SAMPLE SURVEY QUALITY: ISSUES AND EXAMPLES FROM AN AGRICULTURAL SURVEY

Ron Fecso, USDA, National Agricultural Statistics Service

INTRODUCTION Industry is experiencing a rapid growth in the acceptance and use of statistical techniques to improve quality and productivity. These techniques can also be used to improve the quality of survey data and the productivity of survey organizations. The fundamentals of statistical quality control in a sample survey organization are outlined, followed by a discussion of quality control procedures in crop yield surveys. A quote by Raj best summarizes the needed approach. "A large-scale survey is an exercise in statistical engineering. Each step in the production line is a potential source of error. The sample unit may not be identified correctly, enumerators may make errors in the field, there may be errors of coding or punching the cards in the office, and so on. Thus, it becomes important to ensure that the production process is under control and that the outgoing quality is acceptable. Sampling methods can play an important part in achieving this."

DEVELOPMENT OF A QUALITY ORGANIZATION To achieve quality in an organization, Tribus and Tsuda identify the changes required in an organization. These topics are purpose, quality in management, manager responsibilities, organizational skills and capabilities, and institutional improvement. This section is a summary of their work.

Considering that a survey statistic is no less constructed from pieces than an automobile or a light bulb, the need for survey quality methodology is as clear as the need for quality in industry. Improved quality of survey data can occur only if there is increased cooperation among all personnel in the survey organization. The cooperation can be spontaneous only if everyone understands, agrees with, and gets the necessary guidance from the statement of the survey organization's purpose.

Besides providing continuity and consistency of purpose, Deming states that the determination of the data users and their needs is a task, essential to quality and productivity, which cannot be delegated by a manager. Thus, the survey system should be thought of as a continuous "do loop":

Determine the need,
Design the survey,
Pre-test the survey/revise the survey design,
Conduct the survey,
Assess the usefulness of the survey information, then
Return to redetermine the need.

"The guiding principle [to quality in management] may be stated simply:

> The problem is not to increase <u>quality</u>. Increasing quality is the <u>answer</u> to the problem."

Survey designers and managers must be aware that there is no trade-off between quality and costs. Increases in quality increase productivity and decrease cost. There is a difference, which must be understood, between the quality of survey design and the quality of survey practices. The false notion that inspection and rejection are the only means to obtain high quality output must be overcome.

Managers' attitudes about quality, need to be improved, especially if they feel that they already know the problems or they feel no need to learn new techniques. Managers must learn that people work in a system, and that the manager works on the system to improve it with their help. Managers can make quality improvements a part of everyone's job if they:

Articulate the purpose of the system,
Define the system,
Define what constitutes "improvement",
Measure quality and quantity of output,
Create a system of continuous improvement,
Lead in problem identification/solution, and
Provide training to fellow workers.

The quality of survey outputs can be determined by:

- The quality of inputs
  Data requirements/objectives/handling
- The design and operation of the survey
  The interaction of sampling theory with data collection; data users with data collection; respondents with data collection; and administrators with survey designers.
- The training and education of all employees
  On-the-job, supervisory, and statistical training.
- Statistical control of quality
  Inspection methods (reinterviews, edits, audits); feedback methods (enumerator school discussions, panels, etc.); materials acceptability (pretesting, research in CATI, etc.)

Managers may delegate many of these issues, but they must accept the responsibility to harmonize the parts of the system. To do so, everyone must understand the difference between a quality survey statistic and a quality survey process. When manufacturing an item repetitively, marketing a high quality product may be possible with enough thorough inspection of the finished product and rejection of items out of specifications. This approach is not as effective as the development of a high-quality process. In surveys, inspection methods are of limited value. The final product, the survey statistics, usually cannot be rejected like a bad item in a lot. The approach to a high-quality survey output is through a high-quality process.

To achieve a high quality process, new skills need to be developed in all personnel: managers and non-managers, statisticians and non-statisticians, etc. Complex surveys are subject to many disturbances (sampling and nonsampling errors) which combine in ways that can only be described and understood through the use of statistics. Sources of variation must be understood. The concepts of bias and variance must be understood from the enumerator to the top administrator. Simple methods of analysis such as flow diagrams, basic statistics, graphical presentation and basic quality control techniques must be taught. The job title "statistician" should not exempt an employee from training in statistics for quality and productivity. Many statisticians have not had formal course work in quality control, and Vardeman and David indicate that current courses and books are "inadequate, and in some cases suffer from inappropriate emphasis". They present a lecture schedule which should be used as a model to develop training in survey organizations. A course proposed for engineering students which could easily be modified for use by survey statisticians is presented by Hogg, et. al. Finally, recommended reading on "Deming's 14 points" is in Joiner.

