

Compilation and Management of Labor Market Statistics for Macroeconomic Analysis
at the U.S. Bureau of Labor Statistics October 2013

Kenneth W. Robertson and John L. Eltinge
U.S. Bureau of Labor Statistics

Abstract

The U.S. Bureau of Labor Statistics (BLS) is the principal Federal agency responsible for measuring labor market activity, working conditions, and price changes in the U.S. economy. Its mission is to collect, analyze, and disseminate essential economic information to support public and private decision-making. BLS serves its user communities by providing data that are objective, timely, and relevant. This paper will provide a brief overview of the compilation and management of labor market statistics at the BLS. This will include a discussion of data collection and data processing practices, with some best practices for publication and dissemination of these key economic data.

1. Key Labor Market Indicators Produced by BLS

The U.S. Bureau of Labor Statistics produces a wide range of data, news releases, and articles to inform government officials, business, academia, students, and others in the public about the state of the U.S. economy. Seven of these news releases have been designated by the U.S. Office of Management and Budget (OMB) as *Principal Federal Economic Indicators* (PFEIs). The PFEIs¹ published by the BLS are

- [Monthly] The Employment Situation, which includes data from two surveys, the Current Population Survey (CPS - the household survey) and the Current Employment Statistics survey (CES - the establishment survey)
- [Monthly] Real Earnings
- [Monthly] Producer Price Indexes (PPI)
- [Monthly] Consumer Price Indexes (CPI)
- [Monthly] Employment Cost Index (ECI)
- [Monthly] U.S. Import and Export Price Indexes
- [Quarterly] Productivity and Costs

These timely economic indicators provide key information to evaluate the state of the U.S. economy. The BLS labor market indicators are, however, not limited to these PFEI news releases. The Bureau also produces a substantial number of non-PFEI data series for the nation, states, metropolitan statistical areas, and counties – this information is highly valued by the many customers that use it.

¹ <http://www.whitehouse.gov/sites/default/files/omb/inforeg/statpolicy/pfei-schedule-of-release-dates-2013.pdf>

Data series, news releases, and documentation for all of these programs are free and available on the BLS website. Together, the BLS' PFEI and non-PFEI data provide an important and timely snapshot of the U.S. economy. These data allow federal, state, and local governments to make informed economic policy decisions. They allow business owners to make informed decisions about business locations, opening and closing of worksites, and about wages, hiring, and layoffs. And they allow private citizens to make informed decisions about areas of study in school, what wages are reasonable to expect when looking for a job, and the state of economy when considering job changes. For examples of BLS data releases, and related supporting materials, see the BLS public website: <http://www.bls.gov/>

2. Standards and Statutes for BLS Operations

2.1 Formal Statutes and Policy Standards

The BLS operates under a number of different statutes and policy standards. The publication *Principles and Practices for a Federal Statistical Agency* (National Research Council, 2013) provides an overview of the general concepts that have led to current standards for U.S. government surveys, and also provides practical guidance following from those concepts. In addition, the principal U.S. government statistical agencies have adopted a joint "Statement of Commitment to Scientific Integrity" based on the *Principles and Practices* book.

More specifically, both data collection and processing at the BLS, and other federal statistical agencies, are based on detailed standards for operating statistical surveys (United States Office of Management and Budget, 2007); guidance on collecting information surveys (United States Office of Management and Budget, 2006); guidance on protecting the confidentiality of respondent data surveys (United States Office of Management and Budget, 2002); standards for classification of industry (United States Census Bureau, 2012), occupation (United States Bureau of Labor Statistics, 2010), geographic area (United States Office of Management and Budget, 2013), and race and ethnicity (United States Office of Management and Budget, 1997); and standards for compiling, disseminating, and evaluating survey data (United States Office of Management and Budget, 1985, 2008).

The BLS also operates under the Confidential Information Protection and Statistical Efficiency Act of 2002 (i.e. Public Law 107-347)². This law prohibits disclosure or release, for non-statistical purposes, of information collected under a pledge of confidentiality. This allows BLS to include a standard pledge of confidentiality with all of its voluntary requests for data:

"The Bureau of Labor Statistics, its employees, agents, and partner statistical agencies, will use the information you provide for statistical purposes only and will hold the information in confidence to the full extent permitted by law. In accordance with the Confidential Information Protection and Statistical Efficiency Act of 2002 (Title 5 of Public Law 107-347) and other applicable Federal laws, your responses will not be disclosed in identifiable form without your informed consent."

² Confidentiality of data collected by BLS for Statistical Purposes - <http://www.bls.gov/bls/confidentiality.htm>

Finally, in keeping with the general principles of statistical transparency, the BLS provides detailed technical information about major programs in the BLS *Handbook of Methods* (United States Bureau of Labor Statistics, 2013). Additional information about methodologies, and reliability is available on the website for each program.

