

**Working Paper No.5**

Topic (ii) Electronic reporting of census data  
and other national experiences

**UNITED NATIONS STATISTICAL COMMISSION and  
ECONOMIC COMMISSION FOR EUROPE  
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION  
STATISTICAL OFFICE OF THE  
EUROPEAN COMMUNITIES (EUROSTAT)**

Joint UNECE/Eurostat Work Session on Electronic Raw Data Reporting  
Geneva, 6-8 November 2006

**ELECTRONIC DATA REPORTING AT THE US BUREAU OF LABOR  
STATISTICS-PROGRESS AND STATUS**

**Michael Searson, Bureau of Labor Statistic, United States**

**INTRODUCTION**

1. The Bureau of Labor Statistics (BLS) is one of several statistical agencies in the United States. It is the principal fact-finding agency for our government in the broad field of labor economics and statistics. The BLS is an independent national statistical agency that collects, processes, analyzes, and disseminates essential statistical data to the American public, the U.S. Congress, other Federal agencies, State and local governments, business, and labor. The purpose of this paper is to note the current status of collecting data in an electronic medium for our surveys and present our efforts to develop strategies designed to significantly increase our electronic collection efforts.

2. As noted above, the United States does not have a unified statistical agency. Within BLS, many of our statistical programs are part of our Federal/State Cooperative Statistical system whereas others are not. The statistical programs within the Federal/State system present a unique challenge in terms of data collection. Since the system's inception more than five decades ago, the states were primarily responsible for these activities with data being collected through a paper collection mode. As most of the firms were locally owned and operated, this approach was certainly appropriate. With a trend toward nationally owned firms with numerous locations (hundreds to thousands), this decentralized method was re-examined. Selected firms in the Federal/State programs were chosen for central collection by BLS staff. The included Federal/State programs are the Current Employment Statistics (CES-monthly); Occupational Employment Statistics (OES-annually); Annual Refiling Survey (ARS-annually) and Multiple Worksite Report (MWR-quarterly)-- components of the Quarterly Census of Employment and Wages program); and, the Occupational Safety and Health Statistics (OSHS-annually) program.

3. The data for the non-Federal/State programs are referred to as the directly collected programs, with BLS staff being totally responsible for these efforts. Included in these programs are the National Compensation Survey (NCS-quarterly), Producer Price Index (PPI-monthly), International Price Program (IPP-monthly), Consumer Price Index (CPI-monthly), and the Job Openings and Labor Turnover Statistics (JOLTS-monthly).

4. It should be noted that the OSHS program is the only survey that employers are required to complete under federal law. The other surveys are all voluntary, with the exception of some Federal/State program surveys whose completion is mandatory under various state laws. The completion of the Multiple Worksite Report form is mandatory in 25 of the 50 states and completion of the CES and OES surveys is required in 5 and 3 states, respectively. These facts are mentioned because it is becoming increasingly difficult to successfully solicit large firms to provide survey data on a strictly voluntary basis. Initially, electronic filing of these data may increase the cost burden on the respondents. This accounts for their reluctance to switch to this medium.

5. To address these issues, the Innovation Board of BLS chartered two teams to study various electronic reporting issues encountered by the surveys mentioned earlier. The purpose of the first team (Electronic Data Collection Expansion Team-EDCET) was to review the current activities of all BLS programs related to electronic data collection and determine whether there were opportunities to expand these methods to other Bureau programs. The second team's (Software/Outsourcing) focus was to determine the feasibility of collecting data for some of these surveys electronically from large employers by concentrating our efforts on the firms who provide payroll and/or human resource software and/or outsourcing services to these large employers. The strategy being that any large firms would be an electronic filing candidate if their provider's system had electronic data capability.

#### **EDCET TEAM**

6. The first team's study revealed that most programs use a variety of methods to collect their survey data, with the amount and methods of electronic data collection varying by program. Most programs still use paper collection and, to a certain extent, phone and on-site personal interviews, mainly to gain the respondent's cooperation. In the phone interviews, most programs are using a Computer Assisted Telephone Inquiry (CATI) system. In the personal interviews conducted on-site for many of our surveys, the data collectors are entering the respondent's data into a laptop computer and then transmitting the data to the survey's database.