Regrettably, the use of statistical reasoning is as resisted in parts of survey organizations as strongly as it often is in industry. When statistical control procedures are being adopted, managers must lead the organization by basing decisions on valid, relevant data and by insisting that all who are involved with the survey process are able to reason statistically and are willing to cooperate with others in devising system improvement methods. Survey organizations, just as manufacturers, do not have enough research staff to study in detail all aspects of their processes. Worse yet, studies are often labelled as "out of touch" with the "operating" organization. To attain "quality", survey organizations need to adopt the philosophy that the enumerators and field office personnel are part of the research team. They are the front line of observation and analysis in developing ideas which can reduce costs, variance, and bias.

Managers need to lead the way to improvement, support improvement efforts, and create a structure that makes continuous improvement a certainty. Managers must strive to make everyone, from respondent and enumerator to top level administrators, partners in the survey improvement process. Workers closest to the survey process should be delegated the responsibility of observing the process, measuring it statistically, and suggesting improvements. Managers must allow changes to be implemented at the lowest possible level. If there is a lack of confidence in the analysis done at an operating level, training must be encouraged. Most importantly, the manager must remain involved to ensure consistency of survey procedures across administrative units.

Managers need to make policy changes which place the responsibility for survey quality on all the workers, not just on supervisory enumerators or other inspectors. Quality control procedures require complete understanding of the entire survey process rather than mass inspection. Finding and understanding the sources of error begins the never-ending process of survey improvement.

The road to improvement should begin with an assessment of the state of "quality" in an organization by encouraging discussion of the following ten issues which are used to award the Japan QC Prize:

1. What does top management say and do? Are these consistent?

2. How is quality control organized, promoted, and managed for the various surveys and the survey organization as a whole?

3. Are employees informed and educated, especially in quality control techniques? How are statistical methods taught? Is education possible and encouraged?

4. How are data gathered and used to evaluate the survey system?

5. How are the analyses of problems prioritized?

6. Are procedures and practices standardized?

7. Are QC charts and information constructed and used? How?

8. How are quality assurance techniques

used in conducting, controlling and maintaining the survey process?

9. Is the survey effective? How is effectiveness measured?

10. Are plans for the future factually based? Is there a consistency between long- and short-range plans?

Soliciting responses to these questions from everyone involved with a survey can provide a summary which serves as a basis to begin improvement.

SAMPLE SURVEY APPLICATION The steps necessary to create a quality survey organization have been briefly presented as well as references for statistical training. The broad range of statistical control techniques cannot be explained here, but some specific applications will be presented to indicate the usefulness of the techniques.

The National Agricultural Statistics Service (NASS), USDA, conducts crop yield surveys in states which are major producers of field crops. The survey data are used to forecast expected yield and production during the growing season and to estimate these values at harvest.

Briefly, the survey design can be described as a multiple step sampling procedure, drawing from an area frame sample to estimate acreage for harvest and subsampling fields and small plots to make measurements related to yield per acre. Detailed information on the area frame design is available in Fecso, Tortora and Vogel. More detail on the crop yield surveys, called objective yield (OY) surveys, can be found in Matthews and in a paper in preparation by Fecso.

QUALITY CONTROL OF THE OY SURVEYS Historic quality control procedures for the OY surveys consisted of the following: Supervisory enumerator visits to the plots (QC) (approximately a 10 percent subsample which included the first sample visited by each enumerator); occasional visits by the field office survey statistician; hand and computer editing of the data; and validation surveys (VS), which are periodic surveys covering a subset of crops and states in a given year with the objective of measuring the overall bias of the survey estimate.

In recent years, the shortcomings of these procedures have been questioned. For example, visits by the supervisory enumerator served mostly as a retraining system; the data was not used to improve the estimates or to estimate biases. Budget and staff reductions have reduced the number of field visits by survey managers. Edits have been changing; new computer edits and local areas beginning to create individualized recording forms have resulted in estimates which may differ from those based on the old editing procedures. Finally, the expensive and administratively burdensome validation survey was receiving increased questioning. A review of the QC/VS methodology was initiated which identified two questions:

1. Are the quality control and validation programs, as they now exist, obtaining information worth the resources utilized?

