AUDITING: THE RELATIONSHIP BETWEEN THE SYSTEMS EXAMINATION AND THE SUBSTANTIVE TESTING OF DETAILS

BY

DAVID J. HATHERLY

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF ACCOUNTANCY IN THE UNIVERSITY OF GLASGOW

DEPARTMENT OF ACCOUNTANCY
FACULTY OF LAW

SEPTEMBER 1980

TABLE OF CONTENTS

SUMMARY

It has long been recognised by the Auditing Profession that an examination of the accounting system and its internal controls allows the auditor, provided the results of the systems examination are favourable, to reduce the extent of the detailed substantive testing. Such an approach is conventionally called a 'systems' approach to the audit. It is also recognised that the auditor does not have to place reliance on internal control if he does not judge such an approach to be cost effective. In this case the auditor maintains the level of substantive testing and the approach is termed a 'substantive' approach.

In this thesis it is argued that the generally recognised role of the systems examination in allowing a reduction of substantive testing is only one of three important audit roles which the systems examination may play in support of the substantive testing. The other two less well recognised roles are:

a) as a source of information to help the auditor select a statistical sampling technique for the substantive testing, and

b) as a source of information to help the auditor assess pieces of evidence used in the substantive testing where the evidence has been processed or generated by the system.

The main body of the thesis consists of a detailed study of each of the three roles.

In the Conclusion it is argued that a recognition of three, rather than one, roles for the systems examination allows a greater understanding of the true distinction between what are conventionally described as the 'systems' and 'substantive' approaches to the audit. Further research is recommended both to substantiate the tentative 'a priori' reasoning of the conclusion and to provide empirical evidence of the extent to which there is a lack of understanding on the part of practising auditors of the three aspects of the relationship between the systems and substantive evidence.

INTRODUCTION

> The auditor must obtain evidence to the extent necessary to provide, inter alia, sufficient assurance as to the completeness, accuracy and validity of the accounting records. For this purpose the principal sources of audit evidence are evaluation/compliance tests of internal controls and substantive tests. In general, the auditor's selected mix of evidence from internal controls and substantive testing will be such as to give the most efficient use of audit resources to achieve the level of assurance required. (1)

This statement usefully serves as the starting point for the thesis which investigates the relationships that exist between the systems and substantive evidence as employed by the external auditor. The thesis assumes that the auditor wishes to establish the credibility of the underlying figures in accounts and that the principal means to such audit verification are an examination of the accounting system and its internal controls on the one hand, and substantive testing on the other, and it sets out to promote better understanding of the variety of important roles which the systems work plays in support of the substantive testing.

The basic arguments of this thesis are expounded in the final section of this introductory chapter. However, first it is desirable to place the issue tackled in the thesis in the wider context of auditing issues, problems and relationships. It is necessary to recognise the following:

1. The "accountability relationship" considered (i.e. that between the directors and shareholders of a limited company) is just one of many such relationships, financial or otherwise, which exist in society.

2. The audit opinions considered (i.e. as to whether the balance sheet gives a true and fair view of the state of affairs and profit or loss account a true and fair view of profit or loss) are just two of five audit opinions required of a limited company's auditor.
3. Establishing the credibility of the underlying figures in the accounts is a necessary, but not sufficient, prerequisite of forming the true and fair view opinions. Other prerequisites are not considered in the thesis.
4. The two sources of audit evidence considered (systems and substantive) are two out of several sources used by the auditor to establish the credibility of accounting figures. The thesis does not consider the interrelationships between other combinations of audit evidence.

At first sight these four factors may appear to seriously limit the relevance of the problem discussed in the thesis. In fact this is not so for the following reasons:

1. A great many audits performed in the United Kingdom are limited company audits;
2. for limited company audits, the majority of audit time relates to the true and fair view rather than other audit opinions;
3. for the true and fair view opinions, much of the audit time is concerned to establish the credibility of the underlying figures, and,
4. when establishing the credibility of the figures, the majority of audit time is concerned with systems and

substantive evidence, rather than other sources of evidence.

Thus, a significant amount of audit resources are affected by the issues discussed in the thesis. The author knows from his contacts with practising auditors that the issue is regarded by them as being significant.

IDENTIFYING THE ACCOUNTABILITY RELATIONSHIP

Whenever one party in society is accountable for his actions to a second party, there exists an "accountability relationship". Should this accountability relationship be essentially a financial one, the performance of the accountability function requires the measurement, generally in money terms, of the financial events conducted. If this measurement is performed or controlled by the accountable party himself, then accountability requires communication of the measurements (in the form of financial statements) to the second party. In these circumstances, accountability is considerably improved through the performance of an examination of the financial statements by an expert person independent of the accountable party, provided the results of this examination are communicated along with the financial statement results. This examination or 'audit' is designed to add credibility to the financial statements prepared by the accountable party.

As far as this thesis is concerned, the only accountability relationship considered is that between the directors of a limited company, incorporated under the Companies Acts, and the shareholders. It does not concern accountability relationships which exist between various groups involved with organisations other than limited companies, or accountability relationships between the directors of a limited company and groups other than its shareholders. The thesis is, therefore, only directly relevant to auditing problems associated with one particular, albeit important, accountability relationship drawn from a much larger population. Although the issues discussed in this thesis may be relevant to auditing problems associated with other accountability relationships, no attempt is made to explore this question.

IDENTIFYING THE AUDIT OPINION

Under the 1967 Companies Act, the auditor of a limited company is required to form an opinion on the following matters:

1. Whether the balance sheet shows a true and fair view of the company's affairs at the balance sheet date.
2. Whether the profit and loss account shows a true and fair view of the company's profit or loss for the financial year.
3. Whether the accounts comply with the Companies Acts 1948 and 1967.
4. Whether proper accounting records have been kept.
5. Whether the company's balance sheet and profit and loss account are in agreement with the accounting records.

This thesis limits itself to consideration of the audit work which is necessary to form the first two of these five opinions (i.e. the 'true and fair view' opinions). The audit work which is necessary in order to form the true and fair view opinions will certainly involve consideration of the Companies Acts 1948 and 1967, the propriety of the accounting records and agreement between financial statements and accounting records. However, there may be additional audit work which is necessary in order to form any of the last three opinions but which is not necessarily required for the true and fair view opinions. It is considered that the question of the most efficient approach to the overall statutory audit of a limited company should be tackled in two stages:

1. What is the most efficient approach in order to form the 'true and fair view' opinions alone?
2. To what extent is it necessary to reconsider the approach identified in 1. above as a consequence of the requirements of the other three audit opinions?

Although there is an interconnection between the two, this thesis does not explicitly answer the second of these two questions, and the matter is left as a subject for further study.

Thus, the thesis is only concerned with auditing problems directly relevant to the formation of two audit opinions from a total population of five required by the 1967 Companies Act. For the most part, it is these true and fair view opinions which generate the majority of the audit work. However, there are occasions when the auditor is unable to form any true and fair view opinion as a result of uncertainty caused by a lack of suitable audit evidence. In such circumstances the auditor is unable to state whether or not, in his opinion, the accounts show a true and fair view, but he may be able to state that proper accounting records have not been kept. The thrust of the audit work is deflected from the true and fair view opinions to the opinion as to proper accounting records. The Companies Act 1967 states that accounting records shall be such as to enable the directors to ensure that any balance sheet or profit and loss account prepared by them give a true and fair view of the company's profit or loss. In circumstances where the auditor cannot ascertain whether the balance sheet or profit and loss account give a true and fair view, there must be a strong likelihood that the directors are not able to ensure that they do, and proper accounting records are not being kept.

This thesis is concerned with the audit of enterprises which have both an accounting system including internal controls (systems evidence) and substantive evidence available to support the financial statements. In these circumstances there is most unlikely to be uncertainty caused by a shortage of audit evidence. This thesis does not consider the implications for the audit of adequate audit evidence being unobtainable.

AN OVERVIEW OF THE AUDIT PROCESS

The auditor's opinion as to the true and fair view of the state of affairs and profit and loss for the year relates to the accounts as a whole and not to the individual items in the accounts. However, the forming of an opinion on the accounts as a whole must be preceded by the forming of an opinion as to the credibility of the figures in the accounts. This thesis is concerned with the audit work necessary to establish credibility in the underlying figures of the accounts. It is not concerned with other audit work which is also necessary before forming an opinion on the accounts as a whole.

Section 14(1) of the 1967 Companies Act specifies that the responsibility for the accounts of a company rests with the directors whilst the auditor's responsibility is to report on those accounts. There is unlikely to be a 'standard' audit approach in circumstances where the directors are considered by the auditor to be unreliable or untrustworthy. However, in circumstances where directors make a serious attempt to discharge their responsibility for the accounts, the audit approaches taken by the major auditing firms in order to establish the credibility of the underlying figures in the accounts are well documented in their audit manuals. Figure 1 (The Audit Process) represents a typical approach. The audit process may be thought of as a logical ordering of the various categories of audit evidence which the auditor uses to establish credibility in accounting figures. This thesis is concerned with the relationship between two stages in the audit process (systems examination and detailed substantive tests) and is not concerned to study other interrelationships.

Audit Objectives

Detailed audit objectives are developed from the auditor's reporting requirements, which for limited companies are specified in the 1967 Companies Acts, and from relevant decided cases. Satisfactory completion of these objectives is a necessary but not, in view of the auditor's responsibility to give an opinion on the accounts as a whole, a sufficient condition for an unqualified audit report. Detailed audit objectives concerned with the correct accounting treatment of a transaction or balance include validity, authority, completeness, valuation, classification and disclosure (for transactions) and existence, ownership, completeness, valuation, classification and disclosure (for balances). (2). However, the development of a complete set of detailed audit objectives is itself a major topic and outside the scope of this thesis. This thesis assumes that an agreed set of detailed audit objectives is in existence and considers,for the subset of objectives concerned with the credibility of accounting figures, alternative approaches to attainment.

CA 67 REPORTING REQUIREMENTS AND DECIDED CASES

DETAILED AUDIT OBJECTIVES

KNOWLEDGE OF BUSINESS

POPULATION STATISTICS

INTER-RELATIONSHIPS BETWEEN ACCOUNTING FIGURES

SYSTEMS EXAMINATION

DETAILED SUBSTANTIVE TESTS

AUDIT OPINION ON CREDIBILITY OF FIGURES

FIGURE 1 : THE AUDIT PROCESS

Knowledge of the Business

Knowledge of the business includes an understanding of the operations of the business and of the financial and economic circumstances in which those operations are conducted. A set of accounts do not show and true and fair view if they do not reflect the operations of the business, and hence such knowledge is essential. Knowledge of operations is also a first step toward an understanding of accounting population statistics, and interrelationships between accounting totals, the types of systems which should be in operation and the type of substantive testing which may be feasible.

Knowledge of the business includes an understanding of the attitudes, motives and character of company management. As previously stated, an assumption underlying the audit process outlined in figure 1 is that management is both able and willing to discharge its responsibility to produce proper accounts. A weak management may not be able to insist on the installation of appropriate controls, whereas a strong but unscrupulous management may override controls and otherwise interfere with audit evidence. It is essential that the auditor carefully examines and considers the character of management and the pressures under which they are operating. (3)

The major source of knowledge of the business is experience gained on prior audits. When an audit is undertaken for the first time it follows that the auditor should make a special attempt to acquire such knowledge. Indeed he probably should do so before accepting the appointment as auditor. Such knowledge may be acquired by the following means:

a) a study of the previous auditor's files

b) plant visits and tours

c) discussion with company management

d) industry and client publications, including financial statements.

<u>Population Statistics</u>

The preparation and study of population statistics is a vital part of the audit process, for the following reasons:

1. It may indicate circumstances where systems examination is uneconomic. Should a small number of items comprise either the entire accounting population or a very large per cent by value, a study of the accounting system may prove uneconomic and the auditor will concentrate on the substantive testing. Generally speaking the inter-relationships between the

systems examination and substantive testing studied in this thesis, are only appropriate where a significant per cent by value of an accounting population is made up of a large number of items.

2. It may indicate circumstances where certain statistical sampling techniques are inappropriate or where substantive testing on a sample basis at all is inappropriate. The accuracy and efficiency of variables sampling techniques can depend upon certain population statistics. For example, mean per unit sampling is inaccurate for skewed populations and inefficient for highly variable populations. Should an accounting population consist of a small number of items it may be decided that any sampling plan is inappropriate and the auditor will test items forming the population on a 100% basis.

3. It may, when combined with a knowledge of business operations, reveal elementary but fundamental errors. For example, a breakdown of the accounting population between major components of the business may reveal a component which has been omitted.

Inter-relationships between Accounting Figures

The relationship between figures in the accounts may be expected, unexpected or known. Where the relationship is in accordance with expectations the auditor may reduce his detailed study of the figures, but where the relationship is unexpected he may increase his detailed study. The auditor may base his expectations upon -

a) a study of the same relationship in past years

b) a study of the same relationship for other firms in the same business.

c) Knowledge of recent operations, and the financial and economic circumstances in which they are conducted.

Perhaps the most commonly used example of a known relationship is that created through the use of the double entry matrix (4). Thus it is possible to test for overstatement of sales by testing for overstatement of debtors, and similarly it is possible to test for understatement of debtors by testing for understatement of sales.

The AICPA (5) suggest two categories of substantive test: (a) tests of details of transactions and balances and (b) analytical review of significant ratios and trends. The latter forms a part of the 'interrelationship between accounting figures' shown in figure 1 , and in figure 1 , the former is entitled 'detailed substantive tests'.

Systems Examination

The installation of suitable internal control procedures is one means by which management satisfy themselves that proper accounts are being prepared. Internal control, therefore, is introduced in the first instance by management to discharge a management responsibility and it may then be used by the auditor as a source of audit evidence. Internal controls ~~primarily~~ cover the authorisation and recording of transactions. ~~although controls over balance sheet items (e.g. physical stocktake procedures, bank reconciliations, ageing and review of debtors etc) may also be important.~~ In addition, through the interrelationship of accounts, an examination of internal control over transactions may play a supportive role in the audit of balance sheet items.

The audit opinions required under the Companies Acts have been previously stated and none of the five opinions directly relates to internal control, but internal control does indirectly provide evidence for the true and fair view opinions and the opinion relating to proper accounting records. Hence, the auditor of a limited company incorporated under the Companies Acts uses his opinion of internal control as indirect evidence in support of the accounts and accounting records, and although he may make comments on internal control to management, this is not a primary function of the audit, since it is not required by the Companies Acts. The discussion in this thesis assumes that an opinion on internal control is not a primary function of the audit and the discussion is not necessarily relevant to audits where this assumption does not hold.

Internal control is indirect evidence and it is also subjective evidence. There are at the present time no generally accepted standards for the evaluation of a system of internal control. Experts can generally agree on the strengths and weaknesses but they tend to disagree on their relative importance. Even if an ideal system could be developed, its operation would still depend upon the people concerned. Any evaluation of internal control requires an evaluation of the people operating the system and this is inevitably a subjective area. As an illustration of this problem consider the nearest possible situation to a perfect system. Suppose a system not only includes many internal checks, reconciliations and divisions of duties but also incorporates an internal audit department which performs all the detailed audit work otherwise conducted by the external auditor. Moreover, the internal audit department produce excellent reports of work performed and conclusions reached and the department has a high level of independence. (For example, the client may be the UK subsidiary of a US corporation and the internal audit department is a part of corporate headquarters staff). In these circumstances would the external auditor be entitled to place reliance on the reports of the internal auditor without reperforming any of the internal auditor's work? The answer is that, for material areas of the audit, he would probably not be so entitled. Even if the external auditor held the internal auditor in the highest possible regard, the users of accounts are likely to expect a measure of reperformance by the external auditor and may regard any external auditor who does not do so, as being negligent. There is at least some recent case law to support this view. In the Continental Vending case (US 1962) it was stated by the judge that 'an auditor should not rely on the mere representations of a client if he can check them himself', and in the Pacific Acceptance case (Australia 1970) it was stated that 'an auditor may properly rely a great deal on enquiries made and explanations sought of the company's staff and management at the appropriate level, but prima facie this is in aid of his vouching and checking procedures and not in substitution for them'.

~~The conclusion is, therefore, that~~ Systems evidence will always need to be supported by substantive testing either because of the inherent limitations of systems evidence and/ or because the users of accounts have an expectation that substantive tests will be performed. (Reperformance by the external auditor of work done by client's staff is regarded by this author as being a substantive test).

In spite of these disadvantages, much auditing literature has suggested that emphasis be given to internal control at the expense of detailed substantive testing. Although the English Institute through the Auditor's Operational Standard (6) now recognise that an auditor may choose not to rely upon internal control if the substantive approach is judged more efficient, the English Institute's 'Statement on Auditing' No. U1 which was superseded by the Auditing Standards stated:

> The auditors should aim to reduce their detailed checking to the minimum consistent with the system of internal control ... Auditors should therefore direct their attention in the first instance to the system of internal control. (7)

The AICPA's second standard of fieldwork states:

> There is to be a proper study and evaluation of existing internal control as a basis for reliance thereon and for the determination of the resultant extent of the tests to which auditing procedures are to be restricted. (8)

Both these statements suggest the auditor should adopt what is known as the 'systems approach' to auditing. They reveal two interrelated but distinct aspects of that approach - emphasis and timing.

Firstly the auditor, as far as possible, <u>reduces</u> his detailed checking and <u>restricts</u> his auditing procedures to that level consistent with or determined by the system of internal control. The systems approach requires the auditor, as far as the system of internal control allows, to maximise his reliance on the examination of internal control. The emphasis is on the system of internal control. However, given its subjective nature, it is inadvisable to rely exclusively on internal control which should be supported by substantive testing.

Following the systems approach the auditor examines internal control <u>in the first instance</u> and <u>as a basis</u> for determination of further tests. The timing is that an examination of internal control precedes output tests. Given the systems approach the auditor is directed, on the basis of his examination of the system to those areas where the system is weak and, hence, the likelihood of error is greatest, and such an approach requires the systems examination to precede substantive output tests.

The distinguishing features of the systems approach may, therefore, be summarised as follows:

1. It gives emphasis to the system of internal control as a source of audit evidence.
2. Given the nature of internal control, it is not possible to rely exclusively on internal control, which must, therefore, be supported by substantive testing.
3. The extent of substantive testing is limited to that considered necessary for the system of internal control.
4. Since the system provides the key to the substantive testing, it is the systems examination which must be performed first.

Detailed Substantive Tests

Detailed substantive tests directly examine items making up the accounting figures on which the auditor must form an opinion, and can be tests of either transactions forming a profit or loss figure or balances forming a balance sheet figure. They provide direct audit evidence and they are also comparatively objective, especially when the sample is chosen at random and statistical techniques are used to provide a measure of the level of uncertainty which results through sampling (sampling error).

However, even if a 100% sample were taken, uncertainty would still exist due to the fact that the bona fides of individual transactions can never be established with certitude (non sampling error). The auditor must attempt to control sampling error through selection of appropriate statistical techniques and sample sizes, and he must control non sampling error through the selection of appropriate audit evidence in support of individual transactions. Both selections are likely to require knowledge of the system of internal accounting control, and the acquisition of this knowledge must precede the performance of the substantive tests.

There is a particular problem in relying on substantive testing alone for audit assurance as to the completeness of accounting records. As far as the completeness objective is concerned it is necessary for substantive testing to be supported by evaluation and compliance testing of internal controls relevant to the completeness objective. For example, substantive testing of despatches to sales invoices is not capable of providing sufficient assurance as to the completeness of sales without a detailed evaluation and compliance testing of controls designed to ensure that no goods are despatched without a despatch note being raised. On the question of completeness, analytical review techniques have an especially important part to play, and since the relationships of analytical review with systems evidence and detailed substantive tests are outside the scope of the thesis, the discussion in the thesis is of lesser relevance to the completeness objective than to other detailed objectives.

Just as there are those who advocate a systems approach, which gives emphasis to a study of the system and its internal controls, there are those who advocate a quite different substantive approach, giving emphasis to detailed substantive tests. In order to obtain background knowledge of the audit approaches adopted by UK auditing firms the author held discussions with twelve of the largest auditing firms in the UK. In each case discussion was with a partner or manager from the technical department of the firm concerned. No attempt was made to conduct a formal survey, but several of the major auditing firms stated in those discussions that they considered their approach to be 'substantive'.

Compared to the systems approach, the substantive approach has been less clearly identified and defined in official pronouncements and the auditing literature. In general, it may be applied to describe all audit strategies where the emphasis is given to the substantive testing of output rather than to the system of internal control which generated that output. The distinguishing features of the substantive approach may be summarised as follows:

1. It gives emphasis to detailed substantive testing as a source of audit evidence.
2. Generally, the performance of detailed substantive tests requires some knowledge of the system of internal control.
3. The extent of systems examination is limited to that considered necessary for the substantive tests.
4. Since substantive testing requires some <u>prior</u> knowledge of the system, it is the systems examination which must be performed first.

THE THESIS

The purpose of this thesis is to examine the relationship between the systems and substantive evidence. This relationship is examined through a study of problems associated with a systems approach to the audit and problems associated with a substantive approach. The principal problem associated with the systems approach is how much evidence to take in support of the figures in the accounts on the basis of the system and, therefore, how much reduction in substantive testing is permissible? To examine this problem it is necessary to examine the nature of internal control as audit evidence.

Chapter 2 of the thesis sets out a view of internal control as audit evidence which is consistent with conventional thinking. It then reviews current techniques for combining the systems and substantive evidence and identifies their operational difficulties.

Basically, these techniques require the selection of confidence level, and therefore sample size, for the substantive test to vary with the auditor's assessment of the system. Such techniques closely follow the ideas of the systems approach since the strength of the system determines the confidence level and hence sample size for the substantive testing.

Although the role of the systems examination as a source of evidence in support of the financial statements is most clearly recognised through a study of the systems approach, it is relevant to all audit approaches which wish to place some audit reliance upon the system. In addition to the role of systems examination as a source of audit evidence to be combined with the substantive evidence, this thesis identifies and examines two other important roles for the systems examination in its relationship with the substantive testing. These two roles are more easily recognised through a study of the substantive approach and its associated problems than through a study of the systems approach.

Nevertheless, the two roles apply to all audit approaches (including the systems approach) which involve some reliance on substantive testing.

<u>Role No. 1</u>

The auditor's traditional concern with the systems assessment as the determinant of confidence levels appropriate to the substantive testing has taken attention away from a possibly

equally important connection between the systems assessment and substantive testing. The value of substantive testing is significantly increased when statistical sampling techniques are employed. There is a range of statistical sampling techniques available to the auditor (generally these techniques are categorised as 'variables' or 'monetary unit'). In any given circumstance, a particular technique may be inefficient or misleading and selection of the right technique is enhanced through prior knowledge based to a significant extent on systems examination. The systems examination, therefore, should be regarded as a significant contributor to prior knowledge which enables the auditor to select the appropriate statistical technique for the substantive testing.
Chapters 3 and 4 investigate the importance of a prior systems examination to the selection of an appropriate statistical technique.

Role No. 2

The auditor's traditional concern with the systems assessment as the determinant of confidence levels has taken attention away from the important role of the systems examination in enabling the auditor to assess the quality of evidence used to substantiate individual transactions or balances selected in the sample. Much evidence used to substantiate individual items selected in a substantive sample is itself either processed or generated by the system and cannot, therefore, be properly assessed without a study of the system and its internal controls. The relevance of the systems examination to an assessment of the quality of substantive evidence processed or generated by the accounting system is discussed in Chapter 5.

CHAPTER 2

SYSTEMS EXAMINATION AS A SOURCE OF AUDIT EVIDENCE

This chapter examines the role of the auditor's systems examination from the perspective of the systems approach to the audit. The systems approach attempts to maximise the auditor's reliance on the support given to the financial statement figures by the system. In order to understand the nature of this support it is necessary to review the nature of the systems examination typically conducted by the auditor. This review provides the first section of this chapter. It is found that the auditor's evaluation of internal control is essentially subjective and that consequently it is necessary to combine the audit evidence provided by the systems examination with substantive testing. The second section of this chapter examines the difficulties inherent in current proposals for the combination of these two sources of evidence.