7. Included in the electronic method is the use of mailed forms with optical character recognition capability. The information from the returned forms is then scanned into the survey's database. Some surveys also Fax their survey forms to their respondents and then have the completed form Faxed back to their database. Both the optical character recognition and fax-backed form methods capture an image of the survey form in the event that this information needs to be examined during the editing and review process. In addition, some surveys use a mixed mode. The CES program uses a Touch-Tone Data Entry (TDE) system for the employer's response but an e-mail message for the monthly notification process. The ARS uses a paper collection form but advises the respondent that they can use a TDE system if all of the information preprinted on the form is correct.

8. Confidentiality concerns have generally restricted the use of e-mail to collect data from respondents. Only one program is authorized to use this approach for receiving data

from respondents and that is only after the respondent acknowledges the security issues with this method. Other programs can request an exception to this policy on a case by case basis. No confidential data is allowed to be sent to the respondent via e-mail messages or in attachments.

9. In 1995 BLS established the Electronic Data Interchange (EDI) Center in Chicago, IL to collect data files from the nation's largest employers for the CES and MWR programs using a formal EDI approach. At that time, it was thought that file transfers using EDI were going to be an effective method for employers to provide survey data. Under this approach, firms can transmit data files in any format and provide the data layout with the file transmission. The file layout instructs the receiving computer how to parse the file. To date, no employer has provided the EDI Center with a formal EDI file. This approach was never accepted by US firms as a quick, efficient means to report data for payroll/tax filing purposes, which is the source of most of the data required for these two statistical programs. Consequently, the EDI Center instead relies on a number of proprietary formats. The initial proprietary format was established by the MWR program in 1992, three years before the EDI Center opened.

10. The EDI Center has been quite successful as it is very cost effective for both employers and BLS. Firms may send data files by a variety of methods: e-mail (subject to the restrictions stated earlier), File Transfer Protocol (FTP), postal mail, delivery service and a file transfer upload using our Internet Data Collection Facility (IDCF), which will be described later in this paper. The firms are required to use the proprietary formats using a flat file (ASCII) mode. Other types of files (dat., Excel, etc.) are acceptable and are converted into the flat file mode before processing by the EDI Center staff.

11. Data files are then loaded into BLS software systems for editing and dissemination to States. The volume of data collected by the EDI Center has grown significantly since 2000. Each month the EDI Center processes 91 firms for the CES program representing 89,000 establishments with 6.7 million employees. Each quarter, 136 firms supply their MWR data files to the EDI Center, representing 9,000 businesses with 200,000 establishments with 7.5 million employees. The EDI Center now collects 33 percent of all CES and 14 percent of all MWR data. One of the goals of the EDCET team was to explore ways of expanding the EDI Center's collection methods to other BLS programs.

12. BLS established the Internet Data Collection Facility (IDCF) in 1999 and it became operational in 2000. The IDCF was established to provide essential security and gatekeeper functions for respondents seeking to provide data through internet applications for BLS surveys. The IDCF allows respondents to provide data for a specific establishment for a specific survey. In addition, it also provides a secure web environment for large firms wanting to transmit data files for hundreds or thousands of establishments for a BLS survey. Many of the firms using the IDCF file upload feature are using this secure method to transmit their data to BLS for the CES and MWR programs. These data are then transferred to the EDI Center server for processing. Thus, the file upload feature is predominantly being used by firms providing data for the CES and MWR programs.

13. Since the announcement of the opening of the IDCF, there has been considerable interest in internet data collection among the different BLS programs. Almost every survey has at least investigated the feasibility of collecting data using the IDCF Facility and has drawn different conclusions as to its effectiveness. The staff of the ARS concluded that that it would take a respondent longer to register with the IDCF than it would it would take to answer the questions on the survey form, thus increasing the respondent burden. Recently however, the IDCF has addressed this issue. Currently four surveys are using the IDCF for internet data collection; OSH, IPP and CES, and the MWR. The OSHS program has seen the number of respondents in its annual survey rise each year from 10,517 employers, representing 24,093 establishments in 2002 (the first year) to 29,157 employers, representing 51,882 establishments in 2003 to 29,551 employers, representing 50,707 establishments in 2004 with 53,757 employers, representing 70,783 establishments in 2005. The IPP program started internet data collection in December 2003 and have converted 2,600 respondents to this method. This represents 45.0% of their sample respondents. The CES program initiated collection in early 2002. Since that time, their number of respondents has increased to 7,500, approximately 3.0% of their sample. The MWR survey started their IDCF collection with data for the first quarter of 2006, soliciting respondents in four states through a limited test pilot project. To date, the project has resulted in 38 % of the solicited respondents converting from paper to web collection. Plans are to expand this collection to one/half of the states for data being collected for the fourth quarter for 2006 and the remainder of the states the following quarter. Ultimately, MWR staff plan to offer this option to 33,000 employers within the next 18 months.