2. Do these programs have well defined purposes, are these purposes useful, and are they needed?

The need to address the concept of OY survey quality as a whole is evident from a review of the validation survey and OY research reports (Fecso). A summary of OY survey concerns includes limitations of the validation design, data problems, needed enumerator training and various nonsampling errors. A literature review was conducted to determine how others have treated survey quality issues.

The literature review of survey quality methodology began with Fellegi's four objectives of survey evaluation.

1. Measure the accuracy of survey results.

2. Analyze sources of error with a view to subsequent improvement.

3. Evaluate alternative methods of survey design.

4. Institute continuing control procedures.

While reviewing other papers dealing with survey quality, most were found to, in general, measure some source of survey error or profile known sources of error. There are many excellent examples: Bailer and Lanpher, Wright, Zarkovich, and the Proceedings of the Census Research Conference to name a few. Rarely was survey quality specifically defined, therefore, definitions from the "Glossary and Tables for Statistical Quality Control," prepared by the American Society for Quality Control are modified as follows:

"SURVEY QUALITY--The totality of features and characteristics of a survey that bears on its ability to satisfy a given need."

"SURVEY QUALITY CONTROL--The overall system of activities whose purpose is to provide a quality of survey that meets the needs of users, also, the use of such a system. The aim of quality control is to provide quality that is satisfactory, adequate, dependable, and economic. The overall system involves integrating the quality aspects of several related steps, including the proper specification of what is wanted; production to meet the full intent of the specifications; inspection to determine whether the resulting product or service is in accordance with the specifications and review of usage to provide the revision of the specification."

Note that quality control is the conformance to requirements and should not be confused with the misconception that quality is elegance, luxury, perfection, etc.

An excellent start towards a formal basis for comprehensive quality control was written by Dalenius. The first of two components in Dalenius' formal basis, measuring the usefulness of a survey, considers traditional QC concerns such as accuracy, cost, relevance, etc., as well as topics more specific to surveys such as privacy protection and wealth of detail. The second component, a comprehensive QC program, has two subprograms: accounting for the control of the survey design and accounting for the control of survey operations.

It is becoming obvious that the current QC/VS program does not provide the desired control. Validation cannot identify parts of the process which are a problem, does not address much about alternative designs, and is not a continuing procedure. The QC work is a continuing procedure, but it is more of a training device than a process to detect errors and improve the survey.

The validation survey has one major goal--to measure the difference between the objective yield crop cutting and the farmer's harvest. Basically, the validation surveys have clearly shown that the difference between the OY crop cutting and farmer's harvest is not equal to zero for all crops, for all years, or for all states. Since the validation surveys have answered the major question for which they were designed, we must ask what purpose would they have in the future?

As an application of statistical decision making, statistical control consists of two parts which are approximated in the OY survey as follows:

$$\text{Statistical Control} = \text{Acceptance Sampling (VS)} + \text{Process Control (QC)}$$

which is the basic structure of statistical control as used in an industrial organization. Our difference is in terminology. We use "validation" in lieu of "acceptance sampling," and "quality control" in lieu of "process control". To understand how validation should fit into the statistical control of the yield survey, consider the three points required in statistical control development:

1. Conformance to requirements.
2. Acceptance sampling to evaluate past performance.
3. Process control to evaluate future performance.

Our first, and as will be shown, major concern when statistically evaluating a survey is to ask the question--Do we have survey requirements for which conformance can be measured? We estimate the standard error of the OY survey estimates and have "acceptable" levels for the standard error. The standard error is not sufficient since we know bias is important in OY surveys. Thus, let's consider the obvious, mean square error (MSE), since it follows our training in total survey design and is the usual quantitive procedure for evaluating the "worth" or "reliability" of a survey.

When combining total survey design with process control, we must think of the money spent on the survey as more than the resources involved with data collection. The survey costs should be thought of as the sum of costs for survey design (including the frame), sample size, validation, training, quality control, and nonsampling error studies. These costs need to be compared with the MSE rather than the variance alone.