2.2 Methodological Background and Ongoing Research to Inform Technical Standards

The abovementioned technical standards are informed by a large body of methodological literature. In brief, that literature has arisen from efforts to evaluate and balance a wide range of measures of data quality, cost and risk. Technical work has tended to focus primarily on a complex set of multivariate measures of quality. For example, Brackstone (1999) presents a quality assessment framework that encompasses six dimensions: accuracy, timeliness, relevance, interpretability, coherence and accessibility. The first criterion – accuracy – has received extensive quantitative study by methodologists, and has led to the development of detailed “total survey error” models. These models have been based on in-depth theoretical and empirical assessment of several conceptually distinct error components associated, respectively, with imperfect frame coverage; sampling error; nonresponse effects; measurement errors; and processing effects. For general background on total survey error models, see, e.g., Andersen et al. (1979), Biemer and Lyberg (2003), Groves (1989), Lessler and Kalsbeek (1992), Weisberg (2005), and references cited therein.

Survey costs have received some attention in the general methodological literature, e.g., Groves (1989), Karr and Last (2006), Laflamme (2011) and references cited therein; but receive additional attention in agency- and survey-specific studies. Assessment of risks arising in survey processes generally have been studied on a case-specific basis, but can have an important effect on practical decisions related to many components of survey work, including frame development, sample design, fieldwork procedures, data management, and the design and management of production systems.

3. Management of Data Collection at BLS

3.1 Data Collection

Data at BLS are collected using a variety of modes. For example, the Current Employment Statistics survey collects data via Computer Assisted Telephone Interviews (CATI), Electronic Data Interchange (EDI), Touchtone Data Entry (TDE), Web, and in limited numbers we collect data by FAX, Spreadsheets, and email. EDI collects data for companies that have hundreds (or thousands) of worksites. In EDI we work with the company to identify a file format that is easily provided from their computer systems that includes the variables that we need. BLS then creates a company-specific decoder program that reformats the electronic data from that company so that it will feed into our systems. The TDE system is older technology, but a significant number of smaller businesses still prefer to self-report using the phone. For surveys collecting only a few data elements this is still a viable technology. More about CES innovations in data collection can be found in Robertson and Hatch-Maxfield (2012). Managing a multi-mode data collection operation can be very complex. Decisions must be made with respect to optimizing the

response / quality / cost tradeoffs. For CES, the highest quality mode, and the mode that returns the highest response rate is CATI; it is also the most costly mode of data collection for the survey. Therefore, businesses that have a history of on-time quality reporting are quickly transitioned to self-reporting where possible, and other businesses are kept on CATI to the extent of the available resources.

3.1.1 Data Collection and Quality Assurance

Response Rates

BLS programs incorporate a number of quality assurance policies and procedures related to data collection. First among these are the response rate requirement from the Office of Management and Budget. The response rate for data collections that will serve as a sampling frame (e.g. a business register) are required to be 95 percent or higher. Survey response rates are required to be 80 percent or higher, and survey item response rates are required to be 70 percent or higher. Any collection of data that falls below these thresholds must plan and execute a nonresponse bias analysis. This requirement ensures that the survey results are representative of the target population and can be used with confidence to make informed decisions.

Questionnaire Design

Questionnaire design is an important and complex topic. BLS maintains a core staff of research statisticians and cognitive scientists to assist with statistical issues. Several of these staff have expertise in the design of questionnaires, and understand the impact that a change in wording or question order can have on the results. When possible, BLS also enlists academic experts in this field to evaluate and improve the quality of its questionnaires and its data collection processes. For example, the Current Employment Statistics survey recently enlisted the assistance of Dr. Don Dillman (Dillman et al., 2009), who is recognized internationally as a major contributor to the development of modern mail, telephone, and internet survey methods, to evaluate and redesign the questionnaire for the CES survey (Harrell and Park, 2012).

Training

Training, especially periodically repeated training, is a powerful tool to ensure high quality data collection. In order for data to be of high quality, the interviewers must understand the subject matter. They must also understand and follow the protocols for asking the questions the right way and in the right sequence. Interviewers must also be able to help the respondent understand what information is requested if questions arise, and they must be able to record responses correctly. All of these issues are critical to a successful data collection, and they contribute to the quality of the data collected. Training on converting reluctant respondents into participants is also a valuable and cost-effective tool to ensure high response rates. BLS ensures that appropriate periodic training is a component of the data collection process for each of its programs.

Monitoring data collection

It is of course very important to have processes and procedures in place to monitor data collection activities. Procedures should be in place to prevent curb-stoning (i.e. a data collector making up data) and if possible statistical routines might be in place to detect this. For repeated surveys, actual collection over

time should be tracked against previous data collection results. Collection rates of individual data collectors and teams of data collectors can be monitored against targeted goals, with corrective action taken if substantial gaps are found. For first-time or one-time surveys, actual response can be compared against targets derived from similar surveys conducted in the past.

