STANDARDS FOR STATISTICAL SURVEYS FOR PRODUCING AND DISSEMINATING SURVEY ESTIMATES IN THE U. S. STATISTICAL SYSTEM

Kevin Cecco; Statistics of Income Division, U.S. Internal Revenue Service
500 N. Capitol St. Washington, DC 20001
Kevin.Cecco@IRS.GOV
and
Stephen Cohen; Bureau of Labor Statistics; 2 Massachusetts Ave. NE, Suite 1950;
Washington, DC 20212;
Cohen.Steve@BLS.GOV

Key words: Quality of Statistical Systems; Current Best Methods, Minimum Standards, and Recommended Practices

Introduction

The United States statistical system is highly decentralized, consisting of more than seventy agencies and units that engage in some statistical activities—ten of these agencies have statistical activities as their principal mission. While this structure provides substantial benefits in ensuring responsiveness to program needs, it makes it more difficult to ensure that agencies are uniformly applying sound statistical techniques and best practices. The U.S. Office of Management and Budget (OMB) is charged with developing and overseeing the implementation of government-wide policies, standards, and guidelines for Federal statistical programs.

The current standards for surveys and dissemination of survey estimates were issued by OMB as *Directive 1 and 2* respectively more than 20 years ago. OMB formed an interagency committee consisting of members from eight agencies that have statistical activities as their principal mission. Under the auspices of the Federal Committee on Statistical Methodology (FCSM), this committee was charged with developing recommendations for updating and revising standards for statistical surveys. The agencies represented on the committee included those that produce a wide variety of economic and demographic statistics including economic activity, employment, health, education, agriculture, energy, tax and transportation. Some of the agencies represented on the committee have their own detailed written standards for their agency's statistical surveys, while others did not have formal documentation or standards that applied across programs within their agency. The existing agency written standards served as the initial structure that was used for updating and revising the statistical standards.

This paper will describe the process the interagency committee went through in updating and revising the statistical standards as well as provide a detailed discussion of standards covering the topics such as data processing, survey estimation, data analysis, and data dissemination. Another paper by Seastrom and Madans (2004) will serve as a parallel paper, addressing the other standards pertaining to survey planning and data collection. The committee

worked diligently to create a common set of standards that will assure the quality of surveys conducted and sponsored by U.S. Federal agencies while recognizing and accommodating some differences in practices among different organizations. Examples of differences include: (1) some agencies conduct many large scale surveys across various topics while other agencies are more focused on specific topics (e.g., health, agriculture, education, energy); (2) certain agencies conduct many demographic surveys while other agencies are almost exclusively focused on economic surveys; and (3) some agencies use contractors to conduct all their surveys while other agencies rely on Federal staff to conduct their surveys.

Objectives of the New Standards

The objective of the new standards is to ensure that agencies adhere to a consistent set of statistical practices when collecting and disseminating data from statistical surveys. Revising the standards also provided the opportunity to bring the standards up to date—ensuring the standards reflect our present views of statistical rigor and quality. The proposed standards for producing and disseminating survey estimates contain 13 specific standards designed for general application to U.S. Federal statistical survey activities. Each standard is accompanied by guidelines that represent best practices that may be useful in fulfilling the goals of the standards. These standards and guidelines for U.S. Federal statistical surveys support agencies in achieving the Information Quality Guidelines requirements for ensuring and maximizing the quality, objectivity, utility, and integrity of information disseminated. These elements, and related standards and guidelines, combine as described below, to ensure that information disseminated by U.S. Federal agencies are useful, accurate, reliable, unbiased, and secure.

Under the OMB quality guidelines, quality is an encompassing term comprising utility, objectivity, and integrity.

Utility refers to the usefulness of the information to its intended users. The usefulness of information disseminated by U.S. Federal agencies should be considered from the perspective subject matter specific users, researchers, policymakers, and the public. Utility is achieved by staying informed of information needs and developing new products and services where appropriate.

Objectivity refers to whether information is accurate, reliable, unbiased, and is presented in an accurate, clear, and unbiased manner. It involves both the content of the information and the presentation of the information. This includes complete, accurate, and easily understood documentation of the source of the information, with a description of the sources of any errors that may affect the quality of the data, when appropriate. Objectivity is achieved by using reliable information sources and appropriate techniques to prepare information products.

Integrity refers to the security or protection of information from unauthorized access or revision. Integrity ensures that the information is

not compromised through corruption or falsification.

U.S. Federal agencies have a number of statutory and administrative provisions governing the protection of information. The Privacy Act, the Computer Security Act of 1987, the Freedom of Information Act, OMB Circulars A-123, A-127, and A-130, Federal Policy for the Protection of Human Subjects, Government Information Security Reform Act, and the E-Government Act of 2002 are examples that may affect the protection of data for all U.S. Federal agencies.

