

## FOCUSING ON QUALITY IN ESTABLISHMENT SURVEYS

Thomas J. Plewes, U.S. Bureau of Labor Statistics  
Room 2919, 441 G Street, N.W., Washington, D.C. 20210

KEY WORDS: surveys, quality,  
establishments, error

The trend toward increasing reliance on the results of Federal Government surveys of business establishments in public and private decisionmaking has, quite naturally, been accompanied by increasing interest in the quality of those data. For example, every nook and cranny of the monthly employment statistics, which have been eagerly awaited by financial and market intermediaries, have been examined.

Both the profession and the popular press have exhibited an elevated level of interest in data collected from establishments. 1/ Series that provide key information on employment and wages, sales, prices, agriculture and energy production, money supplies, and many other aspects of the working of the economic and social order, which are collected from businesses, and compiled and published by a large number of Federal Government agencies, have been scrutinized as never before. The results have not always been complementary.

While the interest in the quality of the data from business sources has recently intensified, it should be recognized that it is not because the surveys are newly available. In fact, the establishment-based data series have been around for a long time, some of them continuous since the early part of this century, and many predating household surveys. Even with this long history, there are some difficulties in fully answering the question, "How good are the data?" In contrast with household surveys, for which a rich literature has emerged over the past 5 decades, very little in the way of theoretical or evaluative work on survey error has been published for establishment surveys.

The comparative shortage of generic literature should not be surprising in view of the Federal Government's intentional laissez-faire approach to establishment surveys. Surveys that cover the economic and social waterfront in various periodicities collected under widely different circumstance by at least 9 different agencies cannot be expected to march to the same drummer.

There is strength in the diversity of approaches to establishment surveys. Still, this autonomy has resulted in a situation unique to establishment surveys. Today, there is no clear "lead" agency that sets the basis for establishment survey design and practice in the U.S.--a role the Bureau of the

Census has played in regard to household surveys. As a result of both the decentralization of responsibility and the way in which the surveys grew and matured, there are few commonly-accepted approaches to the design, collection, estimation, analysis, and publication of establishment surveys. Establishment surveys abound in rich variety, with little standardization of design, practice, and procedure.

This is not to say that Federal agencies do not work hard to insure that the surveys they conduct are carried out in the most professional and efficient manner that is possible, given the resources available. There is a serious dedication on the part of government statisticians to insure that their products are of the highest possible quality. Indeed, the Federal agencies struggle to assure quality not only because they want to, but because they are obliged to by the Office of Management and Budget's clearance process. However, both the agency personnel that have responsibility for the establishment surveys and the OMB staff that reviews the requests for surveys operate in a vacuum, without benefit of key design information that would be available in an error profile.

In the view of many, the collectors and reviewers, and more importantly, the users of data from establishments would be able to understand the sources of error in the surveys and censuses with a collection of information in one place. By a sharing of information on methods for dealing with or overcoming those error sources, the survey operations should be improved.

### A Government-wide Look at Quality

With the overdue need for a review of establishment survey practices in mind, in November 1985, the Federal Committee on Statistical Methodology established a Subcommittee on Measurement of Quality in Establishment Surveys to document, profile and discuss the topic of quality in Federal surveys of establishments. The Subcommittee's parent was organized by OMB in 1975 to investigate methodological issues in Federal statistics. Members of the Committee, selected by OMB on the basis of their individual expertise and interest in statistical methods, serve in their personal capacities rather than agency representatives. The Committee conducts its work through subcommittees, such as

the current group, which prepares a report that reflects the individual and collective ideas of the subcommittee members. Through a process of review and comment, the parent Committee assist the subcommittee in its work.

The Subcommittee on Measurement of Quality in Establishment Surveys set the following goals for its report:

-- To document current understanding of the meaning of quality in establishment surveys;

-- To discuss establishment surveys in terms of sampling and nonsampling error;

-- To identify approaches and practices to be considered by users and designers of establishment surveys; and

-- To profile current practices in the areas of controlling and measuring survey quality.

In order to narrow its work, the Subcommittee limited its scope to ongoing surveys of private sector establishments. One-time surveys, special studies, and exclusively government surveys were excluded. The aim of the subcommittee was to organize and document, not to develop a stand-alone primer or a theoretical work.