A CONVENTIONAL APPROACH TO THE SYSTEMS EXAMINATION

The systems examination is conveniently divided into four stages; ascertainment, identification, compliance and evaluation.

Ascertainment

This is the first stage and is normally achieved by a combination of the internal control questionnaire (I.C.Q.) and a "walk-through" of one of each kind of transaction. A walkthrough should be regarded not as a test of the operation of a client's system, but rather as a means of ensuring that effective communication has taken place between the auditor and client personnel who are explaining the procedures. However, a combination of I.C.Q. and "walkthrough" for ascertaining the system of internal control can be criticised as inadequate on at least two grounds. Firstly, there is the dynamic nature of internal control. "Internal control is people", and the pressures which motivate people may change. In reviewing internal control, the auditor may fail to realise that states of internal control during the period under review, are undergoing continual change due to the influence of people within the organisation. Secondly, and this is connected with the previous point, there will exist in any organisation involving people an informal as well as formal system of internal control. Conventional techniques are particularly inadequate for ascertaining the informal system of internal control in the area of relationships between personnel, and for this purpose a positional analysis of internal control may be needed. (9)

Thus, if the auditor only uses the techniques of I.C.Q. and the "walkthrough" he runs two distinct risks.

1. He may ascertain only one of several internal control systems operating during the period.
2. He may ascertain the formal internal control system rather than the informal system which actually operates.

<u>Internal Control is People</u>

To borrow a phrase, "internal control is people". A system of internal control is operated by people who are expected to perform and report in the manner prescribed by the system. But, the pressures which motivate the people in the "system" may change sufficiently that they cease to act in prescribed fashion, whereupon the internal control may lose its effectiveness.

It is difficult to evaluate the people who operate internal control procedures.

The behaviour of people in an organisation is influenced by the management style of the organisation. For instance, enterprises have been designated "mechanistic" or "organismic", dependent upon the amount of freedom of action which individuals within the organisation possess. Essentially mechanistic enterprises have rigid rules, clearly defined job specifications and hierachy of authority, whereas organismic enterprises represent a more flexible approach. Behaviour differs greatly between these differing environments with significant implications for the system of internal control. Given the systems approach, it is imperative that professional auditors are trained to a much higher standard than at present in the area of management style and its implications.

In practice, certain assumptions about the motivation of people are made by the auditor, and it follows that where these assumptions do not hold, reliance on internal control must be reassessed. The following have been identified (10) as some of the assumptions frequently made :

a) the threat of prompt exposure will deter an individual from committing fraud;

b) if there are proper organisational controls segregating important functions, people will recognise and report irregularities which come to their attention;

c) the probability of collusion is low since people will be deterred by the possible consequences of being rejected in an attempt to propose fraudulent activities;

d) personal relationships within the organisation will not influence or conflict with "formal" relationships established by the system;

e) people will perform only those functions required by the organisational structure;

f) the existence of documentary evidence that an action has taken place implies that the action has in fact been performed;

g) employees have sufficient training, experience, time and motivation to perform their duties competently and conscientiously.

In the context of <u>ascertaining</u> the system of internal control, assumptions (b), (c), (d), (e) and (g) are all relevant.

<u>Identification</u>

The second stage in the systems examination is to identify those individual controls within the system which are relevant to the audit objectives.

An internal control is difficult to define and in a general sense it includes any procedure designed to ensure that an organisation meets its objectives. A financial accounting internal control is any procedure designed to ensure that the financial accounts meet their objectives (ie provide a true and fair view of the economic activity of the organisation) and it is these financial controls with which the auditor is primarily concerned. Other internal controls form part of the auditor's background knowledge of the client and the client's business and, whilst important, they are not as directly relevant as financial controls, to the audit opinion.

Financial controls might include the following :

a) devising an appropriate and properly integrated system of accounts and records.

b) determining the form of general financial supervision and control by management, using such means as budgetary control, regular interim accounts of suitable frequency, and special reports.

c) ensuring that adequate precautions are taken to safeguard (and if necessary to duplicate and store separately) important records.

d) engaging, training and allocating to specific duties management and staff competent to fulfill their responsibilities.

and this suggestion is illustrated in figure 2.

Compliance

The third stage in the systems examination is to test that relevant internal controls are being applied as prescribed. A walkthrough of one of each kind of transaction may be sufficient for the purpose of assuring that effective communication has taken place between the auditor and client's staff. It is not sufficient evidence on which to assume compliance with the ascertained system of internal control, and it has been suggested (11) that for the auditor to arrive at an evaluation of internal control on the basis of completing a questionnaire and observing a few transactions is tantamount to the statistician's describing a population on the basis of a sample of one. The auditor must make a statistically supportable test of compliance, although sample size can be reduced by incorporating the auditor's subjective prior expectations in a Bayesian analysis. A frequent application of statistical sampling in the leading UK firms has been attribute sampling (normally a form of acceptance sampling) to give a maximum

ELEMENTS OF INTERNAL CONTROL

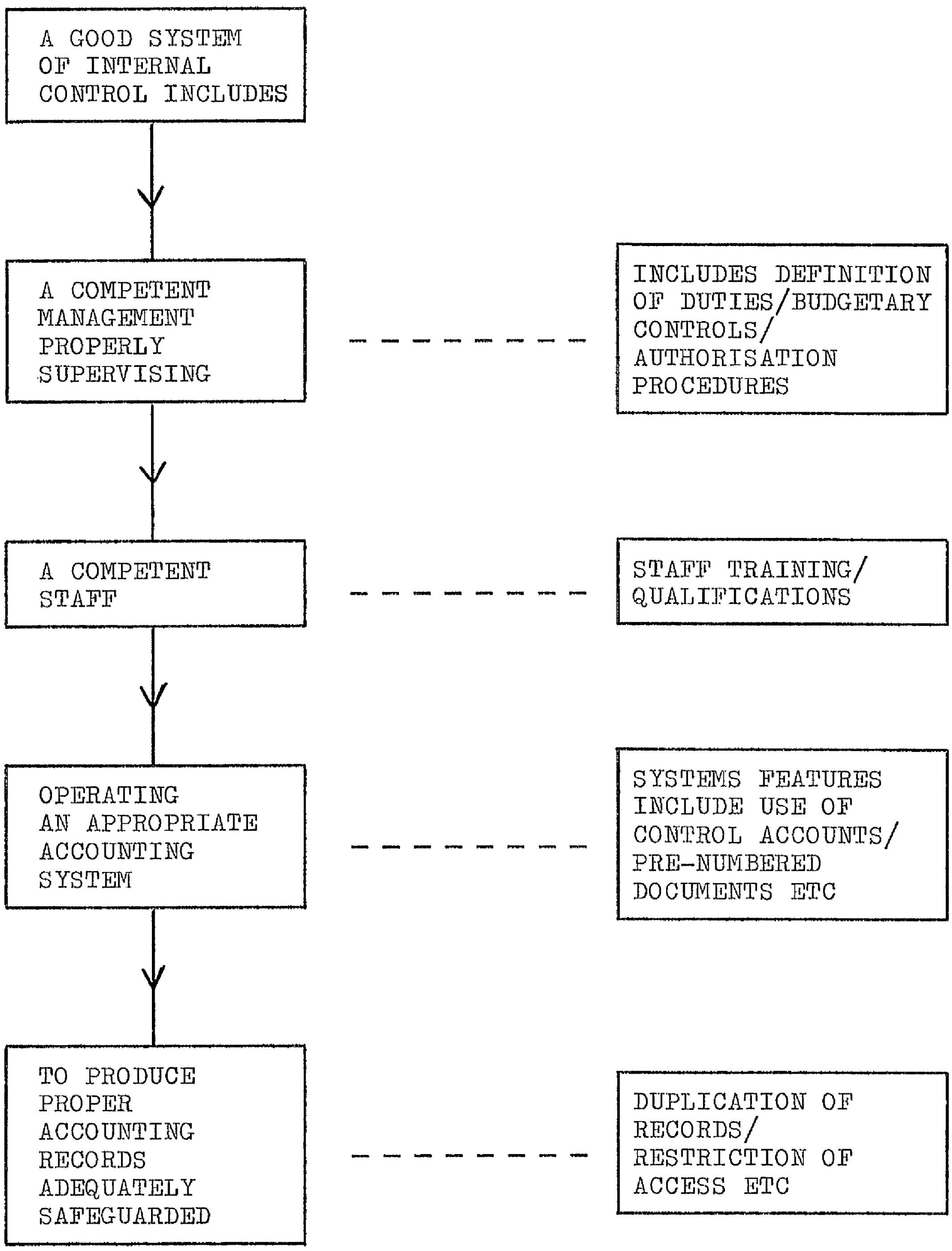


FIGURE 2

potential rate of systems error, although it is doubtful whether such maximum potential error rates are of great use to the auditor since they are not expressed in value terms. Still further difficulties are experienced by those using sampling methods to test compliance since -

1) Many controls (such as segregation of duties) do not leave an audit trail

and

2) where an audit trail is left, it is not conclusive evidence that the control procedure has been performed (eg a signature evidencing arithmetic check of an invoice).

For the second of the above problems it is sometimes suggested that the auditor himself should reperform the control procedure (eg the auditor himself reperforms the arithmetic of the invoice). However, such a reperformance by the auditor can be regarded as a substantive test rather than a compliance test and justification for this view is given in an article by this author (12).

Establishing compliance is essential to the systems approach and yet in many cases, satisfactory compliance evidence may be hard to find, especially if reperformance is to be regarded as substantive.

Evaluation

Evaluation of internal control is the final stage and there are two possibilities. Either an overall subjective assessment can be made or the auditor can employ a more precise analysis. The argument in favour of a precise analysis is made by Mautz and Mini.

> In analysing internal control, the auditor must deal with specifics, not with generalities. The auditor must determine whether specific weaknesses exist, the irregularities thereby permitted, and the specific modifications of his program called for by these conditions. In this way many of the problems associated with the overall, more subjective approach to internal control evaluation would be eliminated. For instance, if the situations by which irregularities are permitted are analysed in a logical manner, the circumstances under which two auditors would disagree as to the presence or absence of weaknesses in a given system of internal control should be rare. Personal standards could, undoubtedly, affect an auditor's assessment of the seriousness of a given weakness (probable irregularity), but not his conclusion as to the presence of that weakness.(13)

Also favouring a precise analysis the A.I.C.P.A.'s Committee on Auditing Procedure states :

> A conceptually logical approach to the auditor's evaluation of accounting control, which focuses directly on the purpose of preventing or detecting material errors and irregularities in financial statements, is to apply the following steps in considering each significant class of transactions and related assets involved in audit:
>
> a) Consider the types of errors and irregularities that could occur;
>
> b) Determine the accounting control procedures that should prevent or detect such errors and irregularities;
>
> c) Determine whether the necessary procedures are prescribed and are being followed satisfactorily;
>
> d) Evaluate any weaknesses - ie types of potential errors and irregularities not covered by existing control procedures - to determine their effect on (1) the nature, timing, or extent of auditing procedures to be applied, and (2) suggestions to be made to the client.(14)

Cushing (15) has suggested a mathematical approach which is consistent with the A.I.C.P.A. recommendation. In his single error - single control case Cushing states that the probability of error after application of the control is :

$$(1 - P_1)(1 - P_2) + (1 - P_1)P_2(1 - P_4) + P_1(1 - P_3)(1 - P_5)$$

where -

P_1 = the probability that the process is correctly executed prior to administering the control procedure.

P_2 = the probability that the control step will detect and signal an error given that one exists.

P_3 = the probability that the control step will not signal an error given that none exists.

P_4 = the probability that the correction step will correct an error given that one exists and has been signalled.

P_5 = the probability that a failure of the control step will be detected and no correction made, given that the control signals an error when none exists.

The mathematical approach suggested by Cushing provides a conceptual framework in which the auditor can evaluate internal controls in terms of numeric error rates. The people problem can be illustrated by the problem in assessing P_4, the probability that the correction step will correct an error given that one exists and has been signalled. Assuming a low prior error rate $(1-P_1)$ the auditor's sample is unlikely to include many errors and almost certainly will not include sufficient errors to enable an empirically based statistical conclusion about P_4. In these circumstances, the auditor must largely base his assessment of P_4 on subjective evaluation of the people operating the procedure.

<u>Numeric error rates and monetary error rates</u>

Following Cushing, it is conceptually possible to evaluate internal control in terms of the number of likely errors. Generally, one internal control system operates on transactions of widely varying monetary sizes. (Occasionally a separate internal control system will operate, either formally or informally, for large transactions). If an internal control system is perfect it will permit no errors of any size. Numeric and monetary error rates are zero. As soon as an internal control system is slightly less than perfect and allows, say, just a single error, the problems begin. The numeric error rate is known (it is 1 divided by the number of transactions in the population), but with one exception, the probability distribution for monetary error rates is unknown. The exception is when -

(1) the single error is known to be an omission of a transaction, and

(2) the population is known.

In this circumstance the monetary error will take each transaction size with a probability corresponding to the relative frequency of the transaction size in the population. In all other circumstances the numeric error rate can only be translated into a monetary rate subjectively. The auditor uses his judgement and experience to determine what he thinks is the size of the monetary error permitted by the internal control system. It follows that <u>for all less than perfect internal control systems, numeric error rates can only be translated into monetary error rates subjectively and the assessment of internal control is subjective rather than objective evidence</u>. Kaplan states that an audit population is actually a mix of two quite different populations; one consisting of a large number of correct items; the other a much smaller population of items in error (16).

Hence, the need to estimate two different parameters in an audit population - the error rate and the distribution of errors for items found to be in error (17). An examination of internal control only provides information for the first of these two parameters.

Qualitative features of compliance errors

The first stage in a precise analysis of internal control is to consider the types of error and irregularity which could occur. The internal control system is then analysed for each error type separately. A classification by error type is a classification according to the qualitative features of errors, and an exhaustive list of possible error types was prepared by Mautz and Sharaf (18). Knowledge of the qualitative characteristics of errors can assist the auditor in his subjective translation of numeric error rates into money error rates.

For example,

1. Errors due to misunderstandings are likely to be more substantial than errors due to lack of care.
2. Errors of computation are likely to take a wide range of values, whereas errors in an original count are likely to have a smaller range of values.
3. Errors of omission and overstatement have a theoretical limit (the size of thetransaction), whereas there is no theoretical limit for understatements.

The extent to which this kind of knowledge can assist the auditor is limited, and, hence, an examination of internal control provides relatively <u>weak</u> subjective evidence.

If it is possible for the auditor to express his assessment of internal control in money terms he is unlikely to hold his assessment with any great conviction.

THE SUBJECTIVE NATURE OF THE SYSTEMS EXAMINATION: A SUMMARY

The following difficulties with the systems examination have been identified:

1. Ascertainment. Techniques used to ascertain the system of internal control are susceptible to error because of the dependence of internal control upon the people operating the system. The auditor is, in general, forced to make assumptions about the people in the system and these assumptions may not hold for a part of the year.
2. Compliance. If the auditor is relying upon the system as audit evidence then establishing compliance is essential. In many cases, however, convincing compliance evidence is not available to the auditor. Reperformance of control procedures by the auditor may not provide suitable compliance evidence since the items rechecked were probably correct before they entered the control.
3. Evaluation. Even the most precise evaluation of internal controls (for all less than perfect internal controls) is not capable of expressing a result in monetary, rather than numeric, terms. A result in numeric terms is of only limited value to the auditor.

For each of these reasons the auditor's examination of the system of internal control is a subjective process.

COMBINING THE SYSTEM EXAMINATION AND SUBSTANTIVE TESTS

If the system of internal control is poor, although the probability of the output of the system being in error is increased, it is not conclusive evidence that the output is in error. A poor system, in the sense that it does not have intrinsic design features which permit the auditor to rely upon it as audit evidence, is quite capable of producing accurate output since a poor system may fail to detect an error once made but does not affect the number of errors available for detection. In terms of Cushing's analysis the control procedure does not directly affect P_1, the probability that the process is correctly executed prior to administering the control procedure.

To take a simple example, a system which had no segregation of responsibilities would be a poor system from the audit point of view, but is quite capable of producing accurate output if the employee concerned is both capable and honest.

A good system on the other hand, although reducing the probability of inaccurate output, is not conclusive evidence of the accuracy of that output. A good system should reduce the number of errors, but not necessarily their value. It follows that internal control, whether good or bad, and whatever the rigour and extent of the audit examination, is by itself inconclusive evidence of the output and must be combined with evidence provided by substantive testing.

Three approaches to the problem of combining the assessment of internal control with the sampling evidence of monetary errors are considered:

1. Arbitrary variation of confidence levels of a sample design as a function of the state of internal control.
2. The A.I.C.P.A. approach given in Statement on Auditing Standards No. 1, Section 320, Appendix B, paragraph 35 (the A.I.C.P.A. approach).
3. Utilisation of a Bayesian approach to sample design.

Arbitrary Variation

There are several examples of arbitrary variation. Willingham and Carmichael suggest the following (19).

Relationship of Internal Control Effectiveness to Reliability

Evaluation of Internal Control	Confidence Level
Excellent	75%
Good	90%
Fair	95%
Poor	99%

Commenting on this approach Smith says:

> Although this approach does accomplish the purpose of inversely relating sample size and internal control, it is difficult to accept as a logical approach. No logical basis has been determined for setting the confidence level correlation with different states of internal control (20).

A more cautious example of an arbitrary relationship is given in The Arthur Young Journal (21):

	Evaluation of Internal Controls		
	Excellent	Fair	Weak or Non-existent
Confidence Level	80 to 90%	95 to 99%	97.5 to 99%

A third example is related by McRae (22):

Evaluation of Internal Control System	Confidence Level
Very good	63%
Average	86%
Rather poor	95%

This table is linked to the "MUS" system as follows:

Confidence Level	Confidence Factor
63%	1.0
86%	2.0
95%	3.0

This approach incorporates the idea that it is imprudent, even when internal control is excellent, to select a confidence level giving a confidence factor of less than 1. If the confidence factor is less than 1, the sampling interval (J) is greater than the maximum potential error (MPE) and hence it is possible for a material item to be untested. Concepts such as the sampling interval (J) and maximum potential error (MPE) are discussed in Meikle (23). Although there is reasoning behind this suggestion it is, nevertheless, an arbitrary approach. All arbitrary approaches can be criticised on two grounds:

1. What is meant by an "excellent" system of internal control, a "good" system etc? A 10% probability of material error might be regarded as "excellent" by one person, but only as "good" by another. If the purpose of the arbitrary approach is to promote uniformity within a firm (albeit uniform

application of an arbitrary procedure) then it could fail because of the subjective nature of the terms "excellent", "good", etc.

2. The linking of internal control evaluation and confidence level is not only arbitrary but it is also a relationship with no direct linking and, therefore, an extremely difficult relationship for the auditor to visualise. This is in contrast to the Bayesian approach, discussed later, where the subjective element is a prior probability distribution more easily visualised by the auditor.

A.I.C.P.A. Approach

The A.I.C.P.A. have suggested that internal control and substantive tests can be linked according to the mathematical law of joint probability (see table 1).

Table 1

Probability of Internal Control Allowing Material Error	Confidence Level	Overall Confidence
(1)	(2)	(3)
70%	93%	95%
50%	90%	95%
30%	83%	95%
10%	50%	95%

Column 1 is the auditor's subjective assessment of the probability of the internal control system permitting sufficient errors to cause the account to be incorrect by a material amount. It is not the probability of an individual transaction being in error by a material amount. Column 2 is the required confidence level from the substantive test, in order that the overall confidence in the accounts figure should be 95% (column 3).

More correctly, rather than column (2) being the confidence level employed in the substantive test it should be regarded as 1 - B (the beta risk associated with the substantive test). Only in the case of dollar unit sampling (as opposed to variables sampling) will 1 - B and the confidence level be the same figure. The A.I.C.P.A. approach recognises that the risk of failing to detect a material error is based on two separate risks -

1. the risk of internal control permitting a material error.
2. the risk of the sampling plan failing to conclude that there is a material error. (Beta risk)

The A.I.C.P.A. procedure gives an overall confidence which is consistent with both the subjective assessment of internal control and the objective evaluation provided by the substantive test. In this respect it may be regarded as a Bayesian approach. However, unlike the Bayesian approach, alternative hypotheses and the auditor's entire probability distribution are not considered. It should also be noted that three of the twenty-one members of the Committee on Auditing Procedure dissented from the inclusion of the A.I.C.P.A. approach in S.A.P. No. 54, (later consolidated into Statement on Auditing

Standards No. 1.) They felt it may be misunderstood in application as prescribing a generally appropriate relationship between internal control and substantive tests. Clearly, they did not feel that the A.I.C.P.A. approach was generally appropriate. Although their precise reasons were not given, the failure to consider the auditor's entire probability distribution in Bayesian analysis is a technical weakness. In addition, the approach suffers from operational problems.

Operational difficulty of A.I.C.P.A. Approach

It can be seen from table 1 that when internal control is very poor (eg 70% probability of material error) very little of the overall confidence is drawn from the assessment of internal control. As the assessment of internal control improves more of the overall confidence is drawn from the internal control and at an increasing rate. A shift in column (1) from 70% to 50% causes a shift in column (2) from 93% to 90%. However, a shift in column (1) from 30% to 10% causes a decrease in column (2) from 83% to 50%. This problem can be represented mathematically as follows:

i = probability of internal control allowing a material error

s = confidence level for substantive test

$$i(1 - s) = 0.05$$
$$i - is = 0.05$$

Differentiating with respect to i

$$1 - i\frac{ds}{di} - s = 0$$

$$\frac{ds}{di} = \frac{1 - s}{i}$$

But $\quad s = \dfrac{i - 0.05}{i}$

$$\frac{ds}{di} = \frac{1}{i} - \left(\frac{i - 0.05}{i^2}\right)$$

$$= \frac{0.05}{i^2}$$

Thus,

i	$\frac{ds}{di}$
0.7	0.10
0.5	0.20
0.3	0.56
0.1	5.00

It can be shown, therefore, that the confidence from sampling becomes increasingly sensitive to shifts in the auditor's assessment of internal control as internal control improves. Let us consider the following possibilities.

1. Internal control is poor (eg 70% probability of material error). In this case the sampling confidence is not sensitive ($\frac{ds}{di} = 0.10$) to the auditor's assessment of internal control. The auditor does not have to be precise in his assessment. However, in this case nearly all the overall confidence is drawn from the sample and internal control only plays a small part.

2. Internal control is good (eg 10% probability of material error). In this case the sampling confidence is sensitive to the auditor's assessment of internal control. ($\frac{ds}{di} = 5.0$). This high level of sensitivity requires the auditor to be precise in his assessment of internal control. Given the

nature of internal control this requirement is likely to render the A.I.C.P.A. approach inoperational.

To summarise the argument,

When sampling confidence is insensitive to the assessment of internal control, internal control is unimportant. When internal control is important, the sampling confidence is highly sensitive to the assessment of internal control. This leads to operational difficulties. This argument is covered in a published note contributed by this author (24).

It is doubtful whether an auditor should use an assessment of internal control better than S = 0.3 (30% probability of material error). In this case $\frac{ds}{di} = 0.56$. This would mean a reluctance to use sampling confidences lower than 83%.

A.I.C.P.A. Approach: Conclusion

Although it is correct to say that the probability of material error is based on the probability of internal control allowing material error and the probability of the sample test failing to recognise material error, the A.I.C.P.A. approach does not provide a satisfactory detailed method of combining the probabilities.

The Bayesian Approach

There are two ways in which the Bayesian approach can be employed to link internal control and sampling evidence.

1. Indirectly via the "pre-audit sample" concept to examine possible confidence levels for a conventional sampling plan.

2. Directly, ie the auditor formulates the prior probability distribution of error and combines this with the sample result in a full Bayesian exercise.