Many surveys will have a supervisor specifically tasked with responsibility for the data collection. This supervisor would be likely to monitor data collection activities on a daily basis. This supervisor would then share management reports on the status of the data collection with the survey manager. The periodicity of management reports monitoring data collection is typically dependent upon the data collection time frame. A monthly survey with a two-week data collection period should have daily reports so that the survey manager and the data collection supervisor can quickly identify problems and take corrective action. A twelve-month data collection period may suggest a report to the survey manager every few weeks. The periodicity should be such that corrective action can be taken in time to have a successfully completed data collection process.

The capability to monitor data collection varies across BLS' programs. Some programs have very mature monitoring programs, which include the ability to listen in on random or specific interviewer calls, refined management reports, and refined routines to detect curb-stoning. Other programs have these capabilities in various states of development.

Editing

Editing of data is a critical component of data quality. Much thought should go into the within-collection unit, cross-sectional, and longitudinal relationships for the data elements collected. For example, a within-unit relationship might be that the number of women employees within a single business can't be greater than the number of all employees at that business. A cross-sectional relationship that might elicit a review of the data is if the ratio of production employees at a particular business is substantially less than the average ratio for all businesses within a selected industry (e.g. manufacturing). A longitudinal relationship that might draw attention is if a particular business reports that it has 15 employees one month, and 155 employees the next month. Business reports that fail one or more of these relationships can be ranked as to the severity of the failure, with follow-ups to data collection and additional reviews targeted at the worst cases first as a way to control resources directed to this activity.

3.1.2 Data breaches

Microdata breaches

Microdata breaches are usually isolated occurrences. For example, a single envelope may be accidentally mailed to a company with another companies request for data, or a single email is incorrectly addressed. At BLS, each of these occurrences is documented, and procedures are reexamined to ensure that appropriate steps are being taken to protect against these microdata breaches. To the extent possible these breaches should be kept to a minimum to guard the agencies reputation as a protector of persons and business' confidential reports.

Macro-data breaches

Macro-data breaches, or a release of estimates prior to the scheduled release data and time, are very serious. Breaches of this nature can have a debilitating impact on the reputation of an agency. One of the key features of a federal statistical agency is the need to provide data to the public in an unbiased manner; this includes providing data to all of the public at the same time. Breaches of this type are, thankfully, rare at BLS. When they do occur, BLS has a management oversight group that is tasked with documenting the occurrence, and to make recommendations to mitigate a reoccurrence.

When a macro-data breach occurs, BLS staff have very specific instructions on what to do. The first step is to inform a supervisor of the occurrence – this notification goes then to the senior agency managers. A quick assessment is made to determine if the breach can be fully contained (for example, the data might be directed early to a single person at another federal agency by mistake). If containment is not an option then a decision is made about how to get the data to the rest of the public as soon as possible, with as little disruption as possible to the media, the financial markets, and other public users of the data.

4. Management of Data Processing at BLS

4.1 Data Processing

4.1.1 Systems design and change control

Processing systems at BLS range the gamut from older mainframe systems (which are close to retirement) to newer state-of-the-art systems. Many of the newer systems utilize a client-server architecture, with a web-browser based client interface. This architecture works very well in the distributed processing environment that many BLS programs use, and ensures that all users are in fact using the same version of the software for processing. In a distributed processing environment this also allows for a mostly client-agnostic configuration of the local hardware and operating system.

A major feature of all major systems at BLS is Change Management. There is a two-part approach to managing established systems. First, the production system is housed in a Production environment. New or updated modules for the system are developed in a development environment. And the new or updated modules are tested in a Test environment. These three environments may be three separate hardware servers, but are more likely to be three virtualized servers on one or two hardware servers. The second part of this approach is the requirement to manage change to the system; this is done using a Change Control Board (CCB). This Board usually consists of senior program managers, management of the programming (Information Technology - IT) staff, and senior staff representing each major group of system users. The CCB reviews and prioritizes proposed system changes. Approved modifications are passed to the IT staff for development. When developed, a team of testers that includes end-users of the system test all existing functionality and the new functionality to ensure that the system is working correctly. After that, a signoff from the CCB allows the updated system to replace the current production system.

4.1.2 Systems access and security

Systems access is controlled in many BLS systems, usually by controlling access both to the servers that the systems reside on, and by controlling access to the system itself. Additional security is required for any data designated by the OMB as a Principle Federal Economic Indicator. Those data are required to have additional security, which includes additional constraints on physical access to the servers that house the data. For BLS, the servers that house PFEI data are located in a secure room with physical access limited to essential personnel.

4.1.3 Processing schedules, training, and documentation

Data processing for most BLS programs is a very structured operation. Written documentation is required that outlines each step in the production process. The step-by-step processing instructions are accompanied by production calendars which indicate when each step is to take place. The calendars are typically available to all affected staff on an internal website.

Processing documentation is developed along with a new system. As initial users utilize the system, they are asked to document the steps they take to perform each job. After that, other staff are asked to perform the job with the experienced staff, but using only the written documentation. Any omission of steps in the written documentation is thereby identified and then corrected. This process continues until a new staff member can perform the job with minimal training.