14. All of the programs involved in the web collection have observed numerous benefits---an increase in data quality due to automatic on-line editing and the users' ability to provide comments to questioned data and correct data after it was initially submitted. The programs have also noted an increase in response rates and a reduction in turnaround time for some surveys. The need to contact survey respondents to clarify reporting issues has decreased, thus lowering the overall respondent burden. BLS staff have noted that, in general, the website appears to be popular with respondents and transitioning surveys from paper to web based collection has been very successful. The only exception to this trend appears to be when transitioning a respondent from TDE to Web, as TDE appears to work well when collecting small amounts of data. The EDCET team is planning to research new options---fillable e-mail forms, expanded use of unencrypted email, and the use of a web form with no prior reported data for that firm. The last item is viewed as possibly being less time consuming for the respondents. Overall, the team also noted that web collection "in the short run, the economies of scale have not yet been achieved; the marginal cost of increasing participation in web collection is not that significant". This situation may change in the near future as more programs are added to this data collection mode and as the number of firms who are providing data for the programs listed above expand significantly.

## **SOFTWARE/OUTSOURCING TEAM**

15. About 18 months ago, the BLS Innovation Board chartered a team to develop an agency-wide strategy for BLS to contact payroll/human resources (P/HR) outsourcing firms and P/HR software vendors to gain their cooperation in supplying data to multiple BLS surveys. The team was directed to develop procedures for identifying the BLS surveys most likely to benefit from such an approach along with the major P/HR firms in the US to target for electronic reporting. The team was also directed to carry out a pilot test of the procedures with at least one P/HR outsourcing firm and at least one P/HR software developer firm. The programs included in this project were the CES and MWR, whose data elements are oriented predominantly toward employment counts, hours worked and wages paid and the OES, NCS and JOLTS programs, whose data elements are more focused on the number of persons employed by occupation, occupational wage rates, benefits provided, and labor turnover records. Information for the first two programs are most likely stored in payroll/tax databases whereas the latter set of programs' data are more likely stored in the human resource database. The premise being that if either type of firm included this electronic data reporting service for their clients or included it in their software programs, all users of these firms' services or software programs could potentially provide data for all programs in an electronic medium in an efficient manner.

16. This approach was based on the successful strategy employed by the MWR program staff since the survey was initiated in 1991. As noted earlier, this program developed a standardized format for employers to provide MWR data through a magnetic/electronic medium in 1992. During a Response Analysis Survey of respondents to the QCEW and CES programs, conducted during 1993, BLS staff became aware of the extremely large number of businesses who either purchased their payroll/tax filing software from a vendor or used the services of payroll/tax filing outsourcing firms. This outsourcing was found to be especially prevalent among the larger firms, a majority of whom would be solicited or potentially required by state law to provide MWR data. Consequently, the MWR staff aggressively pursued the development of relationships with software and/or outsourcing firms. Simply stated, the plan was to spend our limited resources trying to convince these firms to provide the capability for electronic reporting of the MWR data in their systems rather than trying to contact each individual employer that uses their systems or their outsourcing firms or those employers that developed and maintained their own payroll/tax filing systems.