Pre-audit sample

The pre-audit sample concept is based on the fact that when the prior probability distribution is a beta distribution, it combines with the results of a binomial sample to give a posterior distribution which is itself a beta distribution. The beta distribution is, in general,

$$f(E) = KE^{r-1}(1-E)^{n-r-1}$$

and when $r = 1$ and $n = 2$

$$f(E) = K$$

Thus, the situation of equal priors (the prior probability distribution is level) is described by a beta distribution with $r = 1$ and $n = 2$. It follows that any beta probability distribution can be regarded as a Bayesian combination of level priors and a binomial sample. This binomial sample is termed the "pre-audit sample".

Let us assume -

1. An audit sample reveals no errors

and

2. The auditor feels he has a posterior confidence (ie that derived from the sample and internal control) of 99%.

Then the following results are implied by the "pre-audit sample " concept:

Confidence chosen for sampling plan	Percentage saving in sample size due to reliance on internal control
99%	0%
95	35
80	65
63	78

If a sampling confidence of 63% is chosen, the auditor is using his assessment of internal control to justify a 78% reduction in his sampling work given equal priors. Teitlebaum (25) has commented that in view of the weak subjective nature of internal control evidence, 78% may be regarded as far too high and, hence, a sampling confidence of 63% as far too low. The pre-audit sample concept gives useful insights into the problem of selecting confidence levels, but does not directly provide the auditor with a technique for combining sampling and non-sampling evidence. Such a technique is the full Bayesian exercise.

Full Bayesian Exercise

Presumably, an auditor who uses dollar unit sampling could specify a subjective prior probability distribution of error and combine this with his objective sample result by means of Bayes theorem. Many authors have suggested the value of Bayesian techniques in auditing, including Kraft (26), Sorensen (27) and Tracy (28) It is interesting to note that even though these pioneering articles are around ten years old, this author knows of no formal and explicit involvement of

Bayesian techniques in auditing at the present time. However, Bayesian techniques are implicitly employed on a wide scale. As a simple example of the Bayesian approach, assume the following prior distribution -

Reported value £1M

Overstated by	Probability
£ 5000	0.05
15000	0.10
25000	0.30
30000	0.25
35000	0.20
40000	0.10
	1.00

Suppose also that in a random sample of 100 pounds, 0 errors are found. (The example, for illustrative purposes assumes 0 errors are found. Further illustration of the Bayesian calculations may be found in any of the references just mentioned). In these circumstances the results of a Bayesian analysis and the results provided by the A.I.C.P.A. procedure may be compared. Columns (9) and (10) of table 2 give the results of the Bayesian and A.I.C.P.A. methods respectively and it can be seen from these two columns, that whatever error rate (other than 0.005) is considered material, the results given by the two methods differ significantly.

Psychological conservatism

One of the major contributions of the Bayesian approach is that it provides the only posterior probability distribution which is consistent with the prior probability distribution and

(1) Error Rate	(2) Priors	(3) Cumulative Priors	(4) Likelihoods	(5) Posteriors assuming level priors (4) x scaling factor
0.005	0.05	1.00	0.6058	0.6060
0.015	0.10	0.95	0.2206	0.2208
0.025	0.30	0.85	0.0795	0.0797
0.030	0.25	0.55	0.0476	0.0478
0.035	0.20	0.30	0.0284	0.0286
0.040	0.10	0.10	0.0169	0.0171
			0.9988	1.0000

(6) Cumulative likelihoods	(7) (2) x (4)	(8) Bayesian Posteriors (7) x scale factor	(9) Bayesian Cumulative Posteriors	(10) (3) x (6)
1.0000	0.03029	0.3172	1.0000	1.0000
0.3940	0.02206	0.2311	0.6828	0.3743
0.1732	0.02385	0.2498	0.4517	0.1472
0.0935	0.01190	0.1247	0.2019	0.0514
0.0457	0.00568	0.0595	0.0772	0.0137
0.0171	0.00169	0.0177	0.0177	0.0017
	0.09547	1.0000		

TABLE 2

CALCULATION OF RESULTS GIVEN BY THE BAYESIAN AND AICPA METHODS the sample results. Of course, man can judgementally revise his prior probabilities in the light of sample evidence, but experiments conducted by psychologists demonstrate that man is not a competent intuitive statistician. For example, in an experiment by Phillips and Edwards (29), subjects, when acting intuitively, consistently failed to revise their prior probabilities to the extent prescribed by Bayes Theorem. People intuitively extract less meaning from sample evidence than Bayesian techniques would suggest and this phenomenum is termed "conservatism". It is, therefore, undesirable that auditors intuitively combine objective and subjective probabilities since this will result in conservatism and inefficiency.

The design of Phillips and Edwards' experiments was along the following lines : consider 10 bags each containing 100 poker chips, but with r bags containing predominantly red chips and (10 - r) bags containing predominantly blue chips. Predominantly red bags contain p red chips and q blue chips whilst predominantly blue bags contain q red chips and p blue chips ($p > q$). One of the 10 bags is selected at random and from the chosen bag the subject is allowed to sample 20 chips (with replacement). From his knowledge of the sample result and the design of the experiment, the subject was asked to give his intuitive assessment of the probability of the chosen bag being predominantly red or predominantly blue. These intuitive probabilities were compared with the objective probabilities calculated by Bayes Theorem. The experiment was repeated using different subjects and different figures for r, p and q; the subjects were five volunteer University of Michigan undergraduates. These experiments did not concern auditors or auditing, but a recent experiment by Corless (30) also found evidence of conservatism, this time, when a comparison was made between the Bayesian revised and judgementally revised distributions of auditors. It is this psychological conservatism which causes the Bayesian method to be more efficient than judgemental methods and Corless concludes that the use of Bayesian techniques may enable the auditor to achieve a given level of assurance with a smaller sample size.

Significant difficulties of the Bayesian Approach

There are two primary reasons why an explicit Bayesian approach does not at the present time provide the auditor with a practical solution. They are :

1. Bayesian revised distributions are unreliable when the sampling and non-sampling (internal control evaluation) evidence contradict.
2. There is considerable doubt whether satisfactory prior probability distributions can be elicited from auditors.

Contradictory evidence

When the error rate is material and unexpected, the prior probabilities could be much lower than the probabilities computed from the sample result. Bayes theorem combines these two probabilities mathematically to give the posterior probabilities consistent with the priors and the sample result.

However, following the sample result the auditor may no longer be confident of his prior distribution, and it follows that he should not rely on the Bayesian revised distribution. Corless comments (in Corless' experiment 10 per cent is regarded as a material error rate):

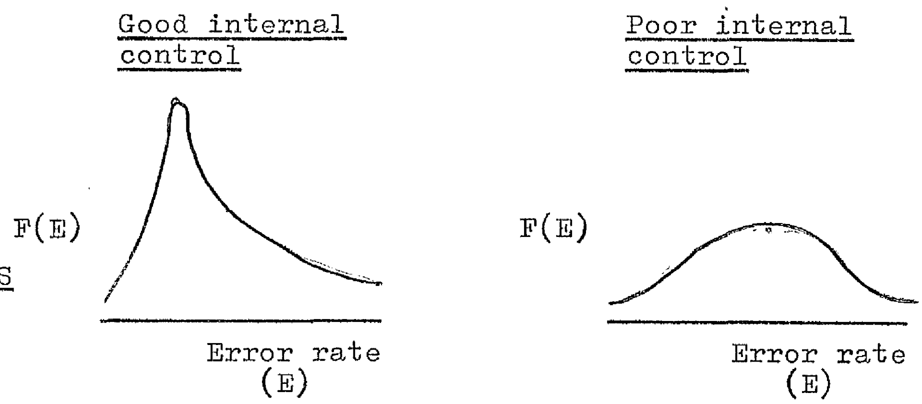
> It may be that the judgementally revised distributions are a better reflection of the meaning of the audit evidence than the Bayesian revised distributions when the sampling evidence contradicts the non-sampling evidence. That is, it seems rational that as an auditor obtains evidence which contradicts other evidence he has obtained, he becomes less certain as to what the error rate is; he does not become more certain that the error rate is somewhere between the smaller error rates suggested by the non-sampling evidence and the 10 per cent sample error rate (31).

When the sampling evidence contradicts the non-sampling evidence, it may be reasonable, given the subjective nature of the non-sampling evidence, that the auditor should seek to give far greater weight to the sampling rather than the non-sampling evidence, and this weighting is not easily facilitated by the Bayesian approach.

<u>Eliciting Prior Probability Distributions</u>

A Bayesian revised distribution can only be as reliable as the subjective prior probability distribution on which it is (in part) based. In most audit circumstances it is likely that appropriate parameters can be chosen for the beta distribution ($f(E) = KE^{r-1}\ (1-E)^{n-r-1}$) so as to satisfactorily represent the continuous prior probability distribution of the auditor.

EXAMPLES
OF BETA
DISTRIBUTIONS



Given a suitable choice of parameters the beta distribution can be made to peak close to zero and tail off to the right (this representing a siutation of good internal control) or parameters can be chosen so as to produce a much flatter curve representing poor internal control.

Corless' Experiment

The feasibility of obtaining reliable prior probability distributions from auditors was investigated in the experiment by Corless. 88 Certified Public Accountants participated in this audit experiment; they each received two audit cases concerning payroll, one with strong and the other with weak internal control; a set of questions designed to investigate their prior probabilities for payroll error; a hypothetical set of sample results and a second set of questions designed to investigate their judgementally revised probabilities in each case.

In the first place Corless was simply concerned to discover whether auditors would be willing and able to provide answers to questions designed to investigate their prior probabilities for payroll error. It was found that auditors were willing, in this experimental study, to provide answers from which prior probability distributions could be constructed. By answering one question, the auditor specified the values of the first quartile, median and third quartile of a probability distribution. From this information, the experimentor was able to construct a beta distribution and compute the values of its mean and variance. Some doubt must remain, however, whether auditors would be quite so willing to release the necessary answers in a non-experimental situation. Corless concludes:

> It must also be realised that auditors may be less willing to specify information about prior distributions for real audit situations when they are aware that these prior distributions will be used in determining the kind of professional opinion to be rendered (or whether to gather more evidence).
>
> Nevertheless, the fact that all the auditors who participated in this study specified all information requested strongly suggests that auditors are willing to specify information from which prior distributions can be constructed (32).

In the second place, Corless was concerned to discover the shape of the prior probability distributions obtained from auditors. It was found that whilst most of the prior distributions had the majority of the probability concentrated on small error rates, there was considerable variability among the distributions assessed by different auditors for each audit case. Hence, auditors were <u>externally inconsistent</u> in the sense that the shapes of their individual beta distributions did not closely correspond with each other. To some extent this variation amongst auditors must be expected since prior probability distributions are subjective and will vary from individual to individual. However, it might have been hoped that individuals with with the same professional training would possess prior probability distributions which were broadly similar. Corless suggests that some of the variability between auditors would decrease in a non-experimental situation.

> It is possible that different auditors read different things into the audit cases when they are not certain as to the meaning of some of the facts given in the cases (whereas in real audit situations the auditor could resolve these uncertainties). (33)

The variability of the prior probability distributions tended to throw doubt on their reliability. Corless was concerned to examine how well the prior probability distributions obtained really reflected the auditors' beliefs about the audit population, and a sub-experiment was organised. When providing information for the construction of a prior distribution the auditor was asked two questions. By answering one question, the auditor specified the values of the first quartile, median and third quartile of a probability distribution and from this information the experimentor was able to construct a beta probability distribution. By answering the other question the auditor specified the probability that the actual error rate was contained in each of the following intervals: zero to 0.01; 0.01 to 0.02; 0.02 to 0.05; 0.05 to 0.10, and over 0.10, and from this information the experimentor was able to construct a discrete distribution. In Corless' experiment the level of consistency between the two distributions was taken to indicate how well the distributions reflected the auditor's belief. In fact, Corless found considerable inconsistencies between the two distributions. Hence, auditors were <u>internally inconsistent</u> in the sense that when giving two different sets of information from which prior distributions could be constructed, they gave two sets of information which were inconsistent.

Corless also examined the effect of the auditor's audit and statistical experience on his choice of prior probability distribution. He found, rather surprisingly, that the extent of audit or statistical experience was not a factor, whereas the kind of audit experience did influence the auditor's choice.

Corless reported the following (expected) results :-

> It was found that the prior distributions had more probability concentrated on small amounts of error when the internal control is stronger. It was also observed that auditors whose clients have stronger internal control tend to assess prior distributions having more probability concentrated on small amounts of error than do auditors whose clients have weaker internal control. (34)

The following result, however, might not be expected :-

> Differences in the amount of audit experience and differences in the statistical background of the auditor were <u>not</u> found to affect the assessment of prior distributions. (35)

In conclusion, therefore, Corless found that auditors are willing to specify information from which prior distributions can be constructed, but there remain serious doubts as to how reliable and effective these distributions might be.

COMBINING THE SYSTEM EXAMINATION AND SUBSTANTIVE TESTING: A SUMMARY

Three approaches to the problem of combining the assessment of internal control with sampling evidence of monetary errors have been considered:

1. Arbitrary variation of confidence levels of a sample design as a function of the state of internal control.

2. The A.I.C.P.A. approach given in Statement on Auditing Standards, No.1, Section 320, Appendix B, paragraph 35.
3. A Bayesian approach.

Arbitrary Variation

The first of these possibilities is found unsatisfactory for the following reasons:

1. It is difficult to provide any definitive criteria as to what constitutes an "excellent", "good", "poor" (or whatever other categories are permitted) system of internal control. Without such criteria the arbitrary variation method is likely to degenerate into a method dependent upon the judgement of the individual auditor as to the appropriate description of the internal control. Comprehensive guidelines as to what constitutes each category of internal control are desirable if the method is to be at all convincing.
2. Even if such comprehensive guidelines exist, the linkage between each category of internal control and related confidence level is essentially arbitrary and with little or no logical support.

The A.I.C.P.A. Approach

The second possibility is found unsatisfactory for the following reasons:

1. As internal control improves the selection of the sampling confidence level becomes more and more sensitive to the assessment of internal control. Given the subjective nature of internal control assessments, this leads to operational difficulties.

2. The result of applying the A.I.C.P.A. procedure can diverge significantly from the result obtained by applying a Bayesian exercise in the same circumstances. Given that the Bayesian approach has the more rigorous mathematical underpinning, this divergence casts doubt on the adequacy of the A.I.C.P.A. procedure.

The Bayesian Approach

Finally, the Bayesian approach, whilst promising, suffers from two crucial drawbacks:

1. Bayesian revised distributions are unreliable when the evidence of the internal control assessment and substantive test results contradict.
2. There is considerable doubt whether satisfactory prior probability distributions can be elicited from auditors.

Thus the investigation conducted in this chapter suggests that there is no satisfactory method of combining the internal control assessment and substantive test results. This, together with the subjective nature of internal control highlight the problems of the systems approach which seeks to maximise reliance on the systems examination and assessment.

CONCLUSION

This chapter has demonstrated the subjectivity of systems evidence, which is shown to result from three distinct causes:

a) techniques used to ascertain the system of internal control are susceptible to error because of the dependence of internal control upon the people operating the system (a "people" problem).

b) In many cases convincing compliance evidence is not available to the auditor (a "compliance" problem).

c) Even the most precise evaluation of internal controls is not capable of expressing a result in monetary, rather than numeric terms (an "evaluation" problem).

Since the subjectivity results from no less than three distinct causes, each one of which can be a major factor, systems evidence is likely to be subjective in all but a small minority of circumstances.

It is the relative subjectivity of the systems evidence which requires it to be combined with substantive evidence. As stated in the introductory chapter, the principal problem associated with the systems approach is how much evidence to take in support of the figures in the accounts on the basis of the systems and therefore how much reduction in substantive testing is permissible given the subjective nature of systems evidence? This chapter demonstrates that the audit profession is some way from solving this problem.

Three approaches to the problem of combining the assessment of internal control with sampling evidence of monetary errors have been studied and each is found wanting:

1) arbitrary variation of confidence levels of a sample design as a function of the state of internal control.
2) The A.I.C.P.A. approach given in Statement on Auditing Standards No 1, Section 320, Appendix B, paragraph 35.
3) A Bayesian approach.

At the present time, therefore, there is no totally satisfactory formal method of combining the systems and substantive evidence in existence. The traditional role of systems evaluation is to enable a reduction in levels of detailed substantive tests. In a sense, therefore, the traditional role of systems evaluation is to serve as a source of substantive evidence. This chapter has highlighted the difficulties involved in the treatment of systems evaluation as a form of substantive evidence.

CHAPTER 3

MONETARY UNIT SAMPLING AND PRIOR INFORMATION

The problems of the substantive approach to the audit are twofold :

1) how many sample items to include in the substantive testing?

2) what evidence to choose to substantiate the sample items chosen?

The first of these questions is brought into focus by a study of those statistical sampling techniques presently used by the auditor to evaluate sample results. Both this and the following chapter examine problems associated with the use of various statistical sampling techniques. This chapter examines those techniques commonly referred to as 'monetary unit sampling' and the following chapter examines 'variables sampling' plans. Both examinations are addressed to the two propositions that :

1. There is no single statistical sampling technique which is appropriate in all circumstances.
2. The chances of making the proper selection of statistical sampling plan improve as a result of examination of the accounting system and its internal controls.

The first proposition

Turning to the first proposition, there are two reasons why any particular sampling plan may be inappropriate :

1) the sampling scheme can be used in circumstances where it may produce inaccurate or misleading results.

For example, the confidence level stated by the sampling plan may be x%, but the true confidence level is lower at y%. The two confidence levels are sometimes called the 'nominal' and 'real' confidence levels and the danger from the auditor's point of view arises whenever nominal confidence is above real confidence. Use of the sampling plan in such circumstances provides the auditor with a false sense of security.

2) The sampling plan can be used in circumstances where it produces results <u>inefficiently</u>.

For example, the confidence level stated by the sampling plan may be x%, but the true confidence level is higher at y%. This causes the auditor to select larger sample sizes than ought to be necessary to achieve acceptable confidence levels. Where nominal confidence is lower than real confidence, the sampling plan is said to be 'conservative'. Although undesirable this is clearly less dangerous than the 'misleading results' case.

If, for each available sampling plan in turn, it can be established that there exists a set of realistic circumstances in which the plan gives the auditor misleading results or the plan is unacceptably inefficient, then the first of the two propositions to which this chapter is addressed, is established. This chapter identifies such circumstances for the class of statistical sampling techniques known as monetary unit sampling.

<u>The second proposition</u>

Two reasons can be identified as to why any sampling plan may require the auditor to obtain prior information. They are -

1) information necessary to enable the plan to be operated without producing misleading or inefficient results.
2) information necessary to avoid the sampling scheme from being used in circumstances where it produces misleading or inefficient results.

In the first category positive use is made of the prior information to enable the sampling scheme to operate satisfactorily. In the second case negative use is made of the prior information which simply identifies the fact that the particular sampling plan cannot be used. The two categories could, therefore, be termed positive and negative prior information. An example of positive prior information is knowledge of the value of each population item when used as the basis of stratification for stratified variables sampling. An example of negative prior information is knowledge of the large variation of values of sample items when used to exclude the use of mean per unit sampling.

This chapter identifies and classifies the prior information required before the selection and operation of monetary unit sampling schemes, and it is shown that, to a significant degree, this prior information (whether positive or negative) comes from a study of the accounting system and its internal controls. If this is the case then it follows that the chances of making the proper selection of statistical technique improve as a result of examination of the accounting system.

Monetary Unit Sampling

A monetary unit sampling plan has been widely employed for many years by Deloitte Haskins and Sells (The H & S plan) and a slightly different version has been adopted by Clarkson Gordon in Canada and Arthur Young McClelland Moores in the United Kingdom. This latter plan is called dollar unit sampling (DUS). Mathematicians have been largely responsible for the development of both plans and this thesis makes no attempt to explain or justify their theoretical basis. There was virtually no discussion of monetary sampling in the auditing literature until publication of details of the H & S plan in a study by Meikle (36) and an important paper on DUS by Teitlebaum (37). Monetary unit sampling has now, however, found its way into established texts (38) and is widely used in practice, there being an estimated 100,000 annual applications in the United States.

In spite of its many advantages, there are a number of difficulties involved, any one of which could lead to misleading or inefficient results:

1) there is a problem in devising a suitable scheme for <u>errors of understatement</u>.

2) in certain circumstances monetary unit sampling gives <u>indeterminate results</u>.

3) monetary unit sampling has been criticised for not allowing the auditor to give more weight to <u>error prone areas</u>.

In order to establish the second of the two propositions, this chapter investigates the prior information necessary to avoid each of these three difficulties. In so far as these difficulties are unavoidable (even with the availability of prior information i.e. only negative prior information is available) then the first proposition is also established.

ERRORS OF UNDERSTATEMENT

Understatements and omissions are the achilles heel of monetary unit sampling, since although the theoretical limit for the overstatement of an item in the population is the size of the item, there is no theoretical limit to the possible understatement. Moreover, the greater the understatement the less the chance of the item being selected in a CMA sample and an item which is reported as zero (an omission) has no chance at all of selection. One school of thought therefore, is that MUS is completely unsuited to the problem of understatements and that its use should be limited to tests for overstatements. Teitlebaum does not regard this as a serious problem since he argues that the audit objective is to express an opinion on reported book values rather than the population of true values (39). DUS is certainly compatible with this audit objective, but it is an audit objective of doubtful validity. In this author's view an opinion on reported book values should be with reference to the population of true values. The audit objective is to discover overstatements and understatements, although the auditor's ability to detect overstatements is certainly greater than his ability to detect understatements, and perhaps this fact should be given formal recognition in the audit report.

The redeeming factor for MUS is that many audit tests for understatement can be re-organised into tests for overstatement by means of a 'reciprocal' population.

The 'double entry' matrix

		Primary Error	Corollary Error A	L	I	E
Assets	(A)	U	O	U	U	O
Liabilities	(L)	U	U	O	O	U
Income	(I)	U	U	O	O	U
Expenses	(E)	U	O	U	U	O

In many cases where the primary objective is to test for understatement, it is possible, as a result of the double entry system, to find a 'reciprocal' population to test for overstatement. This is shown by the 'double entry' matrix, although there are 4 possible situations where there is no reciprocal population. They are -

1. Assets (U), Liabilities (U)
2. Assets (U), Income (U)
3. Expenses (U), Liabilities (U)
4. Expenses (U), Income (U)

These four cases are extracted from the double entry matrix as being the combinations which involve understatements only.

Source Documents

Another approach to the problem of finding a reciprocal population is to test the accounting system's source documents for overstatement. An illustration is given by Stronge and Stewart:

> If the object is to test for understatement of the population, then there is little point in selecting from what is already recorded. The selection must be made from some 'reciprocal' population, the overstatement of which may imply the understatement of the population of primary interest. For example, one

way of testing for the understatement of sales is to test despatches for overstatement (a despatch may be regarded as overstated if the sale has not been billed). (40)

Similarly, an overstatement of goods received notes can imply an understatement of assets or expenses. A frequent problem with source documents, however, is that they are not expressed in money units and this makes CMA sampling inoperable.

Schemes for including understatements

In spite of the conceptual difficulties, monetary unit sampling plans do evaluate understatements without recourse to a reciprocal population. The following paragraphs briefly outline and discuss various ways in which understatements are evaluated.

With monetary unit sampling, the number of overstated pounds found in the sample may be used:

(1) to give an estimate of the number of overstated pounds in the population. This estimate is termed the most likely error (MLE) for overstatement.

(2) to give, with a stated level of confidence, a maximum figure for the potential overstatement in the population. This figure is termed the upper error limit (UEL) for overstatement.

(3) to give, with a stated level of confidence, a minimum figure for the potential overstatement in the population. This figure is termed the lower error limit (LEL) for overstatement.

Similarly, based on the understated pounds in the sample, it is possible to calculate the MLE, UEL and LEL for understatement.