Application of Measures Across Various Statistical Agencies

The application of standards over the wide diversity of U.S. Federal statistical activities, of course, requires judgment considering such factors as the uses of the resulting information and resources. Some surveys are extremely large undertakings requiring millions of dollars and the resulting general-purpose statistics have significant, far-reaching effects. (Examples of major U.S. Federal information programs, many based on statistical surveys, are listed in the Principal Economic Indicators.) Other statistical activities may be more limited and focused on specific program areas (e.g., customer satisfaction surveys, program evaluations, and research).

U.S. Federal agencies are encouraged to adhere to all standards for every statistical survey. However, standards cannot be applied uniformly or precisely in every situation. Consideration must be given to the importance of the uses of the information as well as the fitness of the information to those uses. In addition, agencies must evaluate the potential improvement in data quality that would arise from adherence to the standard if funding or other contingencies make it impossible for all standards to be met. An agency must consider these standards and guidelines and apply them efficiently and effectively to achieve the goal of information quality. However, the provision of standards and guidelines cannot substitute for agency judgment about the most appropriate expenditure of funds.

In instances where the strict application of a standard is impractical or infeasible, the agency needs to consider alternative methods of achieving the standard's intent. The agency is also to include in the standard documentation for the survey, or in an easily accessible public venue, the reasons why the standard could not be met and what actions the agency has taken or will take to address any resulting issues. If the information is not provided as part of standard documentation, it should be provided within one year of the release of the standard documentation.

Finally, the proposed standards and guidelines are not designed to be completely exhaustive of all efforts that an agency may undertake to ensure the quality of its statistical information. Agencies are also encouraged to develop additional, more detailed standards focused on their specific statistical activities.

The Framework for Standards for U.S. Federal Statistical Agencies

A common framework for the standards was first chosen based upon topic areas that had been endorsed by all of the principal statistical agencies as part of the 2002 Information Quality Guidelines. The framework covered the following chapters:

1. Development of Concepts, Methods, and Design
2. Collection of Data
3. Processing and Editing of Data
4. Production of Estimates and Projections
5. Data Analysis
6. Review of Procedures
7. Data Dissemination

This paper will detail the standards and guidelines associated with the final five items mentioned above (processing and editing of data, production of estimates and projections, data analysis, review of procedures, and data dissemination). The Seastrom and Madans (2004) paper covers the standards and guidelines associated with the first two items mentioned above. A paper by Arends et al (2004) presented at the Eurostat/European Conference on Quality and Methodology in Official Statistics in Mainz, Germany covers the process the committee used to develop the standards.

The Standards for Processing, Estimating and Disseminating Statistics from Surveys

The proposed standards on processing, estimating and disseminating statistics from surveys follow. We have maintained the same numbering schema that will be used in the published standards.

The principle theme underlying these five groups of standards is that care should be exercised in processing data to minimize errors. The effect of all types of errors, sampling and nonsampling error, should be documented and products published should be clearly written and carefully reviewed for errors. Data releases should document the methods used to construct the estimates with all limitations specified so that a researcher will know the constraints in use of the data.

Section 3: Processing and Editing of Data

3.1 Data Editing: Collected data must be edited appropriately, based on prior knowledge, to mitigate or correct detectable errors.

3.2 Nonresponse Analysis and Response Rate Calculation: Unit and item nonresponse must be appropriately measured, adjusted for, reported, and analyzed to assess the impact on data quality and to inform users. Response rates must be computed using standard formulas to measure the proportion of the eligible sample that is represented by the responding units in each study.

3.3 Coding: Codes must be added to collected data to identify aspects of data quality from the collection (e.g., missing data) to allow users to appropriately analyze the data. Codes added to convert collected text information into a form that permits immediate analysis must use standardized codes, when available, to enhance comparability.

3.4 Data Protection: Safeguards must be taken throughout the production process to ensure that survey data are handled to avoid disclosure.

3.5 Evaluation: Agencies must evaluate the quality of the data and make the evaluation public to allow users to interpret results of analyses, and to help system designers of recurring surveys focus improvement efforts.

Section 4: Production of Estimates and Projections

4.1 Developing Estimates and Projections: Direct survey-based estimates, and model-based estimates and projections that use survey data must use accepted theory and methods. Error estimates must be calculated and used to assess if estimates are appropriate for particular uses. Evaluations must be planned to assess the quality of the projections.

Section 5: Data Analysis

5.1 Analysis and Report Planning: An analysis plan that uses survey data must be developed prior to the start of a specific analysis to ensure that statistical tests are used appropriately and that adequate resources are available to complete the analysis.

5.2 Inference and Comparisons: Statements of comparisons and other statistical conclusions taken from the data must be based on acceptable statistical practice.