17. The advantages to this MWR strategy are numerous. First, the number of employers that would need to be contacted about the electronic reporting option would be dramatically reduced. Second, the time that each of these employers using the system of a software developer or services of an outsourcing would have to spend on researching, programming, set-up, and testing of changes to their own software would also be reduced. All of these tasks would have been already completed by the software developer or outsourcing firm, thus eliminating a duplication of effort. Simply stated, approaching the software vendor and outsourcing firms was viewed as "having the biggest bang for the buck". For example, if one software firm installs the electronic reporting capability in their system, BLS would potentially be able to collect data electronically from every firm that uses their software package with minimal effort compared to the time and expense of contacting each employer and having them go through the research, cost/benefit analysis, proposal development and testing activities normally incurred by transitioning to electronic reporting.

These statements were validated by a survey of employers that had expressed interest in providing electronic files to the Center but did not complete the process.

18. There are eleven payroll software firms that currently include the capability for electronic reporting of the MWR data in their systems. Discussions are in the initial stages with five other firms to switch their paper MWR product to the electronic medium. In addition, two payroll outsourcing firms offer this service to their clients with additional outsourcing firms recently expressing interest in adding this service for their clients. MWR staff have been successful in convincing these firms to add this feature to their systems for several reasons. However, none of these firms added this capability simply because BLS requested their participation. This electronic reporting feature was added to meet the requests of their clients.

19. MWR staff participate in 2-3 payroll conferences each year that are attended by employers as well as payroll software and outsourcing vendors. These conferences provide an excellent forum to meet and discuss electronic reporting with these parties. MWR staff have a booth in the Exhibit Hall and also have classes on electronic reporting of the MWR data for respondents to attend. In the Exhibit Hall and during the classes, the MWR staff stress that the key to having software or outsourcing vendors add the electronic reporting capability to their system is the clients. The users of the software or outsourced payroll products must be very vocal in wanting this option. Most of these vendors have forums, or "User Groups", which serve to advise the vendors of system problems that the clients have encountered and desired enhancements that the clients are seeking. We stress that the User Group must strongly support this electronic reporting feature. Please note that this strategy may take years to implement and repeated requests of the message are usually necessary.

20. After extensive research, the Software/Outsourcing team identified the largest 18 firms providing either P/HR software or outsourcing services. Based on our RAS in 1993 of the payroll industry, BLS staff were very familiar with most of the payroll software developers and outsourcing firms. Several of these firms were noted as producing all four of these services or software. In fact, it was noted that some of the largest of these firms initially started providing software or outsourcing services in the payroll industry and then started selling their software programs or moved into the HR outsourcing field or vice versa. Likewise, the same observation was noted for some of the HR outsourcing and/or software firms. Some firms had initially only provided software or outsourcing services and then branched out into the other market.

21. The Software/Outsourcing team decided to review all of the data elements for the five different programs (CES, MWR, OES, NCS, and JOLTS) and determine if there was any commonality. The first two programs focus on payroll related data elements whereas the remaining three are more human resource related. In most instances, the data elements did not overlap across all five programs. The CES and MWR programs had already performed this activity over 10 years ago and had developed two proprietary formats for their use. These formats overlap for the first 17 data elements where there is commonality, but they deviate for the remaining data elements for each program. The team decided to use the proprietary format

already established for the CES and MWR programs as a guide and proceeded to add all of the other data elements for the other programs.

22. It was also decided to request that the firms send one electronic file containing all of the data elements each month. The logic being that it would be easier for the firms to set a schedule to send the data by a fixed date each month than try to determine the appropriate time to send the data for each particular program. The latter case would vary as some of the program's data were requested monthly (CES and JOLTS), quarterly (MWR and NCS), and annually (OES with 2 different cycles). The various programs would then retrieve their data elements at the appropriate time from the BLS server that is storing the data. The EDI Center was selected as the logical choice for storing these data.

22. A solicitation letter, signed by the Commissioner of Labor Statistics, and a brochure explaining our project were mailed to each of the identified top 18 firms in January 2006. Unfortunately, none of the firms responded to our initial solicitation letter. The team then attempted to contact the persons to whom the letters had been sent. Many of these follow-ups resulted in firms indicating that they were not interested in participating. More often, team members left messages with the respondents explaining the purpose of the calls and the idea behind the solicitation. For the most part, these individuals never responded to these phone messages.