In both DUS and the H & S plan, overstatements and understatements are separately evaluated to give not one sided precision limits but an appropriate bracket (for both UEL and LEL) around the most likely error. In DUS, to evaluate net overstatement, upper error limit for the overstatement (ignoring understatements) is reduced by the most likely understatement (ignoring overstatements). To evaluate net understatements, the upper error limit for understatement (ignoring overstatements) is reduced by the most likely overstatement (ignoring understatements). The H & S plan adopts a different procedure for overstatement; the upper error limit for overstatement (ignoring understatements) is reduced by the lower error limit for understatement (ignoring overstatements). A third possibility, of course, is to net errors of overstatement and understatement found in the sample and evaluate the net result, but this possibility has little support in the literature.

Support for the DUS procedure is found in a series of tests performed by Teitlebaum (41). The three alternative procedures described above were applied to samples drawn from a series of populations, each with a known frequency of understatements and overstatements which netted to give exactly a material error rate. The tests were repeated for :

a) Various sample selection techniques
b) Various 't' values where t (tainting) = $\frac{\text{error}}{\text{book value}}$ x 100, a concept which allows MUS to deal with partial errors.
c) Various gross error frequencies netting to exactly material error rates
d) Various confidence levels
e) Various sample sizes

Since the frequences of understatements and overstatements in the population are known, the risk of a specified sample evaluation procedure failing to detect the exactly material error rate can be calculated. This true risk is then compared with the stated risk of the evaluation procedure. The conclusion drawn from the tests was -

1. The H & S method is far too conservative. (i.e. true risks are much less than the stated risks)
2. The DUS method is appropriate but minor safeguards may be required. (in nearly all cases true risks were close to stated risks)
3. The method of netting error taintings and evaluating the net result is inappropriate at reasonable confidence levels (80 - 99%).

Hence, the H & S method, though its treatment of overstatement is not inaccurate, ~~it~~ is highly conservative, and since its treatment of understatement is not conservative, the implication is that the auditor is not as worried about understatements as he is about overstatements. Conservatism in financial reporting is no longer regarded by many people as desirable.

For instance, Hendriksen comments:

> Conservatism is, at best, a very poor method of treating the existence of uncertainty in valuation and income. At its worst, it results in a complete distortion of accounting data. (42)

On the basis of Teitlebaum's study, the DUS procedure is to be preferred to the other two possibilities considered.

The Error Limit Assumption

The difference between the upper error limit and the most likely error is termed 'precision'. Precision itself subdivides into basic precision (the precision which would be present even if there were no errors in the sample) and precision gap widening (the increase in precision caused by the discovery of errors in the sample). If there are no errors, precision gap widening is zero and precision equals basic precision. Precision gap widening relates to the errors found in the sample and its size is determined in part, by the size of the errors discovered. On the other hand, basic precision takes into account the fact that there may be errors in the population of a type (i.e. size) not discovered in the sample. Whether the evaluation concerns overstatement or understatement errors, some kind of assumption has to be made about the possible size of errors not discovered. Since the poisson factor for basic precision is normally given a 100% tainting, the underlying assumption concerning errors not discovered in the sample is that they average 100%. For errors of overstatement, which in general have a maximum tainting of 100%, this is a conservative assumption. For errors of understatement where there is no such theoretical maximum, the assumption could be false and lead to misleading results. In order to form a view as to the validity of this, or any alternative assumption concerning understatements, the auditor requires prior information (the sample errors found are no help since the assumption relates to errors not discovered in the sample). However, it is difficult to be definite about what kind of prior information is most relevant. Certainly knowledge of the accounting system is likely to be of some assistance in this regard.

Understatements and Prior Information : A Summary

The discussion concerning the problem of evaluating understatements has revealed the following:

(1) the possibility of monetary unit sampling producing misleading results where the 100% error limit assumption for understatements does not hold.

(2) the possibility, based on Teitlebaum's results, that the H & S plan is inefficient in its treatment of overstatements.

(3) the impossibility of structuring all audit tests so that they are tests for overstatement only. This is a result of there being no reciprocal population given by the double entry matrix in four situations.

    Where source documents are used there is the problem of non-monetary source documents.

The first of these requires knowledge of the possible size of understatements which exist, and in this connection knowledge of the accounting system may be relevant. The second does not concern prior knowledge since the suggestion is that DUS is preferable to the H & S plan in all circumstances. The third may involve prior knowledge of the system where source documents are used. What is clear, however, is that the availability of prior information makes relatively little impact on the problem of understatements and this problem remains a significant limiting factor on the usefulness of all monetary unit sampling plans.

INDETERMINATE RESULTS

In general, a sampling plan can be examined by reference to two risks :

α risk - the risk that the sampling plan causes materially correct financial statements to be rejected.

β risk - the risk that the sampling plan causes materially incorrect statements to be accepted.

The use of the alpha and beta risk concepts in an auditing context was first made by Elliott and Rogers (43).

The auditor should choose a sampling plan which controls not only the beta but also the alpha risk, and it follows that each sampling plan must incorporate 2 decision rules (one for acceptance and the other for rejection). If the sample result does not satisfy conditions for either acceptance or rejection, then the sample result is indeterminate and of little or no use to the auditor. To simplify the discussion it is assumed that only errors of overstatement are possible and ignores the problem of 'netting' overstatements and understatement, discussed in the preceding section.

Acceptance

It is easy to give a decision rule for acceptance of a reported population value: if the upper error limit is less than a material amount the reported value can be accepted. The risk of accepting a materially incorrect population (β risk) cannot be greater than 1 - specified confidence level. It should be possible to persuade clients to alter the original reported value by the amount of errors discovered in the sample, and if this is done the acceptance rule can be modified to include cases where the upper error limit minus discovered errors is less than a material amount:

<u>accept</u> when $UEL - E < M$

where UEL = upper error limit

E = actual discovered errors

M = material amount

<u>Rejection</u>

A decision rule for rejection can be formulated using the lower precision limit for the poisson distribution rather than the upper precision limit. The rejection rule is:

<u>reject</u> when $LEL > M$

where LEL = lower error limit

M = materiality

The risk of rejecting a materially correct population (α risk) is 1 - specified confidence level. In practice, rejection may mean a request to the client to 100% recheck the population, followed by a fresh sample of the revised population. If the client refuses to recheck, the auditor can consider a qualification for the audit report.

<u>Avoiding the Indeterminate Result</u>

Kaplan (44) and Smith (45) have both suggested the same procedure for selecting a sample size which will not give an indeterminate result. It is necessary to distinguish the amount considered material for purposes of accepting a materially incorrect population (M_A) from that considered material for purposes of rejecting a materially correct population (M_R).

β risk = risk of accepting a materially incorrect population where true error $> M_A$

α risk = risk of rejecting a materially correct population where true error $< M_R$

and $M_A > M_R$

The greater the divergence between M_A and M_R, the easier it is for a sample to distinguish between a materially correct and a materially incorrect population. However, M_A and M_R should be pre-set by the auditor independently of this consideration. The choice of M_A and M_R give an upper (Pa) and lower (Pr) limit for the permissible error rate and it is the size of Pa and Pr along with the size of the alpha and beta risk, also specified by the auditor, which determines sample size. As an example of the method consider the following:

Acceptable α risk 10%

Acceptable β risk 5%

Reported population 1m pounds

M_A = £20,000 (Pa = 0.02)

M_R = £ 5,000 (Pr = 0.005)

Table 3 gives the poisson confidence factors (C) for a beta risk of 5%

<u>Table 3</u>

K	C
0	3.00
1	4.74
2	6.30
3	7.75
4	9.15
5	10.5

$\underline{\beta = 0.05}$

Beta risk is controlled by selecting sample size n such that $n\ p_a \geqslant C_k$ and if k or fewer errors are discovered, the population is accepted. Table 4 gives the poisson confidence factors (C) for an alpha risk of 10%.

Table 4

K	C
0	0.105
1	0.532
2	1.10
3	1.75
4	2.43
5	3.15

$\alpha = 0.10$

Alpha risk is controlled by selecting sample size n such that $n\ p_r \leqslant Ck$ and if more than k errors are discovered the population is rejected. It is necessary to choose values for n and k which satisfactorily control both the beta and alpha risks. This can be done by trial and error as follows:

Let $k = 0$

For beta risk	n	0.02	$\geqslant$	3
	n		$\geqslant$	150
For alpha risk	n	0.005	$\leqslant$	0.105
	n		$\leqslant$	21

These conditions are inconsistent.

Let $k = 3$

For beta risk	n	0.02	$\geqslant$	7.75
	n		$\geqslant$	387
For alpha risk	n	0.005	$\leqslant$	1.75
	n		$\leqslant$	350

Again, this is inconsistent.

Let k = 4

For beta risk	$n \quad 0.02$	$\geqslant$	9.15
	n	$\geqslant$	457
For alpha risk	$n \quad 0.005$	$\leqslant$	2.43
	n	$\leqslant$	486

In order to control both alpha and beta risk at their required levels, the auditor must choose a sample size of 457 and accept if he discovers 4 or fewer errors, or reject if he finds more than 4 errors.

Two Stage Approach

It is clear from this example that the use of monetary unit sampling to control both the beta and alpha risks may require sample sizes which are in excess of those commonly used in auditing. In practice the auditor adopts a 2 stage approach:

1. An initial sample is selected to control the beta risk only.
2. In the event of an indeterminate result further audit work is performed on the basis of a study of the errors obtained in the sample.

An initial sample

Sample size is chosen on the basis of expected error rates and the required beta risk. (In the example, where $\beta = 5\%$, a sample size of 150 is required for an expected zero error rate, whereas a sample size of 387 is required for an expected error rate of 0.0077 (3/387)). If the error rate discovered in the sample does not exceed expectations, the auditor can accept the population as the beta risk requirement is satisfied. However, if the error rate in the sample exceeds expectations, the auditor has an indeterminate result since the alpha risk is not controlled and the population cannot, therefore, be rejected. An alternative to varying the initial sample size on the basis of the expected error rate is to select a discovery sample size, whatever the expected error rate. (This is the basis of the H & S plan though not DUS). The auditor can then only accept populations where no errors are discovered in the sample, and the effect is to increase the incidence of indeterminate results, although there is, of course, the benefit of reduced samples.

Further audit work

Further audit work is required whenever indeterminate results are obtained. This work may be a general increase in sample size, but is more likely to be an increase in sample size in the particular area or areas suggested by a study of the errors. The auditor's study will include consideration of whether:

(a) the error only recurs in a particular system.

(b) the error only recurs when the system is operated by a particular individual, or within a particular department or division.

(c) the error only recurs when a particular type of transaction, or input is processed.

(d) the error only recurs during a restricted time period.

Instead of trying to project the errors in the sample throughout the population, the auditor attempts to place boundaries on the possible prevalence of the error and then evaluates the error within those boundaries. Evaluation of the error within boundaries may require a further statistical sampling plan. The case for the two stage approach has been well made by Moriarity (46):

> The auditor's objective in this random sample is to reduce to an acceptable level the probability of the existence of a material undiscovered type of error. If an unanticipated type of error is discovered, the auditor will not simply keep a tally of it, but rather will attempt to determine its cause and how it got through the system undetected. The auditor will determine if it is possible for a sufficient number of this type of error to exist and to total a material discrepancy. If it is possible for this type of error to have a material effect, the auditor will then design a specific sampling plan to search for more instances of this type of error. The auditor does not simply increase the number of random samples to be taken.
>
> If the preceding is an accurate description of the auditor's reaction to the discovery of an error, it is not necessary for auditors to take sufficiently large random samples so that an estimate of the effect of all errors can be derived from the sample. Instead, the purpose of the sample should be to assure that at least one instance of every material type of error in the population occurs in the sample so that the auditor's attention will be directed to the problem area. This, of course, is discovery sampling.
>
> These last comments have argued that statistical estimation should not be the objective when determining the extent of tests of details. However, the argument has not been made that estimation is never useful. Once the auditor has determined that a particular type of error exists in the client's population, he may well wish to design a sample of specific transactions to estimate the total effect of that type of error on the client's financial statements.

Adjustments

Where an indeterminate result is obtained it may, in a few cases, be possible for the client and auditor to avoid further audit work by negotiating an adjustment to the population to a figure acceptable to both parties on the basis of the initial sample results. Of course since the alpha risk is not measured, the adjustment may be taking place to a materially correct population. This is not a problem provided the population figure after adjustment is also materially correct. For an adjustment to be made, two conditions must be satisfied:

(1) The client must be willing to adjust.
Relatively few clients are willing to adjust to the MLE shown by the auditor's sample and from the client's point of view this is quite sensible since the accounts are his responsibility and he ought to have more confidence in the accounting system which has produced the figures than in the auditor's sample results. The instances where the client is willing to adjust are likely to be those where MLE (rather than the UEL) is itself higher than materiality. Where the client is unwilling to adjust, the auditor has no grounds on which to reject the population and should perform further audit work.

(2) The auditor must be willing to adjust.
Where the auditor finds a relatively large number of errors in the sample he is likely to be happier with an adjustment of the MLE (the best estimate of error in the population) than with an adjustment of UEL less Materiality. The latter adjustment would be quite different to the best estimate of error. However, even where the adjustment is based on the MLE the auditor may wish to be cautious. Following an adjustment of the MLE, the accounting figure is generated by the sample rather than by the accounting

system, and high confidence in the sample results is a necessity. In addition, the auditor is much more likely to accept the adjustment of MLE when precision given by the sample results is less than materiality, and he is, therefore, more satisfied that the adjustment is to a materially correct figure. It should be remembered that "most likely" does not say "how likely" and the most likely error rate, although the best available guess, may not meet such precision requirements (i.e. precision less than materiality). The identification of circumstances where both client and auditor are likely to be satisfied with an adjustment of MLE is made in Appendix 1.

For these reasons, adjustment is seen by this author as seldom being a practical alternative to further audit work at the stage of having completed the initial sample. It is much more likely to prove a possibility after further audit work has been completed and the auditor has more thorough evidence on the basis of which he can press for an adjustment.

Diagrammatic Summary of MUS decision rules

A diagrammatic summary of the decision rules discussed in this section is given in figure 3 . The decisions are as follows:

(1) If the auditor adopts the two stage approach, he starts with an initial sample and accepts the population if the decision rule for acceptance ($UEL - E < M_A$) is satisfied. Since only the beta risk is controlled, any failure to satisfy this condition causes an indeterminate result.

(2) Where an indeterminate result is obtained, the auditor is happy that the MLE gives a materially correct estimate, and the client is willing to adjust on that basis, the auditor should accept the adjusted figure.

(3) If the client is unwilling to adjust, the auditor studies the errors in the sample and decides whether further audit work should be general, or limited to specific areas.

(4) Having performed further audit work, the auditor assesses the results and decides whether he is now in a position to accept or reject the population.

(5) If the auditor rejects the population he may be satisfied with an adjusted figure, provided he has high confidence that the adjusted figure is materially correct.

(6) If a suitable adjustment cannot be negotiated and the client refuses to recheck the population himself, the auditor may be left with no alternative to a qualification in his audit report.

Meikle (47) in his comments on post evaluation states that the auditor faced with unacceptable precision has four choices:

(a) Request the client to check the population 100%, and correct all errors located; then resample (RECHECK)

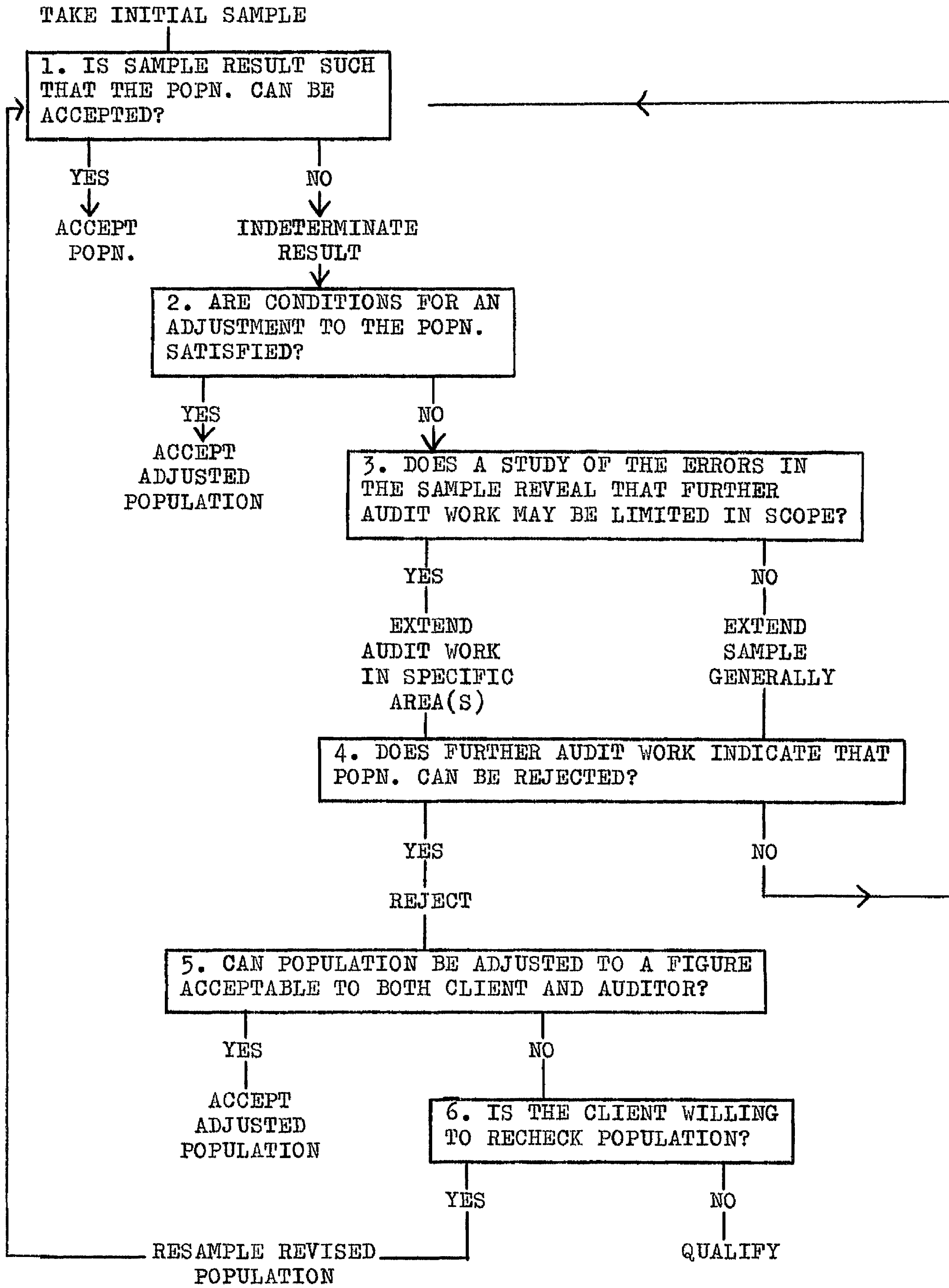
(b) Request an adjustment to the financial statements in an amount sufficient to reduce the adjusted precision to an acceptable level (ADJUSTMENT)

(c) Perform additional work as appropriate in the circumstances (FURTHER WORK)

(d) Consider the appropriateness of an unqualified opinion (QUALIFY)

DIAGRAMMATIC SUMMARY OF MUS DECISION RULES:

FIGURE 3



Meikle does not, however, suggest a logical sequence to these options, which are seen in figure 3 as being:

ADJUST (on basis of initial sample) → FURTHER WORK → ADJUST (on basis of further work) → RECHECK → QUALIFY.

The high error rate problem

It can be seen from the discussion in this section that there are two serious problems in using monetary unit sampling for populations with high error rates:

(1) If the auditor wishes to reject the population he needs to control the alpha risk, and this can only be done with large samples.

(2) If the auditor wishes to seek an adjustment on the basis of the sample results he will, in many cases, have difficulty in determining an adjustment which is acceptable to both the client and himself. (See appendix 1.)

For these reasons monetary unit sampling is of only limited assistance in sampling of high error rate populations and in such circumstances the auditor may prefer to use a variables technique. Variables techniques are discussed in the next chapter. Although they are often inappropriate for low error rate populations they frequently out-perform monetary unit sampling for high error rate populations. On the basis of his study of the errors found in an initial monetary unit sample, the auditor may, where the number of errors discovered is relatively large, wish to :

(1) perform further audit work in a specific area by conducting a variables sample in that area

or

(2) perform further audit work generally but to transfer to a variables sampling plan, rather than extend the initial monetary unit sample.

In the second case it is clear that the auditor would have been advised to use a variables sampling plan from the outset had he examined the systems of internal accounting control and the examination indicate the likelihood of high error rates. A study of the system can be useful in determining the circumstances in which a monetary unit or variables sampling plan is the appropriate technique.

Even in the first case a study of the systems might be preferred to an initial monetary sample as the preliminary survey which directs the auditor's attention to the high error risk area.

Indeterminate results and the need for prior information

The discussion concerning the problem of indeterminate results has revealed the following:

(1) Because of the difficulties involved in either rejecting a population or making an adjustment to a population on the basis of the sample result, monetary unit sampling is generally unsuited to high error rate populations. A prior study of the accounting system is valuable in so far as it enables the auditor to determine in advance the likelihood of high error rates and hence avoid the inefficiency of having to switch midstream from an MUS to a variables plan.

(2) Because of the large sample sizes required to control the alpha risk, the efficient use of monetary unit sampling generally involves an initial sample (which controls only the beta risk) followed by appropriate further audit work

where necessary. This further audit work includes, among other matters, consideration of whether:

a) the error only recurs in a particular system

b) the error only recurs when the system is operated by a particular individual

c) the error only recurs when a particular type of transaction is processed.

The three considerations require prior knowledge of system boundaries (case a), the personnel operating the system (case b) and the types of data processed (case c).

ERROR PRONE AREAS

Monetary unit sampling has been criticised for not allowing the auditor to give more weight to those areas where the risk of error is highest. For example, Smith comments that "the advocates of CMA sampling proclaim its advantages too loudly"(48) since it "does not make explicit use of any knowledge about key areas that might be error prone" (49) and Goodfellow, Loebbecke and Neter (50) argue in their criticism of monetary unit sampling that since large audit units have a greater chance of selection than small units, MUS cannot be the most effective sample design if most of the errors are in the small items. Monetary unit sampling can be shown to be equivalent to upper boundary priced physical unit sampling with an infinite degree of stratification and, therefore, automatically stratifies the population by value, but <u>not</u> by error risks (51).

The discussion which follows investigates this criticism of monetary unit sampling and finds the criticism valid in the sense that the efficiency of MUS may be improved if the error

prone section of the population is separately identified and given separate consideration in the sample. In this discussion the process of splitting the population into two sections according to risk characteristics is termed "segmentation", and the ratio of the population in one segment to the population in the other is termed "the segmentation ratio". The term is also used in this way by Kinney (52).
If the population can be divided into two sections or segments (one error prone and the other not) there are two ways in which this ~~this~~ could be given consideration in the monetary sample:

(a) total sample size is allocated between segments proportionately according to the number of population units in each, and separate random samples are selected. (Strategy 2)

(b) total sample size is allocated between segments according to their risk characteristics. A proportionately larger sample is allocated to the higher risk segment and separate random samples are selected. (Strategy 3)

Strategy 2 may be thought of as a "passive" and strategy 3 as an "active" strategy. To demonstrate the validity of the criticism of MUS, both the passive and active strategies are compared with a conventional monetary unit sample of the same size. Neither DUS nor the H & S plan give special consideration to error prone sections of the population: the sample is one unrestricted sample from the entire population and this may be thought of as a "conventional" strategy (strategy 1). It is assumed that whichever strategy is employed, the sample is to serve as an initial discovery sample as outlined in the previous section on "indeterminate results". A comparison between the efficiency of the strategies (conventional, active and passive) can therefore be made on the basis of the probability of each strategy failing to detect a single error, given that there is material error in the population. For each strategy the demonstration assumes a sample size of 100 and a population with the characteristics given in figure 4 . For a more rigorous demonstration the process of comparison can be repeated using different sample sizes and population characteristics. The effect of varying the segmentation ratio is studied later in this section but the effect of varying any of the other characteristics is not reported. The study of segmentation is simply to provide a demonstration of how it can improve the efficiency of MUS. A fuller examination of the issue is given elsewhere by this author (53).