#### A Notion of Establishment Survey Error

The Subcommittee issued its report, Quality in Establishment Surveys, early this summer. 3/ The report discusses, in very general terms, the potential sources of error that may affect counts and estimates derived from surveys and censuses of establishments. By classifying these sources of error, the report focuses on practices that are used to improve and measure the quality of establishment data.

To this extent, the approach of the Subcommittee on Measurement of Quality in Establishment Surveys was rather straightforward and fairly conventional. For example, only the more traditional aspects of quality are considered--those that refer to the accuracy of the survey estimate or its closeness to an exact ("true") value. Other aspects of quality such as relevance and timeliness, which the current literature considers to be critical components of a total quality approach from the vantagepoint of the user, are not given equal emphasis.

The report retains the usual distinction between sampling error and non-sampling error as the central dicotomization. 2/ Sampling error is discussed in terms of sample design, estimation, and variance estimation. Nonsampling error is partitioned into five areas--specification error, coverage error, response error, nonresponse error, and processing error. Error is discussed in terms of sources, control, and measurement.

#### Profile of Survey Environment

The core of this study is a profile of the Federal Government's current establishment survey environment. In an attempt to quantify the information presented in the report, the Subcommittee collected data on design, estimation, control, and measurement practices for 55 surveys from 9 Federal agencies. The surveys were selected to include all of the known major ongoing establishment surveys conducted by the Federal Government and thus provide a virtual census of the current establishment survey environment. Agencies included in the profile were: Bureau of Labor Statistics (13 programs); Bureau of the Census (15 programs); Energy Information Administration (9 programs); Federal Reserve Bank (6 programs); National Agricultural Statistics Service (4 programs); Center for Educational Statistics (3 programs); Bureau of Economic Analysis (2 programs); Bureau of Mines (2 programs); and the National Center for Health Statistics (1 program). Information was collected by means of a structured questionnaire.

#### Key Observations

This paper summarizes the key points from the discussion of establishment survey error sources, control, and measurement. In preface, three major points seem to stand out:

-- In general, Federal Government establishment surveys have procedures in place designed to control major known sources of survey error.

-- Error measurements are not extensively derived.

-- Error measurements are rarely published when they have been derived.

While the relative differences in the extent of use of control and measurement can be understood in terms of resource priorities, there does not appear to be a clear reason why error information is not published when available. The limitations in the availability of published error information made it quite difficult for the Subcommittee to collect its information. Hopefully, now that collection has been completed, this report will be more valuable as a reference document.

#### Sample Design and Estimation

Establishments are different from households. Their populations have very skewed distributions, with a few large firms commonly dominating totals for most characteristics of interest. These distributions impact on the frame development and maintenance, sample design, and estimation practices of establishment surveys. Given the importance of large units, extensive resources are devoted to improving frame coverage and content for large units. One-stage, highly stratified designs, with certainty selection of large

establishments are used in the vast majority of establishment surveys profiled.

Roughly one-fifth of the surveys profiled were described as having designs or implementations which do not result in a probability design. These surveys included those for which substitution is allowed for nonresponse, a segment of the target population has no chance of selection, units are selected judgmentally, and other practices are followed that are at variance with probability design practice.

Estimators which do not reflect probability of selection are also commonly used in establishment surveys. The estimators in use may generally be described as model-based, although the model often is implicit, rather than explicitly stated. Imputation techniques are frequently employed because cutoff sampling is a common design practice.

One-fourth of the sample surveys profiled in the data collection by the Subcommittee did not compute variances, and another one-fifth did not publish estimates of sampling error in survey publications. This lack of generation and publication of sampling error information was not seen to be a function of agency practice, since it was not confined to one or two agencies, but rather it appeared to be somewhat correlated with the use of nonprobability-based estimation procedures.

#### Nonsampling Error

Establishment surveys typically seek hard data for which records are available. This is a central characteristic which improves the collection while complicating the interpretation of the data. The collection is improved because the data are "hard", i.e. data of record from which the data of interest are extracted. This contrasts with "soft" data that rely on the memory, opinions, or interpretations of the respondents, as is often the case for household surveys.