22. The Software/Outsourcing team was successful in reaching a second stage with two software developer firms, both of which initially indicated some interest in pursuing the project. One firm stated that they would be interested in the project only if it was to be developed in stages, with the first stage including data for the MWR and CES programs. After further internal discussions within their organization, the software firm's staff informed us, about 4-6 weeks later, that they did not have the resources available to spend on developing these programs.

23. With the other software firm, the team was involved in a one hour teleconference with various representatives to explain the project and answer any questions that these development staffers had concerning the project. The team's original proposal included the development of a single electronic file that these outsourcing and/or software firms could send to BLS each month to meet the needs of the five programs. The software representatives noted that some of the data elements being requested were needed every month (CES and JOLTS) whereas others were requested on a quarterly (MWR and NCS) or annual basis (OES). They also noted that some of these data elements would be stored within different databases in their systems---some within the payroll/tax component (mainly employment counts, hours worked and paid--CES and MWR) and others within human resources (staff turnover and occupational salary and coding information---JOLTS, NCS and OES). Information on administrative type information (physical location address and unique identifiers for each worksite) and benefit level information were also stored in different databases.

24. Consequently, accessing these data bases would be required every month under this proposal, even though many of the data elements would only be used by BLS staff four times

during the year or annually. The original idea was that it would be easier for these firms to provide all data elements each month, as opposed to programming the different data elements and their time requirements for each program. If all of the information came in each month, BLS staff would simply use the data when required by their program. The software firm's staff explained, however, that each time they were required to access the different databases, there would be a significant increase in the costs of providing these data. They decided that since many of these data elements were not needed each month, this was not a justifiable cost. It would be better to program the electronic report to be prepared when it was actually needed. They stated that having all of the data elements within the same file format was not a problem.

25. Shortly thereafter, the team pursued a different approach with this software developer, mainly adding the electronic data collection of the CES program to their system. This software developer had already included the MWR electronic filing about 5-6 years ago. Since these program's data elements are very similar in content and scope, it was logical to try a scaled-back approach with them than risk their non-participation in the project. Several months ago, the team was informed by their software contact, that they would include the CES program data in their new system. Unfortunately, the new system will not be available until 2010.

26. To a certain extent, the responses from these 18 firms mirrored the CES and MWR experience soliciting firms to change reporting formats over the past 10 years. We have concluded that these firms are not going to add these electronic reporting options as a service for their clients or in their software for their users unless their clients present a strong demand for their inclusion. The main reason that the clients demand these products is their need to meet state law reporting requirements. The MWR is required to be filed under state laws in twenty-five states. The CES program is required in only five states whereas the OES program in only three states. None of these programs are required under federal law. The main reason that the MWR electronic reporting has been successfully included by 11 payroll systems and is now being offered, at least to a limited extent in some payroll outsourcing firms is the state law requirement. The addition of this capability by the software developers was a direct result of the clients expressing a strong demand for the MWR electronic reporting at their User Conferences. This ultimately led to their inclusion in these systems.

27. The net result of the initial phase of the project is that none of the 18 firms contacted were interested in participating in the electronic reporting project in the next 2-3 years. One firm was interested in adding the CES electronic reporting to their product but not until a new release of the system occurs in 2010.

28. Although the results were not positive, the team did learn some important information about the nature and technical abilities of these firms during the research phase. The team discovered that an independent, non-profit firm was formed several years ago to deal with the exchange of human resource data and information between various parties. Their website notes that their organization is "dedicated to the development and promotion of a standard suite of XML specifications to enable e-business and the automation of human resources-related data exchanges". Their extensive membership list includes at least 8 human



resource software developers or outsourcing firms that were part of our project. Team members have spoken with the main contact at this organization and provided them with the file format and material explaining the purpose of our project. Their representative agreed to review our data elements and to the extent possible, match them with their existing schema to determine any overlap. Discussions are continuing with this organization to determine if this approach will meet our data reporting needs or whether BLS needs to develop an alternative plan. Regardless, the team feels that this approach might be very effective for the human resource related programs---JOLTS, OES and NCS.

29. The team also recently became aware of the Extensible Business Reporting Language (XBRL) approach for transferring large volumes of data between parties on an international level. Most of this information seems oriented toward financial and accounting transactions rather than payroll, tax and employment reporting needs. Even so, this appears to be an approach that should be researched further. Several years ago the Internal Revenue Service and Social Security Administration of the federal government and many of the state governmental revenue collecting and Unemployment Insurance agencies attempted to create an electronic reporting system that all businesses could report their data to using standardized formats acceptable to all parties.