Section 6: Review Procedures

6.1 Review of Information Products: Agencies are responsible for the quality of information that they disseminate and must institute appropriate content/subject matter, statistical, and methodological review procedures to comply with OMB and agency Information Quality Guidelines.

Section 7: Data Dissemination

7.1 Releasing Information: Information intended for the general public must be released according to a dissemination plan that provides access to all users and provides information about any planned and unanticipated data revisions.

7.2 Data Protection and Disclosure Avoidance for Dissemination: All information products must be released in accordance with the survey pledge to the respondents and all applicable Federal legislation.

7.3 Survey Documentation: Survey documentation must include those materials necessary to understand how to properly analyze data from each survey, as well as the information necessary to replicate and evaluate each survey. Survey documentation must be readily accessible to customers, unless it is necessary to restrict access to protect confidentiality.

7.4 Documentation and Release of Public Use Micro Data: Public Use Micro Data documentation must clearly describe how the information is constructed, and provide the metadata necessary for users to access and manipulate the data. Public Use micro data documentation and metadata must be readily accessible to customers.

Discussion:

Each standard has an extensive list of key terms and guidelines to help users comply with the standards. The key terms are defined in a glossary to assist users in interpreting the standard and guidelines. We will review selected guidelines to direct the reader to the subcommittee's logic on steps which, if implemented, would result in optimal compliance with the standard.

Guideline for the evaluation standard

Guideline 3.5.1: Include an evaluation component (see Section 1.1). Review past surveys similar to the one being planned to determine likely sources of error, appropriate evaluation methods, and problems that are likely to be encountered. Address the following areas:

1. Potential sources of error, including:
   - Coverage error (including frame errors),
   - Nonresponse error, and
   - Measurement error, including sources from the instrument, interviewers, and collection process.

2. Data processing error,

3. How sampling and nonsampling error will be measured, including variance estimation and studies to isolate error components,

4. How total mean square error will be assessed,

5. Methods used to reduce nonsampling error in the collected data,

6. Methods used to mitigate nonsampling error after collection,

7. Post-collection analyses of the quality of final estimates. The data and estimates derived from the data should be compared to other independent collections of similar data, if available, and

8. Make evaluation studies public to inform data users.

The intention of the guideline that supports the evaluation standard is to alert the data user to issues they should consider when interpreting the survey results or analyzing public use files from the survey. It should address sources of error that might not be obvious or determinable by an outsider. Examples of errors to be cited include issues such as response problems with certain parts of the study, errors in the collection process, and instrument design issues uncovered during collection. The evaluation should also serve as documentation to assist designers of the survey to focus improvement efforts or assist designers in other agencies of potential pitfalls they should avoid.

Guidelines for developing estimates and projections

Guideline 4.1.1: Develop direct survey estimates according to the following practices:

1. When data from a sample survey are used to calculate population estimates, employ weights appropriate for the sample design. However, an agency may employ an alternative method (e.g., ratio estimators) to calculate population

estimates if the agency has evaluated the alternative method and determined that it leads to acceptable results.

2. Use auxiliary data to improve precision and/or reduce the error associated with direct survey estimates.

3. Calculate variance estimates by a method appropriate to a survey's sample design taking into account probabilities of selection, stratification, clustering, and the effects of nonresponse, post-stratification, and raking. The estimates must reflect any design effect resulting from a complex design.

Guideline 4.1.2: Develop model-based estimates according to accepted theory and practices (e.g., assumptions, mathematical specifications, etc.).

Guideline 4.1.3: Develop projections in accordance with accepted theory and practices (e.g., assumptions, mathematical specifications, etc.).

Guideline 4.1.4: Any model used for developing estimates or projections should be subjected to:

1. Sensitivity analysis to determine if changes in key model inputs cause key model outputs to respond in a sensible fashion, and

2. Validation to analyze a model's performance by comparing the results to available independent information sources**.**

Guideline 4.1.5: Establish criteria for determining when the error (both sampling and nonsampling) associated with a direct survey estimate, model-based estimate, or projection is too large to publicly release the estimate/projection.

Guideline 4.1.6: Document methods and models used to generate estimates and projections to help ensure objectivity, utility, transparency, and reproducibility of the estimates and projections. (For details on documentation, see section 7.3). Also, archive data and models so the estimates/projections can be reproduced.

The guidelines supporting the standard on developing estimates and projections recommend both survey based and model based estimates use accepted theory and methods. However, 3 of the 6 guidelines pertain only to models as they are becoming more prevalent as agencies attempt to utilize data to the maximum extent possible. The emphasis is that accepted theory and practice should be employed at all stages. Moreover, models used for estimates or projections must be evaluated through extensive diagnostic work during the process of model development. This diagnostic work generally will include testing of the main effect and interaction coefficients, and other relevant parameters, in alternative models; graphical analyses such as residual plots and boxplots; and evaluation of model fit through R^2, mean squared error ratios and other omnibus measures. Agencies are encouraged not to publicly release estimates, regardless whether a survey estimate, model or projection if the sampling and/or nonsampling error is

too high.