However, establishing the concepts and definitions to be used in the surveys is no simple matter. When dealing with businesses, special care must be taken to carefully consider the firms' recordkeeping systems, definitions, and data availability to avoid introducing specification error into the data. Typically, agencies do this through requirements reviews or consultations with respondents or trade associations. How well the agencies mesh their specifications with the recordkeeping practices of establishments is difficult to measure. There is currently no single specification error measurement practice used by a large majority of the surveys profiled. Although slightly more than half of the surveys regularly compared

survey results to independent estimates to gain a better understanding of specification error, the independent estimates may have error also that complicates the evaluation.

Establishment surveys commonly use list frames, and thus are subject to the inherent problems associated with list frames--duplication, overcoverage of out-of-scope and out-of-business units, undercoverage of business births, and misclassification of units. In apparent recognition of these potential sources of error, well over half of the surveys profiled regularly used procedures designed to control these error sources, such as updating for structural changes, updating/sampling for births, and internal consistency checks for duplicates. On the coverage error measurement side, little is commonly done except to provide such indirect measures as out-of-business and out-of-scope rates. No direct measurement technique was reported as regularly used by more than half of the surveys.

The fact that data are acquired from records enables subject-matter analysts to identify possible reporting error at the microdata level. As a result, common control procedures for response error include not only those typically in place for household surveys, such as editing for reasonableness, questionnaire pretest, and detailed training/guidelines for interviewers, but also include analyst review of data, and record-keeping practices studies. Outside of the calculation of edit failure rates, little response error measurement is done across surveys.

The control of nonresponse in establishment surveys generally relies upon conventional practices, including unit and item nonresponse followup, and advance notification. However, the skewed nature of the population has led to other widely-used control techniques weighted toward large units which are unique to establishment surveys. These techniques include intensive followup of critical units, central office consolidation of all responses from the same establishment, other special reporting arrangements, and provision of survey publications to respondents. Several indirect measures of nonresponse error, such as unit and item response rates and refusal rates, are commonly generated. Because of the population distribution, weighted response rates are also commonly derived. Very little is done on direct measurement of nonresponse error.

Processing error control procedures do not differ from those in use for household surveys. The identified control procedures were all used by over half of the surveys profiled. The most common measurement produced were edit

failure rates, which are most likely generated due to concern about response error rather than processing error.

#### Quality Improvement

This Subcommittee report refrained from making specific recommendations, trusting that the discussion (which begins today) and profiling of error sources as applied to establishment surveys will give impetus to consideration of survey practices on the kind of case-by-case that is necessary given the vast differences in the establishment survey operations.

Nonetheless, the tenor of the findings can be depicted as critical. The profiles portray a number of key Federal Government surveys with deficiencies in the measurement and documentation of sampling and nonsampling errors, and point to a need to focus additional attention, and resources, on the general improvement of survey practices.

This first government-wide profile has also reminded us of the limitations of our understanding of errors, their sources and the means of reducing or accounting for them. More importantly, a decade after Brooks and Bailar 4/ made a recommendation to improve error profiling by coming to a better understanding of the interaction of the errors, little is known of whether the various sources of errors compound or reduce total error in their interaction. If this profile engenders interest in refining our understanding of errors and their interaction, it will have more than proved its usefulness.

The Subcommittee believes that the amalgam of theory and practice that provides the framework for its report provides a useful tool in a systematic approach to understanding and evaluating quality in establishment surveys. The framework and the report are hoped to constitute a step forward in the process of quantifying and improving the quality of the important surveys of establishments conducted by the Federal government.

1/ Juster, F. Thomas, "The State of U.S. Economic Statistics: Current and Prospective Quality, Policy Needs, and Resources," Conference on Income and Wealth, May 1988; Malabre, Alfred L. Jr., "Shaky Statistics Pose Peril for Forecasters," The Wall Street Journal, May 9, 1988.

2/ Hansen, Morris A, Hurwitz, William N. and Berstad, Max A., "Measurement of Errors in Censuses and Surveys," Bulletin of the International Statistical Institute 38(2), 1961.

3/ Office of Management and Budget, Quality in Establishment Surveys, Statistical Policy Working Paper 15, August 1988.

4/ Brooks, Camilla A. and Bailar, Barbara A., "An Error Profile: Employment as Measured by the Current Population Survey," Statistical Policy Working Paper 3, Office of Management and Budget, September 1978.