Training of new staff members to perform processing jobs is typically accomplished as a four-step process. The first step is to provide the person with a survey overview. This way they understand the goals of the survey, the uses of the data, and who the users of the data are. Another goal of the overview is to provide an understanding of survey operations; how the data are collected, what data elements are collected, how data flow from collection systems into the processing system, how data are processed and reviewed, and how estimates get from the processing system into a dissemination system. The second step is to have the new person shadow an assigned mentor. That is, they watch what the mentor is doing as they perform the processing and review steps. After that, the third step is to have the new person and the mentor change places, and the mentor watches the new staff member go through the processing steps. Once the mentor and new staff member agree that processing has been accomplished correctly, the final step is for the new person to perform the processing steps alone – but usually with a senior staff member reviewing the final results of the production processing.

4.1.4 Management oversight of processing

For most BLS surveys, management is directly involved in monitoring the status of production processing and in reviewing the end results. For data collection, there are many management reports that can be developed and reviewed. The frequency of management review of these reports is typically based on the periodicity of the survey.

For estimation systems, management reports should be available for each major task. For example, weighting adjustment, non-response adjustment, calibration to a population total, and production of totals (or means or other aggregate values).

4.2 Macro-data review

The macro-data (estimates) system should also indicate when review steps are completed, and should include an audit trail to track any changes to the data. For example, if the system allows the reviewer to manually designate a micro-data report as atypical (and thereby change its weighting for estimation) the system should track that change. Similarly, if the estimation software allows direct changes to estimate values the change should be tracked in the system and documented.

4.3 Macro-data editing

Estimates produced using collected data should mostly be self-supporting, but there are clearly cases where the collected sample will not be sufficient to produce an appropriate result. This may occur because, for example, a major employee strike occurs in a business that failed to report to the survey, or similarly a major retailer publicly changed its prices (in a way that other survey data do not show) but failed to report to a survey. In cases like these the estimate may require a direct change to the estimated value, but this change and the reason for it should be recorded. Similarly, the identification of unusual reports should be facilitated by either the data collection system or the estimation system (or both). Respondent follow-up about unusual reports is a standard survey operation. Designating unusual reports as atypical (and changing its weight) should be carefully considered, as changes in weighting can have a dramatic impact on estimates; these cases should be documented. This documentation process will allow the estimate to be recreated in the future, if needed, for research purposes. It also provides management an ability to track interventions in the estimates and gauge their overall frequency and impact.

5. Best Practices for Publication and Dissemination of Economic Data at BLS

5.1 Publications

Publication of economic data by a federal statistical agency should be factual, consistent, and unbiased. While analyses will change as different elements of the economy change, the approach to analysis should remain as consistent as possible to mitigate the appearance of political influence in the production and analysis of data. The usual practice is for BLS to note significant changes – positive or negative – in each industry super-sector as part of the monthly employment situation report. Reporting only positive results would be biased, and would tarnish the public trust in the results. BLS' role is to collect, tabulate, and disseminate data in an unbiased manner. BLS does not get involved in the policy and decision process, other than by providing unbiased data to inform those processes.

BLS staff publish articles covering a wide range of topics, centered on its statistical programs. Papers range from scholarly statistical and economic papers for peer-reviewed journals, papers for statistical and economic conferences, articles for the BLS' Monthly Labor Review publication, and shorter articles meant for the BLS website. A general theme is to inform the public about theoretical and applied aspects of BLS procedures, to inform the public about analytical features of BLS data, to highlight major turning points in the economy, and to document specific aspects of BLS programs.

5.2 Dissemination

Data produced by the BLS are a public good, and the aggregate results are made available to the public free of charge as soon as the data are released. News release schedules are published months in advance - if not for the entire year. Having a pre-announced release schedule allows everyone who is interested in the data to have access to it at the same time. Prior access to the data are provided only under the restricted circumstances described in Statistical Policy Directives #3 and #4 (see the references at end). Access to the data may be given to a few members of the press in a lock-up to allow them to have a news story ready to go at release time. This is done to promote an orderly review of the data, and to foster improved public understanding of the data in the news release.

Questions about surveys and the results typically follow soon after the dissemination of data. BLS maintains a press office as one point of contact for the media and the public. The staff of the press office are trained to answer general questions about all BLS surveys. Detailed questions about analysis of results or methodology are usually answered by the economists or statisticians who work on the survey. The public is encouraged to ask questions about the data and methodology - each BLS program website includes contact information to facilitate this.

In addition to the BLS national office in Washington, D.C., BLS also maintains six full-service Regional Offices. These offices employ staff who coordinate the work BLS does with its State partners to produce State, Metropolitan area, and county data; they manage a large part of the BLS' data collection operations; they perform outreach operations to foster understanding of why the data we collect and produce are important; and they participate in the dissemination of Regional, State, and sub-state data. Each Regional Office produces a number of news releases each year, which focus on the economies of the States within the region, and they serve as a local point of contact for the media and public.