Figure 4 : Population characteristics

1.	Population size	10^5
2.	Population error rate	0.02
3.	Segmentation of the population into two segments is possible –	
	Segment 1 = 0.1×10^5	
	Segment 2 = 0.9×10^5	

The actual distribution of the error of £2,000 in the population is varied between three positions –

1. Spread: in this case the error is spread evenly across the population and each pound, no matter where it is located,

has approximately an equal chance of being in error.

2. Load in segment 1: the error is loaded into segment 1 and all pounds outwith segment 1 are error free.

3. Load in segment 2: the error is loaded into segment 2 and all pounds outwith segment 2 are error free.

No matter how the error is actually distributed, the auditor's prior assessment is assumed to be that segment 1 is error prone. One possible version of the active strategy is, therefore, to allocate 20 sample pounds to segment 1 and 80 to segment 2. This is the version at first used in the demonstration, although it is later varied.

Results: a comparison of the strategies

The resultant probabilities of failure to detect a single error, for the given audit strategy/error distributions are as per table 5 .

ERROR DISTRIBUTION	STRATEGY 1	STRATEGY 2	STRATEGY 3
SPREAD	0.1326	0.1326	0.1326
LOAD IN SEGMENT 1	0.1326	0.1073	0.0115
LOAD IN SEGMENT 2	0.1326	0.1323	0.1656

TABLE 5

NOTE STRATEGY 1 = CONVENTIONAL
STRATEGY 2 = PASSIVE
STRATEGY 3 = ACTIVE

The workings to support the figures in these and other tables provided in this section are given in the second appendix to this chapter. The results demonstrate that, given spread, there is nothing to choose between the three strategies. Given load in segment 1, judged to be the "high risk" segment, strategy 1 is inferior to strategy 2, (0.1326 to 0.1073) which in turn is markedly inferior to strategy 3 (0.1073 to 0.0115). If, however, the prior assessment of segment 1 as being the high risk segment is completely wrong, and the load is, in fact, in segment 2, then strategy 1 is only marginally inferior to strategy 2 (0.1326 to 0.1323), which is in turn superior to strategy 3 (0.1323 to 0.1656). It follows that strategy 1 is either equivalent to, inferior to or marginally inferior to strategy 2, and can be discarded in favour of strategy 2 in all circumstances, assuming the costs of segmentation are not a material factor. A choice between strategies 2 and 3 depends upon the confidence the auditor can attach to his assessment of segment 1 as the high risk area. If he is confident of his assessment strategy 3 is likely to be preferred.

A study of the passive strategy (strategy 2)

In the illustration the population has been split into segments in the ratio 0.1 : 0.9. This section studies the effect on strategy 2 of a variation in this segmentation ratio. Table 6 provides the probability of failing to detect a single error, given strategy 2 (the sample is allocated proportionately to segment size) for segmentation ratios 0.3 : 0.7 and 0.5 : 0.5 in addition to the ratio 0.1 : 0.9.

As segment sizes approach equality (0.5 : 0.5) strategy 2 becomes less effective in circumstances where the error is loaded in segment 1 (0.1073 to 0.1262 to 0.1298), but marginally more effective should the error be loaded in segment 2 (0.1323 to 0.1314 to 0.1298). In spite of this compensating effect it is likely that strategy 2 will be regarded as becoming less effective as the segment sizes equalise. It is however, in all cases preferable to strategy 1.

SEGMENTATION RATIO	STRATEGY 1	STRATEGY 2	
		LOAD SEGMENT 1	LOAD SEGMENT 2
0.1 : 0.9	0.1326	0.1073	0.1323
0.3 : 0.7	0.1326	0.1262	0.1314
0.5 : 0.5	0.1326	0.1298	0.1298

TABLE 6

NOTE STRATEGY 1 = CONVENTIONAL
STRATEGY 2 = PASSIVE

A study of the active strategy (strategy 3)

The version of strategy 3 used in the illustration was to allocate 20 items to segment 1, the high risk area, and 80 items to segment 2. This section studies the effect on strategy 3 of varying this disproportionate allocation. Table 7 provides the probability of failing to detect a single error, assuming an allocation of 16, 32 and 64 items to segment 1 as alternatives to the allocation of 20 items. The segmentation ratio is 0.1 : 0.9 throughout.

ALLOCATION TO SEGMENT 1	STRATEGY 3	
	LOAD SEGMENT 1	LOAD SEGMENT 2
STRATEGY 2 (10)	0.1073	0.1323
16	0.0281	0.1514
20	0.0115	0.1656
32	0.0008	0.2169
64	APPROACHING ZERO	0.4453

TABLE 7

NOTE STRATEGY 3 = ACTIVE
STRATEGY 2 = PASSIVE

The results show that a comparatively small bias toward segment 1 (eg an allocation of 16 rather than 10) causes a very marked improvement (0.1073 to 0.0281), provided segment 1 is error loaded. A large bias toward segment 1 (eg an allocation of 32) causes a marked deterioration (0.1323 to 0.2169) in the effectiveness of strategy 3 should segment 2 be error loaded. It is likely, therefore, that a relatively small bias toward the anticipated high risk area is to be preferred to either no bias (strategy 2) or to a large bias.

Error prone areas and the need for prior information

The calculated results demonstrate that segmentation can increase the efficiency of MUS and that its usefulness can be improved by giving the allocation of the sample a relatively small bias toward the anticipated high risk segment . In order to be able to adopt the passive strategy the auditor must be able to segment the population according to risk characteristics, but he does not need to assess the risks involved, other than to realise that they are different. A major cause of risk differential is that items passing through different internal control systems have come together in one account, or are treated as one population for monetary unit sampling purposes. Monetary unit sampling is increasingly used to sample "global" populations. In this case several different accounts' figures are added together and treated as a single population of pounds for monetary unit sampling purposes (eg the whole of the assets side of the balance sheet may be regarded as one population). This practice increases the importance of segmentation. In such circumstances, segmentation is likely to follow systems' boundaries and prior knowledge of these boundaries is essential. It may also require knowledge of the different personnel operating the system and the different transactions processed by the system since these factors can also cause risk differential. It should be noted that when the auditor does not segment his initial sample and does not achieve an acceptable sample result, he effectively "post-segments" his population through a study of the errors discovered in the sample. (Figure 2 - decision 3). The argument here is in favour of "pre-segmentation" of the initial sample, which may later require revision as a result of a study of the errors found.

In order to be able to adopt the <u>active strategy</u> the auditor must attempt to assess the risks of error in the different segments. The more confident he is of his assessment, the more he may be prepared to bias the sample toward the high risk segments. Clearly such an assessment requires a thorough examination of the system of internal accounting control.

MONETARY UNIT SAMPLING AND THE SECOND PROPOSITION OF THE CHAPTER

The chapter considers three major problems in the application of monetary unit sampling. They are -

1. treatment of understatements
2. frequency of indeterminate results
3. lack of consideration of error prone areas.

For each heading, the chapter investigates the problem and identifies prior information which may be helpful. The results of the investigation are summarised in table 8 . It is clear from this table that the auditor who, for monetary unit sampling, wishes to structure his audit tests (1A), who wishes to avoid misleading results (3A) and who wishes to employ the technique efficiently (2A, 2B, 3B) needs prior information of the accounting system, ranging from mere knowledge of the documents and document flow (1A) through to a full assessment of the system (3A, 3B) including an evaluation of the risk of error.

The second proposition of the chapter states :

The chances of making the proper selection of statistical sampling plan improve as a result of examination of the accounting system and its internal controls.

The following reasons have been identified as to why it is desirable to carry out an examination of the system prior to performance of monetary unit sampling:

1) to recognise appropriate "reciprocal" populations.

2) as a basis for an assumption about the maximum size of understatement errors not found in the sample.

3) to avoid using monetary unit sampling when error rates are high.

4) to enable "post segmentation" when errors are discovered in a monetary unit sample.

5) to enable the auditor to "pre-segment" the population.

TABLE 8

PRIOR KNOWLEDGE OF THE SYSTEM	PROBLEM FOR WHICH REQUIRED
1. Documents and document flow.	A. Understatements - the identification of source documents suitable as a "reciprocal" population. Helps to determine the structure of the audit test.
2. Knowledge of systems boundaries, operating personnel, transactions processed.	A. Indeterminate results - essential for a proper study of errors found in the sample when the two stage approach is employed. The two stage approach reduces the inefficiency of MUS caused by the lack of control of the alpha risk.
	B. Error prone areas - essential to the adoption of the passive strategy where segmentation is employed. Segmentation may reduce the inefficiency of MUS caused by failure to consider error prone areas.
3. A full assessment of the system(s) including evaluation of the risk of error.	A. Understatements - important for a consideration of the error limit assumption applied to understatements. If the assumption is invalid MUS may give misleading results.

3. Continued ...

B. Error prone areas - essential to the adoption of the active strategy where segmentation is employed. The active strategy can lead to greater efficiency than the passive strategy provided the auditor is confident of his prior risk assessment.

C. Indeterminate Results - high error rates will lead to indeterminate results and inefficiency.

MONETARY UNIT SAMPLING AND THE FIRST PROPOSITION OF THE CHAPTER

Understatements

As far as the treatment of understatements is concerned, the prior information available makes little impact in circumstances where it is not possible to structure the audit tests so that they are tests for overstatement only. In these circumstances the auditor is forced to make an error limit assumption for understatement and a study of the system generally gives inadequate guidance as to the possible size, as opposed to frequency, of errors.

Indeterminate results

Monetary unit sampling is shown to be generally inappropriate for populations with high error rates, for the following reasons:

1) If the auditor wishes to reject the population, he needs to control the alpha risk and this can only be done with large samples.

2) If the auditor wishes to seek an adjustment on the basis of the sample results he will, in many cases, have difficulty in determining an adjustment which is acceptable to both the client and himself.

Although prior information in the form of systems evaluation is valuable in detecting possible high error rate populations, such prior information is essentially negative in character and monetary unit sampling's inappropriateness for such populations remains a significant limitation on its usefulness.

Error prone areas

It is shown that monetary unit sampling is more efficient when it takes account of error prone areas through segmentation of the population. Prior knowledge is required to enable efficient segmentation but such prior knowledge is essentially positive in character. It is not considered that a failure to consider error prone areas represents an inevitable limitation on the usefulness of monetary unit sampling plans which will ultimately be developed to incorporate prior knowledge of error prone areas in the evaluation.

The first proposition of the chapter is stated as follows:

There is no single choice of statistical sampling plan which is appropriate in all circumstances.

This chapter has established that there are circumstances in which monetary unit sampling plans, at least, are inappropriate.

APPENDIX 1

Negotiation of an adjustment on the basis of the MLE

The following table is taken from the table of cumulative precision factors employed by Arthur Young McClelland Moores & Co and is based on the poisson distribution.

For 95% Confidence Level

ER	MLE	BP	PGW	UEL
0	0	3	N/A	3
1	1	3	0.75	4.75
2	2	3	1.30	6.30
3	3	3	1.76	7.76
4	4	3	2.16	9.16
5	5	3	2.52	10.52
6	6	3	2.85	11.85
7	7	3	3.15	13.15
8	8	3	3.44	14.44
9	9	3	3.71	15.71
10	10	3	3.97	16.97

ER = error rate in sample

MLE = most likely error in population

BP = basic precision

PGW = precision gap widening

UEL = upper error limit (MLE + BP + PGW)

1. It was suggested in the main body of the chapter that the client is likely to be willing to adjust whenever MLE exceeds materiality. ie when MLE $>$ M

2. It was also suggested that the auditor is likely to allow an adjustment of MLE whenever precision is less than materiality. ie when P $<$ M (P = BP + PGW)

3. Both these conditions are satisfied when

$$P < M < MLE$$

4. The conditions can be studied by slightly rewriting the table (at 95% confidence).

ER	P	MLE	MLE - P
0	3	0	-3
1	3.75	1	-2.75
2	4.30	2	-2.30
3	4.76	3	-1.76
4	5.16	4	-1.16
5	5.52	5	-0.52
6	5.85	6	+0.15
7	6.15	7	+0.85
8	6.44	8	+1.56
9	6.71	9	+2.29
10	6.97	10	+3.03

5. The conditions are impossible to satisfy at error rates of 5 or lower since P $>$ MLE

6. The conditions may possibly be satisfied when ER = 6, although M must lie in a very narrow range between 5.85 and 6 (0.15).

7. The range increases as ER increases and satisfaction of the conditions becomes rather more likely, although the conditions remain highly restrictive.

8. It is concluded that a negotiated adjustment based on the MLE is only possible in a limited number of cases.

9. Even these limited number of cases ignore possible constraints imposed by the LEL.

APPENDIX 2: WORKINGS TO SUPPORT SEGMENTATION STATISTICS

STRATEGY 1

probability of the first sample item being correct

$= 0.98$

probability of first and second sample items being correct

$= 0.98^2$

probability of 100 sample items all being correct

$= (0.98)^{100} = 0.1326$

STRATEGY 2

In general,

N	=	population size
Kn	=	size of segment into which error loaded
n	=	sample size
Kn	=	segment sample size
p	=	population error rate

then error rate in the segment is $\frac{pN}{KN} = \frac{p}{K}$

and probability of failing to detect a single error in a segment sample size of Kn $= (1-\frac{p}{K})^{Kn}$

$K = 0.1, \; p = 0.02, \; n = 100$

$$\left(1 - \frac{0.02}{0.1}\right)^{10} = 0.1073$$

$K = 0.3$

$$\left(1 - \frac{0.02}{0.3}\right)^{30} = 0.1262$$

$K = 0.5$

$$\left(1 - \frac{0.02}{0.5}\right)^{50} = 0.1298$$

$K = 0.7$

$$\left(1 - \frac{0.02}{0.7}\right)^{70} = 0.1314$$

$K = 0.9$

$$\left(1 - \frac{0.02}{0.9}\right)^{90} = 0.1323$$

$K = 1.0$

$$\left(1 - \frac{0.02}{1}\right)^{100} = 0.1326$$

STRATEGY 3

In general,

N = population size

KN = size of segment into which error loaded

n = sample size

s = segment sample size

p = population error rate

then error rate in the segment is $\frac{p}{K}$

and probability of failing to detect a single error in a segment sample size of s = $(1-\frac{p}{K})^s$

$s = 16,\ K = 0.1,\ p = 0.02$

$$(1 - \frac{0.02}{0.1})^{16} = 0.0281$$

$s = 20,\ K = 0.1$

$$(1 - \frac{0.02}{0.1})^{20} = 0.0115$$

$s = 32,\ K = 0.1$

$$(1 - \frac{0.02}{0.1})^{32} = 0.0007922$$

$s = 64,\ K = 0.1$

$$(1 - \frac{0.02}{0.1})^{64} = 0.0000006$$

$s = 84,\ K = 0.9,\ p = 0.02$

$$(1 - \frac{0.02}{0.09})^{84} = 0.1514$$

$s = 80,\ K = 0.9$

$$(1 - \frac{0.02}{0.9})^{80} = 0.1656$$

$s = 68,\ K = 0.9$

$$(1 - \frac{0.02}{0.9})^{68} = 0.2169$$

$s = 36,\ K = 0.9$

$$(1 - \frac{0.02}{0.9})^{36} = 0.4453$$

CHAPTER 4

VARIABLES SAMPLING AND PRIOR INFORMATION

The previous chapter studied the following two propositions in relation to monetary unit sampling:

1. There is no single choice of statistical sampling technique which is appropriate in all circumstances

and

2. The chances of making the proper selection of statistical sampling technique improve as a result of examination of the accounting system and its internal controls.

This chapter extends the study of these two propositions to the use of variables sampling techniques.

The possibility of the accountant or auditor using variables sampling became widely recognised in the early 1960's following publications by Arkin, and by Cyert and Davidson (54). The statistical formulae of the various variables techniques can be found in either of these texts. The techniques differ from monetary unit sampling plans in that they are based on normal distribution theory rather than the poisson distribution. However, Goodfellow, Loebbecke and Neter (55) have referred to monetary unit sampling as combined attributes and variables sampling (CAV) in spite of the fact that MUS has no connection with the normal distribution. It is also the case that monetary unit sampling has been developed much more recently, and in certain respects is specifically designed to cater for the auditor's requirements, whereas variables sampling techniques have been widely used for estimation purposes for many years in all kinds of disciplines.

This chapter reviews several well known articles which examine or question the effectiveness of variables sampling plans and, on the basis of this review, points to the limitations of such plans and to the prior information which may be useful. The articles fall into three groups -

1. those concerned with the beta risk implicit in variables sampling plans.

   Elliott and Rogers (56) demonstrate that decision rules frequently employed by auditors on a rule of thumb basis actually have alarmingly high beta risks. They argue that the auditor should explicitly control the beta risk of the variables sampling plan and suggest a decision strategy to achieve their objective. Clearly, 'rule of thumb' decision rules may cause the auditor to take incorrect decisions. Teitlebaum and Robinson (57) are concerned that the difference between real and nominal beta risk for variables sampling plans may vary according to both the distribution of error through the population and the total amount of error in the population. Again, this may cause the auditor to take incorrect decisions.

2. those concerned with empirical testing of variables sampling plans against known accounting populations.

   By testing a sampling plan repeatedly against known accounting populations, the true (or real) risk of applying the technique (as approximated by the empirical results) can be compared with the stated (or nominal) risk given by the theory. Thus the researcher can examine the likelihood of inaccurate or misleading results (i.e. to what extent real risk $>$ nominal risk) and inefficiency (real risk $<$ nominal risk). Inefficiency can also be examined by studying the size of the standard error of the estimator provided by the sample plan. Kaplan (58) has examined variables techniques by reference to populations generated by known stochastic processes whereas Neter and Loebbecke (59) use populations derived from actual accounting populations.

3. those concerned with <u>sampling objectives</u>.

Ijiri and Kaplan (60) have argued that the observable behaviour of auditors, when for instance selecting a judgement sample, is not to seek solely a representative sample but also to seek a sample which includes errors and high value items. They criticise sampling plans employed by auditors for concentrating on 'representation' at the expense of error prone and high value items. Insofar as Ijiri and Kaplan's criticism concerns a failure to bias sample selection toward error prone items, then it is also relevant as a criticism of monetary unit sampling. Kinney (61), however, points out that sophisticated sampling plans do bias sample selection toward error prone and high value items in a way consistent with the auditor's overall risk and materiality requirements. Kinney demonstrates his point by using a <u>variables</u> sampling plan with the population stratified in two dimensions (size and likely error rate). For this reason the discussion of sampling objectives is included in this (variables sampling) chapter. Kinney could equally well demonstrate his point by using a monetary sampling plan with the population segmented according to likely error rate.

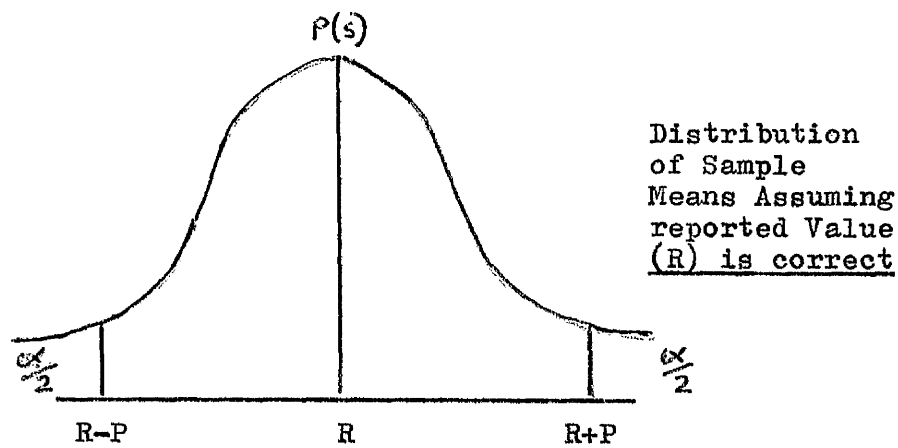
ELLIOTT AND ROGERS

A typical 'rule of thumb' decision rule employed by auditors using variables techniques might be:

1. Let precision (p) equal materiality (M).
2. Let the confidence level (CL) vary according to the assessment of internal control (e.g. 95% for poor and 80% for excellent control).
3. Calculate the statistical confidence interval.
4. If it includes the book value accept the book value as being correct, otherwise reject the book value.

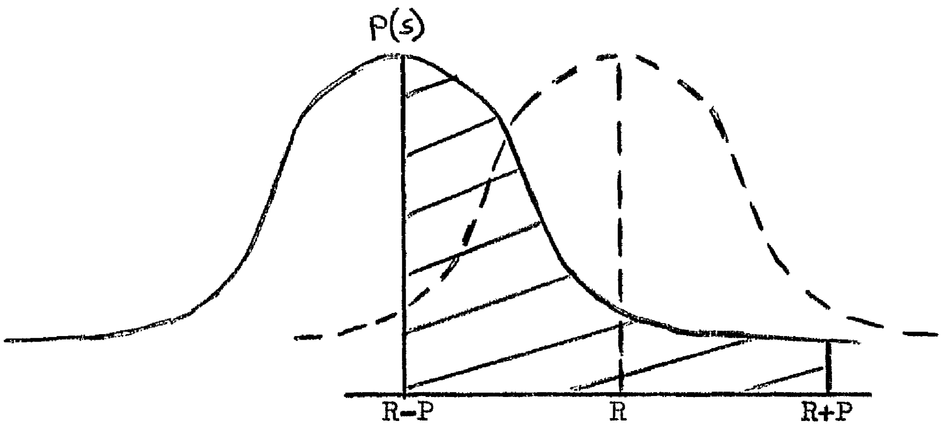
Elliott and Rogers (62) report that many auditors use such decision rules which, although they give an alpha risk of 5% (20% for excellent internal control) give a beta risk of around 50% no matter what the condition of internal control. The risk associated with such a decision rule can be demonstrated as follows:

α risk



If the reported value (R) is correct, the chance of a sample mean lying outside the range (R-P) to (R+P) is 1 - C.L. Hence the chance of a correct reported value lying outside the statistical confidence interval is 1 - C.L. The alpha risk is 1 - C.L.

β risk



Distribution of Sample Means Assuming Reported Value is Incorrect by a Material Amount

If the reported value is materially incorrect, the chance of a sample mean lying inside the range (R-P) to (R+P) equals 0.5 when confidence levels over 80% are used. For example, if a confidence level of 95% is used, P = 1.96 standard deviations, 2P = 3.92 standard deviations and beta = 50%. In most circumstances, therefore, beta risk is 50%. This is unsatisfactory on two counts:

1. it is far too high
2. it does not vary with the confidence level used and adjusting the confidence level is the normal mechanism for allowing for internal control.

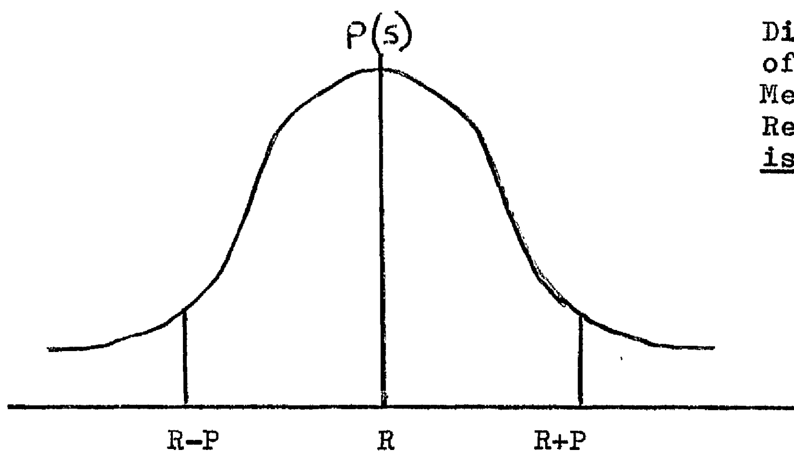
An effective decision strategy

A far better approach is to use a precision which is less than the amount considered material. Elliott and Rogers suggest the following decision rule:

> 1. Let $P = M/(1 + Z^{\beta}/Z^{\alpha/2})$
>
>   Where Zx is the normal table value which includes an area 0.5 - x.
>
> 2. Accept the reported value if it is included in the statistical confidence interval but otherwise reject.