30. Those discussions ultimately led to the development of the “Fed/State Employment Tax (FSET) XML Schema”. Unfortunately, many of the states (Revenue and/or Unemployment Insurance agencies) are not participating in this project, which has resulted in numerous payroll/tax filing software developers and outsourcing firms being reluctant to spending resources on a filing application with limited potential benefit. If 75 % or more of these state agencies accept tax filing information using this standard, then more of these vendors are likely to include the FSET XML schema in their systems. The data elements for the MWR survey are included in this schema. It is reasonable to assume that the data elements, unique to the CES program, could be added to the existing schema if the use of FSET for filing tax reports significantly increases. To date, none of the firms providing data to the EDI Center use this FSET schema format.

31. The Software/Outsourcing team is planning to recommend several options to the Innovation Board. All of these recommendations focus on possibly providing funds to various entities to serve as a catalyst for accomplishing the team’s goals---the inclusion of the electronic reporting of data for these statistical programs in the systems of the outsourcing firms or software providers.

32. The first recommendation is to provide funds to the non-profit XML organization to complete the review of our data elements. Those already defined and mapped would be included in a new taxonomy for statistical reporting for BLS programs. The data elements not included in existing schema would also be defined and mapped to the new taxonomy for statistical reporting. Under this approach, BLS staff could tell potential survey respondents that the data being requested for users of selected software were already identified and tagged in this HR taxonomy, making their extraction from the HR data base far simpler than today. If the software developers or outsourcing firms created a program for their extraction, BLS

staff could advise the respondent to simply run a specific job to create the electronic file required for our statistical programs. All of the software/outsourcing firms that provided this capability could be listed on our website as BLS partners for these surveys. If approved by BLS upper-level management and accepted by the XML organization, this approach could provide a boost in electronic reporting for the NCS, OES and JOLTS programs. This electronic method would be less costly, time consuming, and burdensome to respondents, prompting them to provide the data and facilitate higher data quality.

33. The second recommendation is to re-contact the two software developer firms and perhaps other software payroll developers and ask if they would be interested in a cost-sharing arrangement with BLS to expedite the inclusion of the required programs in their system to facilitate electronic reporting of the MWR and CES data. This proposal was suggested to the two aforementioned firms that initially expressed interest in the project. One firm did not want to pursue this cost-sharing proposal and the other firm is still reviewing the proposal. If this approach is successful, the team recommends pursuing a similar approach with other payroll software developers and outsourcing firms as the XBRL approach has no major governmental support for employment tax reporting purposes at the present time.

34. If the call-backs noted above are unsuccessful, the team recommends that BLS contact firms that specialize in providing output from the larger payroll systems to meet the needs of the CES and MWR programs. The team identified several firms that specialize in writing application programs (known as Application Service Providers) for their clients to meet a specific need. One team member had a specific request from this type of firm to provide CES data elements for their client, a user of one of the software firms in our project. The team feels that if the second software developer that was requested to participate in the cost-sharing proposal declines our offer or if their cost estimate is too high, then BLS should pursue contacting ASPs that specialize in writing programs to access the database of various software clients. Although one specific software firm was used as an example to illustrate this point, the same principle is applicable for all of the payroll software and outsourcing firms on our initial contact list.

35. In conclusion, the BLS is very interested in receiving electronic files from businesses to reduce their reporting burden and costs as well our own. We have explored and employed numerous electronic data collection modes to reduce our costs and increase employer participation in our surveys. Two teams were established to increase the electronic reporting capabilities of all of our programs. Both of these teams advocate BLS expanding their role in working with outside vendors and/or organizations to increase this electronic reporting capability with cost-sharing for developmental work as a potential option.

**The author wishes to acknowledge the assistance of Ms. Emily Thomas of BLS in reviewing this paper and providing numerous constructive suggestions on its content and layout.**

**Note: Any opinions expressed in this paper are those of the author and do not constitute policy of the Bureau of Labor Statistics.**