5.3 Helpful Analyses and Diagnostics: Moving Averages, Standard Errors and Seasonal Adjustment

An agency should periodically review the analyses that the agency presents on various topics, and provide customers of the data with an opportunity for input on them. Ensuring that the analyses being done are helpful for the customers understanding of the data is an important outreach activity. If the data are too volatile due to small sample sizes, the agency might consider a model-based estimator – e.g. a small domain estimator – to provide some stability to the estimate, or the agency might publish a simple moving average with the primary estimate to provide a less volatile perspective of the data. An agency should strive to provide standard errors with estimates, to inform users of the confidence they might place in them.

Seasonally adjusted data can be helpful to customers, in cases where there is a strong seasonal effect present in the data. Removing the normal seasonal movements can help customers identify changes in the economic cycle earlier. When producing seasonally adjusted data, it is advisable to examine the diagnostic outputs to ensure that an appropriate model has been selected, and that the seasonal adjustment process is performing as intended.

5.4 Disclosure Limitation

In keeping with the confidentiality pledges reviewed in Section 2 above, the BLS has an obligation to take appropriate and prudent steps to ensure that data releases do not inadvertently compromise the confidentiality rights of individual respondents to BLS surveys. Views on appropriate confidentiality commitments by statistical agencies generally vary across countries and depend heavily on societal expectations and the applicable legal framework. For the BLS, confidentiality issues receive special attention because almost all of our surveys depend on voluntary cooperation by respondents, and participation in a survey can depend heavily on the extent to which the respondent is fully confident that their data will not lead to inappropriate disclosure risks.

Technical work with disclosure limitation takes place in two relatively distinct areas. The first area centers on disclosure risks arising from the publication of aggregated data in tabular form. Standard examples include publication of estimated employment counts, wage rates or price indices in tables that are defined by the intersection of several classification variables, for, e.g., geography, industry, occupation or product type. Various U.S. government statistical agencies use a range of methodological approaches for tabular disclosure limitation, including suppression of data for specific problematic cells, addition of random error to some data, or replacement of directly observed data with synthetic data.

The second area centers on disclosure risks associated with the use of microdata, i.e., data associated with an individual respondent. In some cases, U.S. government statistical agencies release microdata subject to certain technical restrictions, e.g., topcoding of some sensitive variables, omission of certain design-related information, or use of perturbed or synthetic data. In other cases, statistical agencies allow researchers to have relatively broad access to microdata within a carefully controlled computing environment, subject to constraints on release of the aggregate-level estimates produced by these researchers.

6. Analyzing BLS' Major Data Series – Where is the U.S. Economy Now?

Analyzing BLS data series is, as mentioned earlier, a fact-based and unbiased assessment of the results of our surveys. Decisions about the data to publish, and about the format of publications, are based on extensive consultation with data users. The examples listed below illustrate some of the primary features that are emphasized in several BLS news releases.

6.1 Employment and Unemployment

On October 22, 2013, the BLS reported that total nonfarm payroll rose by 148,000 in September 2013, and the unemployment rate was little changed at 7.2 percent. Temporary help services, typically considered a leading indicator of employment change, added 20,200 jobs. Health care employment added 6,800 jobs; this industry has shown a slight deceleration in job gains per month so far in 2013 (18,500) compared to the monthly job gain in 2012 (26,700). In September, employment increased in construction (20,000), wholesale trade (16,100), and transportation and warehousing (23,400).

In September, the number of long-term unemployed (persons without jobs for 27 weeks or more) was little changed; these individuals accounted for 36.9 percent of the unemployed. The number of persons

employed part time for economic reasons (sometimes referred to as involuntary part-time workers) was unchanged at 7.9 million; these individuals were working part time because their hours had been cut back or because they were unable to find a full-time job. In addition to the usually reported unemployment rate, the BLS also publishes a number of alternative measures of labor underutilization (Table A-15 in The Employment Situation news release). These measures provide rates based on alternative definitions. The broadest measure, referred to as U-6, is the total unemployed, plus all persons marginally attached to the labor force, plus total employed part time for economic reasons, as a percent of the civilian labor force plus all persons marginally attached to the labor force. In September 2013, the U-6 measure was 13.6 percent, little changed from August 2013 (13.7 percent).

On October 24, 2013, the BLS reported that there were 3.9 million job openings on the last business day of August, little changed from July. The hires rate (3.3 percent) and separations rate (3.2 percent) also were little changed in August.

6.2 Prices

On September 17, 2013, the BLS reported that the Consumer Price Index for All Urban Consumers (CPI-U) increased 0.1 percent in August on a seasonally adjusted basis. Over the last 12 months, the all items index increased 1.5 percent before seasonal adjustment.

6.3 Wages & Earnings

On July 31, 2013, the BLS reported that compensation costs for civilian workers increased 0.5 percent, seasonally adjusted, for the 3-month period ending June 2013, following a similar 3-month percent increase in March. Wages and salaries rose 0.4 percent in the June quarter, similar to the 0.5 percent increase for the previous period.