The derivation of this decision rule is as follows:

<u>α risk</u>

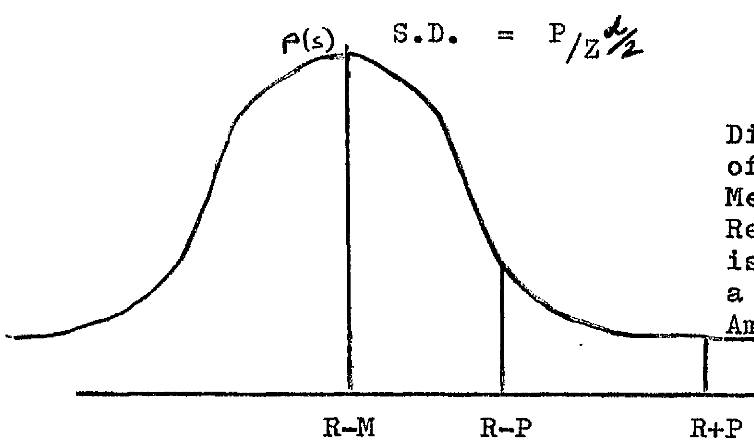


Distribution of Sample Means Assuming Reported Value is Correct

As before, the alpha risk is 1 - C.L.

<u>β risk</u>

$$Z^{\alpha/2} \times \text{S.D.} = P$$
$$\text{S.D.} = P/Z^{\alpha/2}$$



Distribution of Sample Means Assuming Reported Value is Incorrect by a Material Amount

The chance of a sample mean falling inside the range (R-P) to (R+P) is β

$$\text{where } Z = \frac{(R-P) - (R-M)}{S.D.}$$

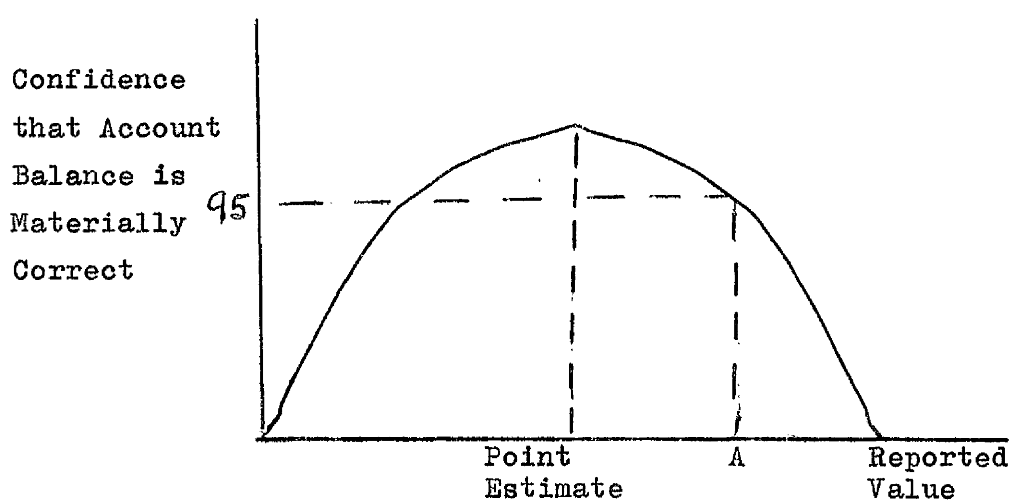
$$= \frac{(M-P)}{P} \; Z_{\alpha/2}$$

$$P = (M-P) \; \frac{Z_{\alpha/2}}{Z_\beta}$$

$$P = M/(1+ Z_\beta / Z_{\alpha/2})$$

Rejection

If application of the decision rules leads to the 'reject' decision the auditor may decide that an adjustment would be acceptable to him. The best estimate of the correct value is the point estimate. However, in many cases the auditor can accept a figure between the point estimate and the reported value. Suppose the auditor wishes to be 95% confident that the account balance is materially correct. The auditor can accept the figure A as shown in the graph.



If the original α is less than 0.05 and/or the original β is less than 0.50, there will be a range of values all of which have at least a 95% confidence of being materially correct (63). Bearing in mind Bayesian statistics, the most logical choice from this range of values is A.

<u>The need for computerisation</u>

There are three reasons why Elliott and Roger's approach is only feasible when computerised.

1. Sample size is calculated on the basis of an expected population variability. Whenever the population variability, as estimated from the sample results, differs from that expected, it is necessary to adjust one or more parameters (α, β, M) or to increase sample size. The necessary adjustments can only be easily computed by computer.
2. In order to keep α, β and sample size to satisfactory levels, stratification is desirable. Again, this can only easily be done by computer.
3. Experience with manual applications suggests that non-sampling errors are frequent. Non-sampling errors can be reduced by computerisation.

To summarise, Elliott and Rogers' decision strategy is as follows:

1. Choose the desired confidence level (C.L. %) so that (100 - C.L.)% corresponds to the alpha risk;
2. Choose the appropriate precision P in relation to materiality M so as to achieve the desired beta risk. $\left(P = M / \left(1 + Z_\beta / Z_{\alpha/2}\right)\right)$

3. Use the decision rule : if the statistical confidence interval includes the reported value, accept; if the statistical confidence interval does not include the reported value, reject.

TEITLEBAUM AND ROBINSON

Teitlebaum and Robinson consider the beta risk for any particular sampling plan at three levels:

Level 1

Given a population with a particular total error (u) provided by a particular pattern of errors, the first level of beta risk (β^{1}) is the probability that a sample does not provide a maximum potential error in excess of the actual error (u).

Level 2

The second level is more general since the β^{1} risks are considered for all possible patterns of errors in the population which total u or more. The second level of beta risk (β^{11}) is the largest value of all these β^{1} risks.

Level 3

The third level is still more general since β^{11} risks are considered for all possible values of u, the actual total error in the population. The third and most general level of beta risk (β^{111}) is the largest of all the β^{11} risks.

β^{111} is the beta risk associated with the use of a sampling plan; β^{11} is the beta risk associated with the particular audit application of the sampling plan where u is equal to materiality. β^{1} is the beta risk associated with the particular audit application to a population with a particular pattern of errors. It should be noted that β^{11} is the beta risk described in the previous section. However, the 'true' beta risk of a sampling plan may be regarded as β^{111}. This is the maximum chance of stating an upper error limit of less than the true actual error value present in the population, <u>however the errors might be distributed in that population</u> and <u>whatever their total value</u>. Since β^{111} is the largest of all β^{11} and β^{11} the largest of all β^{1} it follows that $\beta^{111} > \beta^{11} > \beta^{1}$. If a population can be found with a particular pattern of errors and a particular total error (u) which provides a beta risk (β^{1}) which is greater than the stated (nominal) beta risk of the sampling plan, it follows that the true beta risk (β^{111}) must also be greater than the stated risk, and the plan is invalid for general application. Teitlebaum and Robinson provide such populations to demonstrate the invalidity for general use of stratified sampling and ratio estimation techniques.

In table 9 the population is divided into fifteen strata and stratified sampling suggests the auditor can be 95% confident that the maximum potential overstatement is $6800. (materiality). Stratum four comprises 1296 items around a class mark of $26-25. If stratum four is seeded with $(6800/26.25) = 259$ fictitious items the auditor will fail to detect a fictitious item in his stratum sample of 6 items in about $\left(1 - \frac{259}{1296}\right)^6 = 26\%$ of the time. Hence, a population can be found with a particular pattern of errors and a particular total error ($6800 = materiality) such that the beta risk (β^{1}) of stratified sampling is in excess of the stated risk (5% in this example).

The true beta risk (β 111) must, therefore, be in excess of the stated risk and stratified sampling may be invalidated for general use.

Stratum No	Item count	Value range ($)	Sample Size
1	1,219	0 - 10.50	6
2	1,699	10.50 - 16.50	6
3	1,576	16.50 - 22.50	6
4	1,296	22.50 - 30.00	6
5	923	30.00 - 43.50	8
6	500	43.50 - 63.00	6
7	348	63.00 - 90.00	6
8	294	90.00 - 135.00	8
9	187	135.00 - 195.00	7
10	103	195.00 - 285.00	5
11	68	285.00 - 435.00	6
12	44	435.00 - 675.00	7
13	24	675.00 - 945.00	4
14	16	945.00 - 1,545.00	7
15	12	1,545.00 - 6,945.00	12
	8,309		100

Total value $379,131

TABLE 9

KAPLAN

Kaplan considers a class of auxiliary information estimators which are all described by the following general formula:

$$\hat{Y} = \bar{y} + z\,(\bar{X} - \bar{x})$$

where $\bar{Y}$ = mean true value of population

$\bar{y}$ = mean true value of sampled items

$\bar{X}$ = mean recorded value of population

$\bar{x}$ = mean recorded value of sampled items

z = either a constant or a function of the sample values

For example, direct extension is described by the general formula when $Z = 0$. (i.e. $\hat{Y} = \bar{y}$). The mean true value of sampled items is taken as the estimate of the mean true value of the population. Differences sampling is derived from the general formula by setting $Z = 1$. (i.e. $\hat{Y} = \bar{X} + (\bar{y} - \bar{x})$). The estimate is the mean recorded value adjusted by the difference between the mean true value and mean recorded value of sample items. The standard ratio estimator is derived by setting $Z = \bar{y}/\bar{x} = r$ (i.e. $\hat{Y} = r\bar{X}$). Thus, the estimate is the ratio (r) applied to the mean recorded value of the population. Since Z is computed from sample data, the ratio estimator is biased.

For the class of auxiliary information estimators considered by Kaplan, an estimator of the variance of the mean estimate is given by -

$$S^2_{\hat{Y}} = \frac{1 - f}{n}\, S^2_u$$

where N = number of items in population

n = number of items in sample

f = $^{n}/N$, the sampling fraction

$$S_u^2 = \frac{1}{n-1}\left[\sum_1^n (yj - \bar{y})^2 - 2Z\sum_1^n (yj - \bar{y})(xj - \bar{x}) + Z^2\sum_1^n (xj - \bar{x})^2\right]$$

The simulation experiments were performed using populations generated by known stochastic processes in order to test members of the class of auxiliary information estimators. His results are summarised as follows:

1. As expected, the estimator provided by direct extension (unstratified) had a far higher variance than those provided by differences sampling or ratio estimation.
2. No problem appeared with the bias of the ratio estimator. Its results, even for small samples, were virtually identical to those of the unbiased differences sampling estimator.
3. For differences sampling and ratio estimators, when the estimate of the true value was close to the book value, (i.e. there were few errors in the population) the standard deviation of the estimate was unusually low, and the results given by differences sampling were, as a result, misleading. Differences sampling and ratio estimates give misleading results when errors are infrequent, especially when errors are large and infrequent. These techniques should only be used when a 'qualifying' number of errors exist in the population. The problem is caused by the correlation between the sample mean and standard deviation.

In the extreme case no errors in the sample gives a standard deviation of zero and the result given by differences sampling is that the auditor can be certain there are no errors in the population. (clearly this is a false result). It follows that implanting trivial errors into an error free population does not help, since the estimate of the true value remains close to the book value.

Anderson and Leslie also comment on the unsuitability of implanting trivial errors.

> .... an auditor might be tempted to inject artificial errors into the population in an attempt to create the high population frequency in which he knows difference estimation is more reliable. Yet a moment's reflection indicates that it would be absurd to think that injecting artificial errors (and then reconciling them out at the end) could somehow cure the initial problem any more than one can lift oneself up by one's own bootstraps (64).

Kaplan concludes, that direct extension, differences sampling and ratio estimation are all unsuited to the special requirements of the audit situation:

> The basic problem is that classical techniques described in this paper are designed for a homogeneous population. But an auditing population is actually a mixture of two quite different populations; one consists of a large number of correct items; the other is the much smaller population of items in error. Techniques, such as those presented in this paper, which do not explicitly recognise these fundamentally different populations seem inadequate for auditing applications. (65)

NETER AND LOEBBECKE

Neter and Loebbecke have empirically tested a series of statistical estimators against a variety of created accounting populations. The estimators selected include direct extension (mean per unit), auxiliary estimators such as differences and ratio estimator, stratified mean per unit and dollar unit sampling. Unfortunately, Neter and Loebbecke chose to test a particularly conservative version of dollar unit sampling and their empirical testing of that technique does not, therefore, allow a judgement of the dollar unit sampling plan of Teitlebaum (or, indeed of the Haskins and Sells plan) discussed in the previous chapter. Neter and Loebbecke themselves state in the conclusion to their study that there is a need to examine dollar unit sampling estimators which employ less conservative methods. Their results confirm a number of observations already made in this chapter. For instance:-

(1) direct extension is highly imprecise.

(2) direct extension can be unreliable where populations are moderately skewed.

(3) the bias of the ratio estimator is negligible.

(4) the reliability of the difference and ratio estimator is low when the error rate is low.

Perhaps their most interesting result, however, in view of the potential problem identified by Teitlebaum and Robinson is the empirical testing of the stratified mean per unit estimator. Four broad types of population were constructed by Neter and Loebbecke as follows:-

Population	Skewness of Book Values	Error Direction	Size of Individual Errors
1	Very High	Over and under	Small
2	Moderate	Over and under	Large
3	High	Over only	Moderate
4	Moderate	Over only	Large

Within each population type five different populations were constructed each with the same characteristics of skewness, error direction and error size, but with different error rates. The error rates chosen were 0.5%, 1%, 5%, 10% and 30% (70% for population 2) covering the spectrum from low to very high. Thus, in total 20 populations were constructed. Optimal stratified mean per unit sampling procedures for a sample size of 100 and 15 strata were applied 600 times to each population (optimal procedures are, of course, for book values rather than actual values which are unknown to the auditor). The nominal confidence coefficient given by the statistical theory can then be compared with the proportion of the 600 confidence intervals in the sampling experiment which are correct. If they differ substantially, the implication would be that the auditor cannot rely on the nominal confidence coefficient. Neter and Loebbecke's results for a two sided 95.4% nominal confidence coefficient and a one sided 93.3% nominal confidence coefficient are reproduced overleaf: (66)

Actual Percent of Correct Confidence
Intervals With Mean-Per-Unit Estimator for
15 Strata, n = 100

Population	Population Error Percentage .5	1	5	10	30*
	Two-sided, Nominal 95.4%				
1	94.2	94.2	94.7	95.0	95.5
2	96.3	96.2	96.7	96.0	94.7
3	95.5	95.5	95.8	95.7	95.8
4	94.7	94.7	93.3	91.0	95.0
	One-sided Lower, Nominal 93.3%				
1	94.7	94.8	94.7	96.0	95.0
2	93.5	93.7	94.2	93.8	95.0
3	93.5	93.0	93.2	93.8	93.8
4	89.5	89.5	87.5	85.7	90.5

*For population 2, this error percentage is 70

These results show that for populations 1, 2 and 3 the proportions of correct intervals are close to the nominal confidence for both one sided and two sided intervals. Only for population 4 do the two figures diverge to any significant extent.
Neter and Loebbecke repeated their experiment for a different set of one sided and two sided nominal confidence levels and the nature of the results was unchanged. The nature of the results was also unchanged by extending to 20 strata but an increase in sample size to 200 had the effect of improving the result for population 4.

Although Neter and Loebbecke's study gives stratified mean per unit sampling a relatively clean bill of health they did not specifically test the technique for populations with errors loaded in one stratum (the case where Teitlebaum and Robinson suggest the technique may be vulnerable). The strength of Teitlebaum and Robinson's point must depend on the characteristics of accounting populations and error patterns.

IJIRI AND KAPLAN

Ijiri and Kaplan have developed a sampling plan in which it is assumed the auditor has four sampling objectives. The objectives are to obtain a sample which:

1. represents the population (representative sampling)
2. maximises the number of errors found in the sample (corrective sampling)
3. maximises the monetary value of items included in the sample (protective sampling)
4. prevents anticipation by the auditee of the auditor's choice of sample items (preventive sampling)

Ijiri and Kaplan criticise sampling plans (such as estimation and acceptance sampling) which concentrate solely on the representative objective. Their model incorporating all four objectives depends on three types of prior knowledge being available to the auditor:

1. The auditor must know the recorded monetary value of each item in the population. The contribution a selected sample item makes to the 'protection' objective depends upon its monetary value.

2. The auditor must have knowledge enabling him to stratify the population according to the likelihood of error. The contribution a selected sample item makes to the 'correction' objective depends upon its likelihood of being in error.

3. The auditor must know the particular control area in the organisation that generated each item in the population. Meeting the 'prevention' objective depends upon sample items coming from a sufficient variety of control areas.

<u>Outline of the Model</u>

The model uses the following measures:

1. Representative Sampling: the width w of a confidence interval at a given level of significance.

2. Corrective Sampling: the estimated number r of unchecked items in error, i.e. the product of the number of items <u>not</u> in the sample and the proportion of errors in the sample.

3. Protective Sampling: the ratio u of the total dollar value of items <u>not</u> in the sample to the (estimated) total dollar value of the population.

4. Preventive Sampling: the ratio v of the number of areas in which a specified minimum number of items (say, 1 or 2) have <u>not</u> been selected in the sample to the total number of areas.

The model, which relies on nonlinear and goal programming, gives the smallest sample necessary to meet specified values of w, r, u and v.

In outline, the approach is as follows:

1. A sample is designed as if representative sampling were the only objective.
2. Before checking the items in (1) above the sample is examined to see if it meets the corrective objective. If not, additional (high error risk) items are included in the sample.
3. Before checking the items in (1) and (2) above the sample is examined to see if it meets the preventive objective. If not, additional items are chosen from unsampled control areas.
4. Finally, the items in (1), (2) and (3) above are examined from the protective point of view. If their monetary value is insufficient additional (high value) items are chosen.

<u>KINNEY</u>

Kinney is critical of Ijiri and Kaplan's assumptions and, whilst not quarrelling with the representative and preventive objectives, he claims it is illogical for an auditor to seek correction and protection after his representative objective has been met.

> Suppose that for a particular type of audit situation we can determine the minimum level of confidence and precision required under generally accepted auditing standards. If an auditor's sample has met the representative sampling objective and it can be shown that he selects additional higher dollar value and/or higher error rate items, we can ask why he is unwilling to rely on the mathematical statement which he can make .... to the extent that auditors have met generally accepted auditing standards by representative sampling and still extend the sample, the added cost is wasted or at least inefficiently used. (67)

In reply to Kinney, Ijiri and Kaplan stress they do not claim the corrective and protective objectives to be necessary in any statistical sense; these two objectives are the auditor's 'non-statistical' objectives and the fact that they exist can be seen by observing auditor's behaviour. Given that these objectives exist Ijiri and Kaplan show how they can be included in a mathematical model.

> We hoped to make explicit the non-statistical objectives that we felt auditors had in mind while conducting an audit. We attempted to show how these non-statistical objectives could be modelled mathematically, and how a normative model could be developed to achieve all of the auditor's objectives in an efficient way. The policy issue of whether the auditor should pursue these objectives must be debated and it may be more fundamental than the empirical question of whether auditors actually have these objectives in mind when they sample. We do not believe that what is expected of auditors' 'due professional care' in sampling should be identical to that expected of statisticians, insofar as the purpose of auditing is broader than the objectives of statistical sampling. (68)

In the absence of research in the area (a simple and rather inconclusive test was performed by Hubbard and Stawser (69)), Ijiri and Kaplan are forced to rely on their experience of auditors' behaviour to support their corrective and protective assumptions.

This writer's experience is that auditors do pursue these objectives; for example, they are concerned with cut-off tests (high error risk) and with the proportion of a population total formed by the value of items tested. It is intuitively sensible to seek the high error risk and high value items when deploying scarce audit resources.

However, Kinney does not deny that auditors do and should seek correction and protection. The question is, how much correction and how much protection should auditors seek, and the answer is, just as much as necessary to meet the representative objective. Thus, Kinney sees correction and protection as sub-objectives of the representative objective. Correction and protection are consistent with and result from the representative objective. Kinney demonstrates as follows:

1. Stratify population in two dimensions; by reported dollar value and likely error rate.
2. The resulting cells are assumed to be independent and the dollar amount of the error of an item in a cell is assumed to be proportional to the reported dollar value.
3. The expected error is zero for each cell.
4. Given (1) to (3) the standard deviation (s_{ij}) of the error distribution of a cell is given by -

$$s_{ij} = \left[p_i (f D_j)^2\right]^{\frac{1}{2}}$$

where

p_i = probability of an item in the 'i'th error class being in error

f = constant of proportionality

D_j = average reported dollar value of items in the 'j'th dollar value class

5. Using stratified sampling and the Neyman Allocation it can be shown that the number of sample items in each cell (n_{ij}) is given by -

$$n_{ij} = n \frac{n_{ij}\, s_{ij}}{\sum_i \sum_j N_{ij}\, s_{ij}}$$

where

N_{ij} = the number of items in cell (i j)

6. Thus, the auditor will sample more from a cell if the standard deviation of the error distribution is higher than for other cells.
7. A higher standard deviation (sij) results from a higher probability of error (pi) or larger dollar values (Dj) (see 4).

The argument is, therefore, that the auditor should not explicitly seek a specified level of correction or protection, but does explicitly seek a given width of a confidence interval with a stated level of significance. However, in order to meet the latter objective, an efficient sampling plan will, given prior knowledge of money values and likelihood of errors, sample from high error risk and high value items thus giving an appearance of 'correction' and 'protection'.

LIMITATIONS OF VARIABLES SAMPLING

Direct Extension

The work of Kaplan, and Neter and Loebbecke confirms that direct extension can give misleading results when populations are skewed and is inefficient when populations are variable. The implication is that an auditor needs prior knowledge of the following population characteristics:

(1) skewness

(2) variability

The requirement for prior knowledge of these population characteristics detracts from the only advantage that direct extension appears to have over rival sampling plans viz. it does not require knowledge of the reported values of items making up the population. It is unlikely that the auditor will be in a position to ascertain skewness and variability but not in a position to ascertain reported values of population items. The prior knowledge required by direct extension is of population characteristics rather than the accounting system. However, the number of occasions on which this technique proves useful to the auditor is so small that the fact that it does not require prior knowledge of the system is of relatively little significance.

Differences sampling and ratio estimation

The work of Kaplan, and Neter and Loebbecke confirms that the difference and ratio estimators can give misleading results when the error rate in the population is low. Injection of artificial errors into the population is not a solution to this problem and for difference or ratio estimators to be used, there should be 'qualifying errors' in the population. This prior knowledge of the error rate in the accounting population is probably most conveniently obtained through an examination of the accounting system.

Stratified sampling

Neter and Loebbecke give stratified sampling a relatively clean bill of health but Teitlebaum and Robinson (T & R) demonstrate that the technique may give misleading results when the error is loaded into a single stratum. The point of T & R's article is that stratified sampling (their point also applies to other variables sampling techniques: notably ratio estimation) cannot be relied upon to give accurate results no matter what the distribution of error and no matter what the size of error. If the auditor wishes to employ such techniques he should:

(1) be aware of the error distributions and error sizes for which the technique gives inaccurate results,

and

(2) have prior knowledge of the likely distribution of error and size of error in the accounting population under examination.

There is a need for much further research into the circumstances in which various sampling plans give misleading results. Once these circumstances are generally known, the auditor must learn to identify their existence in the particular accounting populations in which he is interested. Whenever the circumstances include error rates, error sizes or error distributions the auditor is likely to find an examination of the accounting system a necessity.

<u>Computerisation</u>

A number of factors indicated by Elliott and Rogers point to the importance of computerised accounting records and computerised sampling procedures (both selection and calculation) where stratified sampling is used:

1. problems associated with sample selection.

The most versatile variables sampling technique is stratified sampling and the procedures for optimal selection of strata boundaries and allocation of the sample between strata are sufficiently complex that they can only easily be performed by computer.

2. problems associated with the evaluation of sample results.

The formulae for computation of the estimator and standard error of the estimator are also complex and the calculations are preferably performed by computer.

3. problems associated with a revision of sample sizes caused by expectations not being realised.

Sample size in each stratum depends upon the standard deviation of sample items drawn from the stratum. Since sample items and their true values cannot be known in advance the auditor must calculate stratum sample sizes on the basis of estimated standard deviations and whenever these estimates are unrealised he must revise sample sizes. The possibility of having to perform revised calculations reinforces the need for computerisation.

The dependence of stratified sampling on computerisation in all but the simplest applications, limits the extent to which the technique is used in practice and, therefore, also limits the significance of the comments made on the need for prior information in connection with the use of this technique.

Error prone areas

Stratified sampling stratifies the population according to the recorded monetary values of population items, but does not stratify according to the likelihood of error of population items. This is one of the criticisms of conventional sampling plans made by Ijiri and Kaplan. However, Kinney demonstrates that a variables sampling plan can stratify the population in two dimensions (by both reported value and likely error rate) and the application of such a plan clearly requires prior knowledge of likely error rates. Again this prior knowledge may best be gained through an examination of the accounting system.

SUMMARY AND CONCLUSION

Table 10 summarises the comments made in this chapter upon the major variables sampling plans considered for use in auditing.

TABLE 10 Comments on variables sampling plans

Direct extension

1. Variable population results in inefficiency.
2. Skewed population causes misleading results.
3. Prior information needs:
   - (a) population variability
   - (b) population skewness
4. Because of 1 above this technique little used in practice.

Differences

1. Requires qualifying number of errors or results are misleading.
2. Requires errors to be of non-trivial amount or results are misleading.
3. Efficiency improved if differences are all of similar size since standard error of estimate is low.
4. Prior information needs:
   - (a) number of errors
   - (b) size of errors
5. Useful technique for giving the size of an 'adjustment' to the accounts.

Stratified

1. There is a possibility, based on the Teitlebaum and Robinson paper that stratified sampling is misleading if errors follow a certain distribution pattern. This possibility requires further research.

2. Computerisation is 'almost' an essential.

3. Prior information needs:

   (a) distribution of errors

   (b) computerisation

4. May be adapted to stratify by error risk as well as size of items.

5. Additional prior information need for 4:

   (c) likelihood of error for individual items.

6. The most versatile of the variables techniques when a computer facility is available.

This chapter has completed the study of the two propositions which were first investigated in the previous chapter:

1. <u>There is no single choice of statistical sampling technique which is appropriate in all circumstances</u>.

The chapter has shown that differences sampling gives misleading results when population error rates are low. Moreover, for differences sampling to be applicable the errors must be non-trivial in amount. Stratified sampling plans are very time consuming when performed manually and Teitlebaum and Robinson have indicated that stratified variables sampling plans may be defective for certain error distributions in the populations. Generally speaking mean per unit sampling is an inefficient technique.

2. <u>The chances of making the proper selection of statistical sampling technique improve as a result of examination of the accounting system and its internal controls</u>.

Two further reasons why the selection of a sampling plan is assisted by prior examination of the system are recognised in this chapter:

1) to avoid using differences sampling when the error rate is low.
2) to improve the efficiency of stratified variables sampling by stratifying according to error risk as well as value.

As with monetary unit sampling it is likely to be the case that efficient and reliable use of a variables sampling plan requires prior knowledge of the error rate, distribution and size and any such requirement in turn is likely to cause the auditor to make careful examination of the accounting system.

CHAPTER 5

THE THEORY OF AUDIT EVIDENCE AND THE NEED FOR SYSTEMS EXAMINATION

This chapter concentrates on the second major problem of the substantive approach to the audit, viz. not how many sample items to include but what evidence to choose to substantiate the sample items chosen? The chapter is divided into the following sections :

1. Identifying the probabilities

The analysis identifies the probabilities which the auditor implicitly assesses when he evaluates the quality of evidence. It starts with the simplest case of a single piece of evidence used to support a single transaction and develops to consider the single transaction - multiple evidence and multiple transaction - multiple evidence cases.

2. The need to study the system

It is argued that, of the probabilities identified in the analysis, those concerning pieces of evidence processed or generated by the system can only be assessed by the auditor if he has first examined that accounting system. Broadly, the distinction between evidence dependent upon the system and evidence independent of the system corresponds to the distinction between authoritative documents prepared 'inside' and 'outside' the enterprise under examination (70).
Examples of audit evidence employed in substantive tests of details are

a) independent of the system - debtors' circulars, creditors' statements etc.

b) dependent upon the system - despatch notes, goods received notes, copy sales invoices etc.

3. The concept of 'independent' evidence

The nature of independence in the context of the theory of audit evidence is examined in terms of the probabilities identified in 1. on the previous page. It is demonstrated that the greater the independence of the marginal piece of evidence from established evidence, the greater the marginal increase in confidence of the auditor when the marginal piece of evidence is established.

IDENTIFYING THE PROBABILITIES

Single Evidence - Single Transaction

The following analysis of the theory of audit evidence is based on Bayes Theorem and was first suggested by Toba (71). It starts with the case of a single piece of evidence to support a single transaction and develops to consider multiple pieces of evidence and multiple transactions.

Consider the equation -

$$P(A/B) = \frac{P(A) \times P(B/A)}{P(B)} \quad \text{(Bayes Theorem)}$$

Suppose B is a single piece of evidence in support of A and assume that A implies B and hence $P(B/A) = 1$.

Bayes' theorem reduces to -

$$P(A/B) = \frac{P(A)}{P(B)}$$

$$\therefore P(A) = P(A/B) \times P(B) \quad \text{(EQUATION 1)}$$

Two aspects of the theory of evidence follow from equation(1):

(1) The existence of B, the audit evidence, is not certain and, hence the prior probability of B(P(B)) is less than 1. Since $P(B) < 1$ it follows from equation (1) that, $P(A/B) > P(A)$.

The probability (P(A/B)) of A after establishing B is greater than the prior probability of A(P(A)). Toba defines B as supporting evidence since establishing B makes A more tenable. The following statement is established:

> As a result of <u>establishing supporting evidence B, the auditor's confidence in A increases</u>.

(2) Assuming that P(A) remains the same before and after B is established, equation (1) is of the form $xy = c^2$ and describes a rectangular hyperbola. Figure 5 illustrates the rectangular hyperbola graphically.

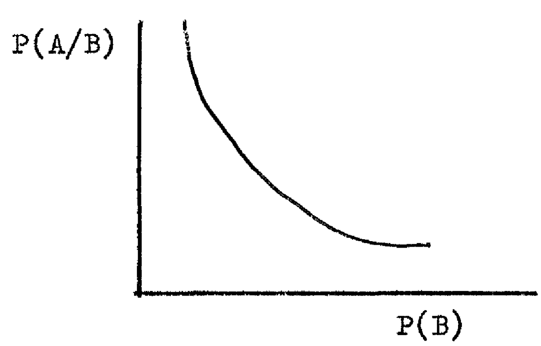


FIGURE 5

It can be seen that P(A/B) is greater, the smaller P(B). The auditor's confidence in A after establishing B, varies inversely with the prior probability of B. The following statement is established:

The increase in the auditor's confidence in A varies inversely with the prior probability of B.

Odds - likelihood form

Toba's analysis of audit evidence is in terms of the simple form of Bayes' theorem. As an alternative, the odds likelihood ratio form can be used. There are only two alternative hypotheses; either A exists or not A ($\bar{A}$) exists and the odds - likelihood form of Bayes Theorem gives

$$\frac{P(A/B)}{P(\bar{A}/B)} = \frac{P(A)}{P(\bar{A})} \times \frac{P(B/A)}{P(B/\bar{A})}$$

Again, suppose B is a single source of evidence in support of A and assume that A implies B and, hence, $P(B/A) = 1$.

The odds - likelihood form of Bayes' theorem reduces to

$$\frac{P(A/B)}{P(\bar{A}/B)} = \frac{P(A)}{P(\bar{A})} \times \frac{1}{P(B/\bar{A})}$$

The relationship between the posterior odds and the prior odds depends on $P(B/\bar{A})$. What is the probability of the evidence B existing, given that A does not exist? If this probability ($P(B/\bar{A})$) is high, then $\frac{1}{P(B/\bar{A})}$ is low and there is not much revision of the odds. If $P(B/\bar{A})$, however, is low, $\frac{1}{P(B/\bar{A})}$ is high and the revision of the odds is much greater. The following statement is established:

The increase in the auditor's confidence in A as a result of establishing B varies inversely with the probability of B given A does not exist ($P(B/\bar{A})$).

It is easier to focus on $P(B/\bar{A})$ than $P(B)$, and, hence, this is a more useful form of Bayes' theorem. For example, if debtors' circulars are used as evidence of debtors' balances, the auditor can, from his experience, form a subjective view of the probability of a confirmation given that the debtor's balance is incorrect. This is more meaningful than simply the prior probability of receiving a confirmation.

If a confirmation is unlikely, given that a debtor's balance does not exist, then a debtor's circularisation provides a valuable source of audit evidence. In a Bayesian analysis the quality of the evidence B can be identified with the conditional probability $P(B/\bar{A})$.

Single Transaction - Multiple Evidence

Toba uses equation (1) to evaluate a single available piece of evidence, and then adapts equation (1) to enable the evaluation of a piece of evidence which is one of several available.

Suppose, instead of a single piece of evidence (B), many pieces of evidence (b_1, b_2, b_3, etc) are available. The auditor has established b_1, b_2 .... bm and is about to establish b_{m+1}.

Equation (1) then becomes:

$$P(A/b_1\ b_2 \ldots\ bm) = P(A/b_1\ b_2 \ldots\ bm{+}1) \times P(bm{+}1/b_1\ b_2 \ldots\ bm) \qquad \text{(EQUATION 2)}$$

Equation (2) can be demonstrated as follows:

Applying equation (1) to the first piece of audit evidence (b_1).

$$P(A) = P(A/b_1) \times P(b_1)$$

$$P(A/b_1) = \frac{P(A)}{P(b1)}$$

to the second piece of audit evidence (b_2)

$$P(A/b_1\ b_2) = \frac{P(A/b_1)}{P(b_2/b_1)}$$

to the m + 1 piece of audit evidence bm+1.

$$P(A/b_1\ b_2\ \ldots\ bm{+}1) = \frac{P(A/b_1\ b_2\ \ldots\ bm)}{P(bm{+}1/b_1\ b_2\ \ldots\ bm)}$$

$$\therefore P(A/b_1\ b_2\ \ldots\ bm) = P(A/b_1\ b_2\ \ldots\ bm{+}1) \times P(bm{+}1/b_1\ b_2\ \ldots\ bm)$$

Toba makes the following points:

1. $P(bm{+}1/b_1 b_2\ \ldots\ bm)$ is less than 1 and, therefore $P(A/b_1 b_2\ \ldots\ bm{+}1)$ is greater than $P(A/b_1 b_2\ \ldots\ bm)$. Hence, the marginal piece of evidence (bm+1) increases the auditor's confidence in A, and it follows that when several pieces of evidence support a proposition confidence in the proposition is increased.
2. If bm+1 is unexpected and $P(bm{+}1/b_1 b_2\ \ldots\ bm)$ is low, the additional confidence resulting from the marginal source of evidence bm+1 is high. As more sources of evidence are established, it is likely that the marginal piece of evidence becomes more expected, and the additional confidence resulting from the marginal piece of evidence declines. This is the principle of diminishing marginal weight of evidence.

Multiple Transaction - Multiple Evidence

The evaluation of a sample result takes place in two stages:

1. evaluation of evidence in support of individual transactions selected in the sample;

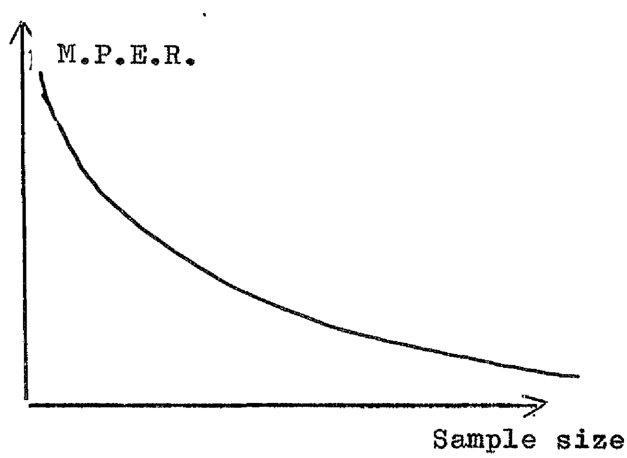
2. evaluation of the sample results. This converts evidence of individual transactions into evidence of an entire account balance.

Toba demonstrates that the second stage is described by equation (2). Suppose an account balance A implies transactions t_1 t_2 t_3 ... t_m t_{m+1} etc. Hence, $P(t_1/A)$, $P(t_2/A)$ etc. = 1. However, establishing a single transaction t_1 is not conclusive evidence (Toba says it does not confirm A) of A. Even establishing a large sample of transactions leaves a possibility that A is untrue. However, the following statements can be made:

1. Establishing the marginal transaction, (t_{m+1}) increases confidence in A.
2. As more transactions are established, the marginal transaction becomes more expected and the additional confidence provided declines.

These statements are consistent with attribute sampling. The graph in figure 6 shows the shape of the curve for the maximum potential error rate given different sample sizes assuming zero errors and a 95% confidence level.

FIGURE 6



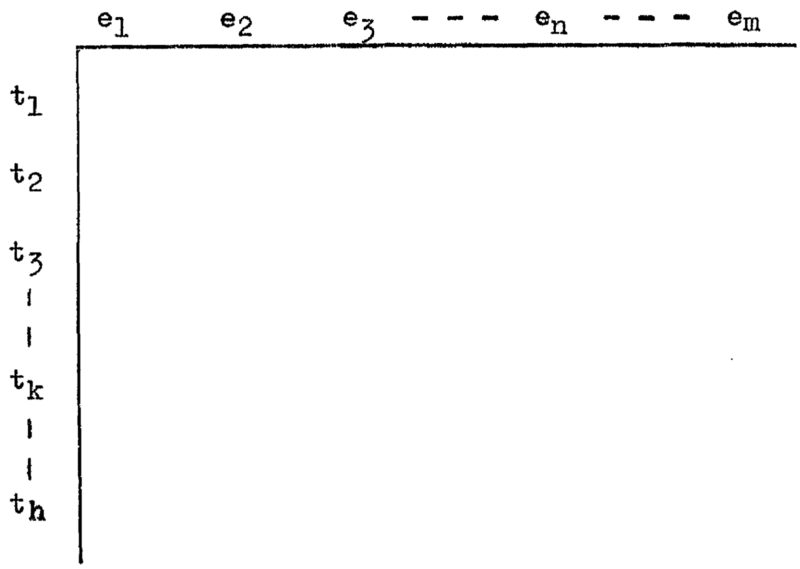
As sample size increases the slope of the curve becomes a lower negative figure indicating that M.P.E.R. falls at a declining rate.

The analysis can be extended to the evaluation of evidence in support of individual transactions (first stage - this application is not considered by Toba). Suppose the existence of a transaction t implies evidence e_1 e_2 ... e_m e_{m+1} etc. Hence, P (e_1/t), $P(e_2/t)$ etc. = 1. However, establishing a single piece of evidence e_1 is not conclusive evidence of t. Even when several pieces of evidence are established some doubt will remain. However, the following statements can be made:

1. Establishing the marginal place of evidence (e_{m+1}) increases confidence in t.
2. As more evidence is established, the marginal evidence becomes more expected and the additional confidence provided, declines.

Reducing the burden of evidence

The auditor is required, not to consult all pieces of evidence in support of each transaction, but to consult sufficient evidence to satisfy himself of the true value. A sufficient number of pieces of evidence is likely to decrease as the number of transactions already tested, increases. Suppose there are m sources of evidence $(e_1 e_2 \ldots e_n \ldots e_m)$ available to support each sample item $(t_1 t_2 \ldots t_k \ldots t_h)$ chosen from a population.



The increase in the auditor's confidence in t_k which results from establishing e_n is then a function of $P(e_n/e_1e_2 \ldots e_{n-1})$. The smaller this probability, the greater the increase in the auditor's confidence when the evidence is established.

Let us assume the order in which the auditor establishes $e_1e_2e_3$ etc. is such that the increase in confidence resulting from e_1 is greater than that from e_2 etc. In this case

$$P(e_1) < P(e_2/e_1) < P(e_3/e_1e_2) \ldots < P(e_n/e_1e_2 \ldots e_{n-1}) \ldots < P(e_m/e_1e_2 \ldots e_{m-1}) \quad - \text{ INEQUALITY 3.}$$

In general, a piece of evidence e_n ceases to contribute to the auditor's confidence as $P(e_n/e_1e_2 \ldots e_{n-1})$ tends to 1. From inequality (3),

$$1-P(e_1) > 1-P(e_2/e_1) > 1-P(e_3/e_1e_2) \ldots > 1-P(e_n/e_1e_2 \ldots e_{n-1}) > 1-P(e_m/e_1e_2 \ldots e_{m-1}) \quad - \text{ INEQUALITY 4.}$$

As the number of sample items for which $e_1e_2 \ldots e_n$ are established, increases, so the auditor revises his assessment of $P(e_n/e_1e_2 \dots e_{n-1})$ upwards. However, it can be seen from inequality (4) that the amount of revision needed for $P(e_1)$ before e_1 ceases to contribute to the auditor's confidence is greater than that needed for $P(e_2/e_1)$ before e_2 ceases to contribute, which is greater than that needed for $P(e_3/e_1e_2)$ before e_3 ceases to contribute etc. Hence, the number of sample items for which e_1 needs to be established is larger than the number of sample items for which e_2 (given e_1) needsto be established, which is larger than the number of sample items for which e_3 (given e_1e_2) need be established, etc.

Figure 7 illustrates the type of approach suggested by this analysis. (In reality the relative cost of establishing each piece of evidence is also an influence on the order in which pieces of evidence are given up by the auditor).

	e_1	e_2	e_3	e_4	e_5
t_1 - t_{10}	x	x	x	x	x
t_{11} - t_{15}	x	x	x	x	
t_{16} - t_{20}	x	x	x		
t_{21} - t_{50}	x	x			
t_{51} - t_{100}	x				

FIGURE 7

There are five available sources of evidence and a sample of 100 items has been chosen. All 5 sources are established for the first 10 items, and none of the evidence contradicts each other. As a result of this the auditor is able to revise his conditional probabilities $P(e_1)$, $P(e_2/e_1)$, $P(e_3/e_1e_2)$, $P(e_4/e_1e_2e_3)$ and $P(e_5/e_1e_2e_3e_4)$ upwards. He is now of the opinion that $P(e_5/e_1e_2e_3e_4)$ is tending to 1 and, hence, no longer makes a contribution. For the next 5 items he considers only 4 sources of evidence. He is then in a position to once again revise his probabilities upwards and to drop out the fourth source of evidence etc.

Variations in the approach do not facilitate this argument. For example, figure 8 reverses the approach illustrated in figure

	e_1	e_2	e_3	e_4	e_5
$t_1 - t_{50}$	x				
$t_{51} - t_{80}$	x	x			
$t_{81} - t_{85}$	x	x	x		
$t_{86} - t_{90}$	x	x	x	x	
$t_{91} - t_{100}$	x	x	x	x	x

FIGURE 8

Neither this reversal, nor any other variation is desirable since it is the experience in the earlier items tested which enables the auditor to revise his probabilities upwards and drop out first e_5, then e_4 etc. In the approach shown in figure 8 there is no experience of $e_2 e_3 e_4$ or e_5 gained in the first 50 transactions.

THE NEED TO STUDY THE SYSTEM

Many of the pieces of evidence employed by the auditor, especially when substantively testing profit or loss items, are either generated or processed by the system. Suppose, for example, that the transactions under investigation are sales, and that the multiple evidence in support of each sales transaction is:

(a) despatch

(b) returned delivery advice

(c) copy sales invoice

(d) stock records entry

(e) cash receipt from customer

In a typical system, the despatch note is generated by goods outwards and used to generate the delivery advice (which travels with the goods) and the sales and copy sales invoices (excluding price and arithmetic). The delivery advice may be returned from the customer with the customer's signature to evidence receipt of goods. The sales and copy sales invoice are processed in the accounts departments (eg price and arithmetic are included) and the despatch note is used to generate an entry reducing stock in stock records. Clearly each of the first four pieces of evidence are both generated and processed by the accounting system.

The auditor must decide which pieces of evidence to use and in what order. Assuming the auditor wishes to use all five pieces of evidence to at least a limited extent, then the ordering may be made in any one of 120 (5!) ways and to choose between these 120 alternatives the auditor must consider an even higher number of conditional probabilities. Clearly an examination of all possibilities is not possible and the auditor is likely to select an order, which although unlikely to be optimal, appears to be reasonably efficient. Suppose he selects the order as follows: despatch $\longrightarrow$ returned delivery advice $\longrightarrow$ copy sales invoice $\longrightarrow$ stock records entry $\longrightarrow$ cash receipt from customer. In such a case the auditor must consider such probabilities as $P(b/a)$ (i.e. the probability of a proper returned delivery advice given a despatch note), $P(c/ab)$ (the probability of a proper copy sales invoice given both a despatch and delivery advice) etc. Only when he assesses such probabilities can the auditor assess the additional confidence which results from establishing each marginal piece of evidence. The assessment of these probabilities requires an examination of the accounting system.

THE CONCEPT OF 'INDEPENDENT' EVIDENCE

In general, the theory of evidence states that when independently derived pieces of evidence support a proposition, confidence in the proposition is increased. What is meant by the concept of 'independent' evidence in this context?

Toba's analysis suggests that when $p(bm+1/b_1b_2 \ldots bm)$ is low, bm+1 makes a significant contribution to the auditor's confidence in proposition A. Hence, given that $b_1b_2 \ldots bm$ have been established, $p(bm+1/b_1b_2 \ldots bm)$ is a measure of the quality of bm+1 as a piece of evidence. However, if there is no prior evidence, $p(bm+1)$ is a measure of the quality of bm+1. The first of these measures the quality of bm+1 as a marginal piece of evidence in the specific circumstances and the second, the quality of bm+1 as a sole piece of evidence.

The difference between these two measures (i.e. between $p(bm+1/b_1b_2 \ldots bm)$ and $p(bm+1)$) can be taken as a measure of the independence of bm+1 from established evidence. If this difference is zero, then -

$$p(bm+1/b_1b_2 \ldots bm) = p(bm+1)$$

and bm+1 is completely independent of established evidence in a strict statistical sense.

However, this strict statistical independence is an unusual occurrence in an audit context. For example, there may be at least five pieces of evidence in support of a receipt of stock -

1. purchase invoice
2. goods received note
3. payment to supplier
4. reconciliation with supplier's statement
5. entry in stock records

To varying degrees, none of this evidence is independent. The entry in the stock records might be made, not from an independent count by the storekeeper, but directly from the goods received note. Payment to a supplier will normally depend upon both an invoice and a goods received note being received in the accounts section concerned. Reconciliation with a supplier's statement depends upon invoices and payments.

In an audit evidence context, there will be a difference between $p(bm+1/b_1b_2 \ldots bm)$ and $p(bm+1)$ and the smaller this difference, the greater the independence of bm+1 from established evidence.

Algebraicly, the position is as follows:

Let Qs = the quality of evidence bm+1 as a sole piece of evidence

Qm = the quality of evidence bm+1 as a marginal piece of evidence

$$\text{then } Qs = 1 - p(bm+1)$$

$$Qm = 1 - p(bm+1/b_1b_2 \ldots bm)$$

$$\text{and } Qs - Qm = p(bm+1/b_1b_2 \ldots bm) - p(bm+1)$$

The quality of bm+1 as a marginal piece of evidence depends upon:

1. its quality as a sole piece of evidence (Qs)
2. its independence from established evidence (Qs - Qm)

Provided it has the ability to contribute as a sole piece of evidence, independent evidence has the ability to contribute as a marginal piece of evidence.

The concept of independent evidence is capable of giving important insights into the use of different pieces of audit evidence. It has, for example, been used by this author to develop a theoretical distinction between compliance tests and substantive tests other than compliance tests (72). Evidence of internal control design reliability and compliance may be used as evidence of substantive conditions (this is the basis of the systems approach). The theory of evidence states that when <u>independent</u> pieces of evidence support a proposition, confidence in the proposition increases. When it is proposed to employ further evidence, other than systems evidence, in support of a proposition as to substantive conditions then such evidence should be independent of the systems evidence (including compliance evidence). Independent evidence of internal control design reliability can be provided by analysis of internal controls and their cumulative effect. What is not always done by the auditor is to select compliance and substantive tests which are consistent with the requirement that evidence of compliance and further evidence of substantive conditions should be independent. The following tests have been suggested to examine the independence of substantive and compliance evidence:

> A substantive test is any test capable of providing additional knowledge of underlying substantive conditions, even though the auditor were to have complete prior knowledge of compliance conditions.
>
> A compliance test is any test capable of providing additional knowledge of underlying compliance conditions, even though the auditor were to have complete prior knowledge of substantive conditions.

The idea is, for example, that if further substantive evidence is independent of compliance evidence, it should be capable of providing additional knowledge of substantive conditions even though, hypothetically, the auditor were already to have perfect evidence (i.e. complete knowledge) of compliance conditions. The author has illustrated the use of these tests of independence by reference to two simple audit procedures (73).

CONCLUSION

Any substantive tests of details take place in two stages:

1. Evaluation of evidence in support of individual transactions selected in the sample.
2. Evaluation of sample results.

One of the advantages of substantive testing as a source of evidence is that the sample results may be objectively calculated using statistical techniques. However, there are problems associated with the selection and operation of sampling plans and these were discussed in earlier chapters. This chapter has considered the problems associated with evaluation of evidence in support of individual transactions selected in the sample. The theory of evidence has been examined in terms of Bayes' theorem to give an understanding of the probabilities implicitly assessed by the auditor when he evaluates evidence. It is shown that where the evidence of transactions (or account balances) is processed or generated by the system (i.e. it is not independent of the system) an examination of the system is necessary for the proper evaluation of the evidence. Hence, this chapter establishes the important role of the systems examination in the assessment and selection of substantive evidence for sample items.

This chapter has also highlighted the importance of the concept of 'independence' in the context of audit evidence combination. Provided it has the ability to contribute as a sole piece of evidence, independent evidence has the ability to contribute as a marginal piece of evidence. Assuming that compliance tests are independent of substantive tests, then the systems examination provides a source of audit evidence independent of the substantive testing. This quality of independence from the substantive testing is likely to make the systems examination a valuable component in the audit approach whenever the internal control design is assessed as reliable.

CHAPTER 6

CONCLUSION

This thesis has studied the relationship between systems and substantive evidence from two different perpectives. Chapter 2 investigated the relationship from the perspective of the systems approach. This chapter reviewed conventional ideas on the nature of systems evidence and on the problems of combining the rather more subjective systems evidence with the rather less subjective substantive testing. It was found that there is at present no satisfactory formal method of combining the evidence, although various procedures are employed in practice for linking the confidence level required from substantive tests with the results of the systems assessment. Chapters 3, 4 and 5 investigated the relationship between systems and substantive evidence from the perspective of the substantive approach. The two key decisions for those adopting the substantive approach are:

1. which sampling technique to adopt?
2. what substantive evidence to employ in support of sample items?

It was found that an understanding of the system is necessary for an efficient solution to both these problems. Whereas the systems approach concentrates on the role of systems examination as a source of audit evidence which supports the financial statements, the substantive approach, therefore, focuses attention on two further important roles for the systems examination.

1. <u>as a source of information for making the choice of statistical sampling technique</u>.

Given the possibly serious consequences of selecting an inappropriate technique, this role of the systems examination deserves understanding and consideration by the auditor. Selection of an appropriate required confidence level based on the systems examination can be no more important than selection of an appropriate statistical technique, since if the statistical technique is inappropriate the nominal confidence provided by the technique may differ significantly from the real confidence level achieved. There is little point in knowing what real confidence the auditor should be looking for from his statistical test if the statistical technique is not capable of measuring the real confidence level.

2. <u>as a source of information for making the selection of substantive evidence to support individual transactions or balances</u>.

The importance of this relationship between the study of the system and an understanding of the quality of the substantive evidence deserves prominence. Selection of an appropriate confidence level is no more important than selection of appropriate audit evidence to substantiate sample items in the substantive test, since there is little point in getting the right number of items in the substantive sample if the evidence substantiating individual sample items is not properly understood.

The thesis, therefore, has identified and studied the following three theoretical roles for the systems examination:

1. systems examination as a source of audit evidence. (Chapter 2)

2. systems examination as a source of information for making the choice of statistical sampling technique. (Chapters 3 and 4)

3. systems examination as a source of information for making the selection of substantive evidence to support individual transactions or balances. (Chapter 5)

Although the first role has been given much coverage in the auditing literature, the second and third roles have been given far less attention.

IMPLICATIONS FOR AUDIT POLICY

The question arises as to what are the implications for audit policy of a recognition of the three, rather than the one, important roles for the systems examination. One contribution of such a recognition is to clarify the distinction between the 'systems' and 'substantive' approaches to the audit. Conventionally the systems approach is regarded as one which places reliance on the system and its internal controls and the substantive approach as one which does not. This is too simple and the distinction between the two approaches should be viewed in terms of <u>how</u> the auditor uses the systems examination rather than in terms of <u>whether or not</u> he uses the systems examination. In the systems approach the systems examination is used primarily as a source of audit evidence supporting the financial statements and, therefore, justifying a reduction in substantive testing. In the substantive approach the significance of the systems examination lies in its assistance with the choice of sampling technique and the understanding it provides of the quality of substantive evidence processed or generated by the system.

It is possible to go further and to suggest that the requirement for 'independent' evidence as explained in Chapter 5 requires the 'systems' and 'substantive' approaches to be defined in such a way that the evidence in support of the financial statements provided by the systems examination and substantive testing is independent.

When the auditor adopts what is, at present, commonly understood to be the systems approach, the systems examination is used first to justify a reduction in the volume of substantive testing but then, quite possibly, again to justify the quality of pieces of evidence processed or generated by the systems and employed to substantiate sample items.

In such circumstances the systems and substantive evidence may not be independent. If such evidence is to be independent then the results of the substantive testing should be capable of supporting the financial statements irrespective of the condition of the accounting system and its internal controls. Should, however, the accounting system be poor, evidence generated by that system may not be capable of substantiating sample items and it follows that any substantive tests which rely on such evidence may not be independent of the systems examination. In other words, if the auditor gets his systems assessment wrong then the substantive test results are also likely to be misleading if they should rely on pieces of evidence processed or generated by the system. To the extent that the substantive testing relies on such internal evidence, the substantive tests and systems examination cannot be regarded as independent. The principal area where substantive testing uses internal evidence is in the substantive testing of transactions. However, internal evidence is also sometimes used to substantiate balance sheet items. For example, despatch notes, invoices etc. are used to substantiate a debt from a customer who does not respond to a debtors circular.

Chapter 5 drew attention to the importance of 'independence' in determining the quality of evidence when it serves as additional or 'marginal' evidence. In chapter 5 it is stated that the quality of a piece of evidence as marginal evidence depends upon:

1. its quality as a sole piece of evidence, and
2. its independence from established evidence.

Independent marginal evidence has the ability to contribute to audit assurance provided it has the ability to contribute as a sole piece of evidence. Non independent marginal evidence will contribute very little to audit assurance even if it has the ability to make a large contribution as a sole piece of evidence.

The implication of any such requirement for 'independence' is that the auditor should pursue one of two policies:

(A) to use the systems examination as a source of audit evidence justifying a reduction in substantive testing; such testing to use only evidence <u>externally generated outside of the accounting system</u>, OR

(B) to use the systems examination to assess the quality of <u>internally</u> generated evidence which is used, together with external evidence in substantiating sample items in substantive tests and to ignore the role of systems examination as a source of audit evidence reducing the volume of substantive testing.

Policy (A) provides a re-definition of the 'systems' approach and Policy (B) a re-definition of the 'substantive' approach. Any policy other than (A) or (B) is likely to risk the auditor taking assurance in the financial statement figures both from the accounting system and from internally generated evidence which lacks independence from the accounting system. In effect policy (A) or (B) is necessary to avoid the danger of the auditor 'double counting' the assurance to be obtained from the system.

Chapter 2 identified the problems of the role of systems examination as a source of audit evidence justifying a reduction in the substantive testing. This difficulty remains for policy (A). It is the problem of satisfactorily combining the systems and substantive evidence. The distinction between policies (A) and (B) may still be thought of in terms of the <u>use</u> to which they put the results of the systems examination. Policy (A) attempts to evaluate the systems evidence in terms of its effect on the likelihood of error in the total accounting population. It is this assessment of the likelihood of error in the total accounting population which determines confidence level and, hence, sample size for the substantive testing. Policy (B), however, attempts to evaluate the systems evidence in terms of its effect on the likelihood of error in the particular items selected from the population for substantive testing. It then relies upon the chosen statistical sampling technique to translate the knowledge concerning the individual population items into knowledge relating to the population as a whole.

FUTURE RESEARCH

This analysis suggests that the 'systems' and 'substantive' approaches should be more clearly defined in order to provide the auditor with a better understanding of his actual policy alternatives. Whether or not this tentative analysis is correct depends upon the outcome of the following future research which needs to be conducted:

1. 'A priori' research to substantiate and develop the analyses of this thesis

In particular such research should further examine:

a) the nature of 'independence' as used in the context of audit evidence and, hence,

b) the significance of the distinction between substantive evidence internally generated by the system and externally generated evidence.

2. Empirical research to assess the level of understanding by practising auditors of the issues addressed in this thesis

A low level of understanding on the part of practising auditors of the relationship between the systems and substantive evidence would support the conclusion that there is a need to clarify the 'systems' and 'substantive' approaches to the audit. The need is recognised for a number of future empirical research studies to discover more about how the auditor currently uses his examination of the accounting system and its internal controls.

Studies would tackle the following issues:

<u>Selection of sampling plans</u>

To what extent do practising auditors:

a) Appreciate the existence of the various sampling plans available?

b) Understand the circumstances in which the various sampling plans give misleading or inefficient results?

c) Appreciate the relationship between the systems examination and the identification of the circumstances in which various sampling plans give misleading or inefficient results?

<u>Selection of Substantive evidence</u>

To what extent do practising auditors:

d) Distinguish between substantive evidence generated or processed internally by the accounting system and evidence generated externally?

e) Consider the possible lack of independence of the systems examination, when used as evidence to support the financial statements, from those substantive tests which employ internally generated evidence?

f) Appreciate the relationship between a study of the system and an understanding of the quality of substantive evidence in support of individual transactions or balances?

REFERENCES

Chapter 1

(1) Statement drawn from 'The Comments of the Audit Practices Committee of The Institute of Chartered Accountants of Scotland, upon the Discussion Drafts of Auditing Standards and Guidelines issued in May 1978 by the Auditing Practices Committee of the Consultative Committee of Accountancy Bodies'. These comments are available for public inspection at the offices of the Scottish Institute.

(2) Arens A.A. and Loebbecke J.K., 'Auditing: An Integrated Approach', Prentice Hall, 1976, pp 191 - 210.

(3) Arens A.A. and Loebbecke J.K., op. cit., see Chapter 4.

(4) For example, J.P. Percy, 'A Case for Statistical Sampling in Auditing', The Accountants Magazine, April 1975, p 140.

(5) A.I.C.P.A., Statement on Auditing Standards No. 1, section 320.70.

(6) Auditing Standards and Guidelines, published by the S.161 Accountancy Bodies, Auditing Standard 101, 'The Auditor's Operational Standard', paragraph 5.

(7) I.C.A.E.W., 'Statement on Auditing U.1', paragraph 9. (Withdrawn April 1980).

(8) A.I.C.P.A., Statement on Auditing Standards No. 1, section 320.01.

Chapter 2

(9) Sweiringa R.J. and Carmichael D.R., 'A Positional Analysis of Internal Control', The Journal of Accountancy, February 1971, pp 34 - 43.

(10) Touche Ross & Co. Canada, Audit Manual, 1975, p.12.

(11) Konrath L.F., 'The CPA's Risk in Evaluating Internal Control', The Journal of Accountancy, October 1971, p.56.

(12) Hatherly D.J., 'Compliance and Substantive Tests - Can the Auditor tell the Difference?', The Accountant's Magazine, September 1978, pp 378 - 380.

(13) Mautz R.K. and Mini D.L., 'Internal Control Evaluation and Audit Program Modification', The Accounting Review, April 1966, p.290.

(14) A.I.C.P.A., Statement on Auditing Standards No. 1, Section 320.65, p.31.

(15) Cushing B.E., 'A Mathematical Approach to the Analysis and Design of Internal Control Systems', The Accounting Review, January 1974, pp 24 -41.

(16) Kaplan R.S., 'Statistical Sampling in Auditing with Auxiliary Information Estimators', Journal of Accounting Research, Autumn 1973, p. 257

(17) Kaplan R.S., 'A Stochastic Model for Auditing', Journal of Accounting Research, Spring 1973, p.46.

(18) Mautz R.K. and Sharaf H.A., 'The Philosophy of Auditing', American Accounting Association, 1961, pp 153 - 157.

(19) Willingham J.J. and Carmichael D.R., 'Auditing Concepts and Methods', McGraw Hill, 1975, p.166.

(20) Smith K.A., 'The Relationship of Internal Control Evaluation and Audit Sample Size', The Accounting Review, April 1972, p.262.

(21) Broderick J.C., 'Statistical Sampling', The Arthur Young Journal, Autumn/Winter 1972/73, p.40.

(22) McRae T.W., 'Statistical Sampling for Audit and Control', Wiley, 1974, p.203.

(23) Meikle G.R., 'Statistical Sampling in an Audit Context', C.I.C.A., 1972.

(24) Hatherly D.J., 'Linking Internal Control and Substantive Tests: A Note', Accounting and Business Research, Winter 1975, p.64.

(25) Teitlebaum A.D., 'Dollar Unit Sampling in Auditing', Paper presented to the December 1973 National Meeting of the American Statistical Association, Appendix III, p.39.

(26) Kraft W.H. Jnr, 'Statistical Sampling for Auditors: A New Look', Journal of Accountancy, August 1968, pp 49 - 56.

(27) Sorensen J.E., 'Bayesian Analysis in Auditing', The Accounting Review, July 1969, pp 555 - 561.

(28) Tracy J.A., 'Bayesian Statistical Confidence Intervals for Auditors', Journal of Accountancy, July 1969, pp 41 - 71.

(29) Phillips L. and Edwards W., 'Conservatism in a Simple Probability Task', Journal of Experimental Psychology, September 1966, pp 346 - 354.

(30) Corless J.C., 'Applying Bayesian Statistics in Auditing', The Accounting Review, July 1972, pp 556 - 566.

(31) Ibid p.565

(32) Ibid p.559

(33) Ibid p.560

(34) Ibid p.565

(35) Ibid pp 565, 566.

Chapter 3

(36) Meikle op. cit.

(37) Teitlebaum A.D., op. cit.

(38) Anderson R.J., 'The External Audit', Pitman Publishing, 1977, pp 330 - 377.
And Leslie D.A., Teitlebaum A.D., Anderson R.J., 'Dollar Unit Sampling. A Practical Guide for Auditors', Pitman Publishing, 1980.

(39) Teitlebaum A.D., op. cit., Appendix 3, p.19.

(40) Stronge D. and Stewart T., 'Sampling: The Way to Certainty for the Auditor', Accountancy Age, 11th October 1974, p.10.

(41) Teitlebaum A.D., op.cit., Appendix 3, p.25.

(42) Hendriksen E.S., 'Accounting Theory', Irwin, 1970, p.119.

(43) Elliott R.K. and Rogers J.R., 'Relating Statistical Sampling to Audit Objectives', The Journal of Accountancy, July 1972, p.42.

(44) Kaplan R.S., 'Sample Size Computations for Dollar Unit Sampling', Journal of Accounting Research, Supplement 1975, pp 126 - 133.

(45) Smith T.M.F., 'Statistical Sampling for Accountants', Accountancy Age, 1976, pp 179 - 192.

(46) Moriarity S., 'Discussion of Decision Theory Aspects of Internal Control Design/Compliance and Substantive Tests', Journal of Accounting Research, Supplement 1975, pp 33,34.

(47) Meikle G.R., op. cit., p.31.

(48) Smith T.M.F., op. cit., p.233.

(49) Ibid p.236.

(50) Goodfellow J.L., Loebbecke J.K. and Neter J., 'Some Perspectives on C.A.V. Sampling Plans', CA Magazine, November 1974, p.50.

(51) Anderson R.J. and Teitlebaum A.D., 'Dollar Unit Sampling. A Solution to the Audit Sampling Dilemma', CA Magazine, April 1973, p.36.

(52) Kinney W.R. Jnr., 'Integrating Audit Tests: Regression Analysis and Partitioned Dollar Unit Sampling', CICA/ICCA Symposium, C.I.C.A., April 1978, Paper 3.

(53) Hatherly D.J., 'Segmentation and the Audit Process', Accounting and Business Research, Spring 1979, pp 152 - 156.

Chapter 4

(54) Arkin H., 'Handbook of Sampling for Auditing and Accounting', McGraw Hill, 1963.
And Cyert R.M. and Davidson H.J., 'Statistical Sampling for Accounting Information', Prentice Hall, 1962.

(55) Goodfellow J.L., Loebbecke J.K. and Neter J., 'Some Perspectives on C.A.V. Sampling Plans', CA Magazine, October 1974, pp 23 - 30, and November 1974, pp 46 - 53.

(56) Elliott R.K. and Rogers J.R., op. cit., pp 46 - 60.

(57) Teitlebaum A.D. and Robinson C.F., 'The Real Risks in Audit Sampling', Journal of Accounting Research, Supplement 1975, pp 70 - 91.

(58) Kaplan R.S., 'Statistical Sampling in Auditing with Auxiliary Information Estimators', loc. cit., pp 238 - 258.

(59) Neter J. and Loebbecke J.K., 'Behaviour of Major Statistical Estimators in Sampling Accounting Populations: An Empirical Study', A.I.C.P.A., 1975, pp 11 - 28.

(60) Ijiri Y. and Kaplan R.S., ' A Model for Integrating Sampling Objectives in Auditing', Journal of Accounting Research, Spring 1971, pp 73 - 87.

(61) Kinney W.R., 'The Auditor's Sampling Objectives: Four or Two?', Journal of Accounting Research, Autumn 1972, pp 407 - 412.

(62) Elliott R.K. and Rogers J.R., op. cit., p.48.

(63) Ibid p.54.

(64) Anderson R.J. and Leslie D.A., 'Discussion of Considerations in Choosing Statistical Sampling Procedures in Auditing', Journal of Accounting Research, 1975 Supplement, p.58.

(65) Kaplan R.S., 'Statistical Sampling in Auditing with Auxiliary Information Estimators', loc. cit., p.257.

(66) Neter J. and Loebbecke J.K., op. cit., p.87.

(67) Kinney W.R., op. cit. p.411.

(68) Ijiri Y. and Kaplan R.S., 'The Auditor's Sampling Objective: Four or Two? A Reply'., Journal of Accounting Research, Autumn 1972, pp 415 - 416.

(69) Hubbard T.D. and Strawser R.H., 'A Test of a Model for Integrating Sampling Objectives in Auditing', Journal of Accounting Research, Autumn 1972, pp 404 - 406.

Chapter 5

(70) Mautz R.K. and Sharaf H.A., op. cit., p.86.

(71) Toba Y., 'A General Theory of Evidence as the Conceptual Foundation in Auditing Theory', The Accounting Review, January 1975, pp 7 - 24.

(72) Hatherly D.J., 'Compliance and Substantive Tests - Can the Auditor Tell the Difference?', The Accountant's Magazine, September 1978, pp 378 - 380

(73) Ibid p.379.

ADDENDUM TO THE THESIS

The meaning of the term 'substantive test'

All audit work is performed because the auditor considers that it contributes, directly or indirectly, towards the 'substantiation' of the accounts. The systems examination makes a direct contribution when it allows a reduction in other audit work designed to substantiate the accounts. Generally it is the substantive tests of details that are reduced. (In what are termed 'substantive tests of details' the auditor takes a sample of items making up an account figure and collects evidence to support these sample items). The systems examination makes an indirect contribution to the substantiation of the accounts when it permits, through roles 1 and 2 as described on pages 20 to 22 of the Introduction, greater understanding of the substantive testing of details. Role 1, through the selection of an appropriate statistical sampling technique, can allow greater understanding of sample size and sample selection requirements. Role 2 can allow greater understanding of the quality of evidence used to support sample items.

The 'systems approach' focuses on the ability of the systems examination to directly substantiate the accounts. Following this approach the systems examination itself may be regarded as 'substantive'. This allows substantive tests of details to be reduced and the study made in chapter 5 suggests that such tests should be independent of the systems examination if they are to contribute to audit assurance once the systems examination has been performed. The 'substantive approach' focuses on the ability of the systems examination to aid, or provide greater understanding of, the substantive testing of details. Following this approach the systems examination becomes a prerequisite for, and in a sense, therefore, a part of, the detailed substantive testing.

In general throughout this thesis the term 'substantive test' refers to substantive tests of details. However, as a result of the study made in the thesis it becomes clear that two rather different meanings of substantive tests of details are implied by the two different audit approaches. A systems approach implies substantive tests which are 'independent' of the systems examination. For example, reperformance by the auditor of the arithmetic of an invoice provides acceptable evidence of arithmetical accuracy whether or not the system is operating properly. It is, therefore, a test independent of the system. A substantive approach, however, implies substantive tests whose design is, at least in considerable part, a function of the results of the systems examination. For example, a sales invoice may be acceptable evidence of a sales item because the auditor approves of the system which produces that invoice.

Statistical sampling

The author has assumed for purposes of this thesis that the auditor faced with substantive testing of details will seek to use a statistical sampling technique in order to provide an indication, and hence greater understanding, of the effectiveness of the sample. At present some auditors prefer judgement sampling to any statistical technique. Given the present stage of development and testing of statistical techniques in audit sampling such a decision may be very reasonable. This thesis does not pursue the question of whether, at the present time, statistical or judgement sampling should be used. It does seem probable to this author, however, that more and more experimental testing of statistical techniques against known accounting populations will lead in due course to extensive knowledge of the performance of those techniques. Experiments of this nature have already been conducted and the results are referred to in this thesis. As such knowledge is gained, and provided there are techniques which perform well, it is likely that more auditors will either adopt statistical techniques formally or will seek to informally 'simulate' techniques when conducting judgement sampling. For example, if statistical techniques suggest that samples should be oriented toward high value and high error risk items, then the auditor will seek to do this in his judgement sampling. In either case there would be an important role for the systems examination in helping the auditor to identify probable high error risk items. The role of the systems examination as an aid to substantive testing of details is, therefore, likely to be important whether statistical or judgemental sampling is employed.