

University of Mississippi

eGrove

Proceedings of the University of Kansas
Symposium on Auditing Problems

Deloitte Collection

1-1-1974

Discussant's response to setting standards for statistical sampling in auditing

Lawrence L. Vance

Follow this and additional works at: https://egrove.olemiss.edu/dl_proceedings



Part of the [Accounting Commons](#), and the [Taxation Commons](#)

Recommended Citation

Contemporary auditing problems: Proceedings of the Touche Ross/University of Kansas Symposium on Auditing Problems, pp. 085-088;

This Article is brought to you for free and open access by the Deloitte Collection at eGrove. It has been accepted for inclusion in Proceedings of the University of Kansas Symposium on Auditing Problems by an authorized administrator of eGrove. For more information, please contact egrove@olemiss.edu.

Discussant's Response to Setting Standards for Statistical Sampling in Auditing

Lawrence L. Vance

University of California, Berkeley

John Broderick has raised interesting and important questions about the application of statistical sampling in auditing. The two areas of most concern that he has discussed and which I wish to comment upon are (1) the role and method of evaluation of internal control and (2) the matter of setting standards in general, with particular reference to the precision band in estimates.

Evaluation of Internal Control

Our use of internal control evaluations may be approached in at least two ways, which I refer to as the three-step and two-step methods. In the three-step method, which appears to be Mr. Broderick's preference, one first evaluates internal control on the basis of descriptive material—organization charts, procedure manuals, and conversations with members of the organization who are operating the system. On this basis, one forms a judgment about the apparent quality or effectiveness of the system. The second step in the three-step method is to test the operation of the system with documents and other records that disclose directly the working of the system. We are all aware that the system prescribed on paper and reported as functioning by members of the organization may in fact be distinctly different from the one that the people involved are actually using, and the effectiveness of the system may vary accordingly. The third step is to use the results of the first two steps in determining the "extent of the testing" (to use the traditional phrase) or to set confidence and precision limits for the sampling designed to appraise the bona fides of the accounts. If either the first or second step shows weaknesses in internal control, the confidence level is raised and the precision limits narrowed for the tests of bona fides; if both indicate effective control, these levels can be reduced. Note that this three-fold concept, if applied under a policy of keeping each step distinct, requires separate samples for step two as against step three.

In the two-step approach to the evaluation and use of internal control information, step one is the same as in the three-step procedure. However, the second step proceeds directly to tests of bona fides, and the extent of these tests, or the statistical criteria they are required to meet, are determined by the subjective evaluation made in step one.

The question that arises when we have to choose between these two concepts is this: is it necessary to have an *objective*—not subjective—estimate of the functioning of internal control before we set standards for the test of bona fides?

In making a clear-cut distinction between the three-step and two-step

procedures and in making a similar distinction between arriving at a conclusion about internal control in contrast with testing *bona fides*, I have been ignoring what evidently is the popular practice. This consists of a blending of steps two and three of the three-step method. A judgment is made about the effectiveness of internal control in step one, but it is tentative—it is modified as documents and other records are examined if this examination shows that the system is working less well than the preliminary judgment indicated. It seems to me that this is a reasonable way to proceed, as it eliminates the implied need for separate tests of documents for internal control evaluation.

Testing Statement Amounts

We must remember in this connection that we are presumably working with some kind of estimation procedure. It is possible to use acceptance sampling techniques to decide that internal control is or is not satisfactory, or the records have been kept with sufficient absence of errors so that we are willing to conclude that the records are sufficiently accurate. However, most auditors evidently prefer to think in terms of amount of dollar error when examining financial statement figures. This requires estimation procedures, and gives rise to the combination of confidence level and precision range that Broderick has discussed. Even if we restrict our attention to errors without regard to their magnitude, most auditors may prefer to estimate the percent of error rather than to set a firm accept-reject criterion. This means that there is no restriction to a single sample size or a final sample size when examining records for a combined check on both internal control and *bona fides* using estimation sampling. The advantage of setting a sample size in advance on the basis of whatever method of evaluation of internal control is selected is that the available information can be used to indicate what a likely minimum size is; in other words, economy can be maximized. But once we have information directly from the subject population itself we can, in estimation sampling, calculate the indicated result, and, if this leaves us with too wide a range of precision or too low a confidence level, we can then simply increase the sample size to the point where we have the desired assurance. Because of this possibility, I doubt the need to make a clean separation of tests of internal control and tests of the so-called *bona fides*.