On September 17, 2013, the BLS reported that real average hourly earnings for all employees rose 0.1 percent from July to August, seasonally adjusted. Real average weekly earnings rose 0.4 percent over the month due to the increase in real average hourly earnings and a 0.3 percent increase in the average workweek.

6.4 Productivity

On September 5, 2013, the BLS reported that nonfarm business sector productivity increased at a 2.3 percent annual rate during the second quarter of 2013. The increase in productivity reflects increases of 3.7 percent in output and 1.4 percent in hours worked.

7. Concluding Remarks

One of the most important issues, that has not been addressed yet, is the importance of the staff of a federal statistical agency. The importance of hiring educated, motivated staff cannot be overstated. BLS provides its excellent staff with opportunities to attend University courses to further their education when it contributes to the work of the BLS. Examples include technical training in economics, methodology, information technology, and project management as well as continuing education in non-technical areas

like communication and leadership. Also, staff are encouraged to do independent or team-based research on projects that further the BLS mission, and then to report on the results of the research in appropriate statistical and economic venues. Mentoring of junior staff by more senior staff is encouraged, and in this manner the BLS tradition of excellence is passed on.

8. References

- Andersen, R., J. Kasper and M.R. Frankel (1979). *Total Survey Error*. San Francisco: Jossey-Bass.
- Biemer, P.P. and L.E. Lyberg (2003). *Introduction to Survey Quality*. New York: Wiley.
- Brackstone, G. (1999). Managing data quality in a statistical agency. *Survey Methodology* **25**, 139-149.
- Dillman, Donald A., Jolene D. Smyth and Leah Malani Christian (2009). *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method, Third Edition*. Available through:
[http://www.amazon.com/Internet-Mail-Mixed-Mode-Surveys Tailored/dp/0471698687#reader_0471698687](http://www.amazon.com/Internet-Mail-Mixed-Mode-Surveys-Tailored/dp/0471698687#reader_0471698687)
- Groves, R.M. (1989). *Survey Errors and Survey Costs*. New York: Wiley.
- Groves, Robert M. and Heeringa, Steven G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, 169, 439-457
- Harrell, Louis J. Jr. and Edward Park (2012). *Revisiting the Survey Form: The Effects of Redesigning the Current Employment Statistics Survey's Iconic 1-Page Form with a Booklet Style Form*. Proceedings of the Section on Survey Research Methods, Joint Statistical Meetings. Available through
http://www.amstat.org/sections/SRMS/proceedings/y2012/Files/306931_77657.pdf
- Heeringa, Steven G. and Robert M. Groves (2004). *Responsive Design for Household Surveys*, Available through: <http://www.isr.umich.edu/src/smp/Electronic%20Copies/127.pdf>
- Herzog, T.N., F. J. Scheuren and W. E. Winkler (2007) *Data Quality and Record Linkage Techniques*. New York: Springer-Verlag.
- Karr, A. and M. Last (2006). Survey costs: workshop report and white paper. Available at:
<http://www.niss.org/sites/default/files/surveycost-whitepaper200609-final.pdf>
- Laflamme, F. (2011). Paradata Collection Research for Social Surveys at Statistics Canada. Presentation to the International Total Survey Error Workshop, June, 2011. Available at:
<http://www.niss.org/events/itsew-2011>
- Lessler, J.T. and W. D. Kalsbeek (1992). *Nonsampling error in surveys*. New York: Wiley.

National Research Council (2013). *Principles and Practices For A Federal Statistical Agency Fifth Edition (2013)*. Committee on National Statistics. Constance F. Citro and Miron L. Straf, Editors. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press. Available as a paperback for purchase or as a free download from http://www.nap.edu/catalog.php?record_id=18318

Robertson, Kenneth W. and Julie Hatch-Maxfield (2012). *Data Collection in the U.S. Bureau of Labor Statistics' Current Employment Statistics Survey*. Paper presented at the UN-ECE Seminar on New Frontiers for Statistical Data Collection (2012), Available through: <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2012/mtg2/WP20.pdf>

Statistics Canada (2009). National Quality Assurance Framework. Paper prepared for the 41st Session of the U.N. Statistical Commission.

United States Bureau of Labor Statistics (2010). Standard Occupational Classification, <http://www.bls.gov/soc/>

United States Bureau of Labor Statistics (2013). BLS Handbook of Methods, <http://www.bls.gov/opub/hom/>

United States Census Bureau (2012). North American Industry Classification System, <http://www.census.gov/eos/www/naics/>

United States Office of Management and Budget (1985). Statistical Policy Directive (#3) on Compilation, Release, and Evaluation of Principle Federal Economic Indicators, http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/dir_3_fr_09251985.pdf

United States Office of Management and Budget (1997). Classification of Race and Ethnicity, http://www.whitehouse.gov/omb/fedreg_1997standards/

United States Office of Management and Budget (2002). Statistical Confidentiality, http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/cipsea/cipsea_statute.pdf

United States Office of Management and Budget (2006). Guidance on Statistical Information Collections, http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/pmc_survey_guidance_2006.pdf

United States Office of Management and Budget (2007). Standards and Guidelines for Statistical Surveys, http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/standards_stat_surveys.pdf

United States Office of Management and Budget (2008). Statistical Policy Directive No. 4: Release and Dissemination of Statistical Products Produced by Federal Statistical Agencies, http://www.whitehouse.gov/sites/default/files/omb/assets/omb/fedreg/2008/030708_directive-4.pdf

United States Office of Management and Budget (2013). Classification of Metropolitan Statistical Areas, <http://www.whitehouse.gov/sites/default/files/omb/bulletins/2013/b13-01.pdf>

Weisberg, H.F. (2005). *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: University of Chicago Press.