With traditional methods, we allocate resources for enumeration based on number of samples, allocation of the samples, etc. Now we have to consider a number of possibilities for resource allocation, such as, how to allocate research time, state supervisor's time, etc. Resources can be allocated to control the mean square error based on the relative importance of the sampling error and the bias. Sampling error can be measured in a well designed survey. What is needed is a measure of the overall bias.

OBJECTIVE YIELD BIAS Many biases have been found in the objective yield surveys. Recent and past research has indicated a long list of concerns which can occur in different states, different years, and different crops. To evaluate the bias, we must also consider biases which we haven't detected or considered yet. Considering the bias as the sum of the known and unknown components gives us the opportunity to formulate a statistical control methodology.

Several considerations seem reasonable when assessing the bias. Is the overall bias consistent over the years? Our data is a time series, especially when considered by the users; thus, knowledge of bias-induced level change is important. Are the sources of bias changing? Are there large enough bias changes to deserve extra concern? Are there any needs for procedural changes to account for and level off a bias, or do we only need to monitor the overall level of bias? Finally, are the biases within a specified tolerance?

STATISTICAL CONTROL We must begin thinking about the "performance" of a survey in terms of a tolerable MSE. Consider now our ability to measure conformance to tolerance. We know that the variance can be measured from the survey data, and actions to reduce the variance are possible although constrained by resources. The bias term must also be considered. Using the validation survey to measure the overall bias allows us to compute a MSE. The methodology to measure the overall bias of OY surveys will be proposed later in this paper.

A methodology which uses measures of both terms in the MSE, the bias and variance, is consistent with the concept of total survey design. If the MSE is too large, we can decide upon survey design improvements such as allocation and sample size, or we can increase our expenditures on bias reduction techniques such as increased field inspection during the survey or increased nonsampling error studies.

At this point, we must define low and high sampling error or bias. In general, we have defined low sampling error. Many of our surveys have a desired coefficient of variation (CV) level at the national and the state levels. For example, in area frame surveys we look for a CV of 2 to 4 percent for the acreage of major crops at the national level, and a 3 to 12 percent CV at the state level. Now consider the magnitude of the bias with respect to the sampling error and total survey design decisions. If both sampling error and bias are low, there is no real need to do research on the survey; survey design, knowledge of nonsampling errors, and the quality control procedures that are existing can be considered as working quite well and research resources would be better used for some other less accurate survey. Of course, cost reduction techniques should always be considered. On the other hand, if both sampling error and bias are high, the MSE is likely to be higher than we desire, and resources for research are desirable.

If the variance is very low, we might reduce the sample size and reallocate those resources to control a high bias through an improved quality control program, nonsampling error studies, training, and the like. The converse can also be true. The bias may be acceptably low for some survey while the variance is too high; resources being spent on nonsampling error studies can be shifted to design research, reallocation, etc.

VALIDATION SURVEY In future validation surveys, we would like to have validation samples in all states. This would not only remove a very unpopular aspect of the survey, but would also improve the usefulness of the analysis. A validation survey for the major crops conducted on an annual or semi-annual basis could be used to adjust survey results and determine the official estimate's relationship with the objective yield survey.

The approach to validation surveys suggested here differs from our existing methodology by modeling error components. Suppose we knew the yield, Y_i, for our objective yield sample field. We know that our estimator of state yield, the mean of the sample field yields, is basically unbiased. For our OY survey, we do not have the true yield, Y_i, but a sample estimate, y_i. This estimate can be modeled as follows:

$$y_i = Y_i + B_{(f)i} + B_{(p)i} + B_{(nr)i} + B_{(u)i} + E_i$$

where

$B_{(f)i}$ = The bias due to the estimator y_i, if any.

$B_{(p)i}$ = The bias due to field procedures.

$B_{(nr)i}$ = Partial nonresponse bias.

$B_{(u)i}$ = Other, maybe unknown, bias sources (editing, coding).

E_i = random error due to sampling two plots in the field.

Suppose we have a measure, Y'_i, of the true yield for a subsample of the OY fields (Y'_i = yield weighed at a grain elevator, for an example).

$$Y'_i = Y_i + M_i$$

where

M_i = The "validation" measurement error at subsample field i.

We can compute for each field in the subsample

$$y_i - Y'_i = B_{(f)i} + B_{(p)i} + B_{(nr)i} + B_{(u)i} + E_i - M_i$$

Making two reasonable assumptions:

(1) $E(E_i) = 0$, and (2) $M_i = 0$,

the total bias can be estimated by

$$\text{Bias} = \sum (y_i - Y'_i)/n'$$

where n' is the validation subsample size.