Standards for Precision and Confidence

The second major thrust of Jack Broderick's remarks concerns the setting of standards for precision and confidence, and he has mentioned both the seeming desire for specific standards which some would evidently have the profession as a whole establish, and also the problem faced by each auditor in setting standards for his own work. We are in agreement on the proposition that specific numerical standards should not be set. There are infinite gradations in the quality of internal control and of materiality relative to dollar totals, and it does not seem practical to fix minimums, which are always likely to become maximums. I agree that general standards, expressed as objectives to be achieved as they are in the standards now established by the profession, are the better kind of regulation. The auditor has to tailor his confidence and precision to the complex facts of each case, and to suggest otherwise would likely do more harm than good.

This leaves the matter of confidence and precision levels to be determined specifically by each auditor, and we must ask: what general guidelines are available? We, of course, have the fact that professional statisticians most often use 95% or 99% or approximations of them for confidence levels. Another way of looking at the problem is to ask what percent of the time one is willing to be wrong in order to economize on sample size. And this leads immediately to another question: what are the consequences of being wrong? If the error is in accepting improper statements, either no one may ever know the difference or it may be discovered and there may be a lawsuit for \$1,000,000, or some equally impressive figure.

A practical approach to this decision was formulated for accountants several years ago and involves, as it must, the assigning of subjective probabilities, or expectations.¹ It is also being discussed by Professor Felix in this symposium. I recommend this approach to all auditors. If you are wrong in believing an acceptable set of figures to be materially in error, you, of course, incur the cost of the additional investigation necessary to establish the fact of acceptability. This cost must be built into the calculation just referred to as is the cost of making the opposite mistake.

Since confidence level and the precision range within which sampling results can be expected to fall are tied together, higher confidence for a particular sample means a wider precision and vice versa. To improve one while holding the other constant requires an increase in sample size. The objective is to arrive at a combination that meets our standards with as small a sample as the circumstances permit. How then should we set precision limits? Mr. Broderick has chosen to define precision in terms of half the range or "confidence interval"; standard statistical practice defines it as the whole range, recognizing that this range runs both plus and minus from our estimate of the mean of the population given by our sample. Half the range is, of course, the maximum amount we expect to be off in our estimate in one direction or the other. Mr. Broderick has indicated that some accountants are inclined to set the precision for their estimates at plus or minus a material amount.

I agree with Mr. Broderick that this is too high. My understanding of "material" in accounting usage is that it represents an amount that significantly changes the interpretation of the figure to which it applies. A precision range or confidence interval that runs in either direction from the estimate to the extent of a material amount leaves plenty of room for a book value that deviates from the proper value by a material amount. Presumably we should set confidence intervals at plus or minus a maximum tolerable error; i.e., by an amount that clearly leaves the interpretation of the published figure unaffected. This has to be an amount significantly different from a material amount; one can not set these amounts side by side. For example, if we have an inventory stated in the accounts as costing \$1,000,000, and if we consider \$100,000 material, we might well use plus-or-minus \$25,000 as our confidence interval. If our confidence interval was plus-or-minus \$100,000, and if our estimate was precisely the true amount of the inventory cost—say \$900,000—then obviously the confidence interval would tend to support the overstated book value and very likely do us no good. In other words, I share Mr. Broderick's concern for the tendency to set wide precision limits and high sampling risks as a means of justifying very small samples.

Concluding Remarks

In conclusion, I would like to endorse one more of Broderick's concerns, namely, that auditors need to familiarize themselves with standard statistical theory in order to apply statistical sampling methods effectively. The AICPA has encouraged this with its publication of self-study materials on statistical sampling, but hopefully a thorough grounding will be obtained routinely by students majoring in accounting in college, and its achievement should be the responsibility of the educators in charge of college accounting programs.

Footnotes

1. Harold Bierman, Jr., "Probability, Statistical Decision Theory, and Accounting," *The Accounting Review*, July 1962, pp. 400-405.