Guidelines for review of information products

Guideline 6.1.1: All information products should undergo a content/subject-matter review. Those conducting the review should have appropriate expertise in the subject matter, operation, or statistical program discussed in the document. Among the areas that reviewers should consider are:

1. Subject-matter literature is referenced in the document if appropriate,

2. Information is factually correct, and

3. Information is presented logically, conclusions follow from analysis, and no anomalous findings are ignored.

Guideline 6.1.2: All information products should undergo a statistical and methodological review. Those conducting the review should have appropriate expertise in the methodology described in the document. Among the areas that reviewers should consider are:

1. Review assumptions and limitations for accuracy and appropriateness,

2. Assure that appropriate statistical methods are used and reported,

3. Review calculations and formulas for accuracy and statistical soundness,

4. Review data and presentations of data (e.g., tables) to assure disclosure risk avoidance, as necessary,

5. Review contents, conclusions, and technical recommendations (statistical and operational areas - not policy), to ensure that they are supported by the methodology used, and

6. Ensure that data sources and technical documentation, including data limitations, are included or referenced.

Guideline 6.1.3: Information products disseminated via the Internet should be reviewed for Section 508 compliance for accessibility. Any product that is disseminated via special software onto the Internet is tested for accessibility and interpretability prior to dissemination.

The guidelines associated with the release of information products are to ensure the subject matter as well as statistical aspects of the programs are correctly documented.

Conclusions

The next step for the interagency committee is to make recommendations on how the standards should be used by OMB in the process for approving surveys proposed by U.S. Federal agencies. Under the Paperwork Reduction Act, OMB must approve all agency information collections before agencies can gather information from the public. The committee will make recommendations to OMB on their review process so that it can effectively implement these standards and improve the quality of Federal statistical information. Over the next several months, many U.S. Federal agencies as well as the general public via the Federal Register will review the proposed standards. OMB will consider the comments and ultimately determine

which comments are accepted and incorporated into the final document.

Other critical areas for further work include establishing minimum thresholds for quality measures such as unit nonresponse rates, item nonresponse rates, and survey frame coverage rates. The committee recommends that agencies conduct analyses of the impact on survey estimates if quality measures are below this threshold. These analyses will likely require additional data collection efforts but will allow agencies to evaluate if survey estimates are of sufficient quality to meet user needs or if additional efforts are needed to improve quality.

Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics or Statistics of Income Division, U.S. Internal Revenue Service.

References

Federal Statistical Organizations' Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Disseminated Information, Federal Register notice issued by 13 Federal agencies on June 4, 2002. http://www.eia.doe.gov/smg/Stat-Agency-FR-June4-2002.pdf

OECD Quality Framework (Draft), Geneva, Switzerland, June 2003.

Some Quality Issues During the Last 50 Years, Some Observations, Lars Lyberg and Eva Elvers, Statistics Sweden, 2003.

Statistical Directives 1 and 2, Statistical Policy Office, U.S. Department of Commerce, 1978.

Statistical Policy Working Paper 31, Measuring and Reporting Sources of0 Error in Surveys, U.S. Office of Management and Budget, June 2001. http://www.fcsm.gov/01papers/SPWP31_final.pdf

Statistical Quality Checklist, Government Statistical Service, United Kingdom, 1997.

Standards Manual, U.S. Energy Information Administration, 2002. http://www.eia.doe.gov/smg/Standard.pdf

Statistical Standards, U.S. National Center for Education Statistics, 2002. http://nces.ed.gov/statprog/2002/stdtoc.asp.

Seastrom and Madans, *Proposed Standards for Planning, Design and Conducting Surveys,* Proceedings of the Government Statistics Section, ASA 2004

William Arends (U.S. National Agricultural Statistics Service), Jay Casselberry (U.S. Energy Information Administration), Kevin Cecco (Statistics of Income Division, U.S. Internal Revenue Service), Stephen Cohen (U.S. Bureau of Labor Statistics), Patrick Flanagan (U.S. Bureau of Transportation Statistics), Brian Harris-Kojetin (U.S. Office of Management and Budget), Jennifer Madans (U.S. National Center for Health Statistics), Karen O'Conor and Marilyn Seastrom (U.S. National Center for Education Statistics), and Alan Tupek (U.S. Census Bureau), *Recommendations for Revising Standards for Statistical Surveys for U. S. Federal Agencies* presented at the European Conference for Quality and Methodology in Official Statistics at Mainz Germany, 2004