

DISCUSSION

F. Thomas Juster, University of Michigan
Institute for Social Research, Ann Arbor, MI 48109

Introduction

This study of quality in establishment surveys conducted in the U.S. is the first of its kind since the 1900's. It has distinctive characteristics, some of which are pluses and some minuses. On the plus side, there is an extensive treatment of both sampling and nonsampling error, with issues relating to sample frames, sampling variances and estimation to the underlying universe being treated in the sampling error section, while issues relating to specification errors, coverage errors, response errors, nonresponse errors, and processing errors are treated in the nonsampling section. Most studies of this sort have little to say about nonsampling errors, which are likely to be the most serious of the various types of survey error.

On the minus side, there is no analysis of quality as it relates to outcomes--given the process, what can be said about the quality of the data? Rather, quality is essentially seen as synonymous with process--good processes lead to high quality, poor processes otherwise. Finally, the committee consists entirely of a group of specialists located within the U.S. statistical agents, who are the producers and to a large extent also the consumers of the data. In assessments of this sort, it always seems to me wise to include a few outsiders--people who know quite a lot about the data and how it might be used, but do not have either a stake in the outcome or, as one might put it, an "establishment" view of establishment surveys.

Overall Conclusions

The report can be described as being mildly critical of present practices. It notes that there is too much use of nonprobability sampling, too much use of list frames that are not very current and have too many omissions, too little attention to sampling errors, both in terms of errors not being computed at all, or if computed, not being published. The committee's careful look at procedures for processing and for exercising control generally come to the conclusion that the procedures are what should be expected of a high quality data collection operation. There is no reason to question the validity of those conclusions--the problem is that they are not very exciting, at least from the point of view of an outsider who thinks there is probably a lot that needs to be done here rather than only a little.

Assessment of the Assessors

My general impression is that what has been done by the committee is extremely useful, and that serious users of establishment data owe the committee a vote of thanks. Putting together a report of this sort is not likely to lead to professional kudos, and it takes a lot of painstaking work which is often not given much credit, certainly within the academic community and perhaps even within the statistical establishment.

Having said that, there are some obvious weaknesses. The assessment that has been done is extremely narrow--a basic problem is that the committee only tried to answer questions that could be answered quantitatively, and largely ignored those that did not have easily available quantitative indicia which with to address the quality question. We all do that in studies of this sort, but there is an obvious problem with that approach--there is no guarantee that the most important shortcomings on the quality side are reflected by characteristics of the surveys which are not easily quantifiable, while the less important dimensions of quality, or at least the ones where quality is more uniform, are apt to be ones in which quantitative indicia are available.

Moreover, even on the side of questions that could be asked and answered, the report is narrow in what is presented. While there is extensive discussion of nonresponse processing procedure in terms of the incidence of different procedures, there is no information in the report on nonresponse rates, no information on item nonresponse rates, no discussion of imputation procedures (if any were used) and so on. I would have thought that a crucial indicator of quality is the item nonresponse rate, and an assessment of the procedures used to impute in those cases.

Finally, the report reflects what I have long thought to be the likely problem with establishment surveys in the United States--that the major concern of those producing the data is with an estimate of some aggregate, and that there is much less concern with understanding the structure, and potentially the behavior, of the micro-units that comprise the survey. That is, establishment surveys are largely thought of as contributing to an estimate of some aggregate, not as contributing to an understanding of the behavior of establishments in the U.S. economy.

What Are the Next Steps?

It seems to me that the committee--or some successor committee--could usefully turn to the following topics:

1. Quantify the outcomes: we really do want to know about nonresponse rates, and about item nonresponse for particular variables on the establishment surveys;
2. Identify the major culprits in terms of inferior quality--this report does not cast a single stone at any ongoing survey, and they cannot all be up to the same standard.
3. Perhaps most important of all, turn to issues relating to relevance--are the data being collected sufficient to meet the needs of industry, policy-makers, accounting aggregates where the survey data are the basic source of information, and, last but not least, the community of academic users who are basically

interested in understanding the behavior of establishments.

Topics on Relevance

I have a couple of thoughts on issues that needs to be addressed if this committee, or some successor committee, turns to the relevance issue. They are listed in no particular order of priorities.

1. The relevance of the present set of establishment surveys to current topics of importance needs to be examined. Parts of this report sound much like parts of the report of the American Economic Association Committee on the Quality of Economic Statistics, where the conclusion generally was that the quality of what was being measured was fine, but what was being measured was appropriate to the structure of the 1950s and not necessarily to the structure of the 1980s.
2. For a variety of reasons, we need to pay a lot more attention to the behavior of small establishments. That's partly a bias issue, partly a dynamics issue, and inevitably a cost issue--collecting data from small establishments is a lot more expensive per dollar of whatever is being measured than collecting data from large establishments.
3. The establishment data almost certainly needs to pay a lot more attention to import and export patterns than has been the case historically.
4. We almost certainly need to worry more than we now do about the services industries--these have always been badly measured, they have been of steadily increased importance for several decades, and we have not done much about improving the measurements.
5. Continual concern needs to be given to issues relating to classifications--are product and industry definitions now in use up to the changing characteristics of output?
6. Among establishment surveys, is there consistency of measures relating to product, industry, inputs, and outputs, so that productivity estimates, for example, are consistent?
7. Although economic variables are hard to measure, there are almost certainly some non-economic variables that are important in understanding what goes on in establishments--variables like organizational structure, hierarchy, incentive structures, etc. are likely to have an important influence on economic outcomes, but such measures are rarely (ever?) included in establishment data bases.
8. Is the balance between sample size and variables measured per establishment an optimum one, or are the samples larger than they need to be and the available measures thinner than they should be? The issue here is partly one of priorities--academic users want richness in the set of independent and dependent variables, particularly the former, while other users want large sample estimates of policy-relevant dependent variables and don't care much about the presence of explanatory variables.