Two other considerations need to be accounted for with this approach. First, there is nonresponse in the OY survey (refusals, harvested early, etc). This difference between the sampled and target population can be a source of bias. We plan to use farmer reported yield estimates with a double sampling approach to estimate this bias component. Secondly, there may be validation survey nonresponse, especially when trying to obtain elevator yields for all of the validation subsample. We need the assumption that the biases measured by the validation survey are uncorrelated with the action of obtaining elevator yields. This assumption seems reasonable, but should be tested occasionally. This approach provides three estimates: the OY yield estimate, the validation survey estimates of OY bias, and a nonresponse bias estimate, which seem sufficient for a consistent, between-year series of yield estimates.

CONCLUSION To improve survey quality in general, it is recommended that: (1) a quality control course specific to survey organzations be developed, (2) every manager in a survey organziation be required to know process control methods, and (3) the 10 quality issues of the Japan QC Prize be addressed for all surveys.

This paper presents the survey estimate, adjusted for bias, as a process control method for quality control. Our sampling training taught us that a biased estimator was useful, provided the bias was known. This approach allows an estimate of bias, regardless of the source of bias, provided the estimated bias has a sampling error which is small relative to the bias and the sampling error of the estimator.

REFERENCES

Bailer, B.A. and C.M. Lanpher. Development of Survey Methods to Survey Practices. Amer. Stat. Assoc. Washington, D.C. 1978.

Dalenius, T. "Errors and Other Limitations of Surveys" in Wright, Tommy. STATISTICAL METHODS AND THE IMPROVEMENTS OF DATA QUALITY. Academic Press. 1983

Deming, W.E. QUALITY, PRODUCTIVITY, and COMPETITIVE POSITION. Cambridge, MA., MIT. 1982.

Fecso, R. Evaluating Objective Yield Surveys Using Process Control Techniques. USDA, Nat. Agric. Stat. Serv., in preparation, 1986.

Fecso, R. and V. Johnson. The New California Area Frame: A Statistical Study. USDA, Stat. Rept. Ser. SRS-22. Sept. 1981.

Fecso, R., R.D. Tortora and F.A. Vogel. "Sampling Frames for Agriculture in the United States." J. Official Statistics, forthcoming, No. 4, 1986

Fellegi, I.P. "The Evaluation of the Accuracy of Survey Results: Some Canadian Experiences." Am. Stat. Rev., Vol. 41, No.1, pp 1-41. 1973.

Grant, E.I., and R.S. Leavenworth. STATISTICAL QUALITY CONTROL. Fourth edition, McGraw-Hill, 1972

Hogg, R.V., et. al. "Statistical Education for Engineers: An Initial Task Force Report." THE AMER. STAT., Vol. 39, No.3., 1985.

Joiner, B.L. "The Key Role of Statisticians in the Transformation of North American Industry." THE AMER. STAT., Vol. 39, No.3. 1985.

Matthews, R.V. An Overview of the 1985 Corn, Cotton, Soybean, and Wheat Objective Yield Surveys. USDA, Stat. Rept. Ser., Staff Report. Nov. 1985.

Raj, D. SAMPLING THEORY. McGraw-Hill, 1968.

Tribus, M. and Y. Tsuda. "Creating the Quality Company." Mass. Inst. of Tech.

Tribus, M. "Deming's Redefinition of Management." Mass. Inst. of Tech.

Tribus, M. "Reducing Deming's 14 Points to Practice." Mass. Inst. of Tech. 1984.

U.S. Department of Agriculture. Scope and Methods of the Statistical Reporting Service. Statistical Reporting Service, Misc. Publ. No. 1308, Sept. 1983.

U.S. Department of Commerce. First Annual Research Conference Proceedings. Bureau of the Census. March 1985.

Vardeman, S. and H.T. David. "Statistics for Quality and Productivity: A New Graduate Level Statistics Course." THE AMER. STAT., Vol 38, No.4., 1984.

Wright, T., ed. STATISTICAL METHODS AND IMPROVEMENT OF DATA QUALITY. Academic Press, 1983.

Zarkovich, S.S., QUALITY OF STATISTICAL DATA. Food and Agric. Org. of the U.N. Rome, 1966.