Appendix A: Statement of Commitment to Scientific Integrity by the U.S. Principal Statistical Agencies

In 2012, the U.S. Principal Statistical Agencies issued a joint statement of commitment to scientific integrity, and each of those agencies placed the statement on their respective public websites. The BLS statement is available at: <http://data.bls.gov/cgi-bin/print.pl/bls/integrity.htm> and is reproduced below.



BLS Information

Statement of Commitment to Scientific Integrity by Principal Statistical Agencies

Our Nation relies on the flow of objective, credible statistics to support the decisions of governments, businesses, households, and other organizations. Any loss of trust in the integrity of the Federal statistical system and its products can foster uncertainty about the validity of measures our Nation uses to monitor and assess performance and progress.

Federal statistical agencies (or units) whose principal function is the collection, analysis, and dissemination of information for statistical purposes have set for themselves a high standard of scientific integrity. The following agencies are designated as "principal statistical agencies"^[1]:

- Bureau of Economic Analysis (Commerce Department)
- Bureau of Justice Statistics (Justice Department)
- Bureau of Labor Statistics (Labor Department)
- Bureau of Transportation Statistics (Transportation Department)
- Census Bureau (Commerce Department)
- Economic Research Service (Agriculture Department)
- Energy Information Administration (Energy Department)
- National Agricultural Statistics Service (Agriculture Department)
- National Center for Education Statistics (Education Department)
- National Center for Health Statistics (Health and Human Services Department)
- National Center for Science and Engineering Statistics (National Science Foundation)
- Office of Research, Evaluation, and Statistics (Social Security Administration)
- Statistics of Income Division (Treasury Department)

These agencies embrace a common set of professional standards and operational practices designed to ensure the quality, integrity and credibility of their statistical activities. Implementation of these professional standards involves a wide range of managerial and technical challenges.

Principles and Practices of Statistical Agencies

To address these challenges, the National Research Council of the National Academies (NRC) has developed practical guidance in its publication, [Principles and Practices for a Federal Statistical Agency](#)^[2]

The principal statistical agencies use this volume to guide their strategic planning, daily operations, and interactions with stakeholders. The principal statistical agencies embrace the four fundamental principles articulated in the Fourth Edition:

Principle 1: a Federal statistical agency must be in a position to provide objective information that is relevant to issues of public policy;

Principle 2: a Federal statistical agency must have credibility with those who use its data and information;

Principle 3: a Federal statistical agency must have the trust of those whose information it obtains;

Principle 4: a Federal statistical agency must have a strong position of independence within the government.

Actual and perceived violations of any of these principles undermine the scientific integrity of, and public confidence in, the data produced by principal statistical agencies. Of special note is the emphasis that the NRC publication places on the impartiality and independence of each statistical agency. The NRC discussion of independence includes the following.

- Independence must include separation of the statistical agency from the parts of its department that are responsible for policy-making or law enforcement activities.
- Independence must include control over personnel actions, especially the selection and appointment of qualified professional staff, including senior executive career staff.
- Independence must include the statistical agency having authority for professional decisions over the scope, content, and frequency of data collected; analysis, or publishing of the information; authority to release statistical information without prior clearance; and adherence to predetermined schedules for public release of statistical information.
- Independence must also include the statistical agency's ability to control information technology systems used in collection, storage and dissemination of statistical information, since such control is essential for ensuring adherence to laws and regulations requiring appropriate protection of data collected under a promise of confidentiality.

The principal statistical agencies also subscribe to the 11 practices identified by the NRC as critically important in the application of these principles:

Practice 1: A Clearly Defined and Well-Accepted Mission;

Practice 2: Continual Development of More Useful Data;

Practice 3: Openness about Sources and Limitations of the Data Provided;

Practice 4: Wide Dissemination of Data;

Practice 5: Cooperation with Data Users;

Practice 6: Fair Treatment of Data Providers;

Practice 7: Commitment to Quality and Professional Standards of Practice;

Practice 8: An Active Research Program;

Practice 9: Professional Advancement of Staff;

Practice 10: A Strong Internal and External Evaluation Program;

Practice 11: Coordination and Cooperation with Other Statistical Agencies.

All of these practices are important to achieving and safeguarding scientific integrity. Implementation details of these practices vary across agencies.

Statistical Policy Directives and Standards

The Principles and Practices are closely related to Statistical Policy Directives and other standards issued by the Office of Management and Budget (OMB) in its role as coordinator of the Federal statistical system (44 U.S.C. 3504(e)). Specifically, OMB's directives and standards are designed to preserve and enhance the objectivity, utility, and transparency, in fact and in perception, of the statistical products themselves and the processes used to release and disseminate them. Examples include:

- **Statistical Policy Directive Number 3**^[3], which is intended to preserve the time value of principal economic indicators, strike a balance between timeliness and accuracy, prevent early access to information that may affect financial and commodity markets, and preserve the distinction between the policy-neutral release of data by statistical agencies and their interpretation by policy officials.
- **Statistical Policy Directive Number 4**^[4], which enumerates procedures intended to ensure that statistical data releases adhere to data quality standards through equitable, policy-neutral, and timely release of information to the general public.
- **Standards and Guidelines for Statistical Surveys**^[5], which documents important technical and managerial practices that Federal agencies are required to adhere to, and the level of quality and effort expected in all statistical activities to ensure consistency among and within statistical activities conducted across the Federal Government.

Information Quality Guidelines

The Principles and Practices are also related to the Information Quality Act^[6], which requires OMB and all other Federal agencies to maximize the quality, objectivity, utility, and integrity of information, including statistical information, provided to the public. Through the adoption and implementation of the Government-wide Information Quality Guidelines^[7], each statistical agency:

- maintains its commitment to using the best available science and statistical methods;
- subjects information, models, and analytic results to independent peer review by qualified experts, when appropriate;
- disseminates its data and analytic products with a high degree of transparency about the data and methods to facilitate its reproducibility by qualified third parties;
- ensures that the presentation of information is comprehensive, informative, and understandable.

Using Best Scientific Methods to Ensure Data Quality and Integrity

Finally, scientific methods play a critical role in maximizing the quality, objectivity, and credibility of information collected and disseminated by the principal statistical agencies. Examples of the application of scientific methods include probability sampling designed to avoid biased samples and randomized assignment for assessing the impacts of alternative protocols or question wording. Measures need to be valid and reproducible, and interpreting variations in these measures across data sources requires scientific knowledge of their properties. Combining data from disparate sources, such as surveys and administrative records, is increasingly important given the growing reluctance of respondents to provide data in Federal data collection efforts. These efforts require statistical modeling, as does the provision of statistics on small areas where direct estimates from surveys may be subject to large sampling errors. The procedures, equations, and assumptions, which define these models, must be publicly available to ensure that the information is presented in an accurate, clear, complete, and unbiased manner. Moreover, Federal statistical agencies apply complex statistical methods to the information that is publicly released to protect the confidentiality of data about respondents to Federal surveys.

Methodological improvements and rigorous approaches to data collection and analysis require the application of scientific methods. Computer scientists, demographers, economists, geographers, mathematicians, survey

statisticians, and other scientists are needed for producing high quality, objective statistics from surveys or administrative data. Subject area experts, such as epidemiologists and engineers, are also needed to maximize data quality. Research and methodological innovation are required to continuously improve the quality and scope of our data products while protecting privacy and ensuring confidentiality. All of the above mentioned factors are critically important to ensuring the credibility of Federal statistical agencies.

Conclusion

We reaffirm our commitment to maintaining the highest level of scientific integrity in producing official statistics. Together, the *Principles and Practices, Statistical Policy Directives and Standards, and Information Quality Guidelines* form the foundation for achieving and maintaining scientific integrity within and among the principal statistical agencies.

Notes

^[1]The Office of Management and Budget designates principal statistical agencies. The exact number of units considered principal statistical agencies has evolved over time. The most recent listing of such agencies is available in the Office of Management and Budget's *Statistical Programs of the United States Government*, available at: http://www.whitehouse.gov/sites/default/files/omb/assets/information_and_regulatory_affairs/12statprog.pdf.

^[2]National Research Council. *Principles and Practices for a Federal Statistical Agency: Fourth Edition*. Washington, DC: The National Academies Press, 2009. Available at: http://www.nap.edu/catalog.php?record_id=12564.

^[3]http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/dir_3_fr_09251985.pdf; 38932 Federal Register / Vol. 50, No. 186 / Wednesday, September 25, 1985.

^[4]http://www.whitehouse.gov/sites/default/files/omb/assets/omb/fedreg/2008/030708_directive-4.pdf; 12622 Federal Register / Vol. 73, No. 46 / Friday, March 7, 2008.

^[5]http://www.whitehouse.gov/sites/default/files/omb/assets/omb/inforeg/statpolicy/standards_stat_surveys.pdf; announced in 55522 Federal Register / Vol. 71, No. 184 / Friday September 22, 2006.

^[6]Section 515 of the Treasury and General Government Appropriations Act, 2001 (Pub. L. No. 106-554, 44 U.S.C. § 3516 note).

^[7]These guidelines are available at: <http://www.whitehouse.gov/omb/fedreg/reproducible2.pdf>.

Last Modified Date: May 23, 2012