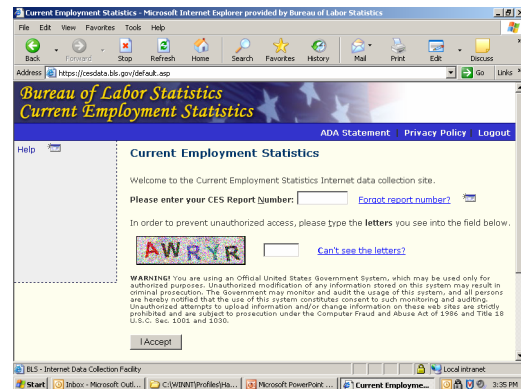
Figure 2. Web-Lite Data Collection Screen



Web-Lite has improved accessibility as well as sufficient security. Transactions are protected with 128 bit Secure Sockets Layer encryption. Collected data are stored on a separate server, and a firewall protects other BLS resources. A Completely Automated Turing Test to Tell Computers and Humans Apart (CAPTCHA) has been added to prevent automated attempts to access the site.

Figure 3 shows the login screen with the CAPTCHA verification. If the CAPTCHA is not readable by the user, a ZIP code option is available. The ZIP code option meets the requirements of the Federal Electronic and Information Technology Accessibility and Compliance Act, which was enacted as the new Section 508 of the Rehabilitation Act of 1973. CAPTCHAs are not accessible to screen readers, so an option to enter a ZIP code was included.

Figure 3. Web-Lite Login Screen.



3.2 Web Data Quality

Table 4 shows the impact of the move from the old web site to the new Web-Lite site. The results are for the initial March 2007 estimate and show that Web-Lite reporters, regardless of the source of the report, have higher item response rates. For example, 4,611 reporters were converted from CATI by our Data Collection Centers (DCCs) to Web-Lite. 52% of these reporters provided Women Workers data. In comparison, 44% of reporters using the old web site provided Women Workers data. The biggest differences between the old web site and new web site were observed with reporters that were converted from Touchtone Data Entry reporting.

The data item response Table 4, in section 2.3, looks at combined data over the February 2007 to June 2007 period. It shows that TDE had higher item response rates for Women Worker, Production Worker Payroll, and Production Worker hour data items, while e-mail and web collection item response rates were higher for the other All Employee related data items.

Table 5 illustrates the impact of switching to the Web-Lite site from the old web site. Regardless of the previous collection source of the Web-Lite report, item response rates for selected data items were higher than obtained using the original web site.

Table 5. Data Item Response Rates, Selected Items, March 2007.

		Women Workers	Gross Month Earnings
	Total	% Received	% Received
DCC to Lite	4,611	52%	45%
State to Lite	51	53%	27%
TDE to Lite	2,331	62%	47%
Old Web to Lite	349	45%	28%
Old Web	1,760	44%	25%

3.3 Summary

The proof of concept in E-mail data collection has yielded a high response rate with a high quality data. The CES program is continuing to investigate alternatives to the current software and is continuing its test of email collection.

References

- Harrell, Louis, R.L. Clayton, and G.S. Werking (1996). "TDE and Beyond: Data Collection on the World Wide Web", *Proceedings of the Section on Survey Research Methods*, American Statistical Association, pp. 768-773.
- Louis Harrell, Richard Rosen, Antonio Gomes, Jason Chute, Hong Yu (2006) "Web versus Email Data Collection: Experience in the Current Employment Statistics Program", *Proceedings of the Section on Survey Research Methods*, American Statistical Association
- Chris Manning, Louis Harrell, and Rick Rosen, (1998) "Web-based Data Collection in the Current Employment Statistics Survey", *Proceedings of the Section on Survey Research Methods*, American Statistical Association.
- Richard Rosen, "Web Evaluation". Presentation at 2005 Federal Computer Assisted Survey Information Collection Conference.
- Training Technologies, Inc.
<http://www.surveytracker.com>.