ESTABLISHMENT SURVEYS: DESIGNING THE SURVEY OPERATIONS OF THE FUTURE

George S. Werking, U.S. Department of Labor, Bureau of Labor Statistics
2 Massachusetts Avenue, N.E., Washington, D.C. 20212

Key Words: E-Mail, Internet, Information Superhighway

Before we can begin a process of re-engineering our survey operations for the future, we must first have a thorough understanding of our primary customers, their operating environment, and their reporting capabilities and flexibilities. For establishment surveys, employers are our primary customers. On voluntary surveys, they must decide whether or not to participate in our survey endeavor, they are the recipients of our complex survey forms and often times confusing definitions and directions, they must devote whatever time and resources that are required to complete our forms and then are faced with follow-back calls when we feel the data they have reported to us cannot possibility be correct.

Over the past decade, we have made many significant improvements in our operations by focusing on the respondent as our customer. We now have tremendous new opportunities for simplifying survey operations for both the respondent and the survey agency. We find ourselves in the middle of a massive transition, or perhaps more accurately a revolution, in the way information is communicated in the business community. The information superhighway is rapidly becoming a reality and this may well lead survey agencies not only to a paper-less collection environment but also to a people-less environment for most of our traditional establishment survey collection and processing operations.

EMPLOYERS OPERATING ENVIRONMENT

Faltering international competitiveness and the subsequent losses in both foreign and domestic markets in the early 1980s forced American businesses to reevaluate their organizational structures, the efficiency of their operations, the quality of their products and their customer focus. This process continues unabated into the 1990s with an ever increasing set of events unfolding in the global economy.

External Forces
Employers are now faced with operating more and more within a global economy. With the introduction of the North American Free Trade Agreement (NAFTA), tariffs are being gradually lowered and employers are being forced to focus on controlling their wage and price structures. The European Community is also attempting to join into a massive new market which when taken as a whole will surpass the U.S. as the world's largest market. China and the Eastern European countries are major new entrants into the world market, each of which will initially have vast markets for goods and services as they build modern economies but correspondingly will also begin to generate competitive products in the future.

Automating Internal Operations
U.S. employers responded to the increasing international competitiveness pressures by downsizing and flattening their organizations, increasing their productivity and controlling their wage and price structures. However, perhaps more importantly, employers responded by investing heavily in computing technology and communications during the 1980s to boost productivity, to link their national and international operations and to provide instantaneous access to critical management information on inventories, personnel, and cash-flow transactions. In 1991, companies for the first time spent more on computing and communications equipment, than on industrial, mining, farm, and construction machines-- as the Industrial Age gave way to the Information Age.

New management philosophies also emerged such as "agile" manufacturing which take advantage of the extensive PC technology and communication investments which have been made. Traditional manufacturing has been based on mass production to drive down fixed costs, however, this also has the negative side-effect of creating large inventories. Under agile manufacturing, much of the product is software controlled and products can be changed overnight to reflect new demands or product improvements. Agile companies stress electronic sharing of information across organizations, alliances with suppliers and customers, lean inventories, and fast responses to changing market conditions.

In the business community, we have entered the age of the information superhighway where instantaneous electronic access and exchange of information is paramount. LANs and INTERNET have blurred

traditional organizational lines creating "virtual enterprises" and provided technology proficient firms with the foundations for a significant competitive edge. These competitive, downsized, highly automated, electronic-information oriented firms are the same ones which we will be contacting to ask that they voluntarily share their information with us and we will need the most flexible, least burdensome approaches to ensure initial acceptance and ongoing cooperation. Our goal is to determine how best to take maximum advantage of the employers new computerized operating environment in re-engineering our survey operations for the future.

SURVEY AGENCIES CURRENT OPERATING ENVIRONMENT

Much like the private sector, survey agencies are also faced with a competition challenge and our reputations and future funding will be based on how successful we are perceived by our customers which include policy makers, the media, and the financial and business communities who use our data. Our products are survey estimates and their quality or useability is based on their relevance within an ever-changing economy, their accuracy, and their timeliness.

To ensure the quality of our products, our survey operations generally include:

- Sample solicitation
- Collection (mailout, mailback, and check-in)
- Nonresponse prompting
- Keypunch and keypunch verification, and
- Screen and edits with reconciliation callbacks

This paper focuses on re-engineering these often times labor-intensive and time-consuming activities to fit into the emerging new streamlined employer operating environment of the future.

Current Survey Operations
Over 80 percent of establishment surveys, on an international basis, are conducted under the "time honored" methods and procedures for mail collection. Many of our current mail survey activities are semi-clerical and labor-intensive in nature such as nonresponse prompting and edit reconciliation and thus, our sample sizes are often limited by the amount of follow-up phone calls we can afford. Additionally, over the years, these mail collection methods and procedures have created a false image that

establishment surveys require many, many months, if not years, to produce and publish survey results.

During the 1980s, in order to address many of the limitations inherent in mail collections operations, we saw a preoccupation with telephone collection methodology. While in earlier years the telephone had been used as a secondary method to mail for prompting and edits, in the 1980s it was catapulted to a primary mode for collection. Methods were developed for CATI collection with discussions of the merits of centralized versus decentralized data collection centers. Methods were developed for computer-based touchtone data collection (TDE) systems for selected surveys and this was immediately followed by equivalent voice recognition (VR) systems. Subsequent research and discussions focused on timeliness of data, significant reductions in direct labor costs for collection, minimal hardware/software investment, and respondent acceptance. FAX transmission also emerged as a mode of collection with encouraging developments for Intelligent Character Recognition (ICR) systems to eliminate the inconvenience of the FAX paper and the subsequent key entry activities.

The 1980s were truly the age of telephone collection, we would share our research results and disseminate our future plans using our new LAN-based E-mail systems which were becoming quickly entrenched in our agencies and also in the private sector. With the broader emerging access to Internet we would share our telephone research with outside agencies and over time, our professional contacts would be equally likely to ask us for our E-mail address as for our telephone or FAX number.

Now, in our day-to-day work environment, E-mail is quickly becoming the preferred choice for communication even over the sophisticated phone systems with voice mail. E-mail is simple, quick, and convenient for both the sender and the receiver. The natural question which then arises is: will this work as simply for survey questionnaires and, if so, what will it mean for our survey operations, organizations, costs, and products.

THE FUTURE: USING E-MAIL WITH INTERNET TO CREATE A PAPER-LESS, PEOPLE-LESS COLLECTION ENVIRONMENT

This paper focuses on the use of E-mail with Internet as the primary future mode of data collection for establishment surveys. With the exception of an

occasional paper on Electronic Data Interchange (EDI), electronic collection, and in particular the use of E-mail, has been largely ignored. However, E-mail embodies all of the strengths of the advanced telephone procedures of the 1980s while at the same time eliminating many of the weaknesses.

It should be noted that E-mail collection, as described below, is quite distinct from what has been termed as EDI. While both would use electronic means (such as Internet) for the transmission of data, the actual respondent-required collection activities are very different under the two electronic approaches. Under EDI the respondent is requested to write software to compute/extract the requested data items and to output transmission records in a mutually agreed upon standard format (i.e., EDIFACT or X-12 standards). Current EDI proposals have been directed towards the collection of large multi-unit firms. However, under the proposed E-mail collection method, the traditional "questionnaire-based" survey approach is preserved and extended to an electronic environment with all traditional labor-intensive and time-consuming mail and telephone activities replaced by their electronic equivalent.

E-mail Survey Operations
Like the automated telephone collection methodologies of TDE and VR, E-mail collection offers major advantages for timeliness, control, and labor cost reduction, however, unlike these procedures, E-mail is far more versatile in the types of surveys and survey operations it can accommodate.

A typical E-mail survey collection cycle of the future would begin with a normal sample control file which now has the respondents E-mail address in addition to the normal respondent contact information of name, address, and phone number. The collection form will be a standard "electronic package" containing an image of the questionnaire, survey instructions, definitions, and any special notes. As the collection cycle approaches, each electronic package is automatically loaded with the respondent specific information, perhaps including data from previous reporting periods, and is E-mailed to the respondent. Under E-mail there are no busy signals, ring but no answer, or data not yet available please call back outcomes. The respondent opens his/her E-mail, fills in the requested data (going to his/her records as required), and E-mails the reply back to the survey agency. Schedules are electronically checked-in and, at predetermined time periods, E-mail nonresponse

reminder packages containing the full original information are automatically sent.

Perhaps more importantly, our current labor-intensive edit and reconciliation operations will also be directly handled under E-mail; allowing for the first time, the respondent to directly review perceived edit failures and correct them as necessary. This will allow the elimination of the large semi-clerical operations of staff poring over reams of computer rejected data and attempting to "correct it" or label it as unusable. The respondent will directly address all edit failure questions either through internal executable edit code in the electronic questionnaire package or through a follow-up E-mail edit-query from the survey agency. Under the E-mail edit-query scenario, as each E-mail questionnaire is returned, it is immediately edited and within the hour a set of user-friendly edit diagnostics are directly E-mailed back to the respondent for verification, correction, or comment code explanation. Under E-mail methodology, most survey data collection operations can be fully automated and the overall process simplified for both the survey agency and the respondent.

E-mail Versatility
Unlike the telephone collection methods of the 1980s, E-mail can accommodate a wide range of surveys and survey operations. The use of telephone collection procedures were often limited by the length and complexity of the questionnaire, the frequency of the collection cycle, and the range of survey operations for which the telephone could be used in a cost effective manner.

Questionnaire Length--CATI was limited to surveys which could be conducted within a 20-minute session and was problematic if respondents needed to refer to their records. TDE and VR were correspondingly limited to the number of items a respondent was willing to push buttons for and the number of questions a respondent was willing to answer to a machine. E-mail, on the other hand, has the ability to accommodate structured questionnaires of any form or length including "form-layout" designs or traditional "question-by-question" designs. The respondent has the ability to refer to records as frequently as needed or to partially complete the questionnaire and return to it at a later time. Under E-mail all of the strengths of mail are merged with the timeliness strengths of the computer assisted telephone procedures.

Survey Frequency--On the surface, E-mail lends itself more naturally to ongoing versus one-time surveys.

Maintaining updated E-mail contact information and conducting periodic cycles for ongoing surveys is easily accomplished for periodic surveys. However, survey agencies can also use the large periodic surveys (which have existing E-mail contacts) as a base for drawing subsamples for also conducting one-time survey supplements, where, if necessary, the E-mail contact can forward the request within the firm to a more appropriate respondent. Large census operations can also obtain current E-mail contact information from the periodic surveys to help automate significant portions of their operations.

Altering Questionnaire Content--E-mail has the same very broad flexibility as traditional mail in easily accommodating content changes (e.g., adding new data items) or conducting periodic survey supplements. The telephone collection operations of CATI and TDE are more limited since they require an immediate answer for the new data item during the interview and this may not be possible if the respondent needs to refer to his/her records. EDI is even more restrictive since it will require each employer to make software changes to produce and output the new data item and may also require a change to the EDIFACT output standard.

Survey Activities--A final added advantage is the ability of E-mail to collapse the various activities of collection, check-in, prompting, and edit follow-back under a single mode of contact, thus simplifying the contact mechanism for both the agency and the respondent.

Cost and Organization Impact
Over the decades we have invested large sums of money to develop and refine the labor-intensive centralized and decentralized operations which help ensure the quality of our estimates, these operations include: collection and collection control, multiple stages and modes of nonresponse follow-up, key entry with verification, and editing with reconciliation of all failures. However, under E-mail reporting, all collection activities can be fully automated and centralized in one room containing a dedicated LAN system. Packages are electronically sent at predetermined dates and information checked-in on a flow basis with edit query packages automatically sent as required. Nonresponse prompting packages are activated also on a fixed schedule as the collection cycle progresses while the respondents handles all key entry and edit verification activities.

The cost-effectiveness of E-mail is a little difficult to fully measure at this time. We do know that many of our labor-intensive activities can be virtually eliminated (Chart 1) and that the time lapse for one-way communication will be reduced from about 5 days under mail to 5 minutes under E-mail (i.e., under one one-thousandth of the mail time). Additionally, we also know that E-mail eliminates the costly telephone line connect-time charges associated with lengthy CATI and TDE interviews since, like FAX, E-mail only requires direct transmission time. However, E-mail transmission costs are unclear since at this time there are no toll booths on the information superhighway.

Product and Customer Service Improvements
The improvements offered by automation and electronic communication will ultimately lead to simplified respondent reporting, more accurate microdata, more timely responses, and improved customer access to our survey products.

Microdata Access--Data quality and access will continue to improve significantly within the establishment. During the 1980s employers began a shift to the use of sophisticated off-the-shelf software and also contracting out for specialized services such as payroll processing. These software packages and services together with electronic communication allow far more accurate and detailed management information to be readily available within the firm. Information which in the past required weeks to become available within the firm is now on-line in a matter of days.

Accuracy--For our surveys, accuracy will also be improved in a number of ways. The microdata from employers will be based more and more on direct computer generated tabulations versus data which are compiled separately from secondary data sources. The respondent will, for the first time, be able to directly respond to all edit queries thus eliminating the need for subjective, atypical treatment by the survey agency of microdata which appear to be out of the normal/expected range. Response rates should also increase since nonresponse prompting can be handled on a far more timely and controlled basis, making the process less vulnerable to publication cut off dates. The unit-cost per schedule will be significantly reduced by the elimination of postage and the many labor-intensive activities under mail and by significantly reducing telephone charges for edit, prompting, and collection calls. This reduction in unit cost can then be redirected towards increased sample size to reduce the level of sampling error for the survey or directed towards other quality-enhancing activities.

Timeliness--Our customers will benefit from more timely data. For some surveys this will mean "final" estimates will be quickly available and thus will eliminate the need for "preliminary" estimates in some surveys, or for others, a reduction in the size of revisions between preliminary and final estimates. Some surveys may be able to publish their data with only a very limited time-lag making the data more relevant to current economic conditions, while others may be able to increase their publication frequency from annual to quarterly or quarterly to monthly.

Customer Service--There will also be many benefits in terms of information dissemination. Our respondents will benefit since we will be able to provide to them, also electronically, a profile of their firm's information against national (or State) industry averages derived from the survey results; examples include employment trend, earnings data and work week hours and overtime. This will allow the firms who participate in the survey to directly follow the performance of their firm against current trends in their industry, thus adding an extra benefit for being a respondent in the survey.

In addition to providing end-result products back to our respondents, electronic communication will provide all users with quick, easy, and cost-effective access to our survey products. Instead of waiting long periods for paper releases (i.e., press releases, periodicals, and bounded volumes), calling or writing for specific tables, or purchasing specialized diskettes; users will have direct menu-driven electronic access to our large longitudinal public-access databases. In the future, this will significantly reduce the labor-intensive overhead associated with our information dissemination activities while providing improved services to the users.

CONCLUSIONS

In our case the future is fairly clear, just as it is equally clear that our mail survey operations are remnants of the past. The information superhighway is not something in the future, with Internet we are traveling on it today. Internet more than doubled in size last year and now has an estimated 20 millions users, of which, approximately 65% are business users. A small Bureau of Labor Statistics (BLS) pilot study (Table 1) indicates that over one-third of all large employers (100+ employees) currently use E-mail in their daily operations. During the mid-1990s these numbers will rapidly expand and we will need to be prepared to transition through a mixed mode mail, telephone, and electronic collection environment. The Bureau's Current Employment Statistics (CES) program will be conducting feasibility tests of E-mail collection over the next 2 years as a future alternative to its current mail, TDE, and VR operations.

During the 1990s, research for mail and telephone survey operations will become relegated to the domain of household surveys, with establishment surveys focusing on direct electronic collection. By the late 1990s, establishment surveys will show a domination by electronic collection with mail and telephone used for only special applications. The information superhighway has now opened up many new dynamic areas of electronic collection research for establishment surveys. The results from this research will ultimately position establishment surveys as the most timely, accurate, and cost effective of the survey operating environments.

Chart 1

DATA COLLECTION COST COMPONENTS
(arrows show direction of recent price change)

Cost Category	Mail (1970s)	Telephone (TDE, VR, and FAX) (1980s)	Electronic Collection (E-mail) (1990s)
LABOR			
mail out	↗	-	-
mail return	↗	-	-
data entry	↗	-	-
edit and edit reconciliation	↗	↗	-
nonresponse follow-up	↗	↗	-
NON-LABOR			
postage	↗	↗	-
telephones	-	↘	-
microcomputers	-	↘	↘

Recent Annual Price Change Factors

Labor	+3.6%	ECI, State and Local Government
Postage	+1.2%	CPI-U, Postage
Telephone	-1.8%	CPI-U, Intrastate toll calls
Microcomputers	-20.9%	PPI, Personal Computers and Workstations

Table 1

AVAILABILITY OF ELECTRONIC MAIL
(n = 4,300 units)

Source: Current Employment Statistics Survey, Bureau of Labor Statistics

Size of Firm		Industry	
0-19	12%	Mining	40%
20-99	15%	Construction	13%
100-249	25%	Manufacturing	23%
250-999	33%	Trade	18%
1000+	43%	Services	20%
		Government	25%
Total	21%	Total	21%

REFERENCES

Holusha, John, "Industry Is Learning to Love Agility", New York Times, 25 May 1994 pp. D1 and D5.

Johnson, Johna T., "The Internet: Corporations Worldwide Make the Connection", Data Communications, April 1994 pp. 67-76.

Stewart, Thomas A., "The Netplex", Fortune Magazine, 7 March 1994 pp. 98-104.

Tetzeli, Rick, "The Internet and Your Business", Fortune Magazine, 7 March 1994 pp. 109-114.

Werking, G.S., A.R. Tupek and R.L. Clayton (1988b), "CATI and Touchtone Self-Response Applications for Establishment Surveys ," *Journal of Official Statistic,